Convergence Properties of
the Runge-Kutta-Chebyshev Method *

J.G. Verwer, W.H. Hundsdorfer, and B.P. Sommeijer
Centre for Mathematics and Computer Science (CWI), P.O. Box 4079, 1009 AB Amsterdam,
The Netherlands

Dedicated to Peter van der Houwen for his numerous contributions
in the field of numerical integration of differential equations

Summary. The Runge-Kutta-Chebyshev method is an s-stage Runge-Kutta method designed for the explicit integration of stiff systems of ordinary differential equations originating from spatial discretization of parabolic partial differential equations (method of lines). The method possesses an extended real stability interval with a length β proportional to s^2. The method can be applied with s arbitrarily large, which is an attractive feature due to the proportionality of β with s^2. The involved stability property here is internal stability. Internal stability has to do with the propagation of errors over the stages within one single integration step. This internal stability property plays an important role in our examination of full convergence properties of a class of 1st and 2nd order schemes. Full convergence means convergence of the fully discrete solution to the solution of the partial differential equation upon simultaneous space-time grid refinement. For a model class of linear problems we prove convergence under the sole condition that the necessary time-step restriction for stability is satisfied. These error bounds are valid for any s and independent of the stiffness of the problem. Numerical examples are given to illustrate the theoretical results.

Subject Classifications: AMS(MOS): 65M20; 65M10; CR: G1.8.

1 Introduction

This paper is devoted to an examination of convergence properties of a class
of Runge-Kutta-Chebyshev (RKC) schemes. These schemes have been designed

---

by van der Houwen and Sommeijer [7] for the *explicit* time integration of stiff systems of ODEs,

$$(1.1) \qquad \dot{U}(t) = F(t, U(t)), \quad 0 < t \le T, \; U(0) \text{ given,}$$

which originate from spatial discretization of *parabolic* partial differential equations (method of lines). For the time being, it is not necessary to define a particular class of parabolic problems or to specify the space discretization technique. The only restrictions for application of the RKC schemes are (i) The eigenvalue spectrum of the Jacobian matrix $\partial F(t, U)/\partial U$ should lie in a narrow strip along the negative axis of the complex plane, and (ii) The Jacobian matrix should 'not deviate too much from a normal matrix'. These two conditions trivially hold if $\partial F(t, U)/\partial U$ is symmetric and negative definite, properties frequently encountered when discretizing elliptic operators.

The RKC method is a typical example of an *explicit, stabilized RK method*. Such a method can be characterized by the fact that it possesses a large number of stages of which almost all serve to obtain good stability properties and only a few to obtain the desired order of consistency (see van der Houwen [5, 6]). The RKC method has been designed such that it possesses an *extended real stability interval*. Its real stability boundary β is in fact proportional to s^2, s being the number of stages, while its main characteristic is that s can be taken arbitrarily large. This is made possible by an intelligent use of Chebyshev polynomials, thus explaining the name of the method. The possibility of using arbitrarily large values for s is of practical relevance due to the fact that the effective real stability boundary β/s linearly increases with s. Hence, in applications it is possible and advantageous to choose the stepsize on the basis of accuracy and to adjust s to meet the demand of (linear) stability (see Sommeijer and Verwer [13] and Sommeijer and van der Houwen [12] for numerical illustrations).

When the problem becomes very stiff, the conditional stability may remain too much of a disadvantage, of course. For example, standard central differencing of the first initial-boundary value problem for the model heat equation

$$(1.2) \qquad u_t = \Delta u, \quad u = u(x, t), \quad x \in \mathbb{R}^d,$$

on a uniform mesh of width h, yields the time-step restriction $\tau\sigma \le \beta$ for stability. Here τ is the stepsize of the RKC method and $\sigma \simeq 4d/h^2$ is the spectral radius of the resulting Jacobian. A typical value for β is $2s^2$, so that, given τ, the number of stages s should be at least $(2\tau d)^{\frac{1}{2}}/h$. On a fine mesh, say $h = 0.01$ in two space dimensions, this yields $200\sqrt{\tau}$ stages per time step, which may be considerable. On the other hand, the explicitness guarantees an easy application to linear and nonlinear problems in two and three space dimensions and, needless to say, offers optimal prospects for implementation on vector/parallel computers. When the stiffness is modest, like for the reaction-diffusion problem

$$(1.3) \qquad u_t = \varepsilon\Delta u + f(u, x, t), \quad u = u(x, t), \quad x \in \mathbb{R}^d,$$

where f represents the reaction term and $0 < \varepsilon \ll 1$, the RKC method offers a very attractive alternative for standard explicit integration schemes and unconditionally stable, implicit ones.

Van der Houwen and Sommeijer [7] have developed a 1st and 2nd order RKC scheme. In this paper we examine both these schemes (the coefficients of our 2nd order scheme slightly differ; they have been taken from Sommeijer and Verwer [13]). While these schemes have been developed along the lines of the classical ODE theory, the purpose of the present examination is to analyse their *full convergence properties*. Full convergence means convergence of the fully discrete solution with respect to the solution of the PDE upon simultaneous refinement of the space-time mesh. For linear PDE problems, whose semi-discretizations take the form

(1.4) $$\dot{U}(t) = MU(t) + g(t), \quad 0 < t \leq T, \ U(0) \text{ given,}$$

with M a *symmetric*, constant coefficient matrix possessing *non-positive eigenvalues*, we prove convergence under the sole condition that the necessary time-step restriction $\tau \sigma(M) \leq \beta$ is satisfied, where $\sigma(M)$ is the spectral radius of M. Of interest is that the derived error bounds are independent of $\sigma(M)$ and valid for arbitrarily large s, the number of stages, thus showing that in applications the best strategy is to have the size of τ determined by the desired accuracy level and s by the stability demand. It is stipulated that this result is quite uncommon for an explicit method. We owe this to the favourable *internal stability* property of the RKC method. Internal stability has to do with the propagation of errors over the stages within one single integration step.

The convergence analysis presented in this paper is akin to the analysis of Sanz-Serna, Verwer and Hundsdorfer [10] and Hundsdorfer and Verwer [9] which, in turn, was inspired by the *B-convergence* analysis from the stiff ODE field (see Dekker and Verwer [2], Ch. 7).

The remainder of the paper consists of five sections. In Sect. 2 we review the 1st and 2nd order scheme under examination. This section is entirely based on [7] and [13]. Section 3 is devoted to the concept of internal stability, while Sect. 4 contains the derivation of local defects for exact PDE solutions. The results of the two latter sections are combined in Sect. 5 so as to obtain bounds for the full global errors. Finally, Sect. 6 presents two numerical examples with the aim of illustrating the convergence behaviour of the method and its performance in practice.

2 Description of the Method

For the ODE system (1.1), the RKC formula considered here is of the form

(2.1) $Y_0 = U_n,$

$Y_1 = Y_0 + \tilde{\mu}_1 \tau F_0,$

$Y_j = \mu_j Y_{j-1} + \nu_j Y_{j-2} + (1 - \mu_j - \nu_j) Y_0 + \tilde{\mu}_j \tau F_{j-1} + \tilde{\gamma}_j \tau F_0 \quad (2 \leq j \leq s),$

$U_{n+1} = Y_s, \quad n = 0, 1, \ldots,$

where $F_j = F(t_n + c_j \tau, Y_j)$; U_n represents the approximation to the exact solution U of (1.1) at time $t = t_n$ and $\tau = t_{n+1} - t_n$ is the stepsize. Throughout it is assumed

that the increment parameters c_j are defined by the integration coefficients μ_j, v_j, $\tilde{\mu}_j$ and $\tilde{\gamma}_j$ in the following way,

$$(2.2) \qquad c_0 = 0, \qquad c_1 = \tilde{\mu}_1,$$

$$c_j = \mu_j c_{j-1} + v_j c_{j-2} + \tilde{\mu}_j + \tilde{\gamma}_j \qquad (2 \leqq j \leqq s).$$

Then, if we bring Y_j in the standard RK form

$$Y_j = U_n + \tau \sum_{l=0}^{j-1} a_{jl} F(t_n + c_l \tau, Y_l) \qquad (0 \leqq j \leqq s),$$

where the coefficients a_{jl} are expressions in μ_j, v_j, $\tilde{\mu}_j$, $\tilde{\gamma}_j$, it is readily seen that the usual condition

$$c_j = \sum_{l=0}^{j-1} a_{jl}$$

is satisfied. Hence, (2.1) is an *explicit, s-stage RK method* and Y_j is an intermediate approximation at the intermediate point $t = t_n + c_j \tau$. Due to the specific recursive nature of the method, as shown in the formula defining Y_j, formula (2.1) is more convenient to work with than the common RK formula. The rationale behind the specific form (2.1) is that this form is easily identified with stable three-term Chebyshev recursions. This will become clear later on. Note that irrespective the number of stages, the number of required storage arrays is maximal 6.

Let us determine the *consistency conditions* (in the classical ODE sense) for order 1 and 2. Suppose $U_n = U(t_n)$, where $U(t)$, $t \geqq t_n$ is a sufficiently smooth solution of (1.1). By definition of c_j it then holds that all Y_j satisfy an expansion

$$(2.3) \qquad Y_j = U(t_n) + c_j \tau \dot{U}(t_n) + X_j \tau^2 U^{(2)}(t_n) + O(\tau^3),$$

where, similar as c_j, X_j is determined by the integration coefficients. Substitution of this expression into (2.1) gives

$$(2.4) \qquad X_0 = X_1 = 0,$$

$$X_j = \mu_j X_{j-1} + v_j X_{j-2} + \tilde{\mu}_j c_{j-1} \qquad (2 \leqq j \leqq s).$$

We conclude that the RKC method is *consistent of order* 1 if

$$(2.5) \qquad\qquad\qquad c_s = 1,$$

and note that the jth stage formula is consistent of order 1 at $t = t_n + c_j \tau$.

It follows from (2.3) that each stage formula is consistent of order 2 at $t = t_n + c_j \tau$, for $2 \leqq j \leqq s$, if $X_j = \frac{1}{2} c_j^2$. In terms of c_j this gives

$$(2.6) \qquad c_2^2 = 2 \tilde{\mu}_2 c_1,$$

$$c_3^2 = \mu_3 c_2^2 + 2 \tilde{\mu}_3 c_2,$$

$$c_j^2 = \mu_j c_{j-1}^2 + v_j c_{j-2}^2 + 2 \tilde{\mu}_j c_{j-1} \qquad (4 \leqq j \leqq s).$$

As pointed out in [13], it is possible to satisfy this condition in a satisfactory way for all $2 \leq j \leq s$. We here adopt this condition and hence the 2nd order scheme derived below has all its stages consistent of order 2 at the intermediate step points $t = t_n + c_j \tau$, except the first one. The original 2nd order scheme from [7] is only consistent of order 2 at the main step points.

For future reference, it is stipulated that the derivation of the current consistency conditions follows the lines of the classical numerical ODE theory [4, 6], as it is based on expanding F-terms. This means that it is tacitly assumed here that F satisfies a Lipschitz condition so that $\tau \|F\| = O(\tau)$. For stiff problems this is unduly restrictive and particularly so for semi-discrete parabolic equations for which $\|F\| \to \infty$ upon grid refinement. In Sect. 4 we will re-examine the consistency properties of the RKC scheme. The derivation presented there is inspired by the B-convergence theory for stiff ODEs, the central theme of which is the derivation of error bounds which do not depend on the stiffness of the problem (see [2], Ch. 7 and [9–11]).

Finally, a natural condition is that all (intermediate) step points lie within the step interval $[t_n, t_{n+1}]$ and increase monotonically with j:

$$(2.7) \qquad 0 = c_0 < c_1 < c_2 < \dots < c_{s-1} < c_s = 1.$$

It will turn out this condition is satisfied for the two selected schemes.

We proceed with the *stability function*. Because the RKC method is an s-stage, explicit RK method, application to the scalar test equation $\dot{U}(t) = \lambda U(t)$ leads to the linear, one-step recursion

$$(2.8) \qquad U_{n+1} = P_s(z)\, U_n, \qquad z = \tau \lambda,$$

where the stability function $P_s : \mathbb{C} \to \mathbb{C}$ is a polynomial of degree s. P_s itself is also defined recursively as follows:

$$(2.9) \qquad P_0(z) = 1, \qquad P_1(z) = 1 + \tilde{\mu}_1 z,$$
$$P_j(z) = (1 - \mu_j - \nu_j) + \tilde{\gamma}_j z + (\mu_j + \tilde{\mu}_j z) P_{j-1}(z) + \nu_j P_{j-2}(z) \qquad (2 \leq j \leq s).$$

In fact, all polynomials P_j are of degree j and satisfy

$$(2.10) \qquad Y_j = P_j(z)\, U_n \qquad (0 \leq j \leq s).$$

Therefore we will also call the intermediate polynomials P_j stability functions, but note that they play no role in the step-by-step stability like P_s.

According to (2.3), each stability function $P_j(z)$ approximates the exponential $e^{c_j z}$ for $z \to 0$ as

$$(2.11) \qquad P_j(z) = 1 + c_j z + X_j z^2 + O(z^3).$$

Hence, each P_j is consistent of order 1 (with the exponential $e^{c_j z}$) and consistent of order 2 if, in addition, $X_j = c_j^2/2$. Substitution of this expansion into (2.9) and equating powers of z then reveals relations (2.2) and condition (2.6). Hence, if we select the coefficients $\mu_j, \nu_j, \tilde{\mu}_j, \tilde{\gamma}_j$ in the recursion (2.9) such that P_j is of order 1 or 2 in the sense of (2.11), then the 1st or 2nd order conditions

associated to expansion (2.3) are automatically satisfied. This is very convenient since it enables us to concentrate entirely on the choice of the stability functions.

The choice of the stability functions P_j is the central issue in the development of the RKC method. This choice underlies two *design rules*:

(I) The coefficients μ_j, v_j, $\tilde{\mu}_j$, $\tilde{v}_j$ in the recursion (2.9) are chosen such that the *real stability boundary*

$$(2.12) \qquad \beta(s) = \max\{-z: z \leq 0, |P_s(z)| \leq 1\}$$

of the genuine stability function P_s is as large as possible, so as to obtain good stability properties for parabolic equations. This requirement leads to the *Chebyshev polynomial of the first kind*

$$T_s(x) = \cos(s \arccos x), \qquad -1 \leq x \leq 1,$$

to which we owe the quadratic increase of $\beta(s)$ with s. For example, within the class of 1st order consistent polynomials, the shifted Chebyshev polynomial

$$(2.13) \qquad P_s(z) = T_s\left(1 + \frac{z}{s^2}\right), \qquad -\beta(s) \leq z \leq 0,$$

yields the largest possible value for $\beta(s)$, viz, $\beta(s) = 2s^2$. The use of this polynomial for numerical integration purposes is extensively discussed in van der Houwen [6]. $\square$

(II) The second design rule has to do with the desirability of applying the method with an arbitrary number of stages which means that, given s, all coefficients μ_j, v_j, $\tilde{\mu}_j$, $\tilde{\gamma}_j$ must be known in analytic form. Further, and this is most important, it should be possible to let s arbitrarily large without severe accumulation of errors within one single step (*internal stability*). The notion of internal stability will be discussed at length in Sect. 3. Here we mention that both these requirements are fulfilled by adjusting the three-term recursion (2.9) for P_j to the known three-term recursion of appropriately chosen shifted Chebyshev polynomials. For example, the polynomials $P_j(z) = T_j(1 + z/s^2)$ satisfy the recursion

$$(2.14) \qquad P_0(z) = 1, \qquad P_1(z) = 1 + \frac{z}{s^2},$$

$$P_j(z) = 2\left(1 + \frac{z}{s^2}\right)P_{j-1}(z) - P_{j-2}(z), \qquad j \geq 2,$$

and adjusting (2.9) gives

$$\tilde{\mu}_1 = 1/s^2, \quad \mu_j = 2, \quad \tilde{\mu}_j = 2/s^2, \quad v_j = -1, \quad \tilde{\gamma}_j = 0 \quad (2 \leq j \leq s).$$

Note that $|P_j(z)| \leq 1$ for all $j \leq s$ as long as z lies within the real stability interval $[-2s^2, 0]$ of the genuine stability function P_s. $\square$

Having outlined these two design rules, we are now ready to specify the stability functions P_j with the associated coefficient sets for the 1st and 2nd order RKC schemes examined in this paper. They all fit in the general form

$$P_j(z) = a_j + b_j T_j(w_0 + w_1 z), \qquad 0 \leq j \leq s, \tag{2.15}$$

where the parameters a_j, b_j, w_0 and w_1 have been chosen in accordance with the design rules (I) and (II). Before specifying them, there is one point left that should be mentioned (to save space we must refer to [6, 7] for more details). This point concerns the parameter w_0. Consider the polynomial (2.13) where $w_0 = 1$. This polynomial alternates between $+1$ and -1, i.e., $|P_s(z)| = 1$ at $s+1$ points $z \in [-\beta, 0]$. It is desirable to introduce some damping in P_s, i.e., to let P_s alternate between values $\simeq 1 - \varepsilon$ and $\simeq -1 + \varepsilon$ for all $z \in [-\beta, 0]$ (with the exception of a small neigbourhood of $z = 0$), where ε is a small positive number. The damping is obtained by choosing $w_0 = w_0(\varepsilon)$, called the *damping parameter*, slightly larger than 1. By introducing this damping in the stability function, we achieve that the *stability region* becomes a long, narrow strip around the negative axis of the complex plane. On the other hand, the real stability boundary slightly decreases [6]. There is practical evidence that with damping the RKC method becomes more robust for nonlinear problems. Finally we wish to note that an early paper on the subject, where a.o. the use of damped Chebyshev polynomials is suggested, is that of Guillou and Lago [3].

The 1st Order Case: RKC1 [7]

$$(2.16) \qquad a_j = 0, \quad b_j = T_j^{-1}(w_0), \quad w_0 = 1 + \frac{\varepsilon}{s^2}, \quad w_1 = \frac{T_s(w_0)}{T_s'(w_0)} \qquad (0 \leq j \leq s).$$

It can be shown that with this choice of parameters

$$\beta(s) \simeq \frac{(w_0 + 1) T_s'(w_0)}{T_s(w_0)} \simeq (2 - \tfrac{4}{3}\varepsilon) s^2 \qquad \text{for } \varepsilon \to 0.$$

A suitable values for ε is 0.05. Since $T_s^{-1}(w_0) \simeq 1 - \varepsilon$, this yields about 5% damping with only a very little decrease in $\beta(s)$, $\beta(s) \simeq 1.90 s^2$. Note that with $\varepsilon = 0$ we recover the polynomials (2.14). Adjusting recursion (2.9) to the current choice for P_j completely defines the general 1st order scheme (2.1):

$$(2.17) \qquad\qquad\qquad \tilde{\mu}_1 = \frac{w_1}{w_0},$$

$$\mu_j = 2 w_0 \frac{b_j}{b_{j-1}}, \quad \nu_j = -\frac{b_j}{b_{j-2}},$$

$$\tilde{\mu}_j = 2 w_1 \frac{b_j}{b_{j-1}}, \quad \tilde{\gamma}_j = 0 \quad (2 \leq j \leq s).$$

Note that each value of s defines a different coefficient set. Also note that $\mu_j + \nu_j = 1$ and that the increment parameters

$$(2.18) \qquad\qquad\qquad c_j = \frac{T_s(w_0)}{T_s'(w_0)} \frac{T_j'(w_0)}{T_j(w_0)} \simeq j^2/s^2$$

satisfy condition (2.7). For more details we refer to [6, 7].

The 2nd Order Case: RKC2 [7, 13]

$$(2.19) \qquad a_j = 1 - b_j\, T_j(w_0), \quad b_j = \frac{T_j''(w_0)}{(T_j'(w_0))^2}, \quad (2 \leqq j \leqq s),$$

$$w_0 = 1 + \frac{\varepsilon}{s^2}, \quad w_1 = \frac{T_s'(w_0)}{T_s''(w_0)},$$

$$a_0 = 1 - b_0, \quad a_1 = 1 - b_1\, w_0, \quad b_0 = b_1 = b_2.$$

For this choice of parameters one can prove that

$$\beta(s) \simeq \frac{(w_0 + 1)\, T_s''(w_0)}{T_s'(w_0)} \simeq \tfrac{2}{3}(s^2 - 1)(1 - \tfrac{2}{15}\varepsilon) \quad \text{for } \varepsilon \to 0.$$

A suitable value for ε is $\tfrac{2}{13}$. This gives about 5% damping ($a_s + b_s \simeq 1 - \tfrac{1}{3}\varepsilon$) with a reduction in $\beta(s)$ of about 2%. The current choice of stability polynomials covers roughly 80% of the optimal real stability interval for 2nd order consistent polynomials (van der Houwen [8]). Adjusting again recursion (2.9) completely defines the general 2nd order scheme (2.1):

$$(2.20) \qquad \tilde{\mu}_1 = b_1\, w_1,$$

$$\mu_j = 2 w_0\, \frac{b_j}{b_{j-1}}, \quad v_j = -\frac{b_j}{b_{j-2}},$$

$$\tilde{\mu}_j = 2 w_1\, \frac{b_j}{b_{j-1}}, \quad \tilde{\gamma}_j = -(1 - b_{j-1}\, T_{j-1}(w_0))\, \tilde{\mu}_j \quad (2 \leqq j \leqq s).$$

The increment parameters are

$$(2.21) \qquad c_1 = \frac{c_2}{T_2'(w_0)} \simeq \frac{c_2}{4}, \quad c_j = \frac{T_s'(w_0)}{T_s''(w_0)}\, \frac{T_j''(w_0)}{T_j'(w_0)} \simeq \frac{j^2 - 1}{s^2 - 1} \quad (2 \leqq j \leqq s)$$

and thus satisfy conditions (2.7). For more details, see [6, 7].

3 Convergence Analysis: Internal Stability

The remainder of the paper is devoted to the full convergence analysis of the schemes defined by the coefficient sets (2.17), (2.20) when applied to the linear problem class (1.4). Hence, it is supposed that the $m \times m$ constant coefficient matrix M is *symmetric* and *possesses nonpositive eigenvalues* $\lambda(M)$. This covers all linear parabolic problems with time-independent coefficients in the elliptic operator. We stipulate that the RKC method is very well applicable to *nonlinear* parabolic problems, provided the spectrum of the Jacobian $F'(t, U)$ is located in a long, narrow strip around the negative axis of the complex plane and $F'(t, U)$ does not 'deviate too much from a normal matrix'. A nonlinear analysis of the RKC method is likely to become very complicated, if feasible at all. Our convergence analysis for the linear problem gives also insight in handling nonlinear problems.

Throughout, $\|\cdot\|$ denotes the common (appropriately weighted) Euclidean norm in $\mathbb{R}^m$, or the associated spectral matrix norm. Recall that, since M is normal, $\|M\| = \sigma(M)$, σ being the spectral radius. Further, for any polynomial $P(z)$, the spectrum of the matrix polynomial $P(\tau M)$ is the set of values $P(\tau\lambda)$, where $\tau\lambda$ runs through the spectrum of τM; $P(\tau M)$ is also normal and

$$\|P(\tau M)\| = \sigma(P(\tau M)) = \max_{\tau\lambda} |P(\tau\lambda)|.$$

By assumption on M,

$$-\tau\sigma(M) \leqq \tau\lambda(M) \leqq \max(\tau\lambda(M)) \leqq 0.$$

Hence, if we select the stepsize τ and the number of stages s such that *the stability condition* $\tau\sigma(M) \leqq \beta$ is satisfied, then $\|P_s(\tau M)\| \leqq 1$ for the genuine stability function P_s of the scheme under consideration.

In the analysis of the RKC scheme (2.1), the notion of internal stability plays an important role. Internal stability is investigated with the *perturbed scheme*

(3.1) $\tilde{Y}_0 = \tilde{U}_n,$

$\tilde{Y}_1 = \tilde{Y}_0 + \tilde{\mu}_1 \tau \tilde{F}_0 + \tilde{r}_1,$

$\tilde{Y}_j = \mu_j \tilde{Y}_{j-1} + v_j \tilde{Y}_{j-2} + (1 - \mu_j - v_j)\tilde{Y}_0 + \tilde{\mu}_j \tau \tilde{F}_{j-1} + \tilde{\gamma}_j \tau \tilde{F}_0 + \tilde{r}_j \quad (2 \leqq j \leqq s),$

$\tilde{U}_{n+1} = \tilde{Y}_s, \quad n = 0, 1, \ldots,$

where now

(3.2) $$\tilde{F}_j \equiv F_j(t + c_j \tau, \tilde{Y}_j) = M\tilde{Y}_j + g(t + c_j \tau)$$

and $\tilde{r}_j$ represents a perturbation introduced at stage j (e.g. round off). Likewise, $\tilde{U}_n$ represents a perturbation of U_n.

Let

(3.3) $$e_n = \tilde{U}_n - U_n, \quad d_j = \tilde{Y}_j - Y_j \quad (0 \leqq j \leqq s)$$

represent the errors introduced by these perturbations. Note that, by definition, $d_0 = e_n$ and $e_{n+1} = d_s$. If we subtract the non-perturbed scheme (2.1) from (3.1), while taking into account the linear system form (1.4), we get the *error scheme*

(3.4) $d_0 = e_n,$

$d_1 = d_0 + \tilde{\mu}_1 \tau M d_0 + \tilde{r}_1,$

$d_j = \mu_j d_{j-1} + v_j d_{j-2} + (1 - \mu_j - v_j)d_0 + \tilde{\mu}_j \tau M d_{j-1} + \tilde{\gamma}_j \tau M d_0 + \tilde{r}_j \quad (2 \leqq j \leqq s),$

$e_{n+1} = d_s, \quad n = 0, 1, \ldots.$

Due to the linearity, d_j can be written as

(3.5) $$d_j = P_j(\tau M) e_n + \sum_{k=1}^{j} Q_{jk}(\tau M) \tilde{r}_k \quad (1 \leqq j \leqq s),$$

where P_j are the previously introduced stability functions (cf. (2.10)) and Q_{jk} are new polynomials of degree $j - k$. Of importance is that these new polynomials determine the propagation of all internal perturbations over the stages within one single integration step. We therefore call them *internal stability functions*. In particular, together with the stability function P_s, the internal stability functions $Q_{s1}, \ldots, Q_{sk}$ occurring in the final stage error formula

$$(3.6) \qquad e_{n+1} = P_s(\tau M)\, e_n + \sum_{k=1}^{s} Q_{sk}(\tau M)\, \tilde{r}_k,$$

determine the error e_{n+1} of U_{n+1}. In order to avoid large contributions $Q_{sk}(\tau M)\, \tilde{r}_k$, the polynomials $Q_{sk}(z)$ should mimic, in some sense, the behaviour of the stability function $P_s(z)$ for all $z = \tau \lambda(M) \in [-\tau\sigma(M), 0]$. This is particularly important in applications where both the number of stages s and the spectral radius $\tau\sigma(M)$ are large.

Let us determine a closed expression for all Q_{jk}. Substitution of (3.5) (put $e_n = 0$ for simplicity) into (3.4), and equating coefficients of $\tilde{r}_k$, reveals after a little inspection that they satisfy a three-term recursion with stage index j as recursion index. This suggests that Q_{jk} is also a shifted Chebyshev polynomial like P_j. More precisely, Q_{jk} exists for $1 \leq k \leq s$, $k \leq j \leq s$; for $k = 1, 2, \ldots, s-2$ there holds

$$(3.7) \qquad Q_{jk}(z) = (\mu_j + \tilde{\mu}_j z)\, Q_{j-1,k}(z) + \nu_j Q_{j-2,k}(z), \qquad j = k+2, \ldots, s,$$

$$Q_{kk}(z) = 1, \quad Q_{k+1,k}(z) = \mu_{k+1} + \tilde{\mu}_{k+1} z,$$

while $Q_{s-1,s-1}(z) = Q_{ss}(z) = 1$ and $Q_{s,s-1}(z) = \mu_s + \tilde{\mu}_s z$.

Inserting the coefficient values of (2.17), (2.20) gives, for $k = 1, 2, \ldots, s-2$,

$$(3.8\,a) \quad Q_{jk}(z) = 2\,\frac{b_j}{b_{j-1}}\,(w_0 + w_1 z)\, Q_{j-1,k}(z) - \frac{b_j}{b_{j-2}}\, Q_{j-2,k}(z), \qquad j = k+2, \ldots, s,$$

$$Q_{kk}(z) = 1, \quad Q_{k+1,k}(z) = 2\,\frac{b_{k+1}}{b_k}\,(w_0 + w_1 z),$$

while

$$(3.8\,b) \qquad Q_{s-1,s-1}(z) = Q_{ss}(z) = 1 \quad \text{and} \quad Q_{s,s-1}(z) = 2\,\frac{b_s}{b_{s-1}}\,(w_0 + w_1 z).$$

It follows that $b_j^{-1}\, Q_{jk}(z)$ satisfies the recursion for the shifted *Chebyshev polynomial of the second kind*, due to the factor 2 occurring in the definition of the second starting value $Q_{k+1,k}$. We thus have

$$(3.9) \qquad Q_{jk}(z) = \frac{b_j}{b_k}\, S_{j-k}(w_0 + w_1 z) \qquad (1 \leq k \leq s,\ k \leq j \leq s),$$

where $S_i(x)$ is the ith degree Chebyshev polynomial of the second kind (in literature usually denoted by $U_i(x)$ [1]).

The error scheme (3.6), that is

$$(3.10) \qquad e_{n+1} = P_s(\tau M)\, e_n + \sum_{k=1}^{s} \frac{b_s}{b_k} S_{s-k}(w_0 I + w_1 \tau M)\, \tilde{r}_k,$$

gives a complete description of the stability of the schemes under examination. To proceed with it, we briefly recall a few properties of the second kind Chebyshev polynomial [1]. As opposed to $T_i(x)$, $S_i(x)$ is not bounded by ± 1 for $-1 \leq x \leq 1$. There holds $S_i(\pm 1) = (\pm 1)^i (i+1)$ and $i+1$ is also the maximal value for $-1 \leq x \leq 1$. On the greater part of this interval, $S_i(x)$ alternates between (approximately) $+1$ and -1. The slope of $S_i(x)$ near $x = 1$ is also larger than that of $T_i(x)$. There holds $S_i'(1) = i(i+1)(i+2)/3$.

Theorem 3.1. *Suppose that τ and s are such that the stability time-step restriction $\tau \sigma(M) \leq \beta$ is satisfied. Then the following error bound is valid,*

$$(3.11) \quad \|e_{n+1}\| \leq \|e_n\| + C \sum_{k=1}^{s} (s-k+1)\|\tilde{r}_k\| \leq \|e_n\| + \tfrac{1}{2} s(s+1)\, C \max_k \|\tilde{r}_k\|,$$

where C is a constant of moderate size independent of M, τ and s.

Proof. The bound $\|P_s(\tau M)\| \leq 1$ is a trivial consequence of the time-step restriction $\tau \sigma(M) \leq \beta(s)$. If $-\beta(s) \leq z = \tau \lambda(M) \leq 0$, then, by definition of w_1, $-1 \leq w_0 + w_1 z \leq w_0$ and $w_0 \geq 1$. Hence, due to the normality of M and the behaviour of $S_{s-k}(w_0 + w_1 z)$,

$$\|S_{s-k}(w_0 I + w_1 \tau M)\| \leq S_{s-k}(w_0).$$

We now distinguish two cases:

(i) No damping ($\varepsilon = 0$). Then $w_0 = 1$ and $S_{s-k}(1) = s - k + 1$. For the 1st order schemes, all coefficients b_k are equal to 1 (see (2.16)) and (3.11) is valid with $C = 1$. For the 2nd order schemes, $b_k = \tfrac{1}{3}(1 - k^{-2})$ for $2 \leq k \leq s$ and $b_1 = b_2$ (see (2.19)); (3.11) is valid with $C = b_s b_1^{-1} \leq \tfrac{4}{3}$.

(ii) Damping ($\varepsilon > 0$). Now $w_0 = 1 + \varepsilon/s^2$, so that in this case $S_{s-k}(w_0)$ and b_k must be expanded at $x = 1$ in terms of εs^{-2}. The negative quadratic s^{-2} then implies that $S_{s-k}(w_0)$ is bounded by

$$(s - k + 1)(1 + C\varepsilon),$$

where C is a constant of moderate size. This also holds for all b_k. The proof is merely technical and a bit lengthy and therefore omitted. It is obtained by properly bounding all higher derivatives of the occurring Chebyshev polynomials at $x = 1$. $\square$

This result shows that within one full RKC step the accumulation of internal perturbations, such as round-off errors, is *independent of the spectrum of M* as long as $\tau \sigma(M) \leq \beta$. As far as rounding errors are concerned, the *quadratic increase with the number of stages* renders no problem. For example, if $s = 1000$, which for a serious application is of course a hypothetical value, the local perturbation is at most $\sim 10^6 \max \|\tilde{r}_i\|$. If the machine precision of the computer is about 14 digits, a common value, this local perturbation still leaves 8 digits for accuracy which for PDEs is more than enough.

Table 3.1. Internal stability test. The machine accuracy of the Alliant is $\simeq 10^{-14}$ (double precision) and of the Macintosh $\simeq 10^{-7}$

h^{-1}	s	Alliant FX/4; $d=14$		Macintosh PC; $d=7$	
		α	α/s^2	α	α/s^2
10	25	0.38	$0.60\ 10^{-3}$	$0.72\ 10^1$	$0.11\ 10^{-1}$
20	50	0.69	$0.28\ 10^{-3}$	$0.33\ 10^2$	$0.13\ 10^{-1}$
40	100	1.29	$0.12\ 10^{-3}$	$0.17\ 10^3$	$0.17\ 10^{-1}$
80	199	9.84	$0.25\ 10^{-3}$	$0.41\ 10^3$	$0.10\ 10^{-1}$
160	397	6.91	$0.44\ 10^{-4}$	$0.30\ 10^4$	$0.19\ 10^{-1}$
320	794	51.76	$0.82\ 10^{-4}$	$0.10\ 10^5$	$0.16\ 10^{-1}$

Van der Houwen and Sommeijer [7] also discuss two different stabilized, explicit RK methods. These two methods possess the same stability function P_s as the RKC method, but show a very strong, spectrum dependent accumulation of internal perturbations (see their numerical experiment). They also conclude that for the RKC scheme the accumulation of internal perturbations is negligible and almost independent of s. Their conclusion is not quite correct since it is based on the assumption that this accumulation is governed by the stability functions P_j, rather than by the internal stability functions Q_{jk}.

By way of illustration, we have carried out a similar experiment with RKC2 (coefficient set (2.19)–(2.21) with $\varepsilon = \frac{2}{13}$). RKC1 shows similar results. It consists of carrying out one integration step from $t = 0$ to $t = T = 1$ with the model system that arises from spatially discretizing the first initial-boundary value problem for

$$u_t = u_{xx} + x(1-x) + 2t, \quad 0 \leq x \leq 1.$$

The boundary and initial values are adjusted to the exact solution $u(x, t) = 1 + tx(1-x)$. Standard central differencing is used so that in exact arithmetic the RKC schemes should return this exact solution (see following sections).

Let $U_{1,j}$ represent the jth component of the approximation vector U_1 at $t = 1$. For two different computers, the errors

$$\alpha = 10^d \max_j |U_{1,j} - u(x_j, 1)|, \quad d = 7 \text{ or } 14,$$

listed in Table 3.1 for several values of the spatial gridsize h, nicely illustrate the limited error growth. We note that $U_{0,j} = 1$, so that the only contributions come from round-off. Apparently, the round-off and the quadratic error growth are more pronounced on the Macintosh than on the Alliant. The number of stages s varies with h such that

$$s = 1 + \text{entier}[(1 + \tau\sigma(M)/0.65)^{\frac{1}{2}}], \quad \tau = 1, \quad \sigma(M) = 4h^{-2}.$$

This choice of s guarantees that $\tau\sigma(M) \leq \beta(s)$. Note that we select s slightly larger than necessary. It is stipulated that s is rather large since $\tau = 1$ for all h. Needless to say that if in a practical application such large values for s are necessary, the costs of the integration are too large and one should use implicit methods.

4 Convergence Analysis: The Local Defects

We continue the convergence analysis with the computation of the local defects which arise if an exact PDE solution is inserted in the Runge-Kutta scheme. Consider the semi-discrete PDE problem [9–11]

$$(4.1) \qquad \dot{u}_h(t) = F(t, u_h(t)) + \alpha_h(t), \qquad 0 < t \leq T, \quad u_h(0) \text{ given,}$$

that is associated to the ODE system (1.1). Hence, $u_h(t)$ represents an *exact* PDE *solution* restricted to some space grid parametrized by h, and $\alpha_h(t)$ is the *local space truncation error* that originates from replacing the original PDE problem by its exact, semi-discrete counterpart (4.1). The derivation of the local defects applies to any initial-boundary value problem whose semi-discretization can be put in the generic form (4.1). In particular, in this section F is allowed to be *nonlinear* and merely smoothness assumptions on $u_h(t)$ will be made (cf. the B-convergence theory).

In the previous section we have introduced the perturbed scheme (3.1) for examining the internal propagation of local, arbitrary perturbations $\tilde{r}_j$. If we set in the perturbed scheme $\tilde{Y}_j$ equal to $u_h(t_n + c_j\tau)$, then the $\tilde{r}_j$ represent residual (local) discretization errors, which will be called the *local defects*. These defects will be denoted by r_j in order to distinguish them from the general $\tilde{r}_j$. The local defects are thus defined by

$$(4.2) \quad u_h(t_n + c_1\tau) = u_h(t_n) + \tilde{\mu}_1\tau F(t_n, u_h(t_n)) + r_1,$$

$$u_h(t_n + c_j\tau) = \mu_j u_h(t_n + c_{j-1}\tau) + v_j u_h(t_n + c_{j-2}\tau) + (1 - \mu_j - v_j) u_h(t_n)$$
$$+ \tilde{\mu}_j \tau F(t_n + c_{j-1}\tau, u_h(t_n + c_{j-1}\tau)) + \tilde{\gamma}_j \tau F(t_n, u_h(t_n)) + r_j$$
$$(2 \leq j \leq s).$$

Let $p \in \mathbb{N}$ and assume $u_h \in C^{p+1}[0, T]$. From (4.1) and the Taylor series expansion of $u_h, \dot{u}_h$ at the *intermediate step point* $t_n + c_{j-1}\tau$, it follows that

$$(4.3) \qquad r_j = \tau\theta_{1j}\dot{u}_h(t_n + c_{j-1}\tau) + \ldots + \tau^p\theta_{pj}u_h^{(p)}(t_n + c_{j-1}\tau) + \tau^{p+1}\rho_j$$
$$+ \tau\tilde{\mu}_j\alpha_h(t_n + c_{j-1}\tau) + \tau\tilde{\gamma}_j\alpha_h(t_n) \qquad (2 \leq j \leq s),$$

where the coefficients θ_{qj} and *remainder term* ρ_j are given by

$$(4.4) \quad \theta_{1j} = (c_j - c_{j-1}) - v_j(c_{j-2} - c_{j-1}) + (1 - \mu_j - v_j)c_{j-1} - \tilde{\mu}_j - \tilde{\gamma}_j,$$

$$\theta_{qj} = \frac{1}{q!}(c_j - c_{j-1})^q - \frac{1}{q!}v_j(c_{j-2} - c_{j-1})^q - \frac{1}{q!}(1 - \mu_j - v_j)(-c_{j-1})^q$$

$$- \frac{1}{(q-1)!}\tilde{\gamma}_j(-c_{j-1})^{q-1} \qquad (2 \leq q \leq p),$$

$$\rho_j = \frac{1}{(p+1)!}(c_j - c_{j-1})^{p+1}u_h^{(p+1)}(*) - \frac{1}{(p+1)!}v_j(c_{j-2} - c_{j-1})^{p+1}u_h^{(p+1)}(*)$$

$$- \frac{1}{(p+1)!}(1 - \mu_j - v_j)(-c_{j-1})^{p+1}u_h^{(p+1)}(*) - \frac{1}{p!}\tilde{\gamma}_j(-c_{j-1})^p u_h^{(p+1)}(*).$$

In the remainder term, $u_h^{(p+1)}$ is evaluated in various points in (t_n, t_{n+1}). The formulas (4.3), (4.4) also hold for $j = 1$ if we set $\mu_1 = 1$, $c_0 = c_{-1} = 0$ and $v_1 = \tilde{\gamma}_1 = 0$.

The coefficients θ_{1j}, θ_{2j} can be written in a more convenient form. We have

$$(4.5) \qquad \theta_{1j} = c_j - \mu_j c_{j-1} - v_j c_{j-2} - \tilde{\mu}_j - \tilde{\gamma}_j.$$

Relations (2.2) then imply that $\theta_{1j} = 0$ $(1 \le j \le s)$ and thus the contribution of the *temporal errors* to all defects r_j is always $O(\tau^2)$. Furthermore, by inserting the expression for $\tilde{\gamma}_j$ that follows from (2.2) into (4.4), we get

$$(4.6) \qquad \theta_{2j} = \tfrac{1}{2}(c_j^2 - \mu_j c_{j-1}^2 - v_j c_{j-2}^2) - \tilde{\mu}_j c_{j-1} \qquad (1 \le j \le s).$$

For the second order scheme, the conditions (2.6) then imply

$$(4.7) \quad \theta_{2j} = 0 \quad (4 \le j \le s), \qquad \theta_{21} = \tfrac{1}{2} c_1^2, \qquad \theta_{22} = -\tfrac{1}{2} \mu_2 c_1^2, \qquad \theta_{23} = -\tfrac{1}{2} v_3 c_1^2.$$

In the next section these results will be used to prove convergence. Formula (4.3) will be applied with $p = 1, 2$ for the 1st and 2nd order schemes, respectively. For the convergence analysis *an upper bound for the remainder terms ρ_j is needed*. For the sake of simplicity, we will here derive such a bound for the undamped schemes $(\varepsilon = 0)$.

Consider the 1st order scheme, defined by (2.16)–(2.18), with $\varepsilon = 0$. There holds, for all $1 \le j \le s$,

$$|c_j - c_{j-1}| = s^{-2}(2j-1) \le 2 s^{-1}, \qquad 1 - \mu_j - v_j = 0 \quad \text{and} \quad \tilde{\gamma}_j = 0.$$

Further, $v_1 = 0$ and $v_j = -1$ for $2 \le j \le s$. Hence, (4.4) with $p = 1$ gives

$$(4.8) \qquad \|\rho_j\| \le 4 s^{-2} \max_{t_n \le t \le t_{n+1}} \|u_h^{(2)}(t)\| \qquad (1 \le j \le s).$$

Next, consider the 2nd order scheme, whose coefficients are given by (2.19)–(2.21), with $\varepsilon = 0$. There holds

$$w_1 = 3/(s^2 - 1), \quad b_0 = b_1 = b_2, \quad c_1 = \tfrac{1}{4} c_2$$

and

$$b_j = \tfrac{1}{3}(1 - j^{-2}), \quad c_j = (s^2 - 1)^{-1}(j^2 - 1) \quad (j \ge 2).$$

It easily follows that there is a constant $C > 0$, independent of j and s, such that

$$|c_j - c_{j-1}| \le C s^{-1}, \quad |\tilde{\gamma}_j| \le C s^{-2} \qquad (1 \le j \le s).$$

Further we have

$$(1 - \mu_j - v_j) c_{j-1} = (b_j^{-1} - 2 b_{j-1}^{-1} + b_{j-2}^{-1}) b_j c_{j-1} \qquad (2 \le j \le s)$$

and it follows that

$$|(1 - \mu_j - v_j) c_{j-1}| \le C s^{-2} \qquad (j = 2, 3, 4)$$

for some $C > 0$. For $j > 4$,

$$|(1 - \mu_j - \nu_j) c_{j-1}| = |3 b_j (s^2 - 1)^{-1} ((j-1)^2 - 1) [\varphi(j) - 2\varphi(j-1) + \varphi(j-2)]|$$
$$\leq 3 b_j (s^2 - 1)^{-1} ((j-1)^2 - 1) \max_{j-2 \leq x \leq j} |\varphi''(x)|$$

with $\varphi(x) = (x^2 - 1)^{-1} x^2$ (for $x \geq 2$). Since

$$|\varphi''(x)| = 2 \varphi(x)^3 (3 + x^{-2}) x^{-4} \leq 16 x^{-4}$$

for $x \geq 2$, there is a $C > 0$ such that

$$|(1 - \mu_j - \nu_j) c_{j-1}| \leq C s^{-2} \qquad (4 \leq j \leq s).$$

Finally, with our convention $\mu_1 = 1$, $\nu_1 = c_0 = 0$, we have $(1 - \mu_j - \nu_j) c_{j-1} = 0$ for $j = 1$. Combining the above inequalities, we obtain for the 2nd order scheme with $p = 2$,

$$(4.9) \qquad \|\rho_j\| \leq C s^{-2} \max_{t_n \leq t \leq t_{n+1}} \|u_h^{(3)}(t)\| \qquad (1 \leq j \leq s),$$

where $C > 0$ is independent of s.

Of interest is to observe that the two bounds (4.8) and (4.9) are proportional to s^{-2}. This means that at each stage the *remainder term* contained in the defect (4.3) diminishes with s^{-2} for increasing s. This is also true for the *spatial error* part in (4.3), i.e.,

$$(4.10) \qquad \|\tilde{\mu}_j \alpha_h(t_n + c_{j-1} \tau) + \tilde{\gamma}_j \alpha_h(t_n)\| \leq C s^{-2} \max_{t_n \leq t \leq t_{n+1}} \|\alpha_h(t)\|,$$

since the coefficients $\tilde{\mu}_j$, $\tilde{\gamma}_j$ are bounded by $C s^{-2}$ with $C > 0$ a constant independent of s and j. We have strong numerical evidence that these results are also valid in case of damping ($\varepsilon > 0$). However, the derivation of the bounds (4.8), (4.9) then becomes rather technical and lengthy, while no more insight is obtained.

5 Convergence Analysis: A Bound for the Full Global Error

The results of the two previous sections are now combined so as to derive a bound for the full global error. Hence we again consider the *linear problem class* (1.4) (cf. Sect. 3) and, for simplicity, restrict ourselves to the *undamped schemes* (cf. Sect. 4). In our analysis, the time step τ and the grid distances in space, parametrized by h, are allowed to tend to zero simultaneously and independently of each other. Usually, convergence for explicit methods applied to parabolic equations requires a stepsize restriction $\tau \sigma(M) \leq \text{const.}$, $\sigma(M) \sim h^{-2}$, due to stability. With the RKC schemes $\tau \sigma(M)$ is allowed to become arbitrarily large, stability being achieved by taking s sufficiently large. This advantage over standard explicit methods is fully justified by our *unconditional convergence* analysis where the assumption $\tau \sigma(M) \leq \beta$ is to be interpreted as a condition on s, rather than as a restriction on τ.

Let $e_n = u_h(t_n) - U_n$ be the *full global error*. For these errors we have (cf. (3.6) or (3.10)) the recursion

$$(5.1) \qquad e_{n+1} = P_s(Z) e_n + \sum_{k=1}^{s} Q_{sk}(Z) r_k,$$

where $Z = \tau M$ and the vectors r_k are the local defects due to discretization. Upper bounds for $\|e_n\|$ will be derived by elaborating this recursion with the help of our estimates of the local defects and our results on internal stability.

In the following, C will denote a positive constant independent of τ, M and s, not necessarily always with the same value.

The RKC1 Method $(\varepsilon = 0)$. Consider the method defined by (2.16)–(2.18) with $\varepsilon = 0$. This method was constructed such that the temporal ODE order is one. With temporal ODE order we mean the order obtained from an analysis where the dimension of the problem, and thus the space grid, is fixed. We will show that we have also temporal order one for any value of $\sigma(M)$ and s, hence for any spatial grid refinement, provided $\tau \sigma(M) \leq \beta$.

Suppose $u_h \in C^2[0, T]$. From the results of Sect. 4 (see (4.8), (4.10)) it directly follows that there is a $C > 0$ such that

$$(5.2) \qquad \|r_k\| \leq C \tau s^{-2} (\tau \max_{t_n \leq t \leq t_{n+1}} \|u_h^{(2)}(t)\| + \max_{t_n \leq t \leq t_{n+1}} \|\alpha_h(t)\|) \qquad (1 \leq k \leq s).$$

Using Theorem 3.1, we then immediately obtain the following bound for the global errors:

Theorem 5.1. *Assume $u_h \in C^2[0, T]$ and $\tau \sigma(M) \leq \beta$. Then the global errors of the undamped RKC1 scheme satisfy*

$$\|e_n\| \leq C(\tau \max_{0 \leq t \leq T} \|u_h^{(2)}(t)\| + \max_{0 \leq t \leq T} \|\alpha_h(t)\|) \qquad (n = 1, 2, \ldots; n\tau \leq T)$$

with a constant $C > 0$ independent of τ, M and s. $\square$

The RKC2 Method $(\varepsilon = 0)$. Consider the method defined by (2.19)–(2.21) with $\varepsilon = 0$. This method was constructed such that the temporal ODE order is two. We will derive an error bound which proves that RKC2 has 'almost' order 2 in time for any value of $\sigma(M)$ and s, provided $\tau \sigma(M) \leq \beta$.

Assume $u_h \in C^3[0, T]$ and $\tau \sigma(M) \leq \beta$. Application of formula (4.3) with $p = 2$ shows that, since $\theta_{1k} = 0$ for all k,

$$(5.3) \qquad \left\| \sum_{k=1}^{s} Q_{sk}(Z) r_k \right\| \leq \tau^2 \left\| \sum_{k=1}^{s} \theta_{2k} Q_{sk}(Z) u_h^{(2)}(t_n + c_{k-1} \tau) \right\|$$

$$+ \sum_{k=1}^{s} \|Q_{sk}(Z)\| (\tau^3 \|\rho_k\| + \tau(\|\tilde{\mu}_k \alpha_h(t_n + c_{k-1} \tau) + \tilde{\gamma}_k \alpha_h(t_n)\|)).$$

Using (4.9)–(4.10) and the inequality $\|Q_{sk}(Z)\| \leq C(s - k + 1)$, as in the proof of Theorem 3.1, the entire last term is bounded by

$$C(\tau^3 \max_{t_n \leq t \leq t_{n+1}} \|u_h^{(3)}(t)\| + \tau \max_{t_n \leq t \leq t_{n+1}} \|\alpha_h(t)\|).$$

Further, the coefficients θ_{2k} are known from (4.7). Using definition (3.9) for Q_{sk}, the relations $b_1 = b_2$, $v_3 = -b_3 b_1^{-1}$ and $\mu_2 = 2$, it then follows that

$$(5.4) \qquad \left\| \sum_{k=1}^{s} \theta_{2k} Q_{sk}(Z) u_h^{(2)}(t_n + c_{k-1} \tau) \right\| \leq \| R(Z) u_h^{(2)}(t_n) \|$$

$$+ \tfrac{1}{2} c_1^2 \| \mu_2 Q_{s2}(Z) [u_h^{(2)}(t_n + c_1 \tau) - u_h^{(2)}(t_n)] \|$$

$$+ \tfrac{1}{2} c_1^2 \| v_3 Q_{s3}(Z) [u_h^{(2)}(t_n + c_2 \tau) - u_h^{(2)}(t_n)] \|,$$

where the polynomial $R(z)$ is given by

$$(5.5) \qquad R(z) = \sum_{k=1}^{3} \theta_{2k} Q_{sk}(z) = \tfrac{1}{2} c_1^2 b_1^{-1} b_s [S_{s-1}(1 + w_1 z)$$

$$- 2 S_{s-2}(1 + w_1 z) + S_{s-3}(1 + w_1 z)].$$

As $c_1 = \tfrac{1}{4} c_2$, $c_2 = 3(s^2 - 1)^{-1}$, we can bound $\| R(Z) \|$ by $C s^{-3}$ for some $C > 0$. Further we have $\| u_h^{(2)}(t_n + c_k \tau) - u_h^{(2)}(t_n) \| \leq c_k \tau \max \| u_h^{(3)}(t) \|$. Inserting all these bounds in (5.3), we obtain

$$(5.6) \qquad \left\| \sum_{k=1}^{s} Q_{sk}(Z) r_k \right\| \leq C s^{-3} \tau^2 \max \| u_h^{(2)}(t) \| + C \tau^3 \max \| u_h^{(3)}(t) \|$$

$$+ C \tau \max \| \alpha_h(t) \|,$$

and by using the recursion (5.1) we arrive at the following result:

Theorem 5.2. *Assume* $u_h \in C^3 [0, T]$ *and* $\tau \sigma(M) \leq \beta$. *Then the global errors of the undamped RKC2 scheme satisfy*

$$\| e_n \| \leq C (s^{-3} \tau \max_{0 \leq t \leq T} \| u_h^{(2)}(t) \| + \tau^2 \max_{0 \leq t \leq T} \| u_h^{(3)}(t) \| + \max_{0 \leq t \leq T} \| \alpha_h(t) \|)$$

(for $n = 1, 2, \ldots$; $n\tau \leq T$) with a constant $C > 0$ independent of τ, M and s.   □

Theorems 5.1 and 5.2 prove convergence of RKC1 and RKC2, respectively, irrespective the size of s or $\tau \sigma(M)$. The analysis also shows that the use of many stages within one single step does not adversely affect the accuracy. The temporal error is merely determined by τ and the smoothness of u_h as a function of t. The spatial error is merely determined by the size of the local space truncation error, the common situation.

Theorem 5.2 shows that RKC2 is of 'almost' temporal order 2 and that with s large order 2 will be observed. Theorem 5.2 does not reveal the classical order 2 for fixed M (fixed space grid) and s. This property can be proved with the above analysis, based on the defects, by noting that $R(0) = 0$. Then, since $z^{-1} R(z)$ is bounded as a function of $z \in [-\beta, 0]$, the estimate

$$(5.7) \qquad \| R(Z) u_h^{(2)}(t_n) \| \leq \| R(Z) \| \, \| u_h^{(2)}(t_n) \|$$

that led to the τ^2-term in (5.6), can be replaced by

$$(5.8) \qquad \| R(Z) u_h^{(2)}(t_n) \| \leq \tau \| Z^{-1} R(Z) \| \, \| M u_h^{(2)}(t_n) \|.$$

This estimate gives an extra τ because $M u_h^{(2)}(t)$ is bounded on a fixed space grid. $M u_h^{(2)}(t)$ is also bounded for $h \to 0$ if $u^{(2)}$ satisfies homogeneous boundary

conditions imposed by the homogeneous operator M (see, e.g., [10]). In that case, estimate (5.7) can also be replaced by (5.8) in the proof of Theorem 5.2, which then leads to 'truly' order 2. We shall not pursue the convergence matter further, since in most applications RKC2 will use values of s for which the error term

$$s^{-3}\tau \max_{0\le t\le T} \|u_h^{(2)}(t)\|$$

is hardly noticeable.

6 Numerical Examples

Our first example problem is Fisher's equation

(6.1) $u_t = u_{xx} + u^2(1-u), \qquad 0 \le x, \ t \le 1,$

with the exact solution $u(x,t) = (1+\exp(v(x-vt)))^{-1}, v = \frac{1}{2}\sqrt{2}$. We use this equation to illustrate the convergence behaviour in a non-model situation. The second derivative is approximated with 2nd order central differences on a uniform grid with gridsize h. The schemes are applied with damping. For RKC1 the damping parameter $\varepsilon = 0.05$ (see (2.16)) and for RKC2 $\varepsilon = \frac{2}{13}$ (see (2.19)). In the experiment we let $\tau = h$ decrease and s is chosen to satisfy the stability condition $\tau\sigma \le \beta(s)$, while s is taken as small as possible. We put $\sigma = 4h^{-2} + 4$ and

(6.2) $s = 1 + \text{entier}\,[(1+\tau\sigma/1.90)^{\frac{1}{2}}]$ for RKC1,
             $s = 1 + \text{entier}\,[(1+\tau\sigma/0.65)^{\frac{1}{2}}]$ for RKC2.

The number 4 in the expression for the spectral radius σ serves as a (conservative) upperbound for the derivative of the inhomogeneous term in (6.1). Note that we select s slightly larger than necessary to satisfy the condition $\tau\sigma \le \beta(s)$.

Table 6.1 lists maximum errors at $t = 1$ for a sequence of $\tau = h$ values. As expected, RKC1 converges with order 1 and RKC2 with order 2. We owe the high level of accuracy to the high degree of smoothness of u.

Our second example problem is a system of two reaction-diffusion equations in $2D$, a type of problem already mentioned in the introduction as being appropriate to be solved with the explicit RKC scheme. The problem is the Brusselator with diffusion given in [4], p. 381:

(6.3) $u_t = 1 + u^2 v - 4.4u + \delta(u_{xx} + u_{yy}),$
             $v_t = 3.4u - u^2 v + \delta(v_{xx} + v_{yy}), \qquad 0 \le x, \ y \le 1, \ t > 0,$

Table 6.1. Convergence test on Fisher's equation

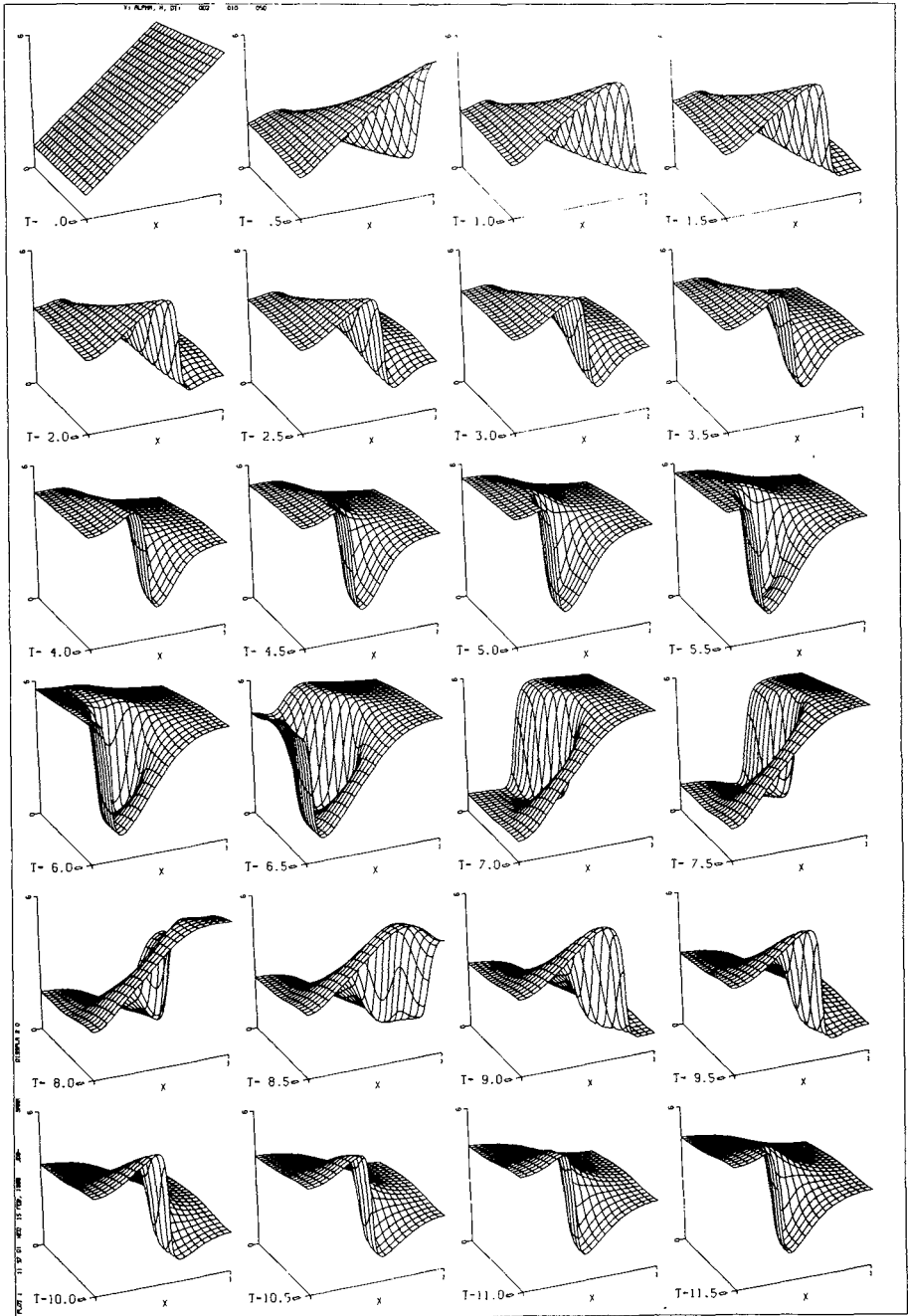| $(\tau=h)^{-1}$	RKC1		RKC2	
	s	error	s	error
5	4	$0.63 \ 10^{-4}$	6	$0.15 \ 10^{-4}$
10	5	$0.26 \ 10^{-4}$	8	$0.25 \ 10^{-5}$
20	7	$0.13 \ 10^{-4}$	12	$0.54 \ 10^{-6}$
40	10	$0.44 \ 10^{-5}$	16	$0.15 \ 10^{-6}$
80	14	$0.21 \ 10^{-5}$	23	$0.33 \ 10^{-7}$
160	19	$0.99 \ 10^{-6}$	32	$0.77 \ 10^{-8}$
320	26	$0.48 \ 10^{-6}$	45	$0.19 \ 10^{-8}$

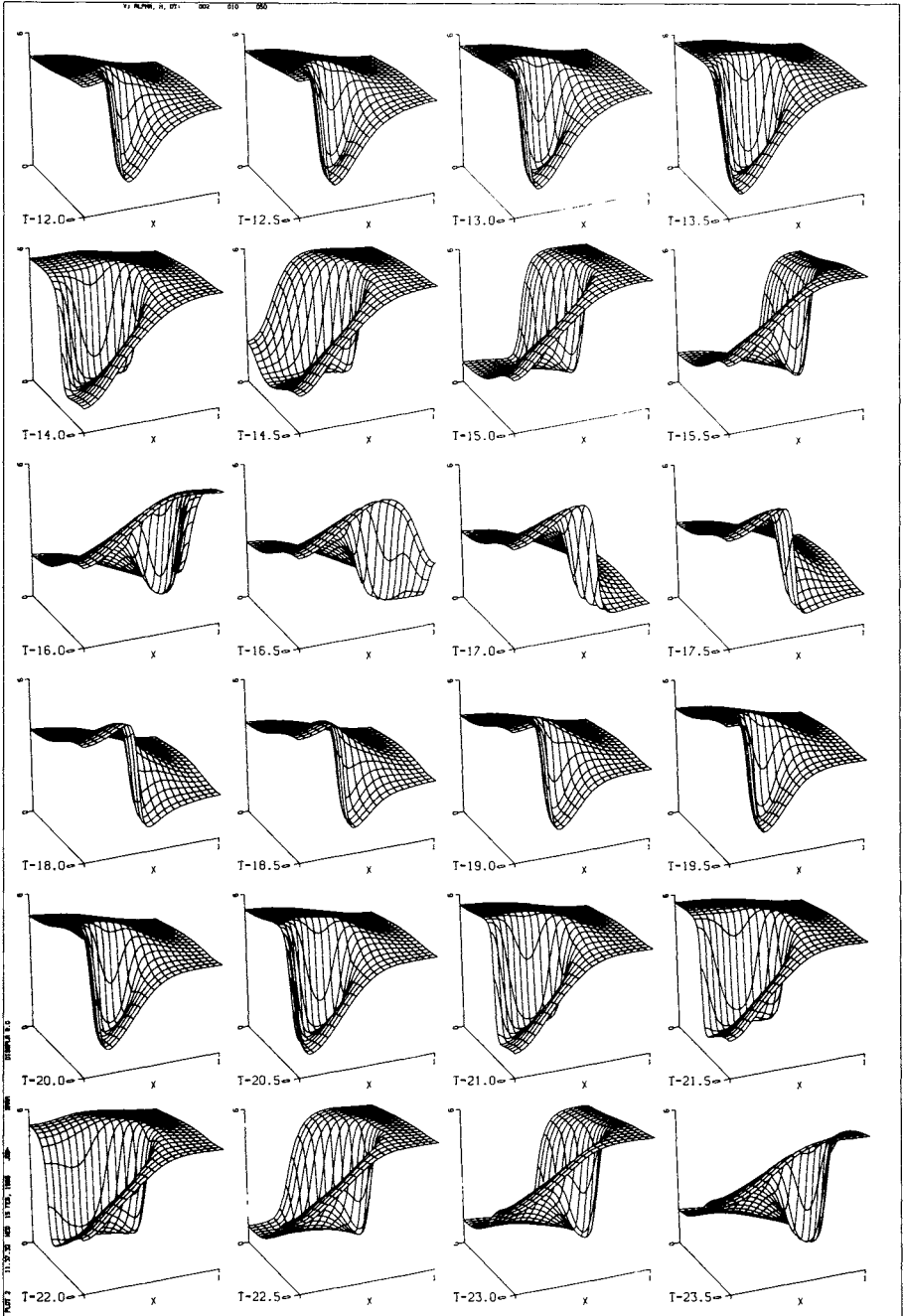**Fig. 6.1** Solution v of (6.3) for $t = 0, 0.5, \ldots, 11.5$

Fig. 6.1 (cont.) Solution v of (6.3) for $t = 12, 12.5, \ldots, 23.5$

subjected to homogeneous Neumann boundary conditions and to the initial conditions $u(x, y, 0) = 0.5 + y$, $v(x, y, 0) = 1 + 5x$. Note that these are not consistent with the boundary conditions. The reaction term and the diffusion parameter δ, here 0.002, are such that the flat initial solutions develop into 'periodically moving' wave fronts with a 'period of about 8'.

Using 2nd order differencing, the PDE is spatially discretized on the uniform mesh

$$\{(x_i, y_j): x_i = (i - 1)\, h, \; y_j = (j - 1)\, h, \; 1 \leqq i, j \leqq N, \; h = 1/(N - 1)\}.$$

Note that this mesh covers the boundary, i.e., we also integrate on the boundary. The additional neighbouring points $x_{-1}, x_{N+1}, y_{-1}, y_{N+1}$ outside the mesh are eliminated with the central difference replacements for the Neumann boundary conditions. We thus obtain a semi-discrete system of dimension $2N^2$.

In our experiment, $N = 101$ resulting in 20402 equations. A fine mesh is needed as the moving fronts are locally rather steep. The spectral radius σ of the Jacobian matrix of the semi-discrete operator is estimated as $\sigma = 10 + 8\delta h^{-2}$ and, like before, the number of stages s per step is determined by (6.2). Using $h = 10^{-2}$ and $\tau = \frac{1}{20}$, we thus get $\sigma = 170$ and $s(\text{RKC2}) = 4$. This small number of stages is due to the small value of δ.

Figure 6.1 shows the evolution of component v, obtained with RKC2 for $h = 10^{-2}$ and $\tau = 0.05$, over the time interval $[0, 23.5]$. The plots are at equidistant time values $t = 0, 0.5, \ldots$ and show the computed solution along every fifth grid line. RKC2 performs very satisfactorily on this problem. The periodic solution behaviour is accurately simulated. The costs of the integration amount to $23.5 \times 20 \times 4 = 1880$ explicit evaluations of the semi-discrete PDE operator.

References

1. Abramowitz, M., Stegun, I.A.: Handbook of mathematical functions. Appl. Math. Ser. **55**, National Bureau of Standards Washington 1964
2. Dekker, K., Verwer, J.G.: Stability of Runge-Kutta methods for stiff nonlinear differential equations. Amsterdam, New York, Oxford: North-Holland 1984
3. Guillou, A., Lago, B.: 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. 1$^{\text{er}}$ Congr. Assoc. Francaise Calcul, AFCAL, Grenoble, pp. 43–56, 1961
4. Hairer, E., Nørsett, S.P., Wanner, G.: Solving ordinary differential equations I, Nonstiff problems. Springer Series in Computational Mathematics, vol. 8, Berlin-Heidelberg-New York: Springer 1987
5. Houwen, P.J. van der: Explicit Runge-Kutta formulas with increased stability boundaries. Numer. Math. **20**, 149–164 (1972)
6. Houwen, P.J. van der: Construction of integration formulas for initial value problems. Amsterdam, New York, Oxford: North-Holland 1977
7. Houwen, P.J. van der, Sommeijer, B.P.: On the internal stability of explicit, m-stage Runge-Kutta methods for large m-values. Z. Angew. Math. Mech. **60**, 479–485 (1980)
8. Houwen, P.J. van der: On the time-integration of parabolic differential equations. In: Watson, G.A. (ed.): Numerical Analysis. Lecture Notes in Mathematics, vol. 912, pp. 157–168. Berlin-Heidelberg-New York: Springer 1982
9. Hundsdorfer, W.H., Verwer, J.G.: Stability and convergence of the Peaceman-Rachford ADI method for initial-boundary value problems. Math. Comput. **53**, 81–102 (1989)

10. Sanz-Serna, J.M., Verwer, J.G., Hundsdorfer, W.H.: Convergence and order reduction of Runge-Kutta schemes applied to evolutionary problems in partial differential equations. Numer. Math. **50**, 405–418 (1987)
11. Sanz-Serna, J.M., Verwer, J.G.: Stability and convergence at the PDE/stiff ODE interface. Appl. Numer. Math. **5**, 117–132 (1989)
12. Sommeijer, B.P., Houwen, P.J. van der: On the economization of stabilized Runge-Kutta methods with applications to parabolic initial value problems. Z. Angew. Math. Mech. **61**, 105–114 (1981)
13. Sommeijer, B.P., Verwer, J.G.: A performance evaluation of a class of Runge-Kutta-Chebyshev methods for solving semi-discrete parabolic differential equations. Report NW 91/80, Centre for Mathematics and Computer Science (CWI), Amsterdam 1980