ELSEVIER

Applied Numerical Mathematics 20 (1996) 261–272

APPLIED NUMERICAL MATHEMATICS

The development of Runge–Kutta methods for partial differential equations

P.J. van der Houwen

CWI, P.O. Box 94079, 1090 GB Amsterdam, Netherlands

Abstract

A widely-used approach in the time integration of initial-value problems for time-dependent partial differential equations (PDEs) is the method of lines. This method transforms the PDE into a system of ordinary differential equations (ODEs) by discretization of the space variables and uses an ODE solver for the time integration. Since ODEs originating from space-discretized PDEs have a special structure, not every ODE solver is appropriate. For example, the well-known fourth-order Runge–Kutta method is highly inefficient if the PDE is parabolic, but it performs often quite satisfactory if the PDE is hyperbolic. In this lecture, we give a survey of the development of ODE methods that are tuned to space-discretized PDEs. Because of the overwhelming number of methods that have been proposed through the years, we confine our considerations to Runge–Kutta type methods. In this contribution to the historical surveys presented at the IMACS 14th World Congress held in July 1994 in Atlanta, we describe work of Crank and Nicolson (1947), Laasonen (1949), Peaceman and Rachford (1955), Yuan' Chzao-Din (1958), Stiefel (1958), Franklin (1959), Guillou and Lago (1960), Metzger (1967), Lomax (1968), Gourlay (1970), Riha (1972), Gentzsch and Schlüter (1978), Vichnevetsky (1983), Kinnmark and Gray (1984), Sonneveld and van Leer (1985), as well as research carried out at CWI.

Keywords: Numerical analysis; Method of lines; Runge–Kutta methods

1. The method of lines

The method of lines transforms initial-boundary value problems for time-dependent partial differential equations (PDEs) into initial-value problems (IVPs) for systems of ordinary differential equations (ODEs). This is achieved by discretization of the space variables using finite difference, finite element or finite volume approximations. The connection of PDEs with systems of ODEs was already known to Lagrange (see the historical notes in the book of Hairer, Nørsett and Wanner [10, p. 25]). In 1759 Lagrange already observed that his mathematical model for the propagation of sound in terms of a system of second-order ODEs is related to d'Alembert's equation $u_{tt} = u_{xx}$ for the vibrating string. However, the actual use of the space-discretized approximation in numerically solving initial-boundary

0168-9274/96/$15.00 © 1996 Elsevier Science B.V. All rights reserved
SSDI 0168-9274(95)00109-3

value problems for PDEs seems to start with Rothe in 1930 [32], and is therefore also, called Rothe's method (see [10, p. 3]).

In this paper, we shall restrict our considerations to the case where the spatial discretization of the PDE leads to an IVP of the form

$$\frac{\mathrm{d}\boldsymbol{y}(t)}{\mathrm{d}t} = \boldsymbol{f}\big(t, \boldsymbol{y}(t)\big), \quad \boldsymbol{y}(t_0) = \boldsymbol{y}_0, \tag{1.1}$$

where t is the time variable and $\boldsymbol{y}_0$ contains the given initial values. Notice that the boundary conditions are lumped into the right-hand side function $\boldsymbol{f}$.

The IVP (1.1) has a number of specific characteristics that play a crucial role in selecting a suitable integrator. Firstly, the system (1.1) can be extremely large, particularly, if it originates from a problem with 2 or 3 spatial dimensions. Secondly, the system is usually extremely stiff (here, (1.1) is considered to be stiff if the solution components corresponding to eigenvalues of the Jacobian $\partial \boldsymbol{f}/\partial \boldsymbol{y}$ that are close to the origin are dominating). Thirdly, the required order of accuracy in time is rather modest (usually not exceeding the order of the spatial discretization, that is, at most order three). Hence, we are led to look for low-order, stiff ODEIVP solvers that are storage economic.

One approach is to look for conventional, general purpose ODEIVP methods that meet these requirements. There are two often used integrators, the second-order trapezoidal rule and the first-order backward Euler method, respectively used by Crank and Nicolson [4] and by Laasonen [24] in their papers of 1947 and 1949 for solving heat flow problems. In the PDE literature, these methods also known as the *Crank–Nicolson* and *Laasonen methods*. An integration method that combines the second-order accuracy of the Crank–Nicolson method and the high stability of the Laasonen method is offered by the two-step method based on backward differentiation (known as the *BDF2 method*). BDF methods were proposed in 1952 by Curtiss and Hirschfelder [3] for solving stiff ODEs and became popular by the papers of Gear in 1967–1968, and in particular by his book [7] of 1971. The Crank–Nicolson, Laasonen and BDF2 methods are applicable to a wide class of space-discretized PDEs (not only heat flow problems) and have comparable computational complexity. In order to solve the implicit relations, one usually applies Newton iteration which leads to a large linear system in each iteration. For one-dimensional problems, these linear systems can be solved by *direct* methods that are in general highly efficient because the band structure of the system can be fully exploited. However, in more than one spatial dimension, direct solution methods usually are out of the question and we have to resort to an *iterative* method. If L_{N} denotes the number of Newton iterations, L_{s} the number of linear system iterations, d the spatial dimension, and Δ the spatial grid size, then the computational complexity of these methods is $\mathrm{O}(L_{\mathrm{N}} L_{\mathrm{s}} \Delta^{-d})$. Often used linear-system-iteration methods are conjugate gradient type methods that require at least $\mathrm{O}(\Delta^{-1/2})$ iterations. Hence, the total computational work involved for integrating the unit time interval with stepsize h is at least $W = \mathrm{O}(L_{\mathrm{s}} h^{-1} \Delta^{-d-1/2})$.

In order to reduce the huge amount of work when integrating higher-dimensional problems, new methods have been developed. The remainder of this paper will be devoted to such methods. Since it is not feasible to present a complete survey, we shall confine ourselves to Runge–Kutta type methods that are tuned to PDEs in two or more spatial dimensions. We shall discuss explicit Runge–Kutta (RK) methods for parabolic and hyperbolic problems (spectrum of the Jacobian $\partial \boldsymbol{f}/\partial \boldsymbol{y}$ along the negative axis and imaginary axis, respectively), and splitting methods represented as RK methods with fractional stages.

2. Explicit Runge–Kutta methods

Consider the s-stage RK method

$$\begin{aligned} \boldsymbol{Y} &= \boldsymbol{e} \otimes \boldsymbol{y}_n + h(A \otimes I)\boldsymbol{F}(t_n\boldsymbol{e} + \boldsymbol{c}h, \boldsymbol{Y}), \\ \boldsymbol{y}_{n+1} &= \boldsymbol{y}_n + h(\boldsymbol{b}^{\mathrm{T}} \otimes I)\boldsymbol{F}(t_n\boldsymbol{e} + \boldsymbol{c}h, \boldsymbol{Y}), \end{aligned} \tag{2.1}$$

where h is the integration step, $\boldsymbol{y}_n$ and $\boldsymbol{y}_{n+1}$ represent approximations to the exact solution vector $\boldsymbol{y}(t)$ at $t = t_n$ and $t = t_{n+1}$, $\otimes$ denotes the Kronecker product, the s-dimensional vector $\boldsymbol{e}$ is the vector with unit entries, I is the identity matrix whose dimension equals that of the IVP, and the s-by-s matrix A and the s-dimensional vectors $\boldsymbol{b}$ and $\boldsymbol{c} := A\boldsymbol{e}$ contain the RK parameters. The s components $\boldsymbol{Y}_i$ of $\boldsymbol{Y}$ represent intermediate approximations to the exact solution values $\boldsymbol{y}(t_n + c_i h)$ and $\boldsymbol{F}(t_n\boldsymbol{e} + \boldsymbol{c}h, \boldsymbol{Y})$ contains the derivative values $(\boldsymbol{f}(t_n + c_i h, \boldsymbol{Y}_i))$. In the following, the dimensions of $\boldsymbol{e}$ and I may vary, but will always be clear from the context in which they appear.

If A is strictly lower triangular, then (2.1) defines an *explicit* RK method (the first method of this type was proposed by Runge [33] about 100 years ago). Explicit RK methods are relatively cheap, provided that the integration step h can be chosen sufficiently large. For stiff ODEs, the step is restricted by a stability condition of the form

$$h < \frac{\beta}{\rho(J_n)}, \quad J_n := \frac{\partial \boldsymbol{f}(t_n, \boldsymbol{y}_n)}{\partial \boldsymbol{y}}, \tag{2.2}$$

where $\rho(J_n)$ is the spectral radius of J_n and β is the so-called stability boundary. In the case of *parabolic* and *hyperbolic* problems, where the Jacobian of the right-hand side function respectively has (more or less) negative and imaginary eigenvalues, β denotes the *real* stability boundary β_{real} or the *imaginary* stability boundary β_{imag} of the RK method. The real stability boundary is defined by the maximum length of the negative interval $(-\beta, 0)$ that is contained in the region where the *stability polynomial* $R_s(z) := 1 + \boldsymbol{b}^{\mathrm{T}}(I - zA)^{-1}\boldsymbol{e}$ assumes values within the unit circle. Similarly, the imaginary stability boundary is defined by the maximum length of the interval $(0, \mathrm{i}\beta)$ where R_s is bounded by 1.

2.1. Conventional RK methods

For conventional RK methods, $R_s(z)$ is given by the Taylor polynomial of degree s in z, that is, the polynomial that coincides with the truncated Taylor expansion of $\exp(z)$ at $z = 0$. Let us first consider the parabolic case. The *real* stability boundary of Taylor polynomials is (approximately) given by (cf. [14, p. 236] and [20,21])

$$\beta_{\mathrm{real}} \approx 0.368(s+1)\sqrt[2(s+1)]{19(s+1)}. \tag{2.3}$$

This approximation is already quite close for $s \geqslant 4$. We conclude from (2.2) and (2.3) that we can take any step we want by choosing s sufficiently large, but these formulas also show that for large s the total number of function calls needed for integrating the unit interval with maximum step $h = \beta_{\mathrm{real}}\rho^{-1}(J_n) \approx 0.368s\rho^{-1}(J_n)$ is given by $N_f \approx 2.7\rho(J_n)$, that is, *independent of* s. Hence, conventional RK methods are as costly as the explicit Euler method (but of course highly accurate as s increases). Since $\boldsymbol{f}$ has $\mathrm{O}(\Delta^{-d})$ components and since for parabolic problems $\rho(J_n) = \mathrm{O}(\Delta^{-2})$, the computational work can be estimated by $W = \mathrm{O}(\Delta^{-d-2})$. This differs by a factor of order

$O(h\Delta^{-3/2})$ from the estimate derived for the Crank–Nicolson, Laasonen and BDF2 methods (when applied to higher-dimensional problems). Usually, this factor is quite large (e.g., if $h = O(\Delta)$), so that conventional RK methods are not the way to solve space-discretized PDEs of *parabolic* type. They are "too costly and too accurate".

Next we consider the hyperbolic case. It happens that for the imaginary stability boundary β_{imag} we do not always obtain nonzero values. If $z^{-(p+1)}[R_s(z) - \exp(z)] \to C_{p+1}$ as $z \to 0$, where p denotes the order of accuracy of the RK method, then it can straightforwardly been shown that β_{imag} is only nonzero if either $C_{p+1}i^p < 0$ for p even or $C_{p+1}i^{p+1} < 0$ for p odd. For the Taylor polynomials this implies that the imaginary stability interval is empty for $p = 1, 2, 5, 6, 9, 10, \ldots$. For the other orders, quite reasonable values are obtained. For example, for $p = 3, 4, 7, 8$, we have $\beta_{\text{imag}} \approx 1.7, 2.8, 1.7, 3.4$. Taking one of these latter methods and assuming that $\rho(J_n) = O(\Delta^{-1})$, the total computational work associated with the unit interval can be estimated by $W = O(\Delta^{-d-1})$. This is a factor of order $O(h\Delta^{-1/2})$ better than the estimate derived for the Crank–Nicolson, Laasonen and BDF2 methods. Hence, unlike the situation for parabolic problems, conventional RK methods seem to be preferrable for *hyperbolic* problems.

2.2. *Parabolic RK methods*

Our conclusion that for parabolic problems explicit RK methods are "too costly and too accurate" suggests sacrifycing accuracy in order to reduce computational costs. By observing that an s-stage RK method of order p with $s > p$ possesses a stability polynomial R_s of the form

$$R_s^{(p)}(z) := \beta_0 + \beta_1 z + \beta_2 z^2 + \cdots + \beta_s z^s, \quad \beta_j = \frac{1}{j!},\ j = 0, \ldots, p, \tag{2.4}$$

where the coefficients β_j, $j = p+1, \ldots, s$, are free parameters, it is natural to use these free parameters for obtaining larger stability boundaries. For parabolic problems, where the eigenvalues of the Jacobian often are along the negative axis, we are led to construct polynomials $R_s^{(p)}(z)$ with increased *real* stability boundary. Having found an appropriate stability polynomial $R_s^{(p)}$, it is always possible to construct an RK method with $R_s^{(p)}$ as its stability polynomial (see, e.g., [14]). Such methods will be called *parabolic RK methods*.

Until now, closed form solutions for the polynomials with maximal real stability boundaries (to be called *optimal polynomials*) are only known for $p = 1$. They are given by the shifted Chebyshev polynomials

$$C_s^{(1)}(z) := T_s\left(1 + \frac{z}{s^2}\right), \quad \beta_{\text{real}} = 2s^2, \tag{2.5}$$

where $T_s(z) := \cos(s \arccos(z))$ denotes the first kind Chebyshev polynomial of degree s. They have been rediscovered in the literature again and again (even in recent years, see, e.g., [2]). As far as I know, they were first mentioned for integrating parabolic equations: in 1958 by Yuan' Chzao-Din in his thesis [41], in 1959 by Franklin in his paper [6] that appeared in the *Journal of Mathematical Physics*, and in 1960 by Guillou and Lago in the proceedings [9] of the first conference of AFCAL (the French Association for Computing). These authors were not aware of each other's work.

For $p \geqslant 2$, only approximate solutions have been constructed. In the thesis of Metzger [28] in 1967, we find numerical approximations for $p \leqslant 4$, $s \leqslant 5$, and in a NASA report of Lomax [27] of 1968,

a general approach for computing the coefficients was indicated. Lomax conjectured that the optimal polynomials satisfy the so-called *equal ripple* property, that is, the optimal polynomial has $s-p$ local extrema $+1$ or -1 (this property was actually proved by Riha [31] in 1972 who also showed the unique existence of the optimal polynomials for all p and all $s > p$). Using the equal ripple property, an iterative method can be constructed for the numerical computation of the coefficients. However, this equal-ripple-iteration method needs rather accurate initial iterates in order to converge. Presumably for this reason, Lomax did not use the equal-ripple-property approach, and instead, computed least squares approximations for $p = 2$ and $s \leqslant 10$. Again, Metzger, Lomax and Riha found their results independently.

At CWI we used the least squares approach of Lomax for generating initial iterates to start the equal-ripple-iteration method. In this way, we computed the optimal stability polynomials, together with their real stability boundaries, for $p \leqslant 4$ and $s \leqslant 10+p$ (tables for the coefficients can be found in [13,14]). These computations indicated that β_{real} increases quadratically with s as s increases. In fact, we found

$$\beta_{\text{real}} \approx \gamma_p s^2 \quad \text{as } s \to \infty, \qquad \gamma_2 = 0.814, \qquad \gamma_3 = 0.489, \qquad \gamma_4 = 0.341. \tag{2.6}$$

The quadratic behaviour is important. It implies that the total number of function calls needed for integrating the unit interval with maximum step $h \approx \gamma_p s^2 \rho^{-1}(J_n)$ is now given by $N_f \approx (\gamma_p s)^{-1}\rho(J_n)$, which is a factor $2.7\gamma_p s$ less than the number of function calls needed for conventional RK methods. Hence, for large values of s, RK methods generated by (2.5) are much cheaper than conventional RK methods, provided that they are available for large values of s. Unfortunately, the numerical computation of the optimal polynomials becomes increasingly more difficult as s increases. This motivated us to look for analytical expressions for nearly optimal polynomials that are valid for arbitrary high values of s. In 1971, Bakker [1] derived in his Master thesis for $p = 2$ and $p = 3$ analytically given polynomials which are quite close approximations to the optimal stability polynomials, in the sense that the stability boundaries are close to the maximal attainable values. These polynomials, to be called the *Bakker polynomials*, are given by

$$B_s^{(2)}(z) = \frac{2s^2+1}{3s^2} + \frac{s^2-1}{3s^2} T_s\left(1 + \frac{3z}{s^2-1}\right), \quad \beta_{\text{real}} \approx \frac{2}{3}(s^2-1), \qquad s > 2, \tag{2.7}$$

$$B_s^{(3)}(z) = 1 + \frac{3\beta^2 - 2(40k^2-1)\beta}{576k^4} - \frac{3\beta^2 - 2(36k^2-1)\beta}{512k^4} T_{2k}\left(1 + \frac{2z}{\beta}\right) \tag{2.8}$$

$$+ \frac{3\beta^2 - 2(4k^2-1)\beta}{4608k^4} T_s\left(1 + \frac{2z}{\beta}\right), \qquad k := \frac{s}{6}, \quad s = 6, 12, 18, \ldots,$$

$$\beta_{\text{real}} \approx \beta := \frac{2}{9}s^2 - 1 + \frac{1}{9}\sqrt{\frac{8s^4 - 60s^2 + 297}{5}} \approx \frac{2}{9}s^2\left(1 + \sqrt{\frac{2}{5}}\right) \approx 0.363 s^2 \quad \text{as } s \to \infty,$$

where again T_s denotes the first kind Chebyshev polynomial of degree s (in addition, Bakker actually proved the quadratic behaviour of the real stability boundaries of the optimal polynomials and obtained lower and upper bounds for γ_p up to $p = 15$). A comparison of (2.6) with (2.7) and (2.8) reveals that the Bakker polynomials respectively possess 80% and 75% of the maximal attainable, asymptotic stability boundary. Later on in 1982, we found for $p = 2$ an even better approximation given by (cf. [17])

$$A_s^{(2)}(z) = \frac{2}{2-z} - \frac{z}{2-z} T_s\left(\cos(\pi/s) + z\frac{1-\cos(\pi/s)}{2}\right), \tag{2.9}$$

$$\beta_{\text{real}} = \frac{2}{[\tan(\pi/2s)]^2} \approx 8\frac{s^2}{\pi^2} \approx 0.810\, s^2 \quad \text{as } s \to \infty.$$

These polynomials are not the optimal ones, but yield 99.5% of the maximal attainable, asymptotic stability boundary!

The parabolic RK methods generated by the analytically given polynomials (2.5), (2.7) and (2.9) enable us to select an integration step h on the basis of accuracy considerations and to adapt the number of stages according to the stability condition $s \approx \sqrt{\gamma_p^{-1} h\rho(J_n)}$. Hence, effectively, we have an *unconditionally stable* method.

As we remarked earlier, given the stability polynomial, many RK methods possessing this stability polynomial are possible. One of the most simple implementations of first-order or second-order RK methods with stability polynomial (2.4) reads

$$\begin{aligned} &\boldsymbol{Y}_i = \boldsymbol{y}_n + a_i h \boldsymbol{f}(t_n + c_{i-1}h, \boldsymbol{Y}_{i-1}), \quad a_i := \frac{\beta_{s-i+2}}{\beta_{s-i+1}}, \ i = 1, \ldots, s, \\ &\boldsymbol{y}_{n+1} = \boldsymbol{y}_n + h\boldsymbol{f}(t_n + h, \boldsymbol{Y}_s), \end{aligned} \tag{2.10}$$

where a_1 is assumed to vanish. This implementation is of the form (2.1) with $\boldsymbol{b} = \boldsymbol{e}_s$ and a matrix A with zero entries except for the lower off-diagonal entries. We shall call (2.10) the *diagonal* implementation. Unfortunately, when we actually applied the diagonal implementation with $s \approx \sqrt{\gamma_p^{-1} h\rho(J_n)}$, it turned out that the numerical solution lost accuracy for larger values of s. On a computer with 14 digits arithmetic, s should not be greater than 12. This is caused by the development of *internal* instabilities within a single step. Just as the step values $\boldsymbol{y}_n$ are required to be stable by imposing the (external) stability condition $h < \beta_{\text{real}}/\rho(J_n)$, we also have to require that the internal values $\boldsymbol{Y}_i$ are stable. In the implementation (2.10), the internal perturbations satisfy the recursion $\Delta\boldsymbol{Y}_i = a_i h J_n \Delta\boldsymbol{Y}_{i-1} = R_{i-1}(hJ_n)\Delta\boldsymbol{Y}_1$, where the so-called *internal stability polynomials* $R_i(z)$ are of degree i in z. This leads to the *internal* stability conditions $h < \alpha_i/\rho(J_n)$, $i = 1, \ldots, s$, where α_i denotes the stability boundary associated with R_i. For large values of s, these conditions are much more restrictive than the external stability condition $h < \beta_{\text{real}}/\rho(J_n)$. As a consequence, the main advantage of the polynomials (2.5), (2.7) and (2.9), viz. that they are available for arbitrarily large values of s, cannot be exploited.

Fortunately, it is possible to avoid, or at least to suppress the internal instabilities, just by choosing another implementation than (2.10). The first attempt to internal stabilization of RK methods with many stages is due to Gentzsch and Schlüter [8] in 1978, who 'rediscovered' the shifted Chebyshev polynomials (2.5) and exploited the fact that these polynomials possess s real zeroes z_i on the negative axis. Although their approach was restricted to linear IVPs, it can directly be extended to nonlinear problems to obtain an RK method of the form

$$\begin{aligned} &\boldsymbol{Y}_1 = \boldsymbol{y}_n, \quad \boldsymbol{Y}_{i+1} = \boldsymbol{Y}_i - \frac{1}{z_i} h\boldsymbol{f}(t_n + c_i h, \boldsymbol{Y}_i), \quad i = 1, \ldots, s-1, \\ &\boldsymbol{y}_{n+1} = \boldsymbol{Y}_s - \frac{1}{z_s} h\boldsymbol{f}(t_n + c_s h, \boldsymbol{Y}_s). \end{aligned} \tag{2.11}$$

This implementation may be interpreted as an RK method that is factorized in a sequence of Euler steps and will be called the *factorized* implementation. If the zeroes z_i are ordered such that $z_i < z_{i+1}$ or $z_i > z_{i+1}$, then the performance the factorized implementation is hardly better than that of the diagonal implementation as s increases. However, Gentzsch and Schlüter reported satisfactory results for extremely large values of s (up to 997) if special orderings of the z_i are used. A disadvantage in actual applications is that a suitable ordering depends on s.

When reading the paper of Gentzsch and Schlüter, we suddenly realized that the problem of internal stabilization was already solved a long time ago by numerical analysts working in *elliptic* PDEs! The spatial discretization of elliptic PDEs leads to the problem of solving linear systems $A\boldsymbol{y} = \boldsymbol{b}$, where A is known to have a negative spectrum in the negative interval $(-\rho(A), 0)$ with $\rho(A)$ large positive. A well-known iterative method for solving such problems is due to Richardson, who proposed in his paper [30] of 1910 the recursion $\boldsymbol{y}_i = \boldsymbol{y}_{i-1} + \alpha_i(A\boldsymbol{y}_{i-1} - \boldsymbol{b})$, where the parameters α_i are chosen such that after s iterations, the polynomial P_s occurring in the error formula $\boldsymbol{y}_s - \boldsymbol{y} = P_i(A)(\boldsymbol{y}_0 - \boldsymbol{y})$ has a small norm in the eigenvalue interval (a, b) of A. Various approaches to achieve this have been proposed. Richardson suggested choosing P_s such that it has uniformly distributed zeros in (a, b), Stiefel proposed to minimize an integral measure of P_s (cf. [36]), but most numerical analysts prefer to minimize the maximum norm of P_s. The latter approach leads to shifted Chebyshev polynomials that are very similar to (2.5). This process is now known as Richardson's method of *first* degree. However, application of this method for large values of s suffers the same internal instability as the method (2.11). Just as Gentzsch and Schlüter, one has tried to improve the stability by special choices of the ordering of the parameters α_i (see, e.g., the experiments of Young [40] in 1954), but a real break-through was due to Stiefel [36] in 1958. He observed that Chebyshev polynomials satisfy a *stable* three-term recursion, so that using a three-term recursion for the iterates $\boldsymbol{y}_i$, rather than the two-term recursion of Richardson, would avoid the instability problem. This two-step iteration method is known as Richardson's method of *second* degree or, in the more recent literature, the *Chebyshev semi-iterative* method. Realizing that the stability polynomials (2.5), (2.7) and (2.9) are also expressions in terms of shifted Chebyshev polynomials, brought us to construct internally stable implementations of the corresponding parabolic RK methods (cf. [15,16]). For the second-order consistent polynomia $A_s^{(2)}$ and $B_s^{(2)}$, it was pointed out by Sommeijer (see [15]) that it is even possible to make the $\boldsymbol{Y}_i$ no only stable, but also second-order accurate approximations to the exact solution at the intermediate points $t_n + c_i h$, $i = 1, \ldots, s$.

The internally stable Runge–Kutta method generated by the Bakker polynomials $B_s^{(2)}$ performs slightly better than the method generated by $A_s^{(2)}$ (its smaller stability boundary is compensated by its smaller error constants). It is a highly efficient integrator for general heat flow problems, particularly for 2D and 3D problems. We called it the *Runge–Kutta–Chebyshev method*, but is could equally well have been called the *Runge–Kutta–Bakker method*. A detailed study of its convergence is presented in [37] and an extensive performance evaluation can be found in [12]. The Runge–Kutta–Chebyshev method has been implemented by Sommeijer as the code RKC and is available through netlib [34].

Another code that is based on stabilized RK methods is the code DUMKA developed by Lebedev and his coworkers of the Institute for Numerical Mathematics of the Russian Academy of Science. They approximate the optimal stability polynomials by so-called Zolotarev polynomials. Like Gentzsch and Schlüter, internal stability is achieved by a special ordering of the stages rather than using recurrence relations. More details can be found in the references [25,26].

Finally, we compare the total computational work of conventional and parabolic RK methods needed for integrating the unit interval with a given step h. Assuming that s is defined by

$$s \approx \sqrt{\gamma_p^{-1} h\rho(J_n)},$$

we find for the stabilized RK methods

$$W = h^{-1} s\, O(\Delta^{-d}) \approx h^{-1}\sqrt{\gamma_p^{-1} h\rho(J_n)}\, O(\Delta^{-d}) \approx O(h^{-1/2}\Delta^{-d-1}).$$

Comparing this estimate with that derived for conventional RK methods, we see that the computational complexity of the stabilized RK methods differ by a factor of order $O(h^{1/2}\Delta^{-1})$. With respect to the Crank–Nicolson, Laasonen and BDF2 methods using conjugate gradient type iteration methods, the stabilized RK methods are at least competitive.

2.3. *Hyperbolic RK methods*

Instead of maximizing the *real* stability boundary of stability polynomials of the form (2.4), we may also maximize the *imaginary* stability boundary, to obtain a *hyperbolic RK method* that should be suitable for integrating hyperbolic problems that have Jacobians with imaginary eigenvalues.

For $p = 1$, the optimal polynomials are given by

$$I_s^{(1)}(z) = (-\mathrm{i})^s \left[\mathrm{i} T_{s-1}\left(\frac{\mathrm{i}z}{s-1}\right) - \left(1 + \frac{z^2}{(s-1)^2}\right) U_{s-2}\left(\frac{\mathrm{i}z}{s-1}\right)\right], \tag{2.12}$$

$$\beta_{\mathrm{imag}} = s - 1, \quad s \geqslant 2,$$

where $U_s(z) := \sin((s+1)\arccos(z))/\sin(\arccos(z))$ denotes the second kind Chebyshev polynomial of degree s. For *odd* values of s, these polynomials were given in 1972 in [13] (a proof can be found in [14]). At the time, it was not realized that (2.12) is also valid for *even* values of s, because in [13] the polynomials $I_s^{(1)}$ were represented in the form

$$I_s^{(1)}(z) = T_k\left(1 + \frac{z^2}{2k^2}\right) + \frac{z}{k}\left(1 + \frac{z^2}{4k^2}\right) U_{k-1}\left(1 + \frac{z^2}{2k^2}\right), \tag{2.13}$$

$$\beta_{\mathrm{imag}} = s - 1, \quad s = 2k + 1, \ k \geqslant 1,$$

which cannot directly be extended to even values of s. It turns out that the odd-degree polynomials are identical to the optimal polynomials corresponding to $p = 2$, i.e., $I_s^{(2)}(z) = I_s^{(1)}(z)$ for s odd.

In 1984 Kinnmark and Gray [22] derived the representation (2.12) which is valid for all values of s. This result was also obtained, independently, by Sonneveld and van Leer [35] in 1985.

Kinnmark and Gray [23] have also derived approximations to the optimal polynomials $I_s^{(3)}$ for s odd and to $I_s^{(4)}$ for s even. These *Kinnmark–Gray polynomials* are given by

$$K_s^{(3)}(z) = \frac{1}{\beta^2 + 1}\left[1 + z + \mathrm{i}^{s-1}\beta^2 T_{s-1}\left(\frac{\mathrm{i}z}{\beta}\right) + \frac{1}{2}\mathrm{i}^{s+2}\beta\left\{(s-2)T_s\left(\frac{\mathrm{i}z}{\beta}\right) - sT_{s-2}\left(\frac{\mathrm{i}z}{\beta}\right)\right\}\right], \tag{2.14}$$

$$\beta_{\mathrm{imag}} = \beta := \sqrt{(s-1)^2 - 1}, \quad \text{odd } s \geqslant 3,$$

and

$$K_s^{(4)}(z) = \sqrt{\frac{1}{\beta^2+1}}\left[\mathrm{i}^{s+1}\beta T_{s-1}\left(\frac{\mathrm{i}z}{\beta}\right) + \frac{1}{2}\mathrm{i}^s\left\{(s-2)T_s\left(\frac{\mathrm{i}z}{\beta}\right) - sT_{s-2}\left(\frac{\mathrm{i}z}{\beta}\right)\right\}\right], \tag{2.15}$$

$$\beta_{\mathrm{imag}} = \beta := \sqrt{(s-1)^2-1}, \quad \text{even } s \geqslant 4.$$

Earlier, in 1983, Vichnevetsky [38] had already proved that $\beta_{\mathrm{imag}} \leqslant s-1$ for all p and s. Hence, this result of Vichnevetsky indicates that the Kinnmark–Gray polynomials are extremely close approximations to the optimal ones. However, it also indicates that, unlike the situation for parabolic problems, *hyperbolic* RK methods are hardly more effective than conventional RK methods with nonempty imaginary stability intervals.

3. Splitting methods

Just as RK methods, splitting methods compute in each step two or more intermediate stages. However, unlike RK methods, these stages are not expressed in the full right-hand side of the PDE, but in *fractions* of the right-hand side. Almost all splitting methods proposed in the literature can be represented in RK format. This approach was followed in [19] to develop a unified treatment of splitting methods and allows a straightforward derivation of the order conditions and stability functions.

Suppose that the right-hand side function in (1.1) is split according to

$$\boldsymbol{f}(t,\boldsymbol{y}(t)) = \sum_{i=1}^{\sigma} \boldsymbol{f}_i(t,\boldsymbol{y}(t)), \tag{3.1}$$

and consider the RK type method

$$\boldsymbol{Y} = \boldsymbol{e}\otimes\boldsymbol{y}_n + h\sum_{k=1}^{\sigma}\left(A^{(k)}\otimes I\right)\boldsymbol{F}_k\left(t_n\boldsymbol{e}+\boldsymbol{c}^{(k)}h,\boldsymbol{Y}\right), \tag{3.2}$$

$$\boldsymbol{y}_{n+1} = \left(\boldsymbol{e}_s^{\mathrm{T}}\otimes I\right)\boldsymbol{Y},$$

where $\boldsymbol{F}_k(t_n\boldsymbol{e}+\boldsymbol{c}^{(k)}h,\boldsymbol{Y})$ contains the derivative values $(\boldsymbol{f}_k(t_n+c_j^{(k)}h,\boldsymbol{Y}_j))$. If $\sigma=1$, then (3.2) reduces to the RK method (2.1) with $\boldsymbol{b}^{\mathrm{T}} = \boldsymbol{e}_s^{\mathrm{T}}A$. The method {(3.1), (3.2)} will be called a *σ-term RKS method with s fractional stages*. RKS methods consist of two components, the right-hand side splitting (3.1) and the splitting scheme (3.2).

Restricting our discussion to first-order and second-order methods and using the compact notation in terms of the matrices $A^{(k)}$, we have first-order accuracy if

$$\boldsymbol{e}_s^{\mathrm{T}}A^{(j)}\boldsymbol{e} = 1, \quad j=1,\ldots,\sigma, \tag{3.3}$$

and second-order accuracy if, in addition,

$$\boldsymbol{e}_s^{\mathrm{T}}A^{(j)}A^{(k)}\boldsymbol{e} = \frac{1}{2}, \quad j,k=1,\ldots,\sigma. \tag{3.4}$$

In principle, the abscissa vectors are defined by $\boldsymbol{c}^{(k)} := A^{(k)}\boldsymbol{e}$. However, in actual computations, the time-dependent parts originating from time-dependent boundary conditions, need a more careful treatment. In this overview, we shall not elaborate on this aspect of splitting methods (see, e.g., [5]).

The linear stability of RKS methods can be analysed by means of the linear test equation

$$y'(t) = \sum_{k=1}^{\sigma} J_k y(t), \tag{3.5}$$

where J_k is the Jacobian matrix $\partial f_k(y_n)/\partial y$. It will be assumed that J_k has its eigenvalues in the left halfplane. Defining $Z_k = hJ_k$, $k = 1, \ldots, \sigma$, we deduce from (3.2)

$$Y = e \otimes y_n + \sum_{k=1}^{\sigma} (A^{(k)} \otimes Z_k) Y = (I - S)^{-1}(e \otimes y_n), \quad S := \sum_{k=1}^{\sigma} (A^{(k)} \otimes Z_k).$$

Hence,

$$y_{n+1} = (e_s^{\mathrm{T}} \otimes I) Y = (e_s^{\mathrm{T}} \otimes I)(I - S)^{-1}(e \otimes y_n) = (e_s^{\mathrm{T}} \otimes I)(I - S)^{-1}(e \otimes I) y_n.$$

Thus, the *stability function* is given by

$$R = (e_s^{\mathrm{T}} \otimes I)\left((I \otimes I) - \sum_{k=1}^{\sigma} (A^{(k)} \otimes Z_k)\right)^{-1} (e \otimes I). \tag{3.6}$$

3.1. Splitting methods as RKS methods

This survey paper is concluded with an example of a family of splitting methods that can be represented as an RKS method. For a more detailed analysis of RKS methods with an application to transport problems in three spatial dimensions, we refer to [18].

Consider the two-term, three-stage splitting scheme defined by

$$A^{(1)} = \begin{pmatrix} 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}, \qquad A^{(2)} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 \end{pmatrix}. \tag{3.7}$$

This scheme is second-order accurate whatever we choose for f_1 and f_2. Presumably, the first splitting method proposed in the literature generated by the splitting scheme (3.7) is the Peaceman–Rachford method [29] of 1955. If (3.7) is applied to a space-discretized, two-dimensional PDE in which the right-hand side f can be split into an x-dependent part f_1 and a y-dependent part f_2, then the so-called ADI (Alternating Direction Implicit) method of Peaceman and Rachford is obtained. Other well-known splitting methods generated by (3.7) are the Hopscotch methods proposed by Gourlay in 1970. These methods are obtained by dividing the grid points on which the PDE is discretized in two groups G_1 and G_2, and by defining f_1 and f_2 such that they vanish on G_1 and G_2, respectively. On rectangular grids, often used examples are the Line Hopscotch and the Odd–Even Hopscotch methods which arise if G_1 and G_2 contain grid points lying on alternating lines and diagonals, respectively.

Acknowledgement

The author is grateful to Dr. B.P. Sommeijer for his interest in this survey paper and for his many comments to improve the presentation of the available material.

References

[1] M. Bakker, Analytic aspects of a minimax problem, Report TN 62, Mathematisch Centrum, Amsterdam (1971) (in Dutch).

[2] K. Burrage, Order and stability properties of explicit multivalue methods, *Appl. Numer. Math.* 1 (1985) 363–379.

[3] C.F. Curtiss and J.O. Hirschfelder, Integration of stiff equations, *Proc. Nat. Acad. Sci. U.S.* 38 (1952) 235–243.

[4] 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 (1947) 50–67.

[5] G. Fairweather and A.R. Mitchell, A new computational procedure for A.D.I. methods, *SIAM J. Numer. Anal.* 4 (1967) 163–170.

[6] J.N. Franklin, Numerical stability in digital and analogue computation for diffusion problems, *J. Math. Phys.* 37 (1959) 305–315.

[7] C.W. Gear, *Numerical Initial Value Problems in Ordinary Differential Equations* (Prentice Hall, Englewood Cliffs, NJ, 1971).

[8] W. Gentzsch and A. Schlüter, On one-step methods with cyclic stepsize changes for solving parabolic differential equations, *Z. Angew. Math. Mech.* 58 (1978) T415–T416 (in German).

[9] A. Guillou and B. Lago, Stability regions of one-step and multistep formulas for differential equations: investigation of formulas with large stability boundaries, 1^{e} Congres de l'Association Francaise de Calcul, AFCAL, Grenoble (1960) 43–56 (in French).

[10] E. Hairer, S.P. Nørsett and G. Wanner, *Solving Ordinary Differential Equations*, I: *Nonstiff Problems* (Springer, Berlin, 1987/91).

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

[12] S. Hofmann, First and second-order Runge–Kutta–Chebyshev methods for the numerical integration of parabolic differential equations and stiff ordinary differential equations, Master's thesis, University of Wuppertal, Germany (1992) (in German).

[13] P.J. van der Houwen, Explicit Runge–Kutta methods with increased stability boundaries, *Numer. Math.* 20 (1972) 149–164.

[14] P.J. van der Houwen, *Construction of Integration Formulas for Initial-Value Problems* (North-Holland, Amsterdam, 1977).

[15] P.J. van der Houwen, On the time integration of parabolic differential equations, in: G.A. Watson, ed., *Numerical Analysis*, Lecture Notes in Mathematics 912 (Springer, New York, 1981) 157–168.

[16] P.J. van der Houwen and B.P. Sommeijer, On the internal stability of explicit m-stage Runge–Kutta methods for large values of m, *Z. Angew. Math. Mech.* 60 (1980) 479–485.

[17] P.J. van der Houwen and B.P. Sommeijer, A special class of multistep Runge–Kutta methods with extended real stability interval, *IMA J. Numer. Anal.* 2 (1982) 183–209.

[18] P.J. van der Houwen and B.P. Sommeijer, Splitting methods for three-dimensional transport models with interaction terms, *J. Scientific Comput.* (submitted).

[19] P.J. van der Houwen and J.G. Verwer, One-step splitting methods for semidiscrete parabolic equations, *Computing* 22 (1979) 291–309.

[20] R. Jeltsch and O. Nevanlinna, Stability and accuracy of time discretizations for initial value problems, Report HTKK-MAT-A187, Helsinki University of Technology (1981).

[21] P. Jeltsch and O. Nevanlinna, Stability of explicit time discretizations for solving initial value problems, *Numer. Math.* 37 (1981) 61–91.

[22] I.P.E. Kinnmark and W.G. Gray, One-step integration methods with maximum stability regions, *Math. Comput. Simulation* 16 (1984) 87–92.

[23] I.P.E. Kinnmark and W.G. Gray, One-step integration methods of third-fourth order accuracy with large hyperbolic stability limits, *Math. Comput. Simulation* 16 (1984) 181–184.
[24] P. Laasonen, On a method for solving the heat flow equation, *Acta Math.* 81 (1949) 309–323 (in German).
[25] V.L. Lebedev, Explicit difference schemes with time-variable steps for solving stiff systems of equations, Preprint No. 177, Dept. of Numerical Mathematics, USSR Acad. Sc., Moscow (1987) (in Russian).
[26] V.L. Lebedev, How to solve stiff systems of equations by explicit difference schemes, in: G.I. Marchuk, ed., *Numerical Methods and Applications* (CRC Press, Ann Arbor, MI, 1994) 45–80.
[27] H. Lomax, On the construction of highly stable, explicit numerical methods for integrating coupled ODEs with parasitic eigenvalues, NASA Technical Note NASAIN D/4547 (1968).
[28] C.L. Metzger, Runge–Kutta methods whose number of stages exceeds their order, These (Troisieme cycle), Université de Grenoble (1967) (in French).
[29] D.W. Peaceman and H.H. Rachford Jr, The numerical solution of parabolic and elliptic differential equations, *J. Soc. Indust. Appl. Math.* 3 (1955) 28–41.
[30] L.F. Richardson, The approximate arithmetical solution by finite differences of physical problems involving differential equations, with an application to the stresses in a masonry dam, *Philos. Trans. Roy. Soc. London Ser. A* 210 (1910) 307–357 and *Proc. Roy. Soc. London Ser. A* 83 (1910) 335–336.
[31] W. Riha, Optimal stability polynomials, *Computing* 9 (1972) 37–43.
[32] E. Rothe, Two-dimensional parabolic boundary value problems as limiting case of one-dimensional boundary value problems, *Math. Ann.* 102 (1930) 650–670 (in German).
[33] C. Runge, On the numerical solution of differential equations, *Math. Ann.* 46 (1895) 167–178 (in German).
[34] B.P. Sommeijer, RKC, a nearly-stiff ODE solver, available through netlib (mail: netlib@ornl.gov, send rkc.f from ode) (1991).
[35] P. Sonneveld and B. van Leer, A minimax problem along the imaginary axis, *Nieuw Archief voor Wiskunde* 3 (4) (1985) 19–22.
[36] E.L. Stiefel, Kernel polynomials in linear algebra and their numerical applications, *Nat. Bur. Stand. Appl. Math. Series* 49 (1958) 1–22.
[37] J.G. Verwer, W.H. Hundsdorfer and B.P. Sommeijer, Convergence properties of the Runge–Kutta–Chebyshev method, *Numer. Math.* 57 (1990) 157–178.
[38] R. Vichnevetsky, New stability theorems concerning one-step numerical methods for ordinary differential equations, *Math. Comput. Simulation* 25 (1983) 199–205.
[39] N.N. Yanenko, *The Method of Fractional Steps* (Springer, Berlin, 1971).
[40] D.M. Young, On Richardson's method for solving linear systems with positive definite matrices, *J. Math. Phys.* 32 (1954) 243–255.
[41] Yuan' Chzao-Din, Some difference schemes for the solution of the first boundary value problem for linear differential equations with partial derivatives, Thesis, Moscow State University (1958) (in Russian).