ELSEVIER

Applied Numerical Mathematics 22 (1996) 113–151

APPLIED NUMERICAL MATHEMATICS

Runge–Kutta methods: some historical notes

J.C. Butcher [a,*], G. Wanner [b,1]

[a] *Department of Mathematics, University of Auckland, Auckland, New Zealand*
[b] *Section de Mathématiques, Université de Genève, Genève, Switzerland*

Abstract

This centenary history of Runge–Kutta methods contains an appreciation of the early work of Runge, Heun, Kutta, and Nyström and a survey of some significant developments of these methods over the last hundred years. In particular, the order conditions, as they are now understood, will be outlined, as will the introduction and practical implementation of implicit Runge–Kutta methods, the use of linear and nonlinear stability analysis in the assessment of Runge–Kutta methods, and the theory and applications of the composition of methods. Of the many further developments that have arisen, most are discussed only briefly; however, the recent interest in Runge–Kutta methods for Hamiltonian systems will be surveyed in more detail.

Keywords: Runge–Kutta methods; History; Order conditions; Stability; Order reduction; Composition; Hamiltonian systems

1. Introduction

The year 1895 was important for many reasons. This was the year of the publication of Hurwitz's stability conditions, the death of Cayley, the 80th birthday of Weierstrass, the birth of Aitken. It was also the year in which the Swiss National Library was founded; Röntgen discovered X-rays; the first radio receiver was built; *La Bohème* was first performed. To us 1895 was especially important as the year that Runge's classic paper was published on the numerical solution of differential equations, which initiated the wealth of research of which the present paper gives only a brief overview.

In the early years of Runge–Kutta methods, and this we can take as up to the time of the paper by Nyström [97], practical applications of Runge–Kutta were confined to problems that could be performed by hand. With the advent of modern computers, however, Runge–Kutta methods took on a new significance. Problems that could be solved became larger and more intricate, automatic error monitoring and control of step size became not only appropriate but even necessary, and most important of all, stiff systems became recognised as a special type of problem class requiring special methods for

* Corresponding author. E-mail: butcher@mat.auckland.ac.nz.
[1] E-mail: gerhard.wanner@math.unige.ch.

0168-9274/96/$15.00 Copyright © 1996 Elsevier Science B.V. All rights reserved
PII S0168-9274(96)00048-7

their solution. Progress on Runge–Kutta methods was spread between theory and technique and the central goal, the production of efficient and reliable differential equation software, became established as the driving force for research efforts.

In addition to the use of Runge–Kutta methods in their traditional role of differential equation solvers, related types of initial value problems have been found to be amenable to solution by Runge–Kutta methods or by Runge–Kutta methods suitably adapted to the more general problem class. Examples of these wider classes of problems are PDE's, Volterra integral equations, delay differential equations, differential–algebraic equations and stochastic differential equations.

With this enormous range of activity to survey, it is possible to dwell in detail on only a limited range of topics. The choice is made not in any claim that these are the most important, or even that they are characteristic of Runge–Kutta related activities in general. They are, however, typical, and some at least are of central importance to the main flow of the subject. Some of them developed as a response to scientific needs while some of them anticipated the applications that would eventually make use of them. Some are the work of a few individuals while some are the work of many people working in widely scattered places. In short, research on Runge–Kutta methods is a vital international effort and because it involves the work of mathematicians, computer scientists, engineers, chemists, physicists and many more, it is also an interdisciplinary activity.

In Section 2 we discuss the early days of Runge–Kutta methods. The next topics we will consider will be work on Runge–Kutta methods from 1950 until 1960. This includes the work of S. Gill [61] and of R.A. Merson [94], work which was not only important in its own right, but which was prophetic in that it pointed the way to later developments. This work, which occupies Section 3, leads naturally to a review of the order of Runge–Kutta methods in Section 4. Implicit methods are introduced in Section 5 and this leads in Section 6 to a discussion of linear stability properties of Runge–Kutta methods, especially in the context of stiff problems. Nonlinear stability and the phenomenon of "order reduction" are the subject of Section 7. Section 8 deals with the composition of Runge–Kutta and the algebraic and analytic consequences of this type of operation. In Section 9 we discuss the use of Runge–Kutta methods in the solution of Hamiltonian problems. Finally we will conclude in Section 10 with at least a passing comment on a number of additional Runge–Kutta issues.

2. Runge–Kutta methods up to Nyström

Runge. Carl David Tolmé Runge was born in 1856 in Bremen, spent much of his childhood in Havana, began his studies in 1876 in Munich and then went to Berlin to work with two of the greatest mathematicians of the time, L. Kronecker and K. Weierstrass. Consequently, his first mathematical researches were in number theory and complex analysis. Later, influenced by his nomination as professor at the "Technische Hochschule" in Hanover (1886) and by his lifelong friendship with many physicists (in particular L. Prandtl, M. Planck, H. Kayser and F. Paschen), he developed his broad interests in spectroscopy and in applications of theoretical mathematics to physical and technical problems. Here, in Hanover, he wrote the paper [103], whose centenary we are celebrating and which appeared during 1895 in the *Mathematische Annalen.* In 1905, F. Klein advocated his nomination as Professor of "Applied Mathematics" (the first in Germany!) at the University of Göttingen, where he was extremely active through his teaching and his textbooks (in particular the famous Runge and König [106] from 1924) to promote numerical analysis in general and *the* Runge–Kutta method in

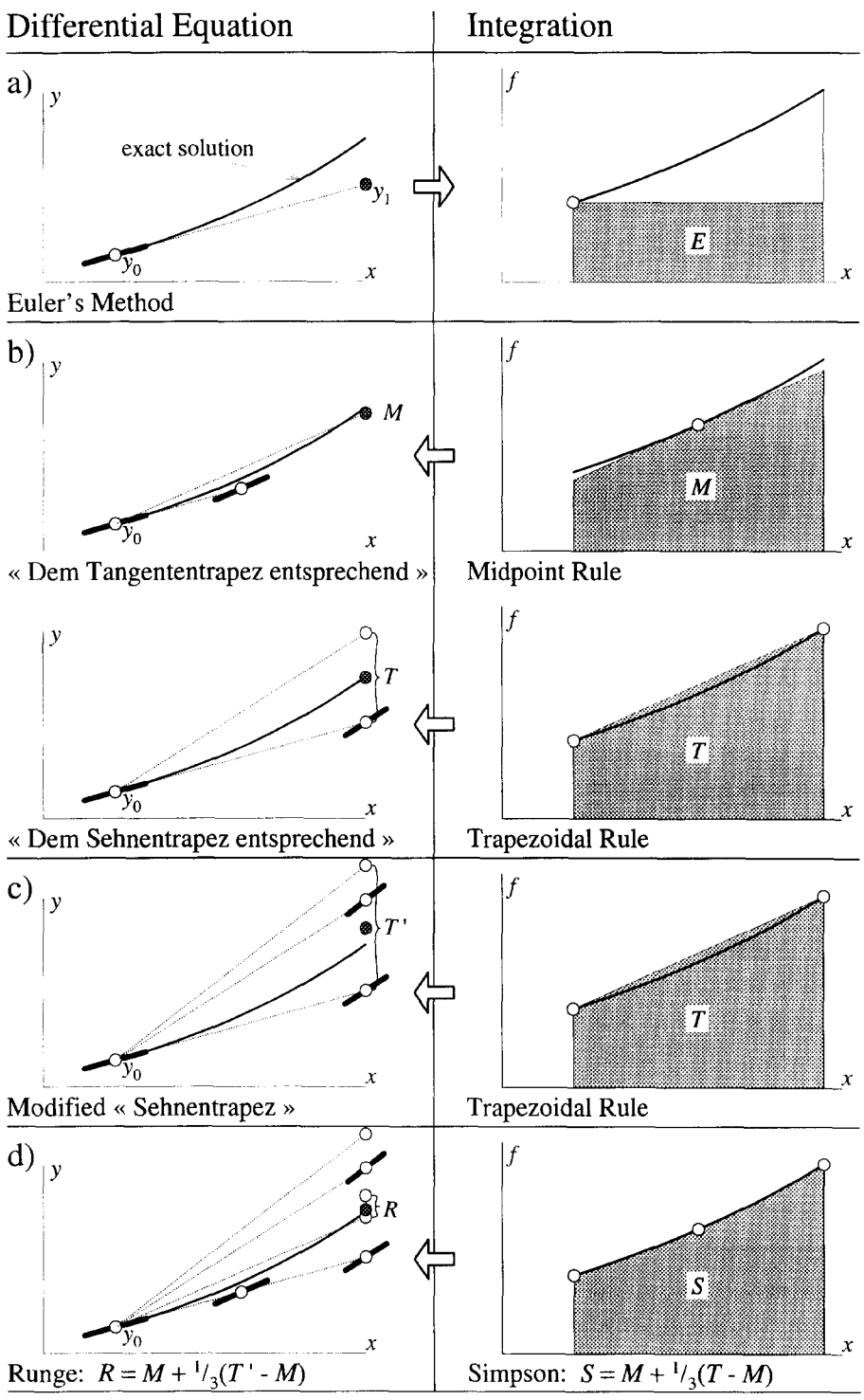


Fig. 1. Runge's motivation.

particular for all sorts of practical applications. For more references see *Encyklopädie der Math. Wiss.*, Vol. II, Teil 3, Hälfte 1 (II,3,1), pp. 148–150; for more details about his life and scientific work see the articles of L. Prandtl, R. Courant and F. Paschen in *Die Naturwissenschaften*, Vol. 15, Berlin, 1927, pp. 227–233.

The ideas of Runge's important discovery are sketched in Fig. 1. His main inspiration stems from the analogy between a differential equation problem

$$y' = f(x,y), \quad y(x_0) = y_0 \tag{1}$$

(left column) and, in the case when f is independent of y, the integration problem

$$y_1 = y_0 + \int_{x_0}^{x_0+h} f(x)\,\mathrm{d}x \tag{2}$$

(right column). He observed (see Fig. 1(a)) that Euler's method [54] gives rise to a rather inefficient approximation of the integral by the area of a rectangle of height $f(x_0)$. Thus, he says, "it is already much better" to extend the Midpoint rule and the Trapezoidal rule to differential equations (Fig. 1(b)) by inserting for the missing y-values the results of Euler steps

$$M := y_1 = y_0 + hf\bigl(x_0 + \tfrac{1}{2}h, y_0 + \tfrac{1}{2}hf(x_0,y_0)\bigr), \tag{3}$$

$$T := y_1 = y_0 + \tfrac{h}{2}\bigl(f(x_0,y_0) + f\bigl(x_0+h, y_0 + hf(x_0,y_0)\bigr)\bigr). \tag{4}$$

He then shows by Taylor expansion, in the same way as all his successors, that for both methods the local error is $O(h^3)$, i.e., the methods are of order 2. His great dream, however, was Simpson's rule with its 4th order accuracy. But the simple translation of the well-known formula $S = M + (T-M)/3$ led to a deception. Taylor series expansion revealed that this expression is only of order 2 if f depends on y. Runge then discovered that a slight modification of the Trapezoidal rule (Fig. 1(c)), by iterating the Euler step a second time, leads with the formula $R = M + (T' - M)/3$ to a third order method (Fig. 1(d)). He also extended his method together with its Taylor expansion to *systems* of equations.

Runge's paper would not have been written by Runge, if he had not concluded it with a practical example, completely worked out in all details and showing how the computations are best displayed on a sheet of paper (see Fig. 2). The equations describe (see also [69, Exercise II.1.2]) the form of a drop exposed to gravity and surface-tension and can be written in the form

$$\frac{\mathrm{d}z}{\mathrm{d}r} = \tan\phi, \quad \frac{\mathrm{d}(\sin\phi)}{\mathrm{d}r} = 2z - \frac{\sin\phi}{r}, \qquad z(0) = -1, \ \sin\phi(0) = 0. \tag{5}$$

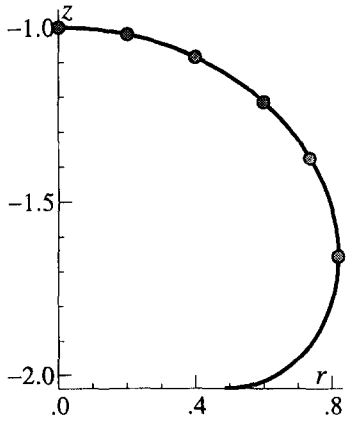
Fig. 3 shows the exact solution together with the points computed by Runge, which can be seen to be extremely accurate.

Heun. Heun (1900) [76] criticized Runge's paper for the curiously inductive way ("auf einem eigentümlichen induktiven Wege") of obtaining the method and argued that the "more general" Gaussian way of thinking should be followed. Thus, a general Gaussian quadrature formula

$$y_1 = y_0 + h\sum_{i=1}^{s} b_i f(x_0 + c_i h), \tag{6}$$

Dem Tangententrapez entsprechend					Dem Sehnentrapez entsprechend				
r	z	$\sin\varphi$	$\tan\varphi$	$2z-\frac{\sin\varphi}{r}$	r	z	$\sin\varphi$	$\tan\varphi$	$2z-\frac{\sin\varphi}{r}$
0	1	0	0	1	0	1	0	0	1
0.1	0	0.1			0.2	0 ←	0.2 ←		
0.1	1	0.1	0.1005	1	0.2	1	0.2	0.2041	1
0.2	0.0201	0.2				0.0408	0.2		
0.2	1.0201	0.2	Tangententrapez			1.0408	0.2	0.2041	1.0816
	1.0204	0.2082	Sehnentrapez						
Differenz	+0.0003	+0.0082				0.0408 ←	0.2163 ←		
$\frac{1}{3}$ Differenz	+0.0001	+0.0027				0.0408	0.4163		
						0.0204	0.2082		
0.2	1.0202	0.2027	0.2070	1.027	0.2	1.0202	0.2027	0.2070	1.027
0.1	0.0207	0.1027			0.2	0.0414 ←	0.2054 ←		
0.3	1.0409	0.3054	0.3207	1.064	0.4	1.0616	0.4081	0.4470	1.103
0.2	0.0641	0.2128				0.0894	0.2206		
0.4	1.0843	0.4155				1.1096	0.4233	0.4672	1.161
	1.0876	0.4215							
	33	60				0.0934 ←	0.2322 ←		
	11	20				0.1348	0.4376		
						0.0674	0.2188		
0.4	1.0854	0.4175							

Fig. 2. Runge's calculation of a drop.



$r_1 = 0.0000 \quad z_1 = -1.0000$
$r_2 = 0.2000 \quad z_2 = -1.0202$
$r_3 = 0.4000 \quad z_3 = -1.0854$
$r_4 = 0.6000 \quad z_4 = -1.2145$
$r_5 = 0.7348 \quad z_5 = -1.3745$
$r_6 = 0.8180 \quad z_6 = -1.6568$

Fig. 3. Runge's numerical results with exact solution.

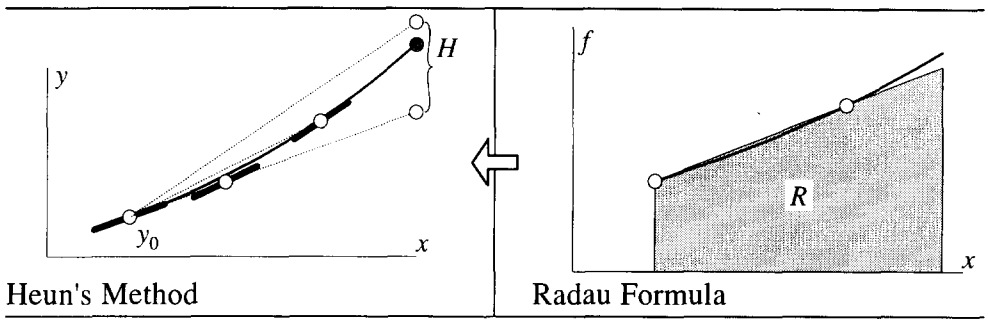


Fig. 4. Heun's method.

could be extended as follows:

$$\begin{aligned} k_1 &= f(x_0, y_0), \\ k_2 &= f(x_0 + c_2 h, y_0 + c_2 h k_1), \\ k_3 &= f(x_0 + c_3 h, y_0 + c_3 h k_2), \\ &\vdots \\ y_1 &= y_0 + h \sum_{i=1}^{s} b_i k_i, \end{aligned} \tag{7}$$

and Taylor expansion would lead to algebraic conditions for the free coefficients in order to make the order as high as possible. Heun obtained in this way formulas of order up to 4 (for $s = 8$), his most famous method is the following of order 3

$$c_2 = 1/3, \qquad c_3 = 2/3, \qquad b_1 = 1/4, \qquad b_2 = 0, \qquad b_3 = 3/4$$

(see Fig. 4). The full accomplishment of Heun's dream, the complete generalization of Gaussian quadrature to differential equations, had to wait for another 64 years.

Kutta. Wilhelm Martin Kutta was born 1867 in Upper Silesia and studied mathematics in Breslau. Details about his life and work (especially his later career in aerodynamics) can be found in the note [12] by R. Bulirsch and M. Breitner. From 1894 until 1909 he was assistant and lecturer at the "Technische Hochschule" in Munich, where the "Aufsatz von Herrn Runge" attracted his attention and the "Abhandlung von Herrn Heun" encouraged him to publish in [88] *his* results. Why not, was his reasoning, allow *all* evaluated derivatives to enter into the calculation of the new evaluation point, that is replace Heun's formulas by

$$\begin{aligned} k_1 &= f(x_0, y_0), \\ k_2 &= f(x_0 + c_2 h, y_0 + a_{21} h k_1), \\ k_3 &= f(x_0 + c_3 h, y_0 + a_{31} h k_1 + a_{32} h k_2), \\ k_4 &= f(x_0 + c_3 h, y_0 + a_{41} h k_1 + a_{42} h k_2 + a_{43} h k_3), \\ &\vdots \\ y_1 &= y_0 + h \sum_{i=1}^{s} b_i k_i. \end{aligned} \tag{8}$$

Der Ansatz ist demnach der folgende: Man stellt auf:

$$\begin{aligned}
\Delta' &= f(x, y)\Delta x,\\
\Delta'' &= f(x + \varkappa\Delta x,\ y + \varkappa\Delta')\Delta x,\\
\Delta''' &= f(x + \lambda\Delta x;\ y + \varrho\Delta'' + (\lambda - \varrho)\Delta')\Delta x,\\
\Delta'''' &= f(x + \mu\Delta x;\ y + \sigma\Delta''' + \tau\Delta'' + (\mu - \sigma - \tau)\Delta')\Delta x,\\
\Delta^{\mathrm{V}} &= f(x + \nu\Delta x,\ y + \varphi\Delta'''' + \chi\Delta''' + \psi\Delta'' + (\nu - \varphi - \chi - \psi)\Delta')\Delta x,\\
&\vdots \qquad\qquad\qquad\qquad \vdots
\end{aligned}$$

und setzt als gewünschte Näherung an:

$$\Delta y = a\Delta' + b\Delta'' + c\Delta''' + d\Delta'''' + e\Delta^{\mathrm{V}} + \cdots.$$

Dabei sind die Gröfsen $\varkappa, \lambda, \mu, \nu \cdots$; $\varrho, \sigma, \tau, \varphi, \chi, \psi \cdots$; $a, b, c, d, e \cdots\cdots$ beliebig verfügbare Zahlenkoeffizienten.

Fig. 5. Kutta's "Ansatz" in the paper [88].

$$\begin{aligned}
\Delta y &= \frac{\Delta' + 3\Delta'' + 3\Delta''' + \Delta''''}{8},\\
\Delta' &= f(x, y)\Delta x,\\
\Delta'' &= f\left(x + \frac{\Delta x}{3},\ y + \tfrac{1}{3}\Delta'\right)\Delta x,\\
\Delta''' &= f\left(x + \tfrac{2}{3}\Delta x,\ y + \frac{3\Delta'' - \Delta'}{3}\right)\Delta x,\\
\Delta'''' &= f(x + \Delta x,\ y + \Delta''' - \Delta'' + \Delta')\Delta x.
\end{aligned}
\qquad
\begin{aligned}
\Delta y &= \frac{\Delta' + 2\Delta'' + 2\Delta''' + \Delta''''}{6},\\
\Delta' &= f(x, y)\Delta x,\\
\Delta'' &= f\left(x + \frac{\Delta x}{2},\ y + \frac{\Delta'}{2}\right)\Delta x,\\
\Delta''' &= f\left(x + \frac{\Delta x}{2},\ y + \frac{\Delta''}{2}\right)\Delta x,\\
\Delta'''' &= f(x + \Delta x,\ y + \Delta''')\Delta x.
\end{aligned}$$

Fig. 6. Kutta's famous methods in [88].

This would allow more freedom in satisfying the required order conditions. See Fig. 5 for a reproduction of Kutta's original "Ansatz". Kutta states, without mentioning their derivation, the order conditions up to order 5 (—for first order equations; systems he treated six years later in [52, p. 92]; see also *Encyklopädie der Math. Wiss.*, Vol. VI,2,24, p. 412) and gives the complete solution for methods up to order 4 ("nach einiger Rechnung"), and some methods of order 5. The following two 4th order methods, which are his most famous results, are shown in Fig. 6. They are usually presented today just by giving the coefficients in a tableau as follows:

$$\begin{array}{c|cccc}
0 & & & & \\
\frac{1}{2} & \frac{1}{2} & & & \\
\frac{1}{2} & 0 & \frac{1}{2} & & \\
1 & 0 & 0 & 1 & \\
\hline
 & \frac{1}{6} & \frac{2}{6} & \frac{2}{6} & \frac{1}{6}
\end{array},
\qquad
\begin{array}{c|cccc}
0 & & & & \\
\frac{1}{3} & \frac{1}{3} & & & \\
\frac{2}{3} & -\frac{1}{3} & 1 & & \\
1 & 1 & -1 & 1 & \\
\hline
 & \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8}
\end{array}. \tag{9}$$

KUTTA gibt noch zwei Formelsysteme an, die sechs Funktionswerte benutzen und eine Approximation fünfter Ordnung ergeben. Aus irgendeinem Grunde ist das zweite von ihnen fehlerhaft; es soll so lauten (wenn wir t, x statt x, y schreiben):

$$\begin{cases} \Delta' x &= f(t,x)\,\Delta t, \\ \Delta'' x &= f\Big(t+\frac{\Delta t}{3},\ x+\frac{\Delta' x}{3}\Big)\Delta t, \\ \Delta''' x &= f\Big(t+\frac{2\,\Delta t}{5}, x+\frac{6\,\Delta'' x+4\,\Delta' x}{25}\Big)\Delta t, \\ \Delta'''' x &= f\Big(t+\Delta t,\ x+\frac{15\,\Delta''' x-12\,\Delta'' x+\Delta' x}{4}\Big)\Delta t, \\ \Delta^{\mathrm{V}} x &= f\Big(t+\frac{2\,\Delta t}{3}, x+\frac{8\,\Delta'''' x-50\,\Delta''' x+90\,\Delta'' x+6\,\Delta' x}{81}\Big)\Delta t, \\ \Delta^{\mathrm{VI}} x &= f\Big(t+\frac{4\,\Delta t}{5}, x+\frac{8\,\Delta'''' x+10\,\Delta''' x+36\,\Delta'' x+6\,\Delta' x}{75}\Big)\Delta t, \\ \Delta x &= \frac{23\,\Delta' x+125\,\Delta''' x-81\,\Delta^{\mathrm{V}} x+125\,\Delta^{\mathrm{VI}} x}{192}. \end{cases}$$

Fig. 7. Nyström's presentation of a 5th order formula [97].

Kutta himself favoured the "three-eighths rule" (the formula on the right), which he found, by evaluating the error constants, to be more precise. Furthermore, as he said, the other formula was inside the framework of Heun's methods, and had just been overlooked by this author. In spite of this stated preference, Runge devoted much of his life to promoting the formula on the left. Surely for its computational simplicity, which Runge judged to be very important ("dieser Vorteil bedeutet jedoch so wenig gegenüber der bequemen Gestalt ..." [106, p. 294]), and perhaps also because it so nicely generalized Simpson's Rule, with which Runge had been struggling for many years.

Nyström. The comprehensive article [97], which E.J. Nyström published in 1925 in the Helsinki *Acta*, described the *state of the art* of the period in numerical methods, one-step as well as multistep for ordinary differential equations. Nyström developed Runge–Kutta theory thoroughly and corrected a fifth-order method of Kutta (see Fig. 7). He then remarked that a *second order* differential system

$$y'' = f(x, y, y'), \tag{10}$$

which theoretically could be transformed into a first order system of double dimension (a procedure used by Runge and Kutta for all their computations), can be integrated more efficiently by special methods, nowadays known as Nyström methods, which take into account the special structure of the equation. Nyström's claim, that his methods were about 25% more efficient, is, however, exaggerated (see [69, Section II.14]), and only becomes true if the differential system (10) contains no first-derivative terms y'.

Rigorous error estimates. Rigorous error estimations and convergence proofs were first given by Cauchy [26] and by G. Coriolis [30]. Coriolis had already discovered Runge's second formula nearly 60 years earlier (see Fig. 8), and gave an estimation of the form

$$e \leqslant Ch^2$$

Examinons maintenant le cas où l'on emploie d'autres équations aux différences pour calculer les valeurs de y. On peut, par exemple, procéder d'une manière analogue à celle qu'on prend pour les intégrales définies, quand on leur substitue l'aire d'un polygone au lieu de la somme des rectangles inscrits. On emploie alors l'équation aux différences

$$\Delta y = [f(x,y) + f(x, y + f(x,y)\Delta x)]\frac{\Delta x}{2}.$$

Fig. 8. Runge's second order formula in Coriolis [30].

on aura pour l'erreur totale sur y_n

$$\delta y_n < \frac{\Delta x^2}{4}\left[\frac{A''}{8} + (P + AQ)Q\right]\left[\frac{\left(1 + Q\Delta x + Q^2\frac{\Delta x^2}{2}\right)^n - 1}{Q + Q^2\frac{\Delta x}{2}}\right];$$

Fig. 9. Rigorous global error estimation in Coriolis [30].

for the global error; that is, he showed that it converged with second order of $h = \Delta x$ (see Fig. 9). Runge [104] then developed ideas for rigorously estimating the error of Kutta's fourth order formula, by suggesting how the Taylor expansions of the true and the numerical solution must be truncated and replaced by error terms of Lagrange type. But the complete elaboration of this program, with an explicit estimation of the type

$$\|e\| \leqslant Ch^4,$$

with explicit formulas for the constant in terms of partial derivatives of f of various orders, has been accomplished only in 1923 by Bieberbach for first order equations, and, after many pages of tedious calculations, in [9] for systems of equations. Even the *statement* of the theorem contains nearly two pages of formulas.

3. Runge–Kutta methods, 1950 to 1960

Three important publications within the years 1950 to 1960 will form the subject matter of this section. The paper by S. Gill [61] is important for several reasons. The first is that it pointed the way to a full analysis of the order conditions by considering the Taylor expansion of the numerical as well as the exact solution, for an autonomous family

$$y'(x) = f(y(x)), \quad y(x_0) = y_0. \tag{11}$$

Even though the expansions applied only to four stage methods and went only as far as fifth powers of the step size h, the analysis was presented in a form that hinted at the general forms of the expansions. Gill's specific aim was a practical one: choose the coefficients of a four stage Runge–Kutta method

so that, not only is the order 4, but the total memory requirement is, for large problems, equal to only four registers per variable. Using the classification of fourth order methods due to Kutta, Gill explored all the cases (except the general case for arbitrary c_2 and c_3) and found that only for one of them could this be achieved. This family is given by

$$\begin{array}{c|cccc} 0 & & & & \\ \frac{1}{2} & \frac{1}{2} & & & \\ \frac{1}{2} & \frac{3\theta}{2+6\theta} & \frac{1}{2+6\theta} & & \\ 1 & 0 & -3\theta & 1+3\theta & \\ \hline & \frac{1}{6} & \frac{1}{3}-\theta & \frac{1}{3}+\theta & \frac{1}{6} \end{array}.$$

The classical method ($\theta = 0$) requires 4 storage registers per variable because at the point in the computation when the third stage derivative is about to be computed, the linear combinations of the initial value for the step and of the first two stage derivatives that are required for (i) the computation of the third stage derivative, (ii) the computation of the second stage derivative and (iii) the final result for the step, are independent, because the matrix

$$\begin{bmatrix} 1 & a_{31} & a_{32} \\ 1 & a_{41} & a_{42} \\ 1 & b_1 & b_2 \end{bmatrix}$$

has non-zero determinant unless $\theta = \pm\sqrt{2}/6$. By analysing the error coefficients in the powers of h^5, Gill concluded that more accuracy should be expected for the choice $\theta = \sqrt{2}/6$, and this is what he proposed.

Gill's paper discusses how the computation in a step should be organized, not only to achieve the advantages of reduced memory requirements, but also to keep the growth of round-off error as low as possible.

The paper by Merson is also motivated by a practical requirement, that of providing along with a computed solution, a usable estimate of the truncation error committed within each step. He proposed a five stage method for which, at least for certain special linear problems, a comparison between quantities computed in the step does provide the required estimate. The estimate itself can be unreliable for many problems but the same has to be said of many more modern methods. Merson's method has been used in many production codes for many years and is still successfully used today.

The mathematical approach of Merson carried Gill's ideas further by attempting a detailed analysis of the Taylor expansions for the exact and approximate solutions. For the exact solution the formula uses elementary differentials, as they are now known. While the same is also true for the approximate solution, Merson does not seem to have been aware of the simple structures that can be used to express this.

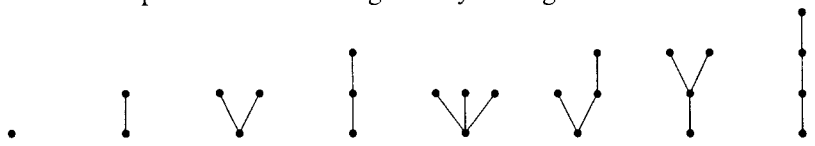
Within the 10 year period we are considering, the famous book by P. Henrici [75] appeared. This attempted, and to a large measure achieved, the elevation of numerical methods for initial value problems from a vague and intuitive collection of methods and techniques, to a systematic mathematical science. The fact that rigorous mathematics is needed to advance this subject is now clearly recognised

and the point hardly needs making today. However, without Henrici's book, the situation may have been quite different. It is not only in the standard of its mathematical exposition that Henrici's book is ahead of its time. If it had appeared even ten years later, the important discoveries made in the intervening period would have had the good fortune to have been expounded in Henrici's masterly fashion.

4. The order conditions

The order conditions as we will describe them here are based on the work of Gill, Merson and one of the present authors (JCB). The first exposition of these conditions, in anything like the form that is commonly used today, was in Butcher [16].

Let T denote the set of "rooted trees" (which we will refer to simply as "trees" in this paper). Thus the trees with up to 4 vertices are given by the figures



where we use the convention of placing the root at the lowest point of each tree. In considering the numerical solution of an autonomous differential equation system

$$y'(x) = f\big(y(x)\big),$$

it is convenient to calculate formulas for the higher derivatives using repeated differentiation and the chain rule. For example, up to third derivatives we find that

$$y''(x) = f'\big(y(x)\big)\big(f\big(y(x)\big)\big), \tag{12}$$

$$y^{(3)}(x) = f''\big(y(x)\big)\big(f\big(y(x)\big), f\big(y(x)\big)\big) + f'\big(y(x)\big)\big(f'\big(y(x)\big)\big(f\big(y(x)\big)\big)\big). \tag{13}$$

For compactness of notation we will write $\boldsymbol{f} = f(y(x))$, $\boldsymbol{f}' = f'(y(x))$, $\boldsymbol{f}'' = f''(y(x))$, ... With this convention we have

$$y'(x) = \boldsymbol{f}, \tag{14}$$

$$y''(x) = \boldsymbol{f}'\boldsymbol{f}, \tag{15}$$

$$y^{(3)}(x) = \boldsymbol{f}''(\boldsymbol{f},\boldsymbol{f}) + \boldsymbol{f}'\boldsymbol{f}'\boldsymbol{f}, \tag{16}$$

$$y^{(4)}(x) = \boldsymbol{f}^{(3)}(\boldsymbol{f},\boldsymbol{f},\boldsymbol{f}) + 3\boldsymbol{f}''(\boldsymbol{f},\boldsymbol{f}'\boldsymbol{f}) + \boldsymbol{f}'\boldsymbol{f}''(\boldsymbol{f},\boldsymbol{f}) + \boldsymbol{f}'\boldsymbol{f}'\boldsymbol{f}'\boldsymbol{f}. \tag{17}$$

It can be seen that the terms occurring in these expressions correspond exactly to the trees of 1, 2, 3 and 4 vertices. To find the term corresponding to a given $t \in T$, simply associate with each vertex the symbol $f^{(k)}$ where k is the (upward) degree for this vertex. The k-linear operator $f^{(k)}$ takes for its arguments the quantities associated with its k successors.

This can be seen from the following diagrams

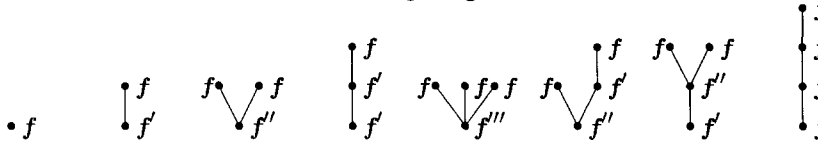


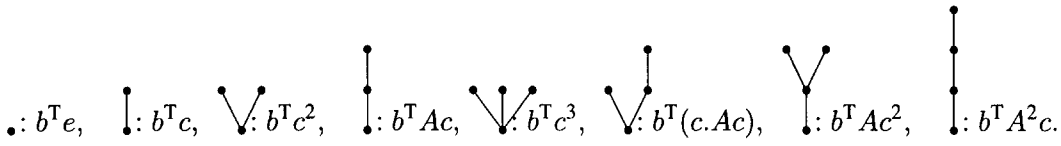
Table 1
Some functions on rooted trees

t	•							
$r(t)$	1	2	3	3	4	4	4	4
$\sigma(t)$	1	1	2	1	6	1	2	1
$\gamma(t)$	1	2	3	6	4	8	12	24

The expressions $\boldsymbol{f}$, $\boldsymbol{f}'\boldsymbol{f}$, ... are known as "elementary differentials" and, by evaluating each of them at the beginning of each time-step, the Taylor series for the exact solution to (11) can be written down. Not only is the exact solution expressible in terms of elementary differentials, but the same is true for the result computed by a Runge–Kutta method.

For each tree t define an expression in terms of the coefficients of the method by associating with each edge of the tree the coefficient matrix A, and associating with each vertex the componentwise product of all upward growing vectors, with the convention that an empty product is the vector e with each component equal to 1. Finally associate with the root of the tree the operation of forming an inner-product with the vector b^{T}.

Thus we have the expressions associated with the first 8 trees, where we have contracted the product $Ae = c$ as follows:



These expressions are known as "elementary weights" and will be denoted here by $\Phi(t)$. Similarly, the elementary weight associated with t will be denoted by $F(t)$, and its value when the argument of each f is replaced with η, by $F(t)(\eta)$.

With this notation we have the following expressions for the formal Taylor expansions of the exact and computed approximations in a step of size h starting from x_0

$$y(x_0) + \sum_{t\in T} \frac{h^{r(t)}}{\sigma(t)\gamma(t)} F(t)\big(y(x_0)\big) \tag{18}$$

and

$$y(x_0) + \sum_{t\in T} \frac{\Phi(t)h^{r(t)}}{\sigma(t)} F(t)\big(y(x_0)\big), \tag{19}$$

respectively. In these expansions, $r(t)$ denotes the "order", that is the number of vertices, of t, $\sigma(t)$ denotes the number of symmetries of t, and $\gamma(t)$, the "density" of t is found by multiplying the order of the subtrees rooted at each vertex of t. To illustrate these quantities, refer to Table 1.

Comparing (18) and (19), we see that these two expansions agree to within $O(h^{p+1})$ if

$$\Phi(t) = \frac{1}{\gamma(t)}, \quad \text{for } r(t) \leqslant p.$$

Note that this condition can be derived, at least for low orders, by considering a single non-autonomous problem

$$y'(x) = f\big(x, y(x)\big)$$

and evaluating the Taylor expansions for the exact and approximate solutions to this problem. However, above order 4, terms of this expansion get confused and insufficient order conditions are found for orders 5 or higher. For the derivation of sixth order methods using this approach, see Huťa [79].

Alternative approaches to order conditions are those of Hairer and Wanner [73] using B-series (see remark at the end of Section 8) as well as that of Albrecht [2], which yields equivalent conditions to those we have given, but expressed in a slightly different form.

We now discuss the possible solution of the algebraic equations for order p for a Runge–Kutta method with s stages. It can be shown that for a solution to exist, it is necessary that $s \geqslant p$ but this is sufficient only for $p \leqslant 4$. In fact for $p > 4$ the minimal number value of $s - p$ increases steadily. If $p = 5, 6$, the minimal value of $s - p$ is 1; if $p = 7$, $s - p$ must be at least 2 and for $p \geqslant 8$, $s - p \geqslant 3$, with equality in the case $p = 8$. Proofs of these "barrier" results have been given in [20,23].

That any order p can be achieved for s sufficiently high, can be shown in various ways. A result of this type is given in Henrici [75] and sharper results follow from a consideration of extrapolation as in the Gragg, Bulirsch and Stoer algorithm [13] and from a consideration of implicit Runge–Kutta methods implemented with a finite number of iterations [120].

Table 2, quoted from [25] and expanded, shows the chronology of attempts to find methods of increasingly high orders with close to the minimal number of stages.

Table 2
Successive derivations of high order Runge–Kutta methods

p	s	Author	Year		Reference
2	2	Coriolis	1837	(Trapezoidal rule method)	[30]
2	2	Runge	1895	(rediscovery of Trapezoidal rule method)	[103]
2	2	Runge	1895	(Midpoint rule method)	[103]
3	4	Runge	1895	(method of Fig. 1)	[103]
3	3	Heun	1900	(method of Fig. 4)	[76]
4	8	Heun	1900		[76]
4	4	Kutta	1901	(methods of Fig. 6)	[88]
5	6	Kutta	1901		[88]
5	6	Nyström	1925	(correction to a method of Kutta)	[97]
6	8	Huťa	1956		[79]
6	7	Butcher	1964		[17]
7	9	Butcher		(known since approximately 1968)	[24]
8	11	Curtis	1970		[37]
8	11	Cooper and Verner	1972	(announced 1969 in J.H. Verner Thesis)	[29]
10	18	Curtis	1975		[38]
10	17	Hairer	1978		[64]

Local error estimation and codes. In terms of the practical application of explicit Runge–Kutta methods for the routine solutions of non-stiff problems, considerable effort has been expended on the design of robust software based on reliable and efficient methods. The first successful codes in this tradition, notably those in the NAG library, were based on Merson's method [94]. At the end of the 1960's, embedded methods due to Fehlberg [55] and England [53] came into use, particularly through implementations by Shampine and Watts (see [114, p. 461]). More recently, Runge–Kutta pairs were constructed by Verner [120] and by Dormand and Prince [49]. The Dormand and Prince pairs optimize the *higher order* formula and thus return to the user the *best* possible result. They make use of the FSAL idea (see [69, 2nd ed., p. 167; 114, p. 370f]). Often-used codes implementing these methods are, e.g., in the RKSUITE of Shampine or the codes DOPRI5 and DOP853 in [69].

Dense output. An important ingredient of a numerical method is a *dense output* or *continuous output* capability. This provides not only approximations at grid points, but also at intermediate points. The first continuous methods appear to have been constructed by Gear [60], for the purpose of starting multistep methods, and by [77]. The modern approach is to use dense output for Runge–Kutta methods in their own right. It is important to avoid interrupting the optimal step size process by too many output points (see Shampine, Watts and Davenport [115]), for the graphical representation of solutions, for event location, Poincaré sections and for the treatment of delay differential equations (see, for example, [69, 2nd ed., Sections II.6 and II.17]).

5. Implicit methods

Implicit Runge–Kutta methods are, today, important tools for the treatment of stiff differential equations and of Hamiltonian systems. The first inventors, however, had totally different motivations for their study. We will discuss these motivations under the three headings (i) Integration, (ii) Collocation and (iii) Simplifying order conditions.

Integration. Cauchy [26], in his 1824 lectures at the *Ecole Polytechnique*, tried to establish rigorous error bounds for Euler's method. One of his ideas was to apply the mean value theorem to the integral

$$y(x_0+h) = y_0 + \int_{x_0}^{x_0+h} f\big(\xi, y(\xi)\big)\,\mathrm{d}\xi, \tag{20}$$

which becomes

$$y_1 = y_0 + hf\big(x_0+\theta h, y_0+\Theta(y_1-y_0)\big) \tag{21}$$

and now represents an implicit equation for the determination of y_1. The values of the θ's (between 0 and 1) are unknown for the exact solution. If we set them arbitrarily, we obtain an equation for the value y_1 which, eventually, majorizes, minorizes, or approximates the true solution. For example we have

$$\begin{aligned}
&\theta=\Theta=0: && y_1 = y_0 + hf(x_0,y_0) && \text{(explicit Euler)},\\
&\theta=\Theta=\frac{1}{2}: && y_1 = y_0 + hf\left(x_0+\frac{h}{2}, \frac{y_0+y_1}{2}\right) && \text{(implicit midpoint)},\\
&\theta=\Theta=1: && y_1 = y_0 + hf(x_1,y_1) && \text{(implicit Euler)}.
\end{aligned} \tag{22}$$

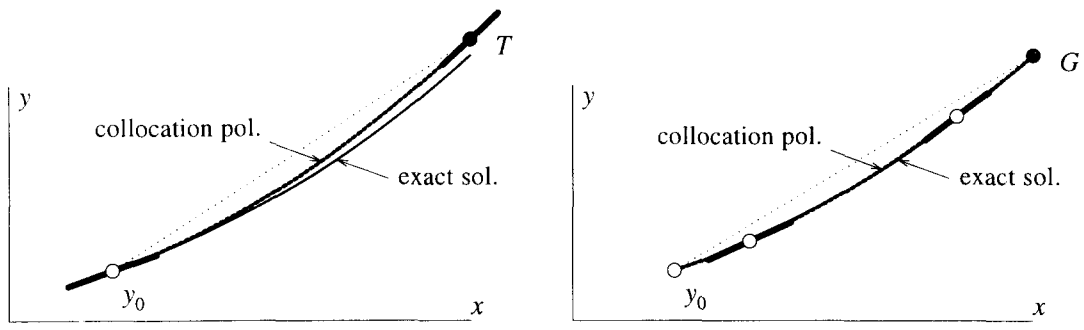


Fig. 10. Trapezoidal rule as collocation (left), collocation at Gaussian nodes (right).

An important class of integration methods was introduced by Adams [1], and became famous through the books of Moulton [95] and Willers (see [122, Section 32, 3rd ed., Section 35]). The simplest of Adams' implicit formulas is a one step method and corresponds to computing the integral in (20) by the *Trapezoidal rule*

$$y_1 = y_0 + \frac{h}{2}\big(f(x_0, y_0) + f(x_1, y_1)\big). \tag{23}$$

Collocation. The idea of this method was originally designed for boundary value problems (see, e.g., [28, Section 10.3]). For the step-by-step integration of ODE's, it was initiated by Hammer and Hollingsworth [74], who were not only interested in "the discrete values found" ("Seemingly by historical accident ..."), but searched also for the "functions pieced together" which represent the solution. They discovered that the trapezoidal rule (23) can be interpreted as generated by a quadratic function "which agrees in direction with that indicated by the differential equation at two points" x_0 and x_1 (see Fig. 10, left). This allows one to "see much-used methods in a new light" and to generalize them by choosing other collocation points, as for example the Gaussian nodes $x_0 + ph$ and $x_0 + qh$ where $p, q = 1/2 \mp 1/\sqrt{12}$ (see Fig. 10, right). It is also suggested that extensions of these results "to higher order integration methods is straightforward", but the way in which these methods are equivalent to implicit Runge–Kutta methods and preserve the full order of the quadrature formulas did not become became clear until some 15 years later (see [5; 63; 69, Section II.7; 123]).

Simplifying order conditions. The first explorers of Runge–Kutta methods realised quite soon that the number of order conditions grows rapidly for higher orders and that orders greater than, say, 6 (37 conditions) or 7 (85 conditions) appeared out of question. Huťa then, in his laborious paper on a 6th order method [79], found a solution of a problem with *more* equations than unknowns. This suggested the existence of some mysterious relations, which would permit a simplification of the order conditions.

After the clear understanding of the structure of the order conditions, Butcher [16] expressed these relations in the form

$$C(\eta): \quad \sum_{j=1}^{s} a_{ij} c_j^{q-1} = \frac{c_i^q}{q}, \quad i = 1, \ldots, s, \ q = 1, \ldots, \eta, \tag{24}$$

and

$$D(\zeta): \quad \sum_{i=1}^{s} b_i c_i^{q-1} a_{ij} = \frac{b_j}{q}\left(1 - c_j^q\right), \quad j = 1, \ldots, s, \; q = 1, \ldots, \zeta. \tag{25}$$

With the help of these conditions, Butcher went from surprise to surprise finding it "remarkable that the choice of 12 independent parameters ... has enabled us to satisfy no less than 17 separate equations", and he also announced "that this situation is capable of extensive generalization" and that "for any value of ν a process of order up to 2ν is possible." These were the *Gauss methods* [18], discovered independently also by Kuntzmann [86], the first of which is the implicit mid-point rule (22). The second, of order 4 (equivalent to the Hammer–Hollingsworth method which we have already mentioned), and the third one of order 6 are

$$\begin{array}{c|cc} \frac{1}{2}-\frac{\sqrt{3}}{6} & \frac{1}{4} & \frac{1}{4}-\frac{\sqrt{3}}{6} \\ \frac{1}{2}+\frac{\sqrt{3}}{6} & \frac{1}{4}+\frac{\sqrt{3}}{6} & \frac{1}{4} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array} \qquad \begin{array}{c|ccc} \frac{1}{2}-\frac{\sqrt{15}}{10} & \frac{5}{36} & \frac{2}{9}-\frac{\sqrt{15}}{15} & \frac{5}{36}-\frac{\sqrt{15}}{30} \\ \frac{1}{2} & \frac{5}{36}+\frac{\sqrt{15}}{24} & \frac{2}{9} & \frac{5}{36}-\frac{\sqrt{15}}{24} \\ \frac{1}{2}+\frac{\sqrt{15}}{10} & \frac{5}{36}+\frac{\sqrt{15}}{30} & \frac{2}{9}+\frac{\sqrt{15}}{15} & \frac{5}{36} \\ \hline & \frac{5}{18} & \frac{4}{9} & \frac{5}{18} \end{array} \tag{26}$$

Since such fully implicit methods appeared quite frightening at that time, Butcher, first in [16], then in full generality in [19], searched also for methods, based on Radau and Lobatto quadrature, where either the first and/or the last stage was explicit, as for example in the method

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 \\ 1 & 0 & 1 & 0 \\ \hline & \frac{1}{6} & \frac{4}{6} & \frac{1}{6} \end{array} . \tag{27}$$

which, being a beautiful extension of Simpson's rule, would certainly have delighted Carl Runge.

Diagonally implicit methods. Another attempt to "civilize" implicit Runge–Kutta methods, strongly advocated in the 70's by several independent articles [3,4,33,87,96], is to search for methods where only the diagonal elements of the matrix and those below them are allowed to be non-zero. Then each stage can be computed one after the other by solving *one* nonlinear system of dimension n each time. A nice formula of this type is the following method of order 4 of Crouzeix and Raviart [35]

$$\begin{array}{c|ccc} \gamma & \gamma & & \\ \frac{1}{2} & \frac{1}{2}-\gamma & \gamma & \\ 1-\gamma & 2\gamma & 1-4\gamma & \gamma \\ \hline & \delta & 1-2\delta & \delta \end{array} , \qquad \gamma = \frac{1}{\sqrt{3}}\cos\left(\frac{\pi}{18}\right) + \frac{1}{2}, \qquad \delta = \frac{1}{6(2\gamma-1)^2}. \tag{28}$$

Such methods are frequently used for the computation of huge systems of equations, often partially discretized partial differential equations.

Implementation costs. Even though implicit Runge–Kutta methods have much better stability properties than linear multistep methods, they have for many years not been as popular as the basis for practical codes, because of their high implementation costs. The reason for these high costs is that for an N dimensional problem there are sN nonlinear equations to solve in evaluating the stages in an s-stage method. If the solution is carried out using an appropriate variant of the Newton method, then the cost can be assessed in two parts (i) the cost of carrying out the LU factorization of the Jacobian matrix for the nonlinear equation scheme and (ii) the cost of the back-substitutions required in each iteration. Assuming that the Jacobian for the differential equation system can be treated as having a constant value J over many steps, and certainly over all stages and all iterations required in a step, then the iteration Jacobian can be written in the form

$$I_s \otimes I_N - hA \otimes J. \tag{29}$$

The cost of factorizing this matrix is approximately s^3N^3.

It is possible to reduce this cost by writing the coefficient matrix in factored form

$$A = T\bar{A}T^{-1},$$

where $\bar{A}$ is in Jordan canonical form. The inverse of (29) is then equal to

$$(T \otimes I_N)(I_s \otimes I_N - h\bar{A} \otimes J)^{-1}(T^{-1} \otimes I_N)$$

and the middle factor can be written in block diagonal form. The additional costs associated with the transformation factors, $(T \otimes I_N)$ and $(T^{-1} \otimes I_N)$ are each equal to approximately s^2N whereas the cost of finding the LU factors of $I_s \otimes I_N - h\bar{A} \otimes J$ is approximately mN^3, where m is the number of distinct eigenvalues of A (with a slightly modified operation count if some of the eigenvalues are complex).

Examples of the use of this approach are in singly-implicit methods [14], where $m = 1$ and in the code RADAU5 [72], where A has 1 real eigenvalue and a single complex conjugate pair of eigenvalues.

Runge–Kutta–Rosenbrock methods. If, say, the equation of a diagonally implicit Runge–Kutta method is solved by only one iteration of Newton's method, one obtains a Runge–Kutta-like expression containing in addition to values of f also the Jacobian J. Such formulas, initiated by Rosenbrock in 1963 [102], and favourably modified over many years, are now called *Rosenbrock* or *linearly implicit methods*. Two theses (by P. Kaps and A. Wolfbrandt) appeared in 1977 and were treating these methods in detail (see also [72, Section II.7]). Successful codes are the extrapolation code METAN1 by Bader and Deuflhard [6] as well as RODAS (see [72]), which is based on a "stiffly accurate" method. Excellent numerical results for RODAS have recently been reported for computations of atmospheric chemistry problems in [109].

6. Linear stability

PDE discretizations. Linear stability theory has its origin in a paper published one year after Runge's death, written by his son-in-law R. Courant together with K. Friedrichs and H. Lewy [31]. This

paper, studying the convergence of difference schemes for partial differential equations, arrived at the famous CFL-condition, which is a condition restricting the step size Δt for achieving stability and hence, convergence. The advantages of *implicit* methods (the trapezoidal rule and the implicit Euler method) for PDE discretizations were then discovered in the 40's independently by Crank and Nicolson [32] and Laasonen [89].

Milne's method. In the ODE literature, the subject was initiated independently by W.S. Loud [92], H. Rutishauser [108] and G. Dahlquist [40], who discovered the so-called *weak instability* of the explicit mid-point rule in its multistep version

$$y_{m+1} = y_{m-1} + 2hf(x_m, y_m). \tag{30}$$

("For this reason this method is not suitable for long-run automatic computation", quoted from [92, p. 47].) The noteworthy first meeting of Dahlquist and Rutishauser 1951 in a hotel room in Freiburg i. Br. is nicely narrated, with many additional references and more details on the stability history, in [44].

Chemical reactions. A third impulse came from numerical calculations of chemical reactions, in particular in 1952 from the paper by Curtiss and Hirschfelder [39], who discovered the BDF methods and their excellent stability properties and coined the expression *stiff* for problems whose widely different time scales resist solution by explicit integration methods. Another influential paper was [101], where the famous "Robertson reaction" was introduced, which helped generations of computing scientists to understand stiffness and to sharpen their software.

Runge–Kutta methods. With the rapid growth of computing experience and the "disappearance of the human inspection of almost every arithmetic result" (quoted from [44, p. 191]), (in)stability phenomena became significant also for Runge–Kutta methods. First studies are due to [92] (without calling it so) and [108]. Another early source is Merson's discussion in [94], where he defended RK methods against the very common opinion, that multistep methods are much more economical and "derive the maximum amount of information from the number of function values calculated". Merson mentioned that "in answer to Dr. Bennett's first question, I have found that the predictor-corrector formulae do not have the wide stability range of R. Kutta processes".

Stability analysis. This was done in [31] by Fourier mode analysis, and in [40,108] by linearization followed (in higher dimensions) by diagonalization of the Jacobian $\partial f/\partial y$. Loud [92] started from the very first with the constant-coefficient case by saying that "if a numerical method is to be of value in solving general differential equations, it should be extremely reliable with the simplest types ...". Then appears a so-called *characteristic equation* containing the eigenvalues λ of $\partial f/\partial y$, whose roots must be smaller than (or equal to) one in modulus, since otherwise "high powers of it may well become large" [92, p. 47]. For one-step methods, this equation is of degree 1 and leads to a rational or polynomial function $R(z)$ where $z = h\lambda$. For stability, $z = h\lambda$ must be in the so-called *stability domain*

$$S := \{z;\ |R(z)| \leqslant 1\}. \tag{31}$$

Such stability domains "did not become a hit until the beginning of the 60's" [44, p. 197], and were, to our knowledge, first sketched by Guillou and Lago [62]. We present in Fig. 11 a drawing from [62] representing a so-called damped *Chebyshev method.* For more details on such methods see the article

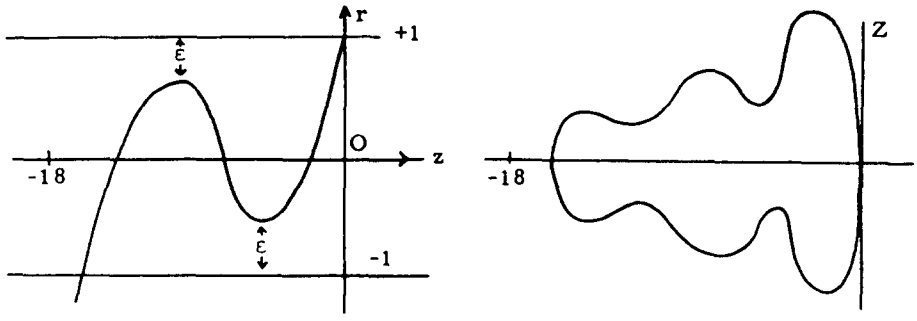


Fig. 11. Stability domains of Chebyshev methods by Guillou and Lago, 1961.

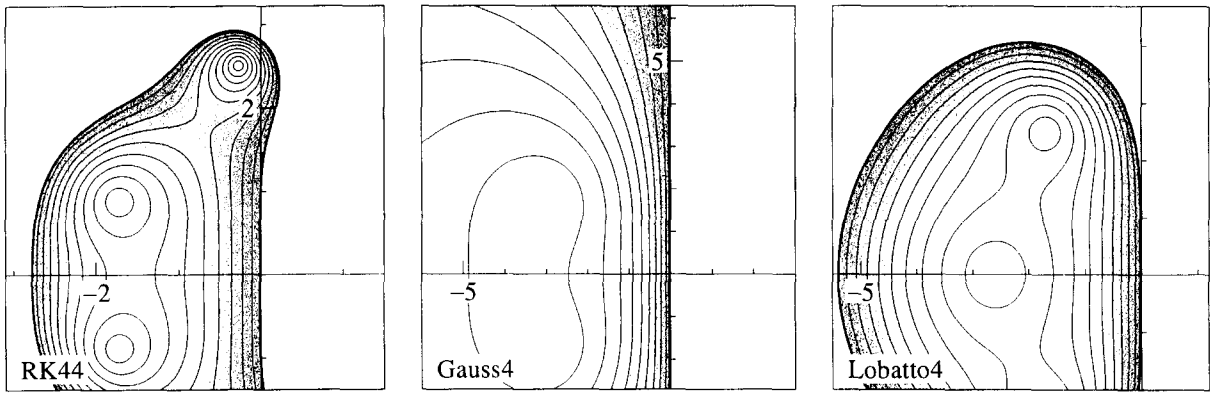


Fig. 12. Stability domains of three RK methods.

of Verwer in this issue. Other examples of stability domains are presented in Fig. 12. We choose the methods (9) (both of these methods have the same stability function), (26) (order 4) and (27), and obtain the stability functions

$$
\begin{aligned}
R(z) &= 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \frac{z^4}{24} && \text{(method RK44 (9))},\\
R(z) &= \frac{1 + z/2 + z^2/12}{1 - z/2 + z^2/12} && \text{(method Gauss4 (26))},\\
R(z) &= \frac{1 + 3z/4 + z^2/4 + z^3/24}{1 - z/4} && \text{(method Lobatto4 (27))}.
\end{aligned}
\tag{32}
$$

It can be observed in Fig. 12 that the stability domain of the 4th order Runge–Kutta methods extends, on the real axis, between $-2.7853 \leqslant h\lambda \leqslant 0$. The 4th order Gauss method of Hammer and Hollingsworth appears to be stable for *all* complex eigenvalues in the negative half-plane, i.e., it is *A-stable*, a notion introduced for multistep methods in the famous paper of Dahlquist [41]. Hammer and Hollingsworth have somehow foreseen this property: "... and we believe it is also the most stable of the three methods". The implicit Lobatto–Simpson formula, however, is only stable for $-\sqrt[3]{40}-2\ (\approx -5.42) \leqslant h\lambda \leqslant 0$.

A-stable Runge–Kutta methods. This subject took off around 1969, apparently influenced by talks of T.E. Hull and C.W. Gear at the IFIP Congress, Edinburgh 1968, and by ideas of J.D. Lawson. The major works were the elegant paper of Axelsson [5] and the impressive thesis of Ehle [51], which was followed by the publication [50]. Firstly, Ehle discovered that Butcher's Gauss-methods had the diagonal Padé approximations $R_{ss}(z)$ to e^z as stability functions. For these it was known that they had the A-stability property [10]. The difficult part of this result, proved by continued fraction methods, was the fact that the denominator had no roots in $\mathcal{C}^-$. Therefore, all Gauss methods (26) are A-stable. Next, as we will see below, Ehle modified Butcher's Radau and Lobatto methods ("Methods of Butcher which are not A-stable") so that they had acceptable stability properties. The original Radau and Lobatto methods, as for the method (27) above, were sufficiently implicit to preclude efficient numerical use, but not implicit enough for adequate stability.

L-stable Runge–Kutta methods. For very stiff problems, the diagonal Padé approximations are still not satisfactory, since their modulus tends to 1 for $|z| \to \infty$, so that "the methods ... give slowly decreasing, and for even n, oscillating terms in the numerical solution" (quoted from [5]). Therefore, rational approximations for which in addition to A-stability we have $R(z) \to 0$ for $|z| \to \infty$ appear more interesting. A large part of [51] is devoted to the proof of the fact that the approximations in the first and second subdiagonal of the Padé table have all the desired properties. In [5], this result is obtained for the first subdiagonal by an elegant application of Rouché's theorem. Ehle then designed various classes of A-stable and L-stable implicit Runge–Kutta methods of Radau and Lobatto type by modifying the derivations of Butcher. The most important are the so-called Radau IIA methods, which are L-stable collocation methods of order $2s-1$ and are equivalent to the methods given in [5]. This class begins with implicit Euler, and the methods of order 3 and 5 are

$$\begin{array}{c|cc} \frac{1}{3} & \frac{5}{12} & -\frac{1}{12} \\ 1 & \frac{3}{4} & \frac{1}{4} \\ \hline & \frac{3}{4} & \frac{1}{4} \end{array} \qquad \begin{array}{c|ccc} \frac{4-\sqrt{6}}{10} & \frac{88-7\sqrt{6}}{360} & \frac{296-169\sqrt{6}}{1800} & \frac{-2+3\sqrt{6}}{225} \\ \frac{4+\sqrt{6}}{10} & \frac{296+169\sqrt{6}}{1800} & \frac{88+7\sqrt{6}}{360} & \frac{-2-3\sqrt{6}}{225} \\ 1 & \frac{16-\sqrt{6}}{36} & \frac{16+\sqrt{6}}{36} & \frac{1}{9} \\ \hline & \frac{16-\sqrt{6}}{36} & \frac{16+\sqrt{6}}{36} & \frac{1}{9} \end{array} \tag{33}$$

A class of methods reproducing the *second* subdiagonal of the Padé table, and thus decreasing still more rapidly for $|z| \to \infty$, are the Lobatto IIIC methods. These have been introduced for $s \leqslant 3$ in [51] and for general s by Chipman [27].

Ehle's conjecture. After having proved that the diagonal and the first two subdiagonals had the A-stability property, Ehle saw that the third and fourth subdiagonal as well as the first column were *not* A-stable. He then left the conjecture that *all* entries below the second subdiagonal were not A-stable. This conjecture fascinated many workers in the 70's, in particular three of them in Geneva. Extensive numerical calculations of the position of the poles and the crossing points of the stability region with the imaginary axis suggested, that there should be a mysterious relation between them.

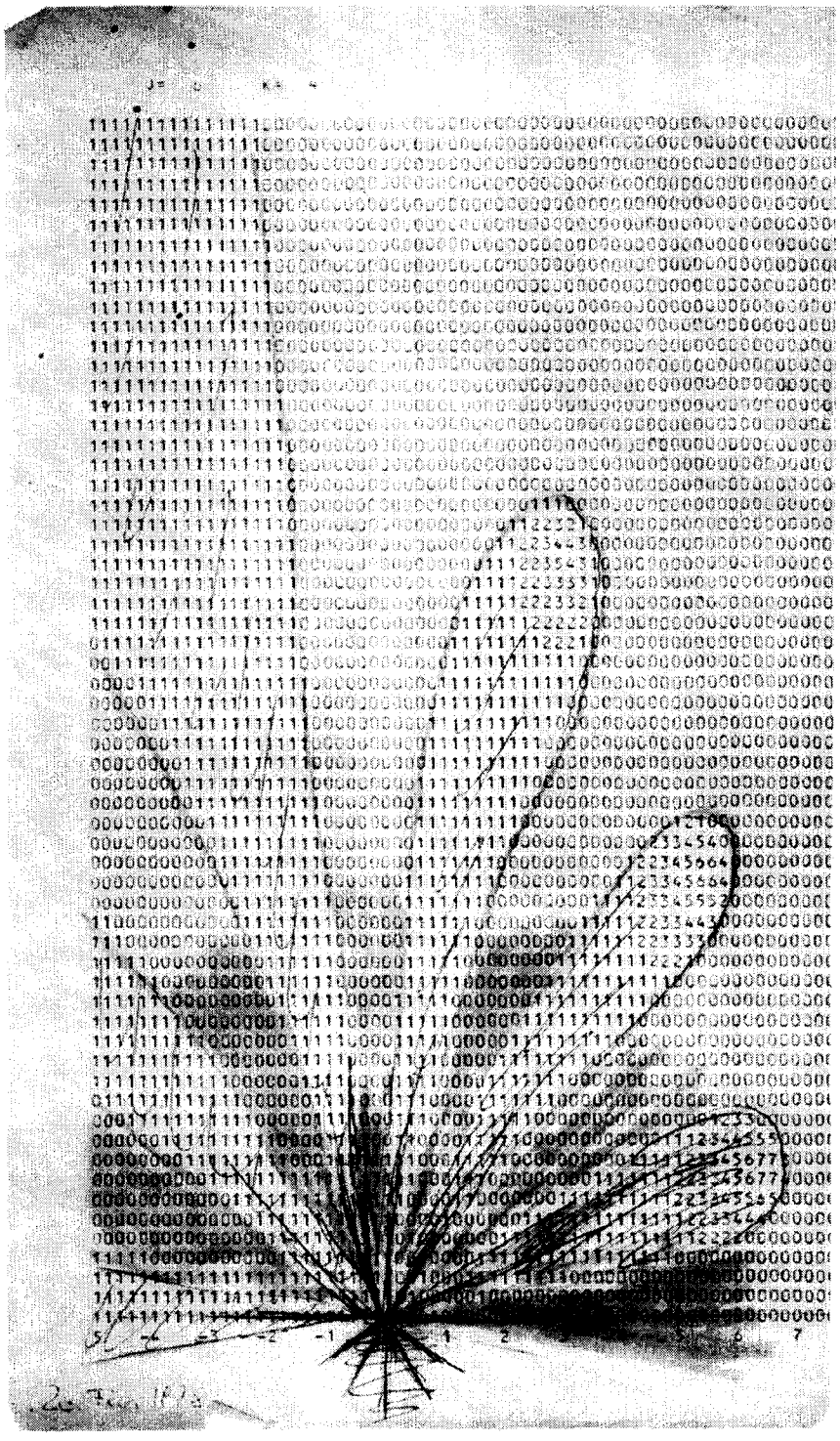


Fig. 13. The first order star.

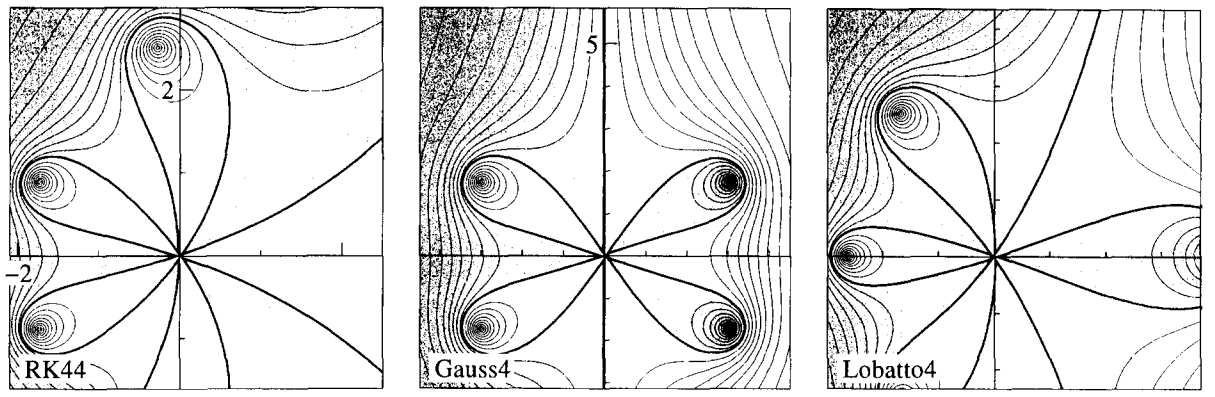


Fig. 14. Order stars of the three RK methods of Fig. 12.

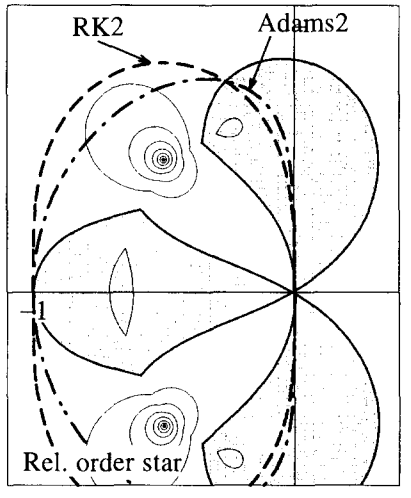


Fig. 15. Relative order star between scaled RK2 and Adams2.

Order stars. The main idea was then: why not cut the surface $|R(z)|$ with $|e^z|$, which is closer to it and might give more information, i.e., define, instead of (31), the order star

$$S := \{z;\ |R(z)| > |e^z|\} = \{z;\ |R(z)/e^z| > 1\}. \tag{34}$$

Fig. 13 shows the first computer output discussed on February 20, 1978 by the three authors of [121]. Fig. 14 shows the order stars which correspond to the stability domains of Fig. 12. Then came the nice surprise, that these order stars not only settled Ehle's conjecture, but also proved all other A-stability results of Ehle. Furthermore, the Daniel-Moore conjecture, a generalization of Dahlquist's "second order barrier" $p \leqslant 2$ for A-stable linear multistep methods, came out by studying order stars on the Riemann surface of the characteristic stability equation. Roughly speaking, the principles are the

following: the fingers coming out from the origin represent the *order* of the method i.e., the numerical precision, the "nails" in the fingers are the poles of $R(z)$ and represent the *numerical work*. One therefore arrives at conclusions of how much order is possible with how much numerical work and that "cheating", by whatever means, is impossible.

The Jeltsch–Nevanlinna theorem. Another surprise with the order stars was the fact that a disc theorem discovered by Jeltsch and Nevanlinna [81] received an elegant interpretation from the *relative* order star, which compares two different stability functions $R_1(z)$ and $R_2(z)$ between each other, i.e., by defining instead of (34)

$$S := \{z;\ |R_1(z)| > |R_2(z)|\} = \{z;\ |R_1(z)/R_2(z)| > 1\}. \tag{35}$$

One of the results of this theory (see [82], with various generalizations in [83]) is that for *explicit* methods, with the *same numerical work* per step, the stability domains are never strictly contained one in the other. This furnishes a theoretical explanation for Merson's observation mentioned above. In Fig. 15 this order star is represented together with the stability domains for explicit Adams2 versus Runge–Kutta2. In order to compensate for the *two* function evaluations per step of RK2, we double the step size for this method, so that $1 + z + z^2/2$ becomes $1 + 2z + 2z^2$, and count this as *two* steps of a method with stability function $R(z) = \sqrt{1 + 2z + 2z^2}$.

7. Nonlinear stability and order reduction

The subjects of this section have been created mainly from the middle of the 70's on, and were crowned in 1984 with the publication of the classic book by Dekker and Verwer [48]. Three main themes constitute this theory: 1. Numerical stability for nonlinear problems; 2. Existence of the numerical solution; 3. Order reduction phenomenon and convergence. The model problem of Prothero and Robinson, and the wider class of singular perturbation problems, have key roles in the understanding of order reduction and in underlining the importance of "stiff accuracy".

Nonlinear stability. Dahlquist, the founder of the modern theory of multistep methods, struggled many years for a successful attack on stability (or, what is perhaps a better word: contractivity) for general nonlinear problems, since multistep methods are not well suited for such a concept. Only the *one-leg* version of these methods, which evaluate at every step the vector field in one point only, allowed him to obtain contractivity results in a certain G-metric and to coin the notion of G-stability. After Dahlquist's memorable 1975 talk in Dundee [42], Butcher [22] immediately applied these ideas to Runge–Kutta methods and proved that the Gauss methods and some A-stable Radau methods were, as he called it, B-stable ("next letter in the alphabet"). A complete algebraic characterization of B-stable methods appeared 1979 in two independent papers: Burrage and Butcher [15] and Crouzeix [34] and was completed 1981 by Hundsdorfer and Spijker [78].

B-stability. The Euclidean norm of the difference between two neighbouring solutions of $y' = f(x, y)$ decreases at least as $Ce^{\nu(x-x_0)}$, wherever

$$\langle f(x, y) - f(x, z), y - z\rangle \leqslant \nu\|y - z\|^2. \tag{36}$$

Here, ν is called "one-sided Lipschitz constant". The smallest value for the above estimate is the so-called "logarithmic norm" of $\partial f/\partial y$ (see, for example, [44, p. 200]). The important fact is that the equation can be arbitrarily stiff with moderate values of ν. For a B-stable method we then require that for $\nu \leqslant 0$, where the distance of two exact solutions does not increase, two neighbouring *numerical* solutions have the same property, i.e.,

$$\|\Delta y_1\| \leqslant \|\Delta y_0\|. \tag{37}$$

The key idea is to extract from the Runge–Kutta formulas the following identity

$$\|\Delta y_1\|^2 = \|\Delta y_0\|^2 + 2h \sum_{i=1}^{s} b_i \langle \Delta f_i, \Delta g_i \rangle - h^2 \sum_{i=1}^{s} \sum_{j=1}^{s} m_{ij} \langle \Delta f_i, \Delta f_j \rangle, \tag{38}$$

from which it can be seen that the "algebraic" conditions

$$\begin{aligned} &\text{(i)} \quad b_i \geqslant 0 \quad \text{for } i = 1, \ldots, s, \\ &\text{(ii)} \quad M = (m_{ij}) = (b_i a_{ij} + b_j a_{ji} - b_i b_j)_{i,j=1}^{s} \text{ is positive semi-definite,} \end{aligned} \tag{39}$$

are *sufficient* conditions for B-stability [15,34]. Two years later, the *necessity* of these conditions for irreducible methods was established in a difficult proof in the "note" [78].

AN-stability. An important precursor to the above result of Hundsdorfer and Spijker was the concept of AN-stability introduced by [15]. This requires numerical stability for all linear nonautonomous equations

$$y' = \lambda(x) y, \quad \operatorname{Re} \lambda(x) \leqslant 0,$$

and leads to the astonishing result that every AN-stable Runge–Kutta method satisfying $c_i \neq c_j$ for $i \neq j$ must satisfy the conditions (39).

The W-transformation. Once the conditions (39) were found, it was interesting to construct Runge–Kutta methods (perhaps, with some specific structure), which satisfied these conditions and possessed a classical order p. With many numerical computations for this question (either on the CDC 3800 in Innsbruck, or on the Univac 1100 in Geneva), the attempt was made to transform the matrices M and A in various ways so as to solve this problem. The surprise was that particularly nice numbers came out of the computer when the matrix M was *not* transformed with the Vandermonde matrix V, as was done previously, but with the matrix W composed by

$$w_{ij} = P_{j-1}(c_i), \quad i = 1, \ldots, s, \ j = 1, \ldots, s, \tag{40}$$

where $P_{j-1}(t)$ are the Legendre polynomials orthogonal on $[0, 1]$. This leads to a similarity transformation of the Runge–Kutta matrix A to a very simple form. For example, for the Gauss methods one obtains

$$W^{-1} A W = \begin{pmatrix} 1/2 & -\xi_1 & & & \\ \xi_1 & 0 & -\xi_2 & & \\ & \xi_2 & \ddots & \ddots & \\ & & \ddots & 0 & -\xi_{s-1} \\ & & & \xi_{s-1} & 0 \end{pmatrix} \quad \text{where } \xi_k = \frac{1}{2\sqrt{4k^2 - 1}}. \tag{41}$$

It was then seen that the structure of this matrix in its first columns or rows was connected to the simplifying assumptions (24) and (25), respectively. Consequently, if one replaces the lower right corner of (41) by an arbitrary positive semi-definite matrix, one obtains all B-stable Runge–Kutta methods of a certain order (see [71; 72, Theorem IV.13.15]).

"Equivalence" of A- and B-stability. Dahlquist [43] proved in 1978 the memorable result that every A-stable one-leg method was also G-stable. However, many A-stable Runge–Kutta methods are not B-stable (such as for example the Lobatto IIIA methods and, for $s = 2$, the trapezoidal rule), and a straightforward extension of Dahlquist's result is not true. The W-transformation together with continued fraction techniques for the stability function then finally allowed it to be shown [65,70] that to every A-stable RK method exists a B-stable RK method with *the same order and the same stability function.*

Existence of numerical solution. For many years no attention was paid to the question of whether or not the algebraic equations associated with an implicit Runge–Kutta method actually have a solution, not only for $h \to 0$, but also for fixed h and arbitrary high stiffness. Pioneering works answering this question around 1980 were [35,36,45]. The original proofs used variants of "Brouwer's fixed point theorem"; an elementary proof has been given in [72, Theorem IV.14.2]. The main ingredient of these existence results is a coercivity condition number of the Runge–Kutta matrix

$$\alpha_0\bigl(A^{-1}\bigr) = \sup_{D>0}\left(\inf_{u\in\mathbb{R}^s}\frac{\langle u, A^{-1}u\rangle_D}{\langle u,u\rangle_D}\right) \tag{42}$$

where $D = \operatorname{diag}(d_1,\dots,d_s)$ with $d_i > 0$ and $\langle u,v\rangle_D = u^{\mathrm{T}}Dv$. Then existence, uniqueness, and stability with respect to perturbations of the numerical solution values can be shown if

$$h\nu < \alpha_0\bigl(A^{-1}\bigr) \tag{43}$$

where ν is the one-sided Lipschitz constant of f.

Computation of $\alpha_0(A^{-1})$. The values of $\alpha_0(A^{-1})$ are now known for many classes of implicit Runge–Kutta methods (see [48, pp. 55–164; 72, Theorem IV.14.5]). The decisive break-through came when Dekker [46] discovered, after many numerical computations, which matrix D brought the expression $DA^{-1} + (DA^{-1})^{\mathrm{T}}$ to diagonal form.

Order reduction and convergence. The phenomenon of the order reduction for stiff equations was discovered in 1974 by Prothero and Robinson [99] with the help of their very instructive model equation

$$y' = \lambda\bigl(y - \varphi(x)\bigr) + \varphi'(x), \quad y(x_0) = \varphi(x_0), \qquad \operatorname{Re}\lambda \leqslant 0 \tag{44}$$

which has $\varphi(x)$ as exact solution and whose numerical solution can be analyzed by elementary calculations. The smooth solution $\varphi(x)$ is, for increasing stiffness $\operatorname{Re}\lambda \to -\infty$, surrounded by wild transients. The evaluation points g_i of the Runge–Kutta solution, however, *leave* this smooth solution and enter into a region where the slopes $f(g_i)$ have nothing to do with the slopes of the smooth solution (see Fig. 16). All this leads to the fact that, with increasing stiffness, the error of the method is no longer related to the one predicted by the classical order analysis. A first corner-stone for a new theory, named theory of "B-convergence" [58], is an estimate for the distance of the points g_i from the smooth solution. Here the conditions (24), which express the so-called "stage order" of the

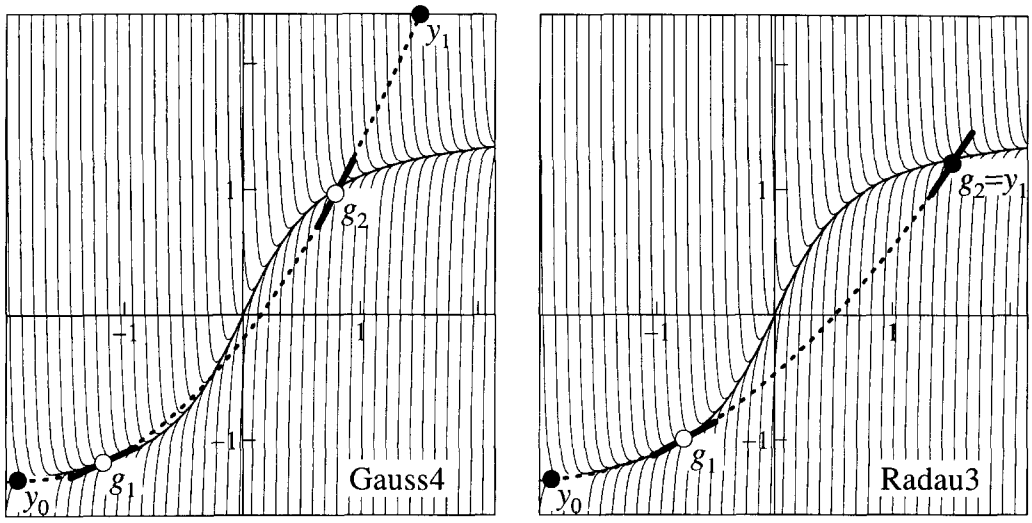


Fig. 16. Order reduction at Prothero–Robinson model.

method, play a crucial role. The corresponding convergence results have then been mainly obtained by the "Vienna group" and by the "Holland group" (for more details see the "Vienna contribution" in this issue, [47,48,59,85]; see also [72, Section IV.15]).

It also becomes clear why DIRK methods, whose first stage is of order 1 only, are often disappointing for high precision and strong stiffness.

Stiffly accurate methods. A second observation to be drawn from Fig. 16 is that the Radau methods, because of $c_s = 1$, suffer less than the Gauss methods. Methods with this property, which for general Runge–Kutta methods becomes

$$a_{sj} = b_j, \quad j = 1, \ldots, s, \tag{45}$$

have been named *stiffly accurate* in [99].

Singular perturbation problems. Another important class of problems for understanding the behaviour of Runge–Kutta methods for stiff problems, is the class of singular perturbation problems

$$y' = f(y, z), \qquad \varepsilon z' = g(y, z), \tag{46}$$

which, for $\varepsilon \to 0$ becomes more and more stiff. For $\varepsilon = 0$ we obtain

$$y' = f(y, z), \qquad 0 = g(y, z), \tag{47}$$

the corresponding *reduced* DAE problem. By analyzing the ε expansions of the true and the numerical solution one can draw conclusions about the convergence properties and order reductions for these problems (see [68], see also [72, Sections VI.1 and VI.2; 2nd ed., Section VI.3]). Once again, the property of being stiffly accurate turns out to be very important.

8. Composition of methods

If two steps are carried out using two, possibly different, Runge–Kutta methods, the functional relationship between the result computed and the input value to the first of the two steps is equivalent to the action of a combined method. Furthermore, it is not necessary to know the details of the two methods to evaluate the elementary weights of the combined method since these can be expressed in terms of the elementary weights of the original two methods. To see how this works, consider methods given by the following tableaux

$$\begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \dots & a_{1s} \\ c_2 & a_{21} & a_{22} & \dots & a_{2s} \\ \vdots & \vdots & \vdots & & \vdots \\ c_s & a_{s1} & a_{s2} & \dots & a_{ss} \\ \hline & b_1 & b_2 & \dots & b_s \end{array} \qquad \begin{array}{c|cccc} \bar{c}_1 & \bar{a}_{11} & \bar{a}_{12} & \dots & \bar{a}_{1\bar{s}} \\ \bar{c}_2 & \bar{a}_{21} & \bar{a}_{22} & \dots & \bar{a}_{2\bar{s}} \\ \vdots & \vdots & \vdots & & \vdots \\ \bar{c}_{\bar{s}} & \bar{a}_{\bar{s}1} & \bar{a}_{\bar{s}2} & \dots & \bar{a}_{\bar{s}\bar{s}} \\ \hline & \bar{b}_1 & \bar{b}_2 & \dots & \bar{b}_{\bar{s}} \end{array} \tag{48}$$

For a given differential equation and step size, let ϕ and $\bar{\phi}$ denote the functions relating the output to the input value for each of these methods, respectively. The composition of these two functions $\bar{\phi} \circ \phi$ is equivalent to a single step of the $s + \bar{s}$ stage method given by the tableau

$$\begin{array}{c|cccc|cccc} c_1 & a_{11} & a_{12} & \dots & a_{1s} & 0 & 0 & \dots & 0 \\ c_2 & a_{21} & a_{22} & \dots & a_{2s} & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ c_s & a_{s1} & a_{s2} & \dots & a_{ss} & 0 & 0 & \dots & 0 \\ d + \bar{c}_1 & b_1 & b_2 & \dots & b_s & \bar{a}_{11} & \bar{a}_{12} & \dots & \bar{a}_{1\bar{s}} \\ d + \bar{c}_2 & b_1 & b_2 & \dots & b_s & \bar{a}_{21} & \bar{a}_{22} & \dots & \bar{a}_{2\bar{s}} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ d + \bar{c}_{\bar{s}} & b_1 & b_2 & \dots & b_s & \bar{a}_{\bar{s}1} & \bar{a}_{\bar{s}2} & \dots & \bar{a}_{\bar{s}\bar{s}} \\ \hline & b_1 & b_2 & \dots & b_s & \bar{b}_1 & \bar{b}_2 & \dots & \bar{b}_{\bar{s}} \end{array} \tag{49}$$

where $d = \sum_{i=1}^{s} b_i$. If we calculate the elementary weights of this combined method for at least the trees of orders 1, 2 and 3, we find that they can be written in terms of the elementary weights for the constituent methods. For the tree t denote the elementary weights of the two methods by $\Phi(t)$ and $\bar{\Phi}(t)$ and for the combined method denote the elementary weight by $\widetilde{\Phi}(t)$. For the trees t_1, t_2, t_3 and t_4 we have

$$
\left.
\begin{aligned}
\widetilde{\Phi}(t_1) &= \sum_{i=1}^{s} b_i + \sum_{i=1}^{\bar{s}} \bar{b}_i && = \Phi(t_1) + \bar{\Phi}(t_1),\\
\widetilde{\Phi}(t_2) &= \sum_{i=1}^{s} b_i c_i + \sum_{i=1}^{\bar{s}} \bar{b}_i (d + \bar{c}_i) && = \Phi(t_2) + \Phi(t_1)\bar{\Phi}(t_1) + \bar{\Phi}(t_2),\\
\widetilde{\Phi}(t_3) &= \sum_{i=1}^{s} b_i c_i^2 + \sum_{i=1}^{\bar{s}} \bar{b}_i (d + \bar{c}_i)^2 && = \Phi(t_3) + \Phi(t_1)^2\bar{\Phi}(t_1) + 2\Phi(t_1)\bar{\Phi}(t_2) + \bar{\Phi}(t_3),\\
\widetilde{\Phi}(t_4) &= \sum_{i,j=1}^{s} b_i a_{ij} c_j + \sum_{i=1}^{\bar{s}} \sum_{j=1}^{s} \bar{b}_i b_j c_j + \sum_{i,j=1}^{\bar{s}} \bar{b}_i \bar{a}_{ij} (d + \bar{c}_j) &&\\
&= \Phi(t_4) + \Phi(t_1)\bar{\Phi}(t_2) + \Phi(t_2)\bar{\Phi}(t_1) + \bar{\Phi}(t_4). &&
\end{aligned}
\right\} \tag{50}
$$

Expressions for $\widetilde{\Phi}(t)$ exist for all t. This is to be expected because the sequence of Φ values characterizes the method. Consider the following three equivalence relations between members of the set of all Runge–Kutta methods.

R_1 Two methods are equivalent if, by deleting unnecessary stages and combining stages that give identical results, they both reduce to the same "irreducible" methods.

R_2 Two methods are equivalent if for any Lipschitz continuous function f, there exists $h_0 > 0$ such that for $0 \leqslant h \leqslant h_0$, the results of computing a single step with each of the two methods are identical.

R_3 For any tree t the value of $\Phi(t)$ is the same for each of the two methods.

It can be shown that $R_1 = R_2 = R_3$. A consequence of this is our remark that the values of $\widetilde{\Phi}$ depend only on the values of Φ and $\bar{\Phi}$. This theory was first presented in [21], and it was shown there that equivalence classes of Runge–Kutta methods form a group under composition of representative members of the classes. Using the representation of equivalence classes by mappings $T \to \mathbb{R}$, where for a particular class each tree t maps to $\Phi(t)$, a homomorphic group can be constructed. We will denote this group by G. For any positive integer p let H_p denote the set of members of G which map all trees with no more than p vertices to 0. It can be shown that H_p is a normal subgroup of G.

Let α and β denote two members of G. Then as we have seen from (50),

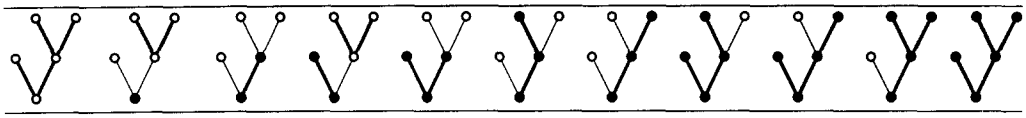
$$
\begin{aligned}
(\alpha\beta)(t_1) &= \alpha_1 + \beta_1,\\
(\alpha\beta)(t_2) &= \alpha_2 + \alpha_1\beta_1 + \beta_2,\\
(\alpha\beta)(t_3) &= \alpha_3 + \alpha_1^2\beta_1 + 2\alpha_1\beta_2 + \beta_3,\\
(\alpha\beta)(t_4) &= \alpha_4 + \alpha_2\beta_1 + \alpha_1\beta_2 + \beta_4,
\end{aligned}
$$

where we have written $\alpha_k = \alpha(t_k)$, $\beta_k = \beta(t_k)$ for $k = 1, 2, \ldots$.

The general formula is expressed recursively in [21] but it has a simple structure as a sum over all "subtrees" u of t. For this subtree, the corresponding term is $\beta(u)$ multiplied by the product of factors $\alpha(v)$, where the v are all trees left over when u is removed from t.

An example of this is illustrated in Table 3, where a formula is constructed for $(\alpha\beta)(t_{11})$ with t_{11} a tree with 5 vertices occurring in the top row of this table. In each column, the part of $t = t_{11}$ corresponding to u is shown with filled-in discs at each vertex. By contrast, the vertices in the various v trees are shown as circular outlines. Thick lines are used for the edges in both the u and v trees

Table 3
An example of the product of two tree mappings



α_{11}	$\alpha_1\alpha_3\beta_1$	$\alpha_1^3\beta_2$	$\alpha_3\beta_2$	$\alpha_1^2\beta_3$	$\alpha_1^2\beta_4$	$\alpha_1^2\beta_4$	$\alpha_1\beta_6$	$\alpha_1\beta_6$	$\alpha_1\beta_7$	β_{11}

and thin lines are used to indicate only where edges existed in the tree t from which the diagrams are derived.

Combining the terms in the last row of this table, we find

$$(\alpha\beta)(t_{11}) = \alpha_{11} + \alpha_1\alpha_3\beta_1 + \alpha_1^3\beta_2 + \alpha_3\beta_2 + \alpha_1^2\beta_3 + 2\alpha_1^2\beta_4 + 2\alpha_1\beta_6 + \alpha_1\beta_7 + \beta_{11}.$$

In addition to the mappings arising as the weight functions of particular Runge–Kutta methods, there are others that could be added to G, such as the limiting Runge–Kutta method formed by allowing the number of stages to become infinite so that the set $\{1, 2, \ldots, s\}$ becomes the interval $[0, 1]$ and the arrays A and b^{T} become linear functions on function spaces. If A is defined in this limiting situation by the formula

$$A(\phi)(\xi) = \int_0^{\xi} \phi(\eta)\,\mathrm{d}\eta,$$

and b^{T} as the linear functional

$$b(\phi) = \int_0^{1} \phi(\eta)\,\mathrm{d}\eta,$$

then the "approximation" computed by this method is the exact solution evaluated at $x_0 + h$. A formal computation of the "elementary weights" for this case gives the values $1/\gamma(t)$ for all t. Denote this by $E(t)$. Using the quotient group G/H_p, the order is p for an arbitrary Runge–Kutta method with elementary weight function α if and only if $\alpha H_p = EH_p$.

The earliest application proposed for the composition formula is in the concept of "effective order". A Runge–Kutta method M has effective order p if there exists a second method N such that the composition of three steps, consisting in turn of N, M and the method which undoes the work of N, is of order p. In terms of the elementary weights of the three methods this can be written as $(\beta\alpha\beta^{-1})H_p = EH_p$. This is a non-trivial generalization of classical order because, for example, in the case of explicit methods with $s = 5$, order 5 is not possible but effective order 5 is possible.

It has been remarked [69] that methods of a given effective order cannot be used as practical algorithms because it is not possible to change step size. In fact it is possible to change from the use of one step size to the next if dependence of some coefficients on the step size ratio is permitted. Generalizing the use of one step methods to a wide class of multivalue methods enables the type of analysis used for effective order to be used in a more general context.

The phenomenon of effective order has been interpreted by Stetter [116] as a possible means of calculating an approximation to the global truncation error.

In contrast to the use of the limiting Runge–Kutta method represented by E as a means of finding a direct route to the order concept, we consider the alternative formulation as presented in [73]. Here, rather than consider the coefficients of the various elementary differentials in the Taylor expansion in a numerical approximation, as being a mapping from trees to real numbers, the series itself is used as an object of study. In this approach, composition of these "B-series", can be viewed as a substitution of one series into the starting point of another B-series. Then the "exact solution" E becomes just the series (18) with all coefficients equal to $1/\gamma(t)$.

9. Hamiltonian symplectic integration

The identity (38), which was used in Section 7 for discussing the preservation of contractivity, can be applied to many more problems of a special structure, for example dissipative problems, gradient problems, conservative problems, orthonormality preserving equations, and Hamiltonian problems. An overview over these subjects, with many references, is given in [117]. We outline here in the following the theory of symplectic methods.

A nice introduction to Hamiltonian integration has recently been published, 99 years after Runge's paper, by Sanz-Serna and Calvo [113]. Some material can also be found in [69, 2nd ed., Section II.16; 111; 112].

Hamiltonian problems are of the form

$$\dot{p}_i = -\frac{\partial H}{\partial q_i}, \quad \dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad i = 1, \ldots, n, \tag{51}$$

where $H(p_1, \ldots, p_n, q_1, \ldots, q_n)$ is "the Hamiltonian". The corresponding flow is *symplectic*, i.e., has the remarkable property that the differential 2-form

$$\omega^2 = \sum_{i=1}^{n} \mathrm{d}p_i \wedge \mathrm{d}q_i \tag{52}$$

is preserved by the flow (Poincaré [98, Vol. III, Section 255, p. 43]). In two dimensions ($n = 1$) this means that the areas of "infinitely small" parallelograms remain everywhere constant (see Fig. 17(a), where the pendulum $H = p^2/2 - \cos q$ is taken as example).

Symplectic Runge–Kutta methods. The first pioneers in symplectic integration were usually working with the Hamilton-Jacobi formalism of "generating functions" (e.g., [56; 107; 113, p. 147]). Around 1988 it was then discovered independently by three authors [90,110,119], that implicit Runge–Kutta methods *are* symplectic, under condition that

$$M = 0, \tag{53}$$

where M is the same matrix as in (39). Indeed, the proof is found by replacing the *scalar* product in (38) by the *exterior* product. This at once made an arsenal of existing methods accessible for Hamiltonian integration. In particular the Gauss methods of order $2s$ are symplectic and with them the implicit midpoint rule (see Fig. 17(d)). On the other hand, the explicit Euler method (as well as all other explicit methods) and the implicit Euler method are not symplectic (Fig. 17, cases (b) and (c)).

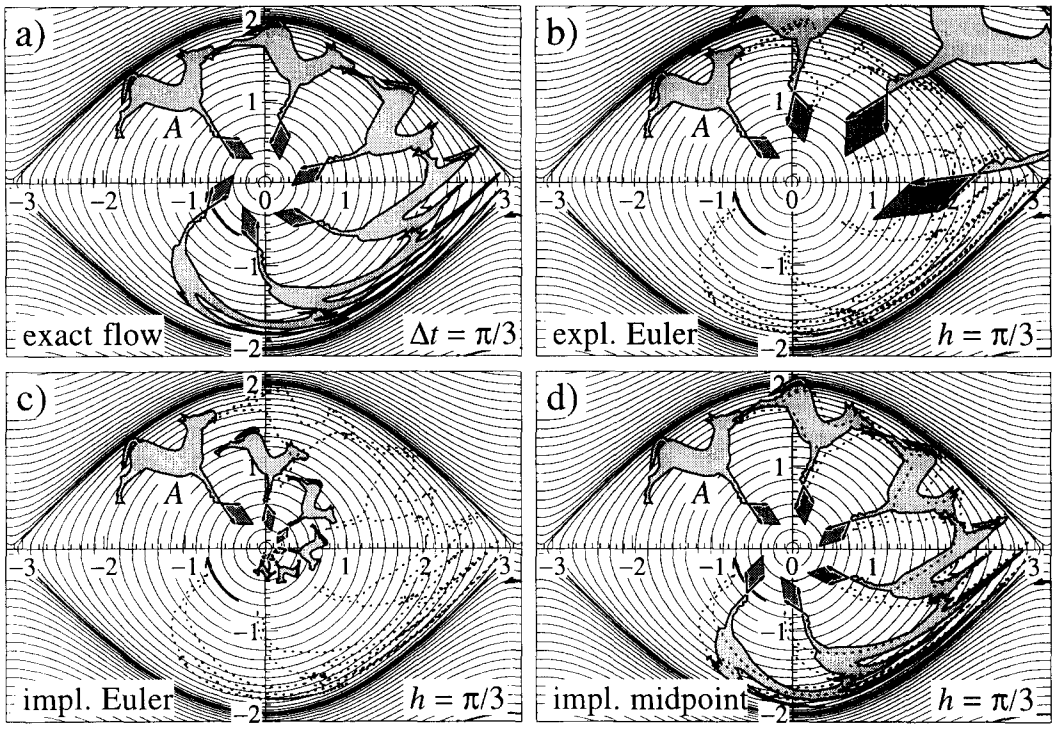


Fig. 17. The pendulum with symplectic and non-symplectic integration.

Partitioned symplectic Runge–Kutta methods. The main battlefield for symplectic Runge–Kutta integration are the *partitioned* methods, which treat the p-part and the q-part of (51) by different coefficient sets b_i, a_{ij} and $\widehat{b}_i$, $\widehat{a}_{ij}$, respectively. These methods are symplectic iff

$$\begin{aligned} &\text{(a)} \quad b_i = \widehat{b}_i, && i = 1, \ldots, s, \\ &\text{(b)} \quad b_i \widehat{a}_{ij} + \widehat{b}_j a_{ji} - b_i \widehat{b}_j = 0, && i, j = 1, \ldots, s. \end{aligned} \tag{54}$$

In particular, the following combinations of explicit–implicit and implicit–explicit Euler

$$\begin{aligned} p_1 &= p_0 - hH_q(p_0, q_1) \\ q_1 &= q_0 + hH_p(p_0, q_1) \end{aligned} \quad \text{and} \quad \begin{aligned} p_1 &= p_0 - hH_q(p_1, q_0) \\ q_1 &= q_0 + hH_p(p_1, q_0) \end{aligned} \tag{55}$$

are both symplectic of order 1. An important class of symplectic partitioned methods is the Lobatto IIIA-IIIB pair, first discussed by [118], which has important applications to *constrained* differential equations [80].

In many cases (e.g., for second order equations and more generally for Hamiltonians of the type $H(p, q) = T(p) + U(q)$) symplectic partitioned methods can be *explicit* (as for example the methods (55)). Such higher order explicit symplectic methods are constructed by solving the order conditions (see, e.g., [69, 2nd ed., pp. 326–332; 113, Section 8.4]), by concatenating with the adjoint method, and by the so-called "composition methods" or "fractional step methods" (see, e.g., [72, 2nd ed., p. 554; 113, Section 12.4]).

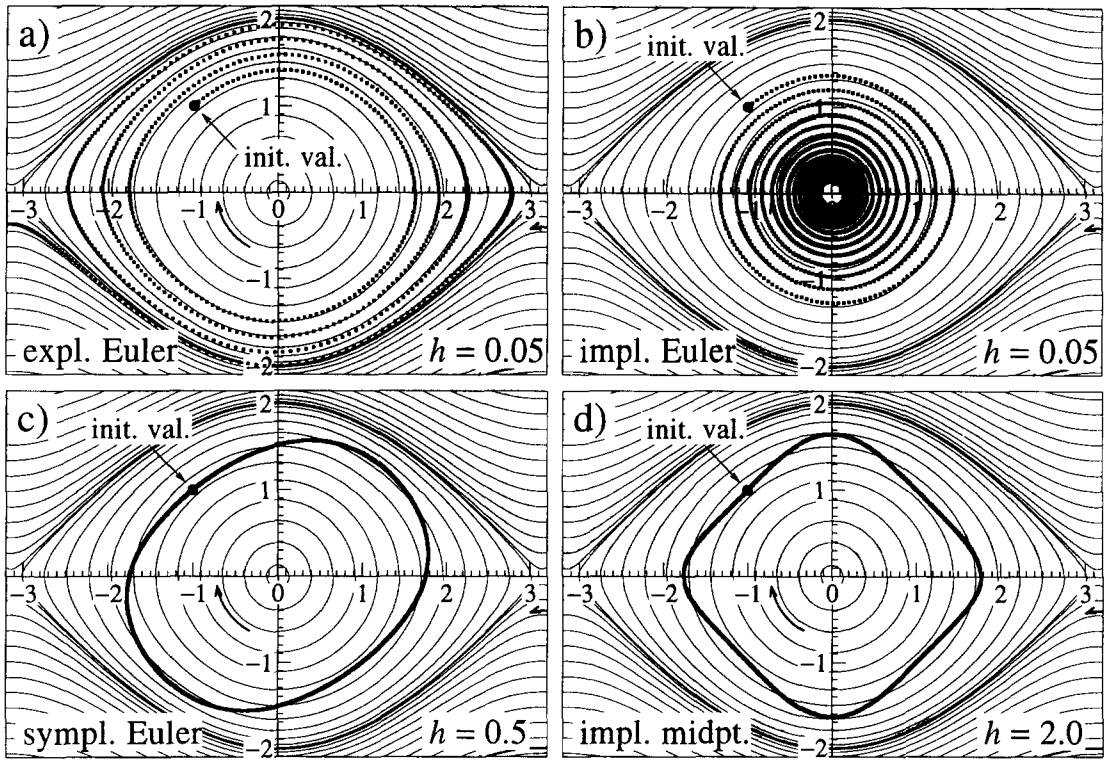


Fig. 18. Symplectic and non-symplectic long-range integration.

Long-range computations. We present in Fig. 18 four different methods applied to the pendulum equation integrated over a long time. Two of the methods (a and b) are not symplectic, the other two (c and d) are, and, in addition, have been applied with a much larger step size. It appears that the symplectic integrators behave much better, while the nonsymplectic methods, after a certain time, spiral to completely wrong positions. At a first view, an explanation for the importance of symplecticity to avoid this phenomenon is not straightforward.

Backward error analysis. The key for a satisfying answer was found around 1992 independently by Sanz-Serna, Yoshida [124] and Feng Kang [57]. The numerical solution is interpreted as the *exact* solution of *another* differential equation. This idea, already used without much ado by Runge in 1908 [105], has been made famous by Wilkinson for the study of errors in linear algebra. In the ODE case this method of analysis usually goes under the name of the "modified equation method" (see, e.g., [113, Section 10.1]). For example, the numerical solution of the left method in (55) applied to the pendulum equation represents the *exact* solution of the *perturbed* Hamiltonian problem with (see [124])

$$\widetilde{H} = \frac{p^2}{2} - \cos q + \frac{h}{2} p \sin q + \frac{h^2}{12}\left(\sin^2 q + p^2 \cos q\right) + \frac{h^3}{12} p \cos q \sin q + \mathrm{O}\left(h^4\right). \tag{56}$$

The contour lines of this Hamiltonian are close to the deformed ovals visible in Fig. 18(c).

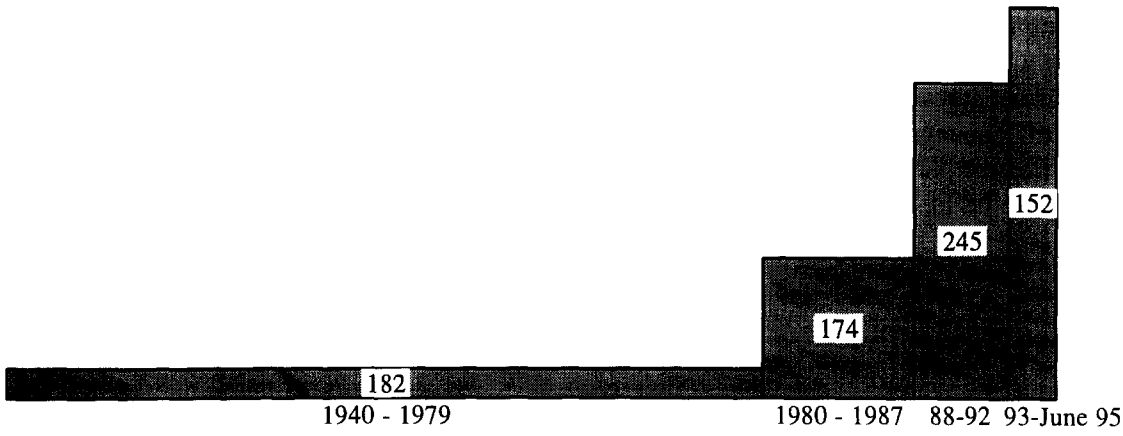


Fig. 19. Number of mathematical reviews with "Runge–Kutta" in the title.

A general theory has been elaborated by Hairer [66]. The perturbed differential equation is expanded as B-series

$$\widetilde{y}' = \beta_1 f(\widetilde{y}) + \beta_2 \frac{h}{2} f'(\widetilde{y}) f(\widetilde{y}) + \cdots = \sum_{\rho(t) \leqslant N} \frac{h^{\rho(t)-1}}{\rho(t)!} \alpha(t)\, b(t)\, F(t)(\widetilde{y}). \tag{57}$$

For the coefficients $b(t)$ there exist tree-algebraic formulas in terms of the B-series of the numerical solution. It then holds that *if the method is symplectic and if the original problem is Hamiltonian, then the perturbed problem* (57) *is Hamiltonian too.*

Long-range error estimations. The series in the perturbed equation (57) are, in general, diverging with increasing N. Therefore it is interesting to find rigorous bounds on the difference of the numerical solution and the exact solution of (57). This has been done independently by Benettin and Giorgilli [8] and Hairer and Lubich [67]. This enables it to be shown, among other results, that the error of the Hamiltonian for symplectic methods remains bounded by $O(h^p)$ for "exponentially long times".

10. Further Runge–Kutta developments

In these historical notes, we have touched on only a small fraction of what has been learnt about Runge–Kutta methods and their possible applications during the last 100 years. In particular, almost nothing has been said about the developments of the last 15 years. We illustrate the impossibility of doing so by presenting in Fig. 19 the number of items of Mathematical Reviews with "Runge–Kutta" in the title (in fact, more than a thousand items over this period have "Runge–Kutta" in the review itself).

Runge–Kutta methods have been adapted to the solution of more general problem classes each of which has been the subject of specialized research in recent years. Some publications relevant to these areas are, for differential–algebraic equations [68] (see also [72, 2nd ed., Chapter VII]) and for delay differential equations [69, 2nd ed., Section II.17] and a forthcoming book by two specialists on this subject, A. Bellen and M. Zennaro. For Volterra integral equations, another subject whose centenary has arrived [7,11] and for stochastic differential equations [84].

In addition to these generalizations of Runge–Kutta to wider problem areas, Runge–Kutta methods have been applied in their own right to partial differential equations through the so-called method of lines. This approach leads to large structured systems of ordinary differential equations which are usually stiff. This approach was first published in [32] (under the name of "Hartree's method") but is of current interest through contributions of the CWI group to Chebyshev methods, reported elsewhere in this special issue, and through convergence results which are independent of the stiffness (see, for example, [48,93]).

Acknowledgements

The authors wish to express their thanks to Joe Flaherty, Ernst Hairer, Christian Lubich and Jan Verwer for their valuable comments on early drafts of this paper.

References

[1] J.C. Adams, Appendix in: F. Bashforth, *An Attempt to Test the Theories of Capillary Action by Comparing the Theoretical and Measured Forms of Drops of Fluid. With an Explanation of the Method of Integration Employed in Constructing the Tables Which Give the Theoretical Form of Such Drops, by J.C. Adams* (Cambridge University Press, Cambridge, 1883).

[2] P. Albrecht, Numerical treatment of O.D.Es.: The theory of A-methods, *Numer. Math.* 17 (1985) 59–87.

[3] R. Alexander, Diagonally implicit Runge–Kutta methods for stiff O.D.E.'s, *SIAM J. Numer. Anal.* 14 (1977) 1006–1021.

[4] R. Alt, Méthodes A-stables pour l'intégration de systèmes différentiels mal conditionnés, Thèse, Univ. Paris VI (1971).

[5] O. Axelsson, A class of A-stable methods, *BIT* 9 (1969) 185–199.

[6] G. Bader and P. Deuflhard, A semi-implicit mid-point rule for stiff systems of ordinary differential equations, *Numer. Math.* 41 (1983) 373–398.

[7] C.T.H. Baker, The state of the art in the numerical treatment of integral equations, in: A. Iserles and M.J.D. Powell, eds., *The State of the Art in Numerical Analysis* (Clarendon Press, Oxford, 1987) 473–509.

[8] G. Benettin and A. Giorgilli, On the Hamiltonian interpolation of near to the identity symplectic mappings with application to symplectic integration algorithms, *J. Statist. Phys.* 74 (1994) 1117–1143.

[9] L. Bieberbach, On the remainder of the Runge–Kutta formula in the theory of ordinary differential equations, *Z. Angew. Math. Phys.* 2 (1951) 233–248.

[10] G. Birkhoff and R.S. Varga, Discretization errors for well-set Cauchy problems I, *J. Math. Phys.* 44 (1965) 1–23.

[11] H. Brunner and P.J. van der Houwen, *The Numerical Solution of Volterra Equations*, CWI Monographs 3 (North-Holland, Amsterdam, 1986).

[12] R. Bulirsch and M. Breitner, Wilhelm Martin Kutta (1867–1944), *DMV Mitteilungen* 2 (1994) 7–8.

[13] R. Bulirsch and J. Stoer, Numerical treatment of ordinary differential equations by extrapolation methods, *Numer. Math.* 8 (1966) 1–13.

[14] K. Burrage, J.C. Butcher and F.H. Chipman, An implementation of singly-implicit Runge–Kutta methods, *BIT* 20 (1980) 326–340.

[15] K. Burrage and J.C. Butcher, Stability criteria for implicit Runge–Kutta methods, *SIAM J. Numer. Anal.* 16 (1979) 46–57.

[16] J.C. Butcher, Coefficients for the study of Runge–Kutta integration processes, *J. Austral. Math. Soc.* 3 (1963) 185–201.
[17] J.C. Butcher, On Runge–Kutta methods of high order, *J. Austral. Math. Soc.* 4 (1964) 179–194.
[18] J.C. Butcher, Implicit Runge–Kutta processes, *Math. Comp.* 18 (1964) 50–64.
[19] J.C. Butcher, Integration processes based on Radau quadrature formulas, *Math. Comp.* 18 (1964) 233–244.
[20] J.C. Butcher, On the attainable order of Runge–Kutta methods, *Math. Comp.* 19 (1965) 408–417.
[21] J.C. Butcher, An algebraic theory of integration methods, *Math. Comp.* 26 (1972) 79–106.
[22] J.C. Butcher, A stability property of implicit Runge–Kutta methods, *BIT* 15 (1975) 358–361.
[23] J.C. Butcher, The non-existence of ten stage eighth order explicit Runge–Kutta methods, *BIT* 25 (1985) 521–540.
[24] J.C. Butcher, *The Numerical Analysis of Ordinary Differential Equations: Runge–Kutta and General Linear Methods* (Wiley, Chichester, 1987).
[25] J.C. Butcher, A history of Runge–Kutta methods, *Appl. Numer. Math.* 20 (1996) 247–260.
[26] A.L. Cauchy, Résumé des Leçons données à l'Ecole Royale Polytechnique. Suite du Calcul Infinitésimal (1824); published: Chr. Gilain and Johnson, eds, *Equations Différentielles Ordinaires* (1981).
[27] F.H. Chipman, A-stable Runge–Kutta processes, *BIT* 11 (1971) 384–388.
[28] L. Collatz, *Numerische Behandlung von Differentialgleichungen*, Grundlehren der Mathematischen Wissenschaften 60 (Springer, New York, 1951); 2nd ed. 1955; 3rd ed. and English translation 1960.
[29] G.J. Cooper and J.H. Verner, Some explicit Runge–Kutta methods of high order, *SIAM J. Numer. Anal.* 9 (1972) 389–405.
[30] G. Coriolis, Mémoire sur le degré d'approximation qu'on obtient pour les valeurs numériques d'une variable qui satisfait à une équation différentielle, en employant pour calculer ces valeurs diverses équations aux différences plus ou moins approchées, *J. Math. Pures Appl. (Liouville)* 2 (1837) 229–244.
[31] R. Courant, K. Friedrichs and H. Lewy, Ueber die partiellen Differenzengleichungen der mathematischen Physik, *Math. Ann.* 100 (1928) 32–74.
[32] J. Crank and P. Nicolson, A practical method for numerical integration of solutions of partial differential equations of heat-conduction type, *Proc. Cambridge Philos. Soc.* 43 (1963) 50–67.
[33] M. Crouzeix, Sur l'approximation des équations différentielles opérationnelles linéaires par de méthodes de Runge–Kutta, Thèse, Univ. Paris VI (1975).
[34] M. Crouzeix, Sur la B-stabilité des méthodes de Runge–Kutta, *Numer. Math.* 32 (1979) 75–82.
[35] M. Crouzeix and P.A. Raviart, Approximation des problèmes d'évolution, Unpublished lecture notes, Université de Rennes (1980).
[36] M. Crouzeix, W.H. Hundsdorfer and M.N. Spijker, On the existence of solutions to the algebraic equations in implicit Runge–Kutta methods, *BIT* 23 (1983) 84–91.
[37] A.R. Curtis, An eighth order Runge–Kutta process with eleven function evaluations per step, *Numer. Math.* 16 (1970) 268–277.
[38] A.R. Curtis, High order explicit Runge–Kutta formulae, their uses, and limitations, *J. Inst. Math. Appl.* 16 (1975) 35–55.
[39] C.F. Curtiss and J.O. Hirschfelder, Integration of stiff equations, *Proc. Nat. Acad. Sci.* U.S.A. 38 (1952) 235–243.
[40] G. Dahlquist, Fehlerabschätzungen bei Differenzenmethoden zur numerischen Integration gewöhnlicher Differentialgleichungen, *Z. Angew. Math. Mech.* 31 (1951) 239–240.
[41] G. Dahlquist, A special stability problem for linear multistep methods, *BIT* 3 (1963) 27–43.
[42] G. Dahlquist, Error analysis for a class of methods for stiff nonlinear initial value problems, in: *Numerical Analysis, Dundee '75,* Lecture Notes in Mathematics 506 (1975) 60–74.
[43] G. Dahlquist, G-stability is equivalent to A-stability, *BIT* 18 (1978) 384–401.
[44] G. Dahlquist, 33 years of numerical instability, part I, *BIT* 25 (1985) 188–204.

[45] K. Dekker, On the iteration error in algebraically stable Runge–Kutta methods, Report NW 138/82, Math. Centrum, Amsterdam (1982).

[46] K. Dekker, Error bounds for the solution to the algebraic equations in Runge–Kutta methods, *BIT* 24 (1984) 347–356.

[47] K. Dekker, J.F.B.M. Kraaijevanger and J. Schneid, On the relation between algebraic stability and B-convergence for Runge–Kutta methods, *Numer. Math.* 57 (1990) 249–262.

[48] K. Dekker and J.G. Verwer, *Stability of Runge–Kutta Methods for Stiff Nonlinear Differential Equations* (North-Holland, Amsterdam, 1984).

[49] J.R. Dormand and P.J. Prince, A family of embedded Runge–Kutta formulae, *J. Comput. Appl. Math.* 6 (1980) 19–26.

[50] B.L. Ehle, High order A-stable methods for the numerical solution of systems of DEs, *BIT* 8 (1968) 276–278.

[51] B.L. Ehle, On Padé approximations to the exponential function and A-stable methods for the numerical solution of initial value problems, Research Report CSRR 2010, Dept. AACS, Univ. of Waterloo, Ontario, Canada (1969).

[52] R. Emden, *Gaskugeln, Anwendung der Mechanischen Wärmetheorie auf Kosmologische und Meteorologische Probleme* (Teubner, Leipzig, 1907).

[53] R. England, Error estimates for Runge–Kutta type solutions to systems of ordinary differential equations, *Comput. J.* 12 (1969) 166–170.

[54] L. Euler, Institutiones Calculi Integralis. Volumen Primum (1768), *Opera Omnia*, Vol. XI, B.G. Teubneri, Lipsiae et Berolini MCMXIII.

[55] E. Fehlberg, Classical fifth-, sixth-, seventh-, and eighth- order Runge–Kutta formulas with step size control, NASA Technical Report 287 (1968); extract published in *Computing* 4 (1969) 93–106.

[56] Feng Kang, On difference schemes and symplectic geometry, *Proceedings 5th Internat. Sympos. Differential Geom. Differential Equations*, Beijing (August 1984) 42–58.

[57] Feng Kang, Formal dynamical systems and numerical algorithms, talk at *Internat. Conf. Comput. Differential Equations Dynamic Syst.*, Beijing (September 1992).

[58] R. Frank, J. Schneid and C.W. Ueberhuber, The concept of B-convergence, *SIAM J. Numer. Anal.* 18 (1981) 753–780.

[59] R. Frank, J. Schneid and C.W. Ueberhuber, Order results for implicit Runge–Kutta methods applied to stiff systems, *SIAM J. Numer. Anal.* 22 (1985) 515–534.

[60] C.W. Gear, Runge–Kutta starters for multistep methods, *ACM Trans. Math. Software* 6 (1980) 263–279.

[61] S. Gill, A process for the step-by-step integration of differential equations in an automatic digital computing machine, *Proc. Cambridge Philos. Soc.* 47 (1951) 95–108.

[62] A. Guillou and B. Lago, Domaine de stabilité associé aux formules d'intégration numérique d'équations différentielles, à pas séparés et à pas liés. Recherche de formules à grand rayon de stabilité, *Ier Congr. Assoc. Fran. Calcul, AFCAL,* Grenoble (September 1960) 43–56.

[63] A. Guillou and J.L. Soulé, La résolution numérique des problèmes différentiels aux conditions initiales par des méthodes de collocation, *R.I.R.O.* No. R-3 (1969) 17–44.

[64] E. Hairer, A Runge–Kutta method of order 10, *J. Inst. Math. Appl.* 21 (1978) 47–59.

[65] E. Hairer, A- and B-stability for Runge–Kutta methods—characterizations and equivalence, *Numer. Math.* 48 (1986) 383–389.

[66] E. Hairer, Backward analysis of numerical integrators and symplectic methods, *Ann. Numer. Math.* 1 (1994) 107–132.

[67] E. Hairer and Ch. Lubich, The life-span of backward error analysis for numerical integrators, *Numer. Math.*, to appear.

[68] E. Hairer, Ch. Lubich and M. Roche, *The Numerical Solution of Differential–Algebraic Systems by Runge–Kutta Methods*, Lecture Notes in Mathematics 1409 (Springer, Berlin, 1989).

[69] E. Hairer, S.P. Nørsett and G. Wanner, *Solving Ordinary Differential Equations* I. *Nonstiff Problems*, Springer Series in Computational Mathematics 8 (Springer, Berlin, 2nd ed., 1993).

[70] E. Hairer and H. Türke, The equivalence of B-stability and A-stability, *BIT* 24 (1984) 520–528.

[71] E. Hairer and G. Wanner, Algebraically stable and implementable Runge–Kutta methods of high order, *SIAM J. Numer. Anal.* 18 (1981) 1098–1108.

[72] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations* II. *Stiff and Differential–Algebraic Problems*, Springer Series in Computational Mathematics 14 (Springer, Berlin, 2nd ed., 1996).

[73] E. Hairer and G. Wanner, On the Butcher group and general multi-value methods, *Computing* 13 (1974) 1–15.

[74] P.C. Hammer and J.W. Hollingsworth, Trapezoidal methods of approximating solutions of differential equations, *MTAC* 9 (1955) 92–96.

[75] P. Henrici, *Discrete Variable Methods in Ordinary Differential Equations* (Wiley, New York, 1962).

[76] K. Heun, Neue Methode zur approximativen Integration der Differentialgleichungen einer unabhängigen Veränderlichen, *Z. Math. Phys.* 45 (1900) 23–38.

[77] M.K. Horn, Fourth and fifth-order scaled Runge–Kutta algorithms for treating dense output, *SIAM J. Numer. Anal.* 20 (1983) 558–568.

[78] W.H. Hundsdorfer and M.N. Spijker, A note on B-stability of Runge–Kutta methods, *Numer. Math.* 36 (1981) 319–331.

[79] A. Huťa, Une amélioration de la méthode de Runge–Kutta–Nyström pour la résolution numérique des équations différentielles du premier ordre, *Acta Math. Univ. Comenian.* 1 (1956) 201–224.

[80] L. Jay, Symplectic partitioned Runge–Kutta methods for constrained Hamiltonian systems, *SIAM J. Numer. Anal.* 33 (1996) 368–387.

[81] R. Jeltsch and O. Nevanlinna, Largest disk of stability of explicit Runge–Kutta methods, *BIT* 18 (1978) 500–502.

[82] R. Jeltsch and O. Nevanlinna, Stability of explicit time discretizations for solving initial value problems, *Numer. Math.* 37 (1981) 61–91; corrigendum: *Numer. Math.* 39, p. 155.

[83] R. Jeltsch and O. Nevanlinna, Stability and accuracy of time discretizations for initial value problems, *Numer. Math.* 40 (1982) 245–296.

[84] P.E. Kloeden and E. Platen, *Numeric Solution of Stochastic Differential Equations*, Applications of Mathematics 23 (Springer, Berlin, 1992; corr. 2nd printing, 1995).

[85] J.F.B.M. Kraaijevanger, B-convergence of the implicit midpoint rule and the trapezoidal rule, *BIT* 25 (1985) 652–666.

[86] J. Kuntzmann, Neuere Entwickelungen der Methode von Runge–Kutta, *Z. Angew. Math. Mech.* 41 (1961) 28–31.

[87] M.A. Kurdi, Stable high order methods for time discretization of stiff differential equations, Thesis, University of California (1974).

[88] W. Kutta, Beitrag zur näherungsweisen Integration totaler Differentialgleichungen, *Zeitschr. Math. Phys.* 46 (1901) 435–453.

[89] P. Laasonen, Über eine Methode zur Lösung der Wärmeleitungsgleichung, *Acta Math.* 81, 309–317.

[90] F.M. Lasagni, Canonical Runge–Kutta methods, *Z. Angew. Math. Phys.* 39 (1988) 952–953.

[91] W. Liniger, Zur Stabilität der numerischen Integrationsmethoden für Differentialgleichungen, Thèse, Université de Lausanne (1956).

[92] W.S. Loud, On the long-run error in the numerical solution of certain differential equations, *J. Math. Phys.* 28(1) (1949) 45–49.

[93] C. Lubich and A. Ostermann, Runge–Kutta methods for parabolic equations and convolution quadrature, *Math. Comp.* 60 (1993) 105–131.
[94] R.H. Merson, An operational method for the study of integration processes, *Proceedings Sympos. Data Processing,* Weapons Research Establishment, Salisbury, Australia (1957) 110-1–110-25.
[95] F.R. Moulton, *New Methods in Exterior Ballistics* (University Chicago Press, Chicago, 1926).
[96] S.P. Nørsett, Semi-explicit Runge–Kutta methods, Report No. 6/74, Dept. of Math., Univ. of Trondheim, Norway (1974).
[97] E.J. Nyström, Ueber die numerische Integration von Differentialgleichungen, *Acta Soc. Sci. Fenn.* 50 (1925) 1–54.
[98] H. Poincaré, *Les Méthodes Nouvelles de la Mécanique Céleste* III (Gauthier-Villars, Paris, 1899).
[99] A. Prothero and A. Robinson, On the stability and accuracy of one-step methods for solving stiff systems of ordinary differential equations, *Math. Comp.* 28 (1974) 145–162.
[100] A. Ralston, Runge–Kutta methods with minimum error bounds, *Math. Comp.* 16 (1962) 431–437; corr. 17, p. 488.
[101] H.H. Robertson, The solution of a set of reaction rate equations, in: J. Walsh, ed., *Numerical Analysis, an Introduction* (Academic Press, New York, 1966) 178–182.
[102] H.H. Rosenbrock, Some general implicit processes for the numerical solution of differential equations, *Computer J.* 5 (1962/63) 329–330.
[103] C. Runge, Ueber die numerische Auflösung von Differentialgleichungen, *Math. Ann.* 46 (1895) 167–178.
[104] C. Runge, Ueber die numerische Auflösung totaler Differentialgleichungen, *Göttinger Nachr.* (1905) 252–257.
[105] C. Runge, Über eine Methode die partielle Differentialgleichung $\Delta u =$ Constans numerisch zu integrieren, *Z. Math. Phys.* 56 (1908) 225–232.
[106] C. Runge and H. König, Vorlesungen über Numerisches Rechnen, Grundlehren der Mathematischen Wissenschaften 11 (Springer, Berlin, 1924).
[107] R.D. Ruth, A canonical integration technique, *IEEE Trans. Nuclear Sci.* 30 (1983) 2669–2671.
[108] H. Rutishauser, Über die Instabilität von Methoden zur Integration von gewöhnlichen Differentialgleichungen, *Z. Angew. Math. Phys.* 3 (1952) 65–74.
[109] A. Sandu, J.G. Verwer, M. van Loon, G.R. Carmichael, F.A. Potra, D. Dabdud and J.H. Seinfeld, Benchmarking stiff ODE solvers for atmospheric chemistry problems I: implicit versus explicit, Report NM-R9603, CWI, Amsterdam (January 1966).
[110] J.M. Sanz-Serna, Runge–Kutta schemes for Hamiltonian systems, *BIT* 28 (1988) 877–883.
[111] J.M. Sanz-Serna, Symplectic integrators for Hamiltonian problems: an overview, *Acta Numer.* 1 (1992) 243–286.
[112] J.M. Sanz-Serna, Geometric integration, in: *Proc. State of the Art in Numer. Anal.* (York, 1996), to appear.
[113] J.M. Sanz-Serna and M.P. Calvo, *Numerical Hamiltonian Problems*, Applied Mathematics and Mathematical Computation 7 (Chapman & Hall, London, 1994).
[114] L.F. Shampine, *Numerical Solution of Ordinary Differential Equations* (Chapman & Hall, London, 1994).
[115] L.F. Shampine, H.A. Watts and S.M. Davenport, Solving nonstiff ordinary differential equations—The state of the art, *SIAM Rev.* 18 (1976) 376–410.
[116] H.J. Stetter, Local estimation of the global discretization error, *SIAM J. Numer. Anal.* 8 (1971) 512–523.
[117] A.M. Stuart and A.R. Humphries, Model problems in numerical stability theory for initial value problems, *SIAM Rev.* 36 (1994) 226–257.
[118] Sun Geng, Symplectic partitioned Runge–Kutta methods, *J. Comput. Math.* 11 (1993) 365–372.
[119] Y.B. Suris, The canonicity of mappings generated by Runge–Kutta type methods when integrating the systems $\ddot{x} = -\partial U/\partial x$, *Zh. Vychisl. Mat. i Mat. Fiz.* 29 (1989) 202–211 (in Russian); English translation in: *Comput. Math. Math. Phys.* 29 (1991) 138–144.

[120] J.H. Verner, Explicit Runge–Kutta methods with estimates of the local truncation error, *SIAM J. Numer. Anal.* 15 (1978) 772–790.
[121] G. Wanner, E. Hairer and S.P. Nørsett, Order stars and stability theorems, *BIT* 18 (1978) 475–489; with an appendix *BIT* 18, 503.
[122] F.A. Willers, *Methoden der Praktischen Analysis* (Göschens Lehrbücherei, Berlin, 1928); later editions 1949, 1957; English translation: (Dover, New York, 1947).
[123] K. Wright, Some relationships between implicit Runge–Kutta collocation and Lanczos τ methods, and their stability properties, *BIT* 10 (1970) 217–227.
[124] H. Yoshida, Recent progress in the theory and application of symplectic integrators, *Celestial Mech. Dynam. Astronom.* 56 (1993) 27–43.