Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

P.J. van der Houwen, B.P. Sommeijer, C.T.H. Baker

On the stability of predictor-corrector methods
for parabolic equations with delay

Department of Numerical Mathematics Report NM-R8410 August

The Centre for Mathematics and Computer Science is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

ON THE STABILITY OF PREDICTOR-CORRECTOR METHODS FOR PARABOLIC EQUATIONS WITH DELAY

P.J. VAN DER HOUWEN, B.P. SOMMEIJER
Centre for Mathematics and Computer Science, Amsterdam

C.T.H. BAKER
University of Manchester

Diffusion problems where the current state depends upon an earlier one give rise to parabolic equations with delay. The efficient numerical solution of classical parabolic equations can be accomplished via methods for stiff differential equations; one such class are predictor-corrector-type methods with extended real stability intervals and with reduced storage requirements. Analogous methods for equations with delay are proposed and analyzed here. Numerical experiments which illustrate the theoretical results are reported.

1980 MATHEMATICAL SUBJECT CLASSIFICATION: 65Q05, 65M20, 35R10.
KEY WORDS & PHRASES: parabolic equations, delay equations, predictor-corrector methods.
NOTE: This report will be submitted for publication elsewhere.

Report NM-R8410
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

1. Introduction

There is an extensive literature on the theory and numerical solution of parabolic equations. The inclusion of a delay in the classical problems of mathematical physics leads to partial differential equations with delay only in time, t. As illustration, consider the *generalized diffusion equation*

$$\frac{\partial}{\partial t}u(t,x) = \alpha^2\frac{\partial^2}{\partial x^2}u(t,x) + f(u(t-\omega,x)) \tag{1.1}$$

$$0 \leqslant x \leqslant 1, \quad t \geqslant \omega$$

with homogeneous Dirichlet conditions on $x = 0$ and $x = 1$ and the prescribed initial function

$$u(t,x) = \phi(t,x) \quad (0 \leqslant t \leqslant \omega,\ 0 \leqslant x \leqslant 1). \tag{1.2}$$

The existence and uniqueness theory for problems of this type has been discussed by Travis and Webb [13], for example, in the case that f is a linear or non-linear scalar-valued function. Cases where the term f is replaced by more general expressions involving the state $u(t-\omega,x)$ also arise, for example (see El'sgol'ts and Norkin [7, pp. 269-272]):

$$\frac{\partial}{\partial t}u(t,x) = \alpha^2\frac{\partial^2}{\partial x^2}u(t,x) + \beta^2\frac{\partial^2}{\partial x^2}u(t-\omega,x); \tag{1.3}$$

the theory of a class of examples of general type is discussed by Artola [1].

Wang [14] provides an example of a realistic system (an automatically controlled furnace): the system is modelled by an equation which falls into the class of problems of the form

$$\frac{\partial}{\partial t}u(t,x) = \mathfrak{L}u(t,x) + f\left(t,x,u(t,x),u(t-\omega^{(1)},x),\ldots,\frac{\partial u}{\partial x}(t,x),\frac{\partial u}{\partial x}(t-\omega^{(2)},x),\ldots\right), \tag{1.4}$$

involving multiple delays, $\omega^{(1)},\omega^{(2)}$, etc., where $\mathfrak{L}$ is a linear operator which is uniformly elliptic in x.

Time delays can enter into diffusion systems in various ways. Wang [15] considers realistic systems in which the delay term is absent from the differential equation but enters into the boundary conditions valid for $x = 0, 1$; $t \geqslant 0$.

In the case $f \equiv 0$ in (1.1), numerical methods for the approximation of $u(t,x)$ can be derived, as is well-known (see also e.g. [10, p. 249]), by semi-discretization in the x-variable and the numerical solution of the resulting ordinary differential equations with respect to time. In the generalized equations considered above, the process of semi-discretization produces (in place of a system of ordinary differential equations) a system of retarded differential equations. Thus, the simplest discretization scheme yields for (1.1) the equations

$$\dot{y}_i(t) = \frac{\alpha^2}{h^2}\left\{y_{i+1}(t) - 2y_i(t) + y_{i-1}(t)\right\} + f(y_i(t-\omega),ih) \tag{1.5}$$

and that for (1.3) yields

$$\dot{y}_i(t) = \frac{\alpha^2}{h^2}\left\{y_{i+1}(t) - 2y_i(t) + y_{i-1}(t)\right\} + \frac{\beta^2}{h^2}\left\{y_{i+1}(t-\omega) - 2y_i(t-\omega) + y_{i-1}(t-\omega)\right\}, \tag{1.6}$$

where $y_i(t) \approx u(t,ih)$, $h = 1/(N+1)$, $i = 1,\ldots,N$.

In the case of a general linear problem, involving one delay, the semi-discretization process yields a system of equations of the form

$$\dot{y}(t) = \Omega_1 y(t) + \Omega_2 y(t-\omega), \tag{1.7}$$

where $y(t) = [y_1(t),\ldots,y_N(t)]^T$. In the cases where the matrices Ω_i are simultaneously diagonisable (which occurs for (1.5) where $f(y_i,ih) = y_i$, and for (1.6)) the study of scalar equations of the form

$$\dot{y}(t) = \omega_1 y(t) + \omega_2 y(t-\omega) \tag{1.8}$$

provides insight concerning the behaviour as $t \to \infty$ of solutions of (1.7). Moreover, (i) in the case (1.5)

where h is small, $f(y)$ is linear in y (or for small y can be approximated by a linear term, and $y(t) \to 0$ as $t \to \infty$) or (ii) in the case (1.6) where $\alpha^2 >> \beta^2$ and h is arbitrary, the model equation (1.8) which is of interest is one in which $|\omega_1| >> |\omega_2|$. In our discussion of stability this case will occupy our attention in particular.

Retarded differential equations are derived here through semi-discretization but also arise directly in their own right in various applications; see Chosky [5], Weiss [16]. It is well-known that the efficient numerical solution of the ordinary differential equations obtained on semi-discretization of (1.1) with $f \equiv 0$ requires numerical methods with large regions of stability. Generalized predictor-corrector methods with extended region of stability have been derived and studied by van der Houwen and Sommeijer [9] with this application in mind. The same authors have [8] adapted their numerical methods to ordinary differential equations with delay, of the general form $\dot{y}(t) = f(t,y(t),y(t-\omega))$, $\omega > 0$. Since such methods are well-suited to the numerical solution of the retarded differential equations obtained on semi-discretization of generalized diffusion equations, we develop the results of [9] with this application in mind.

2. Predictor-corrector methods

In this section we will discuss the construction of numerical methods by reference to a general non-linear system of delay equations involving one delay, that is,

$$\dot{y}(t) = f(t,y(t),y(t-\omega)), \quad \omega \equiv \omega(t,y(t)) \geqslant 0, \quad t \geqslant t_0, \tag{2.1}$$

with $y(t)$ prescribed at (and, if necessary, on an interval to the left of) the point t_0. In the present section we assume only such conditions as ensure smoothness of f and the existence of a unique (smooth) solution $y(t)$. (Later, we assume that the Jacobian matrices of derivatives of $f(t,u,v)$ with respect to u and to v have the same eigensystem and real eigenvalues.)

The methods we describe are predictor-corrector methods for use with formulae discussed by Cryer [6]; they reduce, in the case that the delay term is absent, to methods considered in [9] for the initial-value problem

$$\dot{y}(t) = f(t,y(t)) \quad t \geqslant t_0 \tag{2.1'}$$

with $y(t_0)$ prescribed.

We denote by $\{\rho,\sigma\}$ the implicit linear multistep formula (cf. Lambert [10, pp. 11-40]) with first and second characteristic polynomials

$$\rho(\zeta) = \sum_{i=0}^{k} a_i \zeta^{k-i} \quad \text{and} \quad \sigma(\zeta) = \sum_{i=0}^{k} b_i \zeta^{k-i}.$$

We shall call this formula *the corrector formula.* We assume that the corrector is zero-stable, consistent and of order p [ibid, p.23]. We shall denote by $\{\tilde{\rho},\tilde{\sigma}\}$ a corresponding explicit formula *(the predictor formula,* with $b_0 = 0$), and its order by $\tilde{p}$.

It is necessary to adapt the formulae for (2.1') to permit the treatment of (2.1). Our objective, given a constant integration step $\Delta t > 0$, is to approximate the solution $y(t_n)$ of (2.1) at $t_n = t_0 + n\Delta t$ by $y_n (n = 1,2,3,...)$; for this purpose we shall approximate $y(t_n - \omega_n)$, with $\omega_n = \omega(t_n,y_n)$ using polynomial interpolation on the values $y_j, y_{j-1}, ..., y_{j-l}$ where $t_{j-1} < t_n - \omega_n \leqslant t_j$, $j > 0$. Usually, Hermite interpolation is employed; however, in view of the application to parabolic equations we have in mind, we will use one of those backward differentiation formulae which are highly stable as corrector; consequently, no f-values are stored preventing us using Hermite interpolation. The interpolation formula assumes the form

$$\hat{y}(t_n - \omega_n) = E^{-l}\tau(E,\theta_n)y_j \tag{2.2}$$

wherein $t_n - \omega_n = t_j - \theta_n \Delta t$, $0 \leqslant \theta_n < 1$ and E is the forward shift operator $E\phi_n = \phi_{n+1}$. Here, τ is a polynomial in E whose coefficients depend upon θ_n and (2.2) is merely a symbolic form of Newton's backward formula. Concerning τ we assume $\tau(\xi,0) = \xi^l$, $\tau(1,\theta_n) \equiv 1$; the order of accuracy of (2.2) is $l+1$. We shall assume that the order of the interpolation formula (2.2) is at least that of the method

$\{\rho,\sigma\}$ for (2.1'), i.e. $l \geqslant p$.

The formulae which form the basis for the numerical method for (2.1) now comprise (2.2) and

$$\rho(E)y_n - \Delta t\,\sigma(E)f_n = 0 \quad (n \geqslant 0);$$
$$f_n = \begin{cases} f(t_n, y_n, \hat{y}(t_n - \omega_n)) & (t_n - \omega_n > t_0), \\ f(t_n, y_n, y(t_n - \omega_n)) & (t_n - \omega_n \leqslant t_0); \end{cases} \tag{2.3}$$
$$\omega_n = \omega_n(t_n, y_n).$$

We refer to (2.3) as *the delay-corrector formula.*

Since $b_0 \neq 0$ the formulae (2.3) are certainly implicit. At each integration step it is necessary to solve

$$a_0 y_n - b_0 \Delta t f(t_n, y_n, \hat{y}_n(t_n - \omega_n)) = \Sigma_n, \quad a_0 = 1, \tag{2.4}$$

coupled with (2.2), where Σ_n is computable in terms of values of y_k already computed. From now on we assume $a_0 = 1$, $b_0 > 0$. Observe that $\hat{y}(t_n - \omega_n)$ will, when $\omega_n < \Delta t$, depend on the as yet unknown approximation y_n to $y(t_n)$.

In order to solve (2.4) we use the following predictor-corrector scheme:

$$\begin{cases} y_n^{(0)} := \textit{some initial approximation to the exact solution } \eta_n \textit{ of } (2.4), \\ y_n^{(j)} := \mu_j y_n^{(j-1)} + (1 - \lambda_j - \mu_j) y_n^{(j-2)} + \lambda_j b_0 \Delta t f_n^{(j-1)} + \lambda_j \Sigma_n, \; j = 1,\ldots,m, \\ y_n := y_n^{(m)} \end{cases} \tag{2.5}$$

with $\lambda_1 + \mu_1 = 1$ and λ_j, $\mu_j (j > 1)$ to be determined later. Here

$$f_n^{(j-1)} := f(t_n, y_n^{(j-1)}, \hat{y}(t_n - \omega_n(t_n, y_n^{(j-1)}))).$$

In passing we observe that a conventional predictor-corrector method for (2.1), in $P(EC)^m E$ - mode, is obtained if we choose $\mu_j = 0$, $\lambda_j = 1 (j = 1,2,\ldots,m)$.

The general predictor-corrector method (2.5) will be called a *GPC method;* it falls into a still more general class of methods presented in [9]. For our purposes, (2.5) has sufficient degrees of freedom; our aim is the construction of *GPC* methods which permit the choice of large Δt when applied to (1.8) with $|\omega_1| >> |\omega_2|$, bearing in mind the applicability of such methods to the solution via semi-discretization of a class of parabolic equations with delay.

Remark. In practical computations, the choice of $\{\lambda_j, \mu_j\}$ in (2.5) will be determined by local conditions, but we shall ignore this feature until Section 4.

2.1. The local error

In studying the accuracy of the GPC method (2.5) it is convenient to introduce the *iteration polynomials* $P_j(z)$ generated by

$$P_0(z) = 1, \; P_1(z) = 1 - \lambda_1 + \lambda_1 b_0 z, \tag{2.6}$$
$$P_j(z) = (\mu_j + \lambda_j b_0 z) P_{j-1}(z) + (1 - \lambda_j - \mu_j) P_{j-2}(z), \quad j = 2,3,\ldots,m.$$

Notice that $P_j(1/b_0) = 1$ for all j. The polynomials $P_j(z)$ are uniquely associated with the iteration scheme (2.5).

Furthermore, we need the Jacobian matrix of the right-hand side function f_n. Let $g(t, u, v_1, v_2, \ldots, v_l)$ be the function such that

$$f(t_n, y_n, \hat{y}(t_n - \omega_n)) = g(t_n, y_n, y_{n-1}, \ldots).$$

Recalling that if $\omega_n \equiv \omega(t_n, y_n) < \Delta t$ then $\hat{y}(t_n - \omega_n)$ depends upon y_n, we define the Jacobian matrix

$$Z_n := \Delta t \frac{\partial g}{\partial u}(t_n, \eta_n, y_{n-1}, \ldots), \tag{2.7}$$

where η_n is the exact solution (assumed unique) of the delay-corrector formula (2.3). The local error of the GPC method can be expressed in terms of the corresponding errors of the corrector and predictor formula using the iteration polylnomial $P_m(z)$ and the matrix Z_n.

Theorem 2.1. *If $f(t,u,v)$ and the solution $y(t)$ are sufficiently smooth, then, provided $l \geqslant \max\{p,\tilde{p}\}$,*

$$y_n - y(t_n) = [I - P_m(Z_n)](\eta_n - y(t_n)) + P_m(Z_n)(y_n^{(0)} - y(t_n)) + O(\Delta t^{2p+3} + \Delta t^{2\tilde{p}+3}),$$

where $\tilde{p}$ and p are the orders of accuracy of the predictor formula for $y_n^{(0)}$ and the corrector formula for η_n, respectively, and where we assume $y_j = y(t_j), j < n$. □

Proof. Let the iteration error of the iteration scheme (2.5) be defined by $\epsilon_j := y_n^{(j)} - \eta_n$. Then, by virtue of our assumption of exact back-values (the so-called localizing assumption [10, p. 27]), the local error of (2.5) is given by

$$y_n - y(t_n) = y_n - \eta_n + \eta_n - y(t_n) = \epsilon_m + (\eta_n - y(t_n)), \tag{2.8}$$

that is the sum of the final iteration error and the local error of the corrector formula. The magnitude of the last error can be derived from the literature (e.g. [12]); here we concentrate on the iteration error.

Substitution of $y_n^{(j)} = \eta_n + \epsilon_j$ into (2.5) yields

$$\epsilon_j = \mu_j \epsilon_{j-1} + (1 - \lambda_j - \mu_j)\epsilon_{j-2} - \lambda_j(\eta_n - b_0 \Delta t f_n^{(j-1)} - \Sigma_n),$$
$$f_n^{(j)} = f(t_n, \eta_n + \epsilon_j, \hat{y}_n(t_n - \omega_n(t_n, \eta_n + \epsilon_j))).$$

The expression between parentheses in the recurrence relation for the iteration error is a sort of residual of the corrector equation. The residual term $r_n(y_n^{(j-1)})$ can be written as

$$r_n(y_n^{(j-1)}) = r_n(\eta_n) - b_0[Z_n \epsilon_{j-1} + \Delta t O(\|\epsilon_{j-1}\|^2)].$$

Since $r_n(\eta_n) = 0$ we obtain the recurrence relation

$$\epsilon_j = (\mu_j + \lambda_j b_0 Z_n)\epsilon_{j-1} + (1 - \lambda_j - \mu_j)\epsilon_{j-2} + \Delta t O(\|\epsilon_{j-1}\|^2). \tag{2.9}$$

This relation holds for $j = 1,2,\ldots,m$ (recall that $\lambda_1 + \mu_1 = 1$). For $j = 0$ we have

$$\epsilon_0 = y_n^{(0)} - \eta_n = y_n^{(0)} - y(t_n) + y(t_n) - \eta_n = O(\Delta t^{\tilde{p}+1} + \Delta t^{p+1}). \tag{2.10}$$

Let us state the proposition

$$\epsilon_j = P_j(Z_n)\epsilon_0 + O(\Delta t^s), \quad s := 2 \min\{p,\tilde{p}\} + 3, \quad j \geqslant 0 \tag{2.11}$$

which is either true or false. Since the proposition is true for $j = 0$, we only have to show that (2.11) holds for $j = i+1$ if it holds for $j \leqslant i$. By finite induction it then holds for all j. From (2.9), and (2.11) with $j \leqslant i$ it follows that

$$\begin{aligned}
\epsilon_{i+1} &= (\mu_{i+1} + \lambda_{i+1} b_0 Z_n)\epsilon_i + (1 - \lambda_{i+1} - \mu_{i+1})\epsilon_{i-1} + \Delta t O(\|\epsilon_i\|^2) \\
&= (\mu_{i+1} + \lambda_{i+1} b_0 Z_n)[P_i(Z_n)\epsilon_0 + O(\Delta t^s)] \\
&\quad + (1 - \lambda_{i+1} - \mu_{i+1})[P_{i-1}(Z_n)\epsilon_0 + O(\Delta t^s)] \\
&\quad + \Delta t O(\|\epsilon_0 + \Delta t^s\|^2).
\end{aligned}$$

Finally, by using (2.6) and (2.10) we find that

$$\epsilon_{i+1} = P_{i+1}(Z_n)\epsilon_0 + O(\Delta t^s).$$

Hence, (2.11) holds for all j so that we can write

$$\epsilon_m = P_m(Z_n)\epsilon_0 + O(\Delta t^s) = P_m(Z_n)[y_n^{(0)} - y(t_n) + y(t_n) - \eta_n] + O(\Delta t^s).$$

On substitution into (2.8) we arrive at the result of the theorem. □

From this theorem we immediately conclude that the order of the GPC method is given by $p^* = \min\{p + r, \tilde{p} + \tilde{r}\}$ where r is the multiplicity of the zero at $z = 0$ of $1 - P_m(z)$ and $\tilde{r}$ the multiplicity

of the zero at $z=0$ of $P_m(z)$.

In actual applications the local error $\eta_n - y(t_n)$ of the corrector is usually small in comparison with the local error $y_n^{(0)} - y(t_n)$ of the predictor. Therefore, we will only consider polynomials with $r = 0$ and choose $P_m(z)$ such that $P_m(Z_n)$ decreases the magnitude of the predictor error. If Δt is small, that is $\|Z_n\|$ is small, this can be achieved by choosing $\tilde{r}$ as large as possible. For example, the conventional predictor-corrector method uses $P_m(z) = (b_0 z)^m$ so that $\tilde{r} = m$. However, we want to use relatively large integration steps and consequently (assuming that the order terms in the statement of Theorem 2.1 remain negligible) we should choose $P_m(z)$ such that $|P_m(z)|$ is small in a sufficiently large neighbourhood of the origin on the negative z-axis. As we will see in the stability analysis of the GPC method, the stability condition also requires $|P_m(z)|$ to be small on a negative interval, $[-\beta,0]$ say. Therefore, we postpone the choice of $P_m(z)$ to Section 3.2.

3. Stability theory

We mentioned previously, in Section 1, that the scalar test equation (1.8) provides insight concerning the behaviour as $t \to \infty$ of solutions of (1.7); we consider (1.7) as a linearization of (2.1) with Ω_1 and Ω_2 being defined as the Jacobian matrices $\partial f / \partial u$ and $\partial f / \partial v$ of $f(t,u,v)$, and assuming that Ω_1 and Ω_2 share the same eigensystem. The region in the real (ω_1,ω_2)-plane where the test equation (1.8) has solutions $y(t)$ such that $y(t) \to 0$ as $t \to \infty$, for a given delay ω, will be called the *stability region* corresponding to the delay ω. It can be shown (see e.g. [4, p. 444]) that in the real (ω_1,ω_2)-plane this open region is bounded by the curve

$$\omega_1 = q \operatorname{cotan} \omega q, \quad \omega_2 = -\frac{q}{\sin \omega q}, \tag{3.1}$$

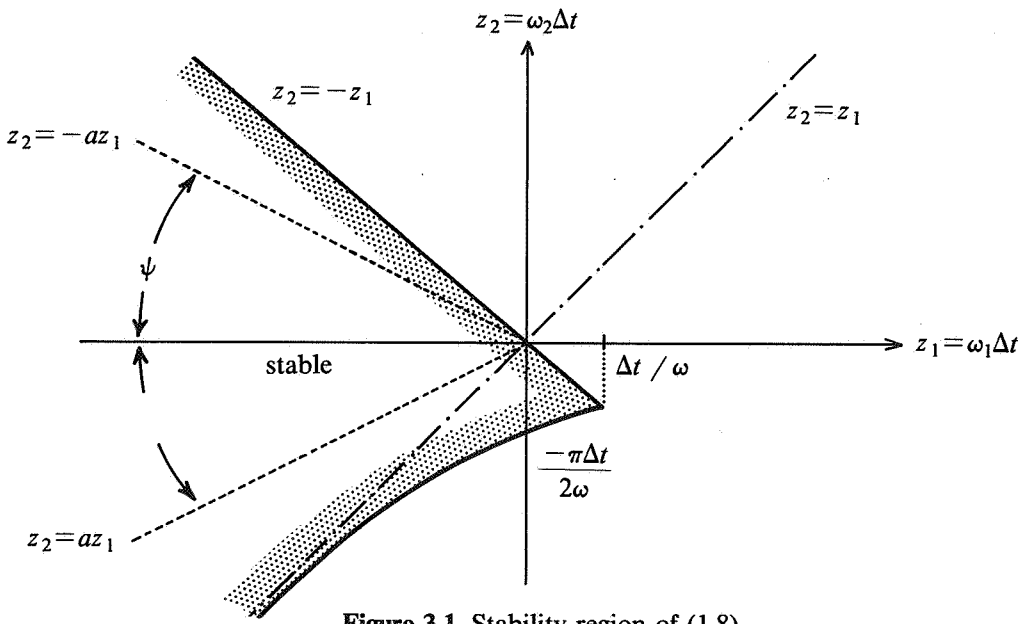

Figure 3.1. Stability region of (1.8).

parametrized by $q, 0 \leqslant q \leqslant \infty$.

In Figure 3.1 these curves are plotted in the $(\omega_1,\omega_2)\Delta t$-plane. To obtain the analytical stability region, which of course cannot have anything to do with Δt, the scaling factor Δt should be removed. However, this factor is included to facilitate comparisons with numerical stability regions, which are used to be plotted in the $(\omega_1\Delta t, \omega_2\Delta t)$-plane.

On the analogy to the definition of the "analytical" stability region (3.1) we define the numerical stability region as the set of points $(\omega_1,\omega_2)\Delta t = (z_1,z_2)$ for which the GPC method when applied to (1.8) yields solutions y_n such that $y_n \to 0$ as $n \to \infty$.

The GPC method is said to be *absolutely stable* for a given point (z_1,z_2) if this point lies in the stability region. (This terminology accords with the usage in e.g. Lambert [10, p. 66], and is referred to as 'strict' in the writings of Baker [2, p. 793] to distinguish it from the weaker definition sometimes encountered). However, for brevity in what follows, we use 'stable at (z_1,z_2)'. If a method is stable at all points in the real infinite wedge $\{(z_1,z_2)|z_1<0,|z_2/z_1|<a=\tan\psi\}$ then the method will be called $P_0(\psi)$-*stable* (see Figure 3.1). If the method is $P_0(\pi/4)$-stable we will briefly refer to P_0-stability [8]. In this connection it should be remarked that Barwell [3] called a numerical method *P-stable* if the numerical stability region contains all complex points (z_1,z_2) with $\mathrm{Re}\, z_1<-|z_2|$. The reader will note, on considering the case $z_2=0$, that $P_0(\psi)$-stable methods collapse in the case of equations with no delay to A_0-stable methods for pure differential equations. In consequence, $P_0(\psi)$-stable methods are necessarily implicit, whilst the GPC methods are explicit. It follows that the best we can expect is that the GPC methods have a region of stability which includes a truncated wedge $\{(z_1,z_2)\in\mathbb{R}^2|-\beta<z_1<0,|z_2/z_1|<\tan\psi\}$. Such methods, with β large, we will term *almost-$P_0(\psi)$ stable.*

We will be particularly interested in almost-$P_0(\psi)$ stable methods with ψ small, because the semi-discrete parabolic delay equations discussed in Section 1 will lead to (z_1,z_2)-points located in a wedge $|z_2/z_1|<\tan\psi$ with small aperture 2ψ. The relevant range $[-\beta,0)$ for z_1 is determined by the Jacobian matrices corresponding to the right-hand side function and the discretization step Δt.

3.1. Derivation of the stability polynomial

Recall that the interpolating polynomial occuring in (2.2) can be written

$$\hat{y}(t_n-\omega_n)=\hat{y}(t_j-\theta_n\Delta t)=E^{-l}\tau(E,\theta_n)y_j,\quad 0\leqslant\theta_n<1, \tag{3.2}$$

where $\tau(\zeta,\theta_n)$ is a polynomial of degree l in ζ with coefficients depending on θ_n. We assume (cf. Cryer [6]) that

$$\tau(\zeta,0)=\zeta^l,\quad \tau(1,\theta_n)=1. \tag{3.3}$$

Furthermore, we will always assume, in what follows, that $\omega_n\geqslant\Delta t$.

Applying the GPC method to the test equation (1.8) and writing

$$\omega=(n-j+\theta)\Delta t=:(\nu+\theta)\Delta t,\quad z_i=\omega_i\Delta t, \tag{3.4}$$

we obtain

$$y_n^{(j)}=\mu_j y_n^{(j-1)}+(1-\lambda_j-\mu_j)y_n^{(j-2)}+\lambda_j b_0\left[z_1 y_n^{(j-1)}+z_2E^{-l}\tau(E,\theta)y_{n-\nu}\right]+\lambda_j\Sigma_n. \tag{3.5}$$

Suppose that the initial approximation $y_n^{(0)}$ is computed by an explicit linear multistep method $\{\tilde{\rho},\tilde{\sigma}\}$, then by repeatedly applying (3.5) we can express $y_n^{(j)}$ in terms of the step vectors $y_{n-1},y_{n-2},y_{n-3},\ldots$. In particular $y_n:=y_n^{(m)}$ can be expressed in terms of preceding step vectors to obtain a linear recurrence relation with constant coefficients. The corresponding characteristic polynomial or *stability polynomial* can be derived in a similar way as given in [9] for the nondelay case. The result is summarized in the following theorem.

Theorem 3.1. *Let the GPC method (2.5) be generated by the $\tilde{k}$-step predictor $\{\tilde{\rho},\tilde{\sigma}\}$, the k-step corrector $\{\rho,\sigma\}$ and the l-step interpolation formula characterized by τ (cf. (3.2)). Then applying this method to the test equation (1.8) leads to the stability polynomial*

$$\begin{aligned}S_\nu(\zeta;z_1,z_2):=\;&\zeta^{\tilde{k}+l+\nu}S(\zeta;z_1)+\gamma_m(z_1)\zeta^{k+l+\nu}\tilde{S}(\zeta;z_1)\\&-z_2\tau(\zeta;\theta)\left[\zeta^{\tilde{k}}\sigma(\zeta)+\gamma_m(z_1)\zeta^k\tilde{\sigma}(\zeta)\right],\quad \nu\geqslant 1,\end{aligned} \tag{3.6}$$

where S and $\tilde{S}$ are the stability polynomials of the corrector and the predictor, respectively, and γ_m is defined by

$$\gamma_m(z_1):=(1-b_0z_1)\frac{P_m(z_1)}{1-P_m(z_1)}. \tag{3.7}$$

Proof. Similar to the derivation of stability polynomials for ODEs (cf. [9]), to which the result collapses on setting $z_2=0$. □

Evidently, the GPC method is *stable* at a point (z_1,z_2) for a given value of θ if $S_\nu(\zeta;z_1,z_2)$ is a Schur polynomial for all $\nu\in\mathbb{Z}_+$ (we will also use the terminology that $S_\nu(\zeta;z_1,z_2)$ is stable at (z_1,z_2)).

A convenient tool in the analysis of (3.6) in the familiar theorem of Rouché: *If $f(z)$ and $g(z)$ are regular on a closed region whose boundary is a closed rectifiable Jordan curve C and $|g(z)|<|f(z)|$ on C, then $f(z)$ and $f(z)+g(z)$ have the same number of zeros inside C.* Thus, *two polynomials $Q(\zeta)$ and $R(\zeta)$ have the same number of zeros within the unit circle if $|R(\zeta)-Q(\zeta)|<|Q(\zeta)|$ on the unit circle.*

Of course, this theorem provides sufficient but not necessary conditions for stability, so that the stability regions obtained may be smaller than the true stability regions (to consider a simple if artificial, case: let $Q(\zeta)=-R(\zeta)$, then $Q(\zeta)$ and $R(\zeta)$ have common roots but the inequality given above is not satisfied). However, as we shall see in Example 3.1, in an actual situation the true stability regions are only marginally larger than what we shall call the "Rouché-stability-regions."

3.2. Stability of the GPC method

Stability plots for the GPC method employing iteration polynomials of the form

$$P_m(z) = \delta T_m\left[c+\frac{c+1}{\beta}\,z\right],\quad \beta := \frac{c+1}{b_0[\cosh(\frac{1}{m}\operatorname{arccosh}\frac{1}{\delta})-c]},\quad c\leqslant 1 \tag{3.8}$$

have been given in [8]; here, T_m denotes the Chebyshev polynomial of the first kind and δ,c are suitably chosen parameters which determine the aperture and the length of stability wedge in the (z_1,z_2)-plane. The choice of the polynomials (3.8) is motivated by the large stability intervals $(-\beta,0)$ which such polynomials generate for GPC methods without delay (cf. [9]).

Since the case $|\omega_2|\ll|\omega_1|$ models an interesting class of problems associated with parabolic equations with delay, we are interested in methods with a long, narrow stability wedge along the negative z_1-axis ($z_1=\omega_1\Delta t$). Therefore, the polynomials (3.8) seem to be a good starting point for constructing efficient GPC methods. We observe that orthogonal polynomials automatically satisfy a three-terms recurrence so that, in this connection, the choice $P_j(z)=\delta T_j(c+z(c+1)/\beta)$ $(j=0,1,\ldots,m-1)$ also automatically satisfies a recurrence of the form (2.6) as we require. Identification of these recursions leads to explicit expressions for the λ_j and μ_j.

The largest stability region, for given δ, is obtained if $c=1$, and in this paper we only consider this case. It should be observed, however, that choosing $c=\cos(\pi/2m)$ yields an iteration polynomial which vanishes at $z=0$ giving rise to an extra damping of the predictor error if Δt is small (see the discussion in Section 2.1). In fact, the stability plots presented in [8] correspond to $c=\cos(\pi/2m)$. These plots were obtained by applying the boundary locus method to the stability polynomial S_ν defined in Theorem 3.1. A disadvantage of this direct approach is (i) we do not know a priori how to choose (δ,m) in order to get a stability wedge of prescribed aperture and length; (ii) we are never sure what is the effect of the delay parameter ν on the stability wedge.

In this section we propose a "computable" approach in obtaining values for (δ,m) which more or less guarantees a stability wedge of prescribed aperture and length for all values of ν.

3.2.1. "Rouché-stability-regions"

The first step is the formulation of a stability condition independent of the delay parameter ν.

Theorem 3.2. *The GPC method (2.4) is stable at the point (z_1,z_2) if it is stable at the point $(z_1,0)$ and if*

$$|z_2| < A_m(z_1,\theta) := \inf_{|\zeta|=1}\left|\frac{S(\zeta,z_1)+\gamma_m(z_1)\zeta^{k-\tilde{k}}\tilde{S}(\zeta,z_1)}{\tau(\zeta,\theta)[\sigma(\zeta)+\gamma_m(z_1)\zeta^{k-\tilde{k}}\tilde{\sigma}(\zeta)]}\right|.\ \square$$

Proof. Applying Rouché's theorem with

$$Q(\zeta) = S(\zeta,z_1)\zeta^{\tilde{k}} + \gamma_m(z_1)\tilde{S}(\zeta,z_1)\zeta^k, \quad R(\zeta) = S_\nu(\zeta;z_1,z_2)$$

the theorem follows immediately. □

In order to obtain a region of stable points (z_1,z_2) we shall determine the stability interval on the z_1-axis for the GPC method, that is the stability interval in the case of a vanishing delay. This special case was studied in [9]. For a GPC method generated by an extrapolation predictor and a backward differentiation corrector (EP-BD pair), δ-values for the polynomial $P_m(z)$ were derived such that the GPC method is stable in the interval $-\beta<z_1<0$ with β defined in (3.8) (we denote stable δ-values for the nondelay case by δ_0). For future reference these values are listed in Table 3.1.

Table 3.1. Stable δ_0-values for GPC methods without delay generated by $EP_{\tilde{k}}-BD_k$ pairs

k	$\tilde{k}=1$	$\tilde{k}=2$	$\tilde{k}=3$	$\tilde{k}=4$	$\tilde{k}=5$	$\tilde{k}=6$	$\tilde{k}=7$
2	1	1/3	1/7				
3	1	1/3	1/7	1/15			
4	.75	1/3	1/7	1/15	1/31		
5	.44	.33	1/7	1/15	1/31	1/63	
6	.13	.07	.07	1/15	.0289	.0147	1/127

Using larger δ-values gives rise to the development of instabilities but not in a severe way. In the numerical experiments reported in Section 4, it will be demonstrated that using larger values will still produce useful results. The reason is that the stability polynomial does not rapidly increase in magnitude if δ increases beyond the nondelay value δ_0. (In contrast, violating the stability condition $z_1 = \omega_1\Delta t > -\beta$, i.e. $\Delta t < \beta / |\omega_1|$, leads to a rapid increase of the magnitude of $P_m(z)$ for $z<-\beta$.) Therefore, the values of δ_0 listed in Table 3.1 should be used as an *indication of the acceptable upper bound for δ in using EP-BD methods;* in actual computation, one may often use much larger values.

In the following example we give the values of the 'effective' length $L(\delta,m)/m$ and the aperture angle $\psi(\delta,m)$ of the wedge contained in the stability region for a fourth-order EP-BD method (see Figure 3.2). We limit ourselves by, from all wedges contained in the stability region, the one with maximal aperture 2ψ. The factor $1/m$ is applied to $L(\delta,m)$ because $O(m)$ operations are employed within each step.

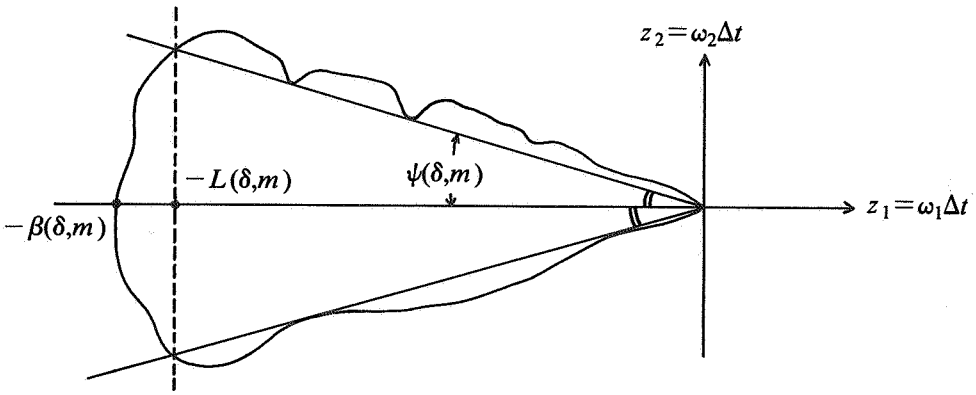

Figure 3.2. Stability wedge defined by $L(\delta,m)$ and $\psi(\delta,m)$.

Example 3.1. By virtue of Theorem 3.2 we can compute estimates of the stability wedge $\{L(\delta,m)/m,\psi(\delta,m)\}$ for any given predictor-corrector pair. For the pair EP_5-BD_4 some results are

listed in Table 3.2; the value of c occurring in (3.8) was set to 1 and $\tau(\zeta,\theta)$ was chosen of degree $l=4$. The values of L are slightly smaller than $\beta(\delta,m)$ given by (3.8). To get some idea about the pessimism due to the estimates obtained through use of Rouche's Theorem we calculated in addition the true stability regions for the parameters given in Table 3.2. Since these regions depend on the value of ν (cf. (3.6)), we made plots for several ν-values and determined the length and aperture of the stability wedge contained in all these stability regions. It turned out that these values (i) hardly depend on the value of ν and (ii) *are only slightly larger* than those listed in Table 3.2.

Table 3.2. EP_5-BD_4 method: $(L\,/\,m,\psi)$-values derived from Theorem 3.2 (ψ is given in degrees).

		$\theta=0$			$\theta=.5$	
δ	$m=2$	$m=8$	$m=64$	$m=2$	$m=8$	$m=64$
.01	(.34,45)	(2.3,42)	(19,35)	(.34,45)	(2.2,35)	(19,23)
$\delta_0=1/31$	(.69,45)	(3.8,16)	(31,.25)	(.69,45)	(3.8,10)	(31,.16)
.10	(1.4,40)	(7.3,.12)	(59,.09)	(1.4,29)	(7.3,.08)	(59,.06)

3.2.2. Choice of the predictor-corrector pair

It is of interest to observe that we are more or less forced to use *extrapolation predictors* if large β-values are desired. To see this we apply Rouche's theorem to the polynomial $S_\nu(\zeta;z_1,0)$ with $Q(\zeta)=S(\zeta,z_1)$ and $R(\zeta)=S_\nu(\zeta;z_1,0)$ to obtain the stability condition.

$$\left|\gamma_m(z_1)\right| < \inf_{|\zeta|=1}\left|\frac{S(\zeta,z_1)}{\tilde{S}(\zeta,z_1)}\right| = \inf_{|\zeta|=1}\left|\frac{\rho(\zeta)-z_1\sigma(\zeta)}{\tilde{\rho}(\zeta)-z_1\tilde{\sigma}(\zeta)}\right|.$$

Since $|\gamma_m(z_1)|$ is proportional to $|z_1|$ for $|z_1|$ large we should choose $\tilde{\sigma}\equiv 0$ in order to get a large stability interval $(-\beta,0)$. Predictor formulas with a vanishing polynomial $\tilde{\sigma}$ are just the extrapolation predictors characterized by $\tilde{\rho}(\zeta)=(\zeta-1)^k$.

In order to choose a suitable corrector $\{\rho,\sigma\}$ we consider the stability wedge (β,ψ) as $\delta\to 0$. (Notice that $\delta=0$ implies $P_m(z)\equiv 0$ (cf. (3.8)), that is the delay-corrector equation is iterated to convergence.) From Theorem 3.2 the following corollary follows:

Corollary 3.1. *Let the corrector formula $\{\rho,\sigma\}$ be $A(\alpha)$-stable then the delay-corrector formula (2.3) is $P_0(\psi)$-stable and*

$$\psi \geqslant \arctan\left[\frac{\sin\alpha}{\tau_1(\theta)}\right],\quad \tau_1(\theta) := \sup_{|\zeta|=1}|\tau(\zeta,\theta)|. \tag{3.9}$$

Proof. We have to show that the stability region of the GPC method with $\delta=0$ (i.e. $\gamma_m(z)\equiv 0$) contains the real infinite wedge $|z_2/z_1| < \tan\psi$ where ψ satisfies (3.9). From Theorem 3.2 with $\delta=0$ and by virtue of the $A(\alpha)$-stability of the generating $\{\rho,\sigma\}$ formula we derive the stability region

$$|z_2| < A_m(z_1,\theta) = \inf_{|\zeta|=1}\left|\frac{1}{\tau(\zeta,\theta)}\left[\frac{\rho}{\sigma}(\zeta)-z_1\right]\right|,\quad z_1 \leqslant 0.$$

Furthermore, it is easily verified that an $A(\alpha)$-stable LM formula satisfies

$$\inf_{|\zeta|=1}\left|\frac{\rho}{\sigma}(\zeta)-z\right| \geqslant |z|\sin\alpha$$

for all $z\leqslant 0$. Thus,

$$A_m(z_1,\theta) \geqslant \frac{1}{\tau_1(\theta)}\inf_{|\zeta|=1}\left|\frac{\rho}{\sigma}(\zeta)-z_1\right| \geqslant \frac{|z_1|\sin\alpha}{\tau_1(\theta)}.$$

The stability region of the delay-corrector formula therefore satisfies

$$|z_2| < \frac{|z_1|\sin\alpha}{\tau_1(\theta)}, \quad z_1 < 0$$

and certainly contains the infinite wedge $|z_2/z_1| < \tan\psi$ for all ψ satisfying (3.9). □

This Corollary shows the strong relation between $P_0(\psi)$-stability of the delay-corrector equation and the $A(\alpha)$-stability of the generating LM method. For parabolic equations with delay (excluding delay equations of type (1.3)), we only need stability wedges with relatively small apertures ψ; therefore we can limit our considerations to generating LM methods $\{\rho,\sigma\}$ that are $A(\alpha)$-stable where α is allowed to be small. Such LM methods are provided by the backward differentiation formulae for which α varies from 90^o for $k=2$ to 18^o for $k=6$ (cf. Lambert [10]).

A second conclusion from Corollary 3.1 is the strong dependency of ψ on $\tau(\zeta,\theta)$. The following example presents values of the lower bound on ψ for various values of θ.

Example 3.2. Consider the delay-corrector formulae generated by the backward differentiation formulas BD_k and by interpolation polynomials τ of degree $l=k$. Then the lower bounds in (3.9) are given in Table 3.3 for a few values of θ. From this Table we see that it is advantageous to select the step size in

Table 3.3. BD_k methods: ψ-lower bounds derived from Corollary 3.1 (ψ is given in degrees).

θ	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$
0	45.0	45.0	43.7	37.9	17.2
.25	39.2	39.2	37.9	32.3	14.1
.5	31.6	31.6	30.5	25.6	10.8
.75	33.0	33.0	31.9	26.8	11.4
1	45.0	45.0	43.7	37.9	17.2

such a way that $t_n-\omega_n$ coincides with a step point (i.e. $\theta=0$); however, if interpolation is performed, at least 2/3 of the maximal ψ-value is obtained.

It would be nice to have the analogue of Corollary 3.1 for the GPC method itself. It is not difficult to find such an analogue for large $|z_1|$ and small values of δ; however, if δ increases the resulting lower bounds become rather pessimistic. Therefore, we also considered upper bounds for ψ. The following corollary of Theorem 3.2 presents these results.

Corollary 3.2. *Let δ_0 be the maximal stable δ-value of the GPC method for nondelay equations, let $\beta=\beta(\delta,m)$ be defined by (3.8) and let $\{\rho,\sigma\}$ be $A(\alpha)$-stable. Then for $-\beta(\delta_0,m) \leqslant z_1 << -1$ the stability region of the GPC method employing an extrapolation predictor $EP_{\tilde{k}}$ is bounded by $|z_2/z_1| < \tan\psi$ where $\psi=\psi(\delta,m)$ satisfies the inequality*

$$\frac{1}{\tau_1(\theta)}\left[\sin\alpha - 2^{\tilde{k}}\frac{b_0\delta}{1-\delta}\operatorname*{Sup}_{|\zeta|=1}\frac{1}{|\sigma(\zeta)|}\right] \leqslant \tan\psi \leqslant \frac{1}{|\tau(-1,\theta)|}\left[1-2^{\tilde{k}}\frac{b_0\delta}{1-\delta}\frac{1}{|\sigma(-1)|}\right] \tag{3.10}$$

with $\tau_1(\theta)$ defined in (3.9), provided that δ is sufficiently small and that the right and left-hand side in (3.10) are positive). □

Proof. For extrapolation predictors we have $\tilde{\rho}(\zeta)=(\zeta-1)^{\tilde{k}}$ and $\tilde{\sigma}(\zeta)=0$ so that the function $A_m(z_1,\theta)$ can be written as

$$A_m(z_1,\theta) = \operatorname*{Inf}_{|\zeta|=1}\left|\frac{1}{\tau(\zeta,\theta)}\left[\frac{\rho}{\sigma}(\zeta)-z_1+\gamma_m(z_1)\zeta^{k-\tilde{k}}\frac{(\zeta-1)^{\tilde{k}}}{\sigma(\zeta)}\right]\right|. \tag{3.11}$$

First we derive the lower bound for ψ. Similar to the derivation of the lower bound (3.9) we obtain

$$A_m(z_1,\theta) \geqslant \frac{1}{\tau_1(\theta)}\operatorname*{Inf}_{|\zeta|=1}\left[\left|\frac{\rho}{\sigma}(\zeta)-z_1\right| - \left|\gamma_m(z_1)\frac{(\zeta-1)^{\tilde{k}}}{\sigma(\zeta)}\right|\right]$$

$$\geqslant \frac{1}{\tau_1(\theta)}\left[|z_1|\sin\alpha - 2^{\tilde{k}}|\gamma_m(z_1)|\operatorname{Sup}_{|\zeta|=1}\frac{1}{|\sigma(\zeta)|}\right].$$

The stability region therefore satisfies

$$|z_2| < \frac{|z_1|}{\tau_1(\theta)}\left[\sin\alpha - 2^{\tilde{k}}\left|\frac{\gamma_m(z_1)}{z_1}\right|\operatorname{Sup}_{|\zeta|=1}\frac{1}{|\sigma(\zeta)|}\right]$$

provided that the right-hand side is positive, that is $|\gamma_m / z_1|$ sufficiently small; this is achieved by choosing δ sufficiently small as can be seen from the result

$$\left|\frac{\gamma_m(z_1)}{z_1}\right| = \left|\frac{1}{z_1} - b_0\right|\left|\frac{P_m(z_1)}{1-P_m(z_1)}\right| \leqslant \left|\frac{1}{z_1} - b_0\right|\frac{\delta}{1-\delta}.$$

Using this upperbound on $|\gamma_m / z_1|$ and assuming $|z_1|$ large we arrive at the left-hand inequality in (3.10).

Next we derive the upper bound on ψ. From (3.11) it follows that for $|z_1| \ll 1$

$$A_m(z_1,\theta) \leqslant \frac{1}{\tau(-1,\theta)}\left|\frac{\rho}{\sigma}(-1) - z_1 + \gamma_m(z_1)\frac{(-1)^k 2^{\tilde{k}}}{\sigma(-1)}\right|$$

$$\approx \frac{|z_1|}{|\tau(-1,\theta)|}\left|1 - \frac{\gamma_m(z_1)}{z_1}\,\frac{(-1)^k 2^{\tilde{k}}}{\sigma(-1)}\right|.$$

Thus, for $|\gamma_m / z_1|$ sufficiently small the stability region satisfies

$$|z_2| \leqslant \frac{|z_1|}{|\tau(-1,\theta)|}\left[1 - \left|\frac{\gamma_m(z_1)}{z_1}\right|\frac{2^{\tilde{k}}}{|\sigma(-1)|}\right]$$

$$\leqslant \frac{|z_1|}{|\tau(-1,\theta)|}\left[1 - \frac{b_0\delta}{1-\delta}\,\frac{2^{\tilde{k}}}{|\sigma(-1)|}\right]$$

which leads to the right-hand inequality in (3.10). □

Example 3.3. For $EP_{\tilde{k}} - BD_k$ methods this Corollary yields

$$\frac{1}{\tau_1(\theta)}\left(\sin\alpha - 2^{\tilde{k}}\frac{\delta}{1-\delta}\right) \leqslant \tan\psi \leqslant \frac{1}{|\tau(-1,\theta)|}\left(1 - 2^{\tilde{k}}\frac{\delta}{1-\delta}\right), \tag{3.12}$$

where it is assumed that δ is sufficiently small to provide positive right- and left-hand sides. Evidently, δ should satisfy the inequality

$$\delta \leqslant \sin\alpha / (2^{\tilde{k}} - \sin\alpha).$$

For the predictor-corrector pair mentioned in Example 3.1, we find for $\delta = .01$, and respectively $\theta = 0$ and $\theta = .5$ the ψ-ranges $32.3^o \leqslant \psi \leqslant 34.1^o$ and $21.3^o \leqslant \psi \leqslant 22.6^o$. For large values of m, these ψ-bounds are in good agreement with the ψ-values given in Table 3.2. (We emphasize that (3.12) has been derived under the assumption that $z_1 \ll -1$.) □

4. Numerical illustrations

A most important aspect of the numerical integration of parabolic equations with delay is the storage requirements. As any algorithm needs an (interpolated) approximation of the delay term, we have to store at least ν arrays of y-vectors (notice that $\nu \approx \omega_n / \Delta t$ may change from step to step). Moreover, the dimension of these y-vectors is usually very large and their storage requires a tremendous computer memory capacity. Therefore, in order to reduce the value of ν, it is of vital importance to be able to

integrate with large time steps. However, large time steps usually have a negative influence both on accuracy and on stability. Consequently, we think that these high-order stabilized predictor-corrector methods are a useful tool for integrating parabolic equations with delay.

To illustrate the performance of these methods we give in this section results obtained by $EP_p - BD_p$ methods. The parameters μ_j and λ_j in the iteration scheme (2.5) are chosen in such a way that the iteration polynomial (3.8) is generated. For a discussion on the implementational details we refer to [9]. In these polynomials the parameter δ, which determines the aperture of the stability wedge, is still kept free and will be varied in the experiments (see also Table 3.1). The choice $c=1$ maximizes the length of the stability wedge. Therefore, we combined a p-step BD corrector with a $(p+1)$-step EP predictor (which are both of order p), resulting in a p-th order $EP-BD$ method (see Theorem 2.1 and the discussion thereafter).

For the sake of comparison we also implemented an adjusted form of the ADI method of Peaceman & Rachford which is well-known for parabolic equations without delay. The ADI method requires a splitting of the function $f(t,y(t),y(t-\omega))$ in (2.1). We assume that f can be written as $f(t,y(t),y(t-\omega)) = f_1(t,y(t),y(t-\omega)) + f_2(t,y(t),y(t-\omega))$, where the functions f_1 and f_2 correspond to the one-dimensional differential operators in x_1- and x_2-direction, respectively. Now, the (nonlinear) ADI method is defined by

$$y^* = y_{n-1} + \tfrac{1}{2}\Delta t f_1(t_{n-1}+\tfrac{1}{2}\Delta t, y^*, \hat{y}_{n-\frac{1}{2}}) + \tfrac{1}{2}\Delta t f_2(t_{n-1}, y_{n-1}, \hat{y}(t_{n-1}-\omega_{n-1})), \tag{4.1}$$

$$y_n = 2y^* - y_{n-1} + \tfrac{1}{2}\Delta t f_2(t_n, y_n, \hat{y}(t_n-\omega_n)) - \tfrac{1}{2}\Delta t f_2(t_{n-1}, y_{n-1}, \hat{y}(t_{n-1}-\omega_{n-1})),$$

where

$$\hat{y}_{n-\frac{1}{2}} = \hat{y}(t_{n-1}+\tfrac{1}{2}\Delta t - \omega(t_{n-1}+\tfrac{1}{2}\Delta t)).$$

The inhomogeneous term, if any, is equally distributed over f_1 and f_2. To compare the accuracy of the various methods we define

$$cd := -\log_{10}(\|\textit{absolute error at the endpoint}\|_\infty), \tag{4.2}$$

denoting the number of correct decimals in the answer.

Our test examples all have an initial ϕ-function (cf. (1.2)) which coincides with the solution, hence no discontinuities in higher derivatives of the solution will arise. Moreover, the solutions are chosen in such a way that the space discretization does not introduce an error, i.e. the solution of the system of ODEs with delay equals the solution of the PDE, restricted to the grid points. This spatial discretization is achieved using standard 5-point molecules on a uniform mesh with mesh size $h = 1/20$.

4.1. A mildly nonlinear example

As a first example consider the parabolic equation, defined on the unit square in the (x_1,x_2)-plane

$$\begin{cases} \dfrac{\partial}{\partial t}u(t,x_1,x_2) = \tfrac{1}{3}\dfrac{(1+x_1+x_2)}{1+t}\left[\dfrac{\partial^2}{\partial x_1^2} + \dfrac{\partial^2}{\partial x_2^2}\right]u^3(t,x_1,x_2) \\ \qquad\qquad -4\dfrac{u^3(t-1,x_1,x_2)}{1+t} + \dfrac{2\pi}{3}(1+x_1+x_2)\cos(2\pi t), \quad 0\leqslant t\leqslant 1, \\ \phi(t,x_1,x_2) = \tfrac{1}{3}(1+x_1+x_2)\sin(2\pi t), \quad t\leqslant 0. \end{cases} \tag{4.3}$$

The solution $u(t,x_1,x_2)$ equals the function ϕ for all t. The Dirichlet boundary conditions are taken from the solution u.

In order to get a stability wedge of sufficient length, that is to have an m-value which is sufficiently large, we must have an estimate of the spectral radius S of the Jacobian matrix $\partial f / \partial y$. We used

$$S_{n-1} = S(\partial f / \partial y))|_{t=t_{n-1}} = 1.1\frac{72}{h^2} \max_{t\in[t_{n-1},t_n]} \frac{\sin^2(2\pi t)}{1+t},$$

where the factor 1.1 is added to obtain a safe upperbound.

In Tables 4.1-4.3 we give the results of the second, fourth and sixth order EP-BD method, respectively for several values of the time step Δt. In these Tables cd is defined in (4.2) and N denotes the total number of iterations, summed over all time steps. Note that the number of iterations per time step is not constant because the spectral radius S_n varies in time. An $\star$ denotes an unstable behaviour. A mutual comparison of the EP-BD methods reveals that the higher order formulae are the more efficient ones.

Table 4.1. (cd / N)-values for the second order $EP-BD$ method; the total number of arrays equals $1/\Delta t + 4$.

δ	$\Delta t = 1/10$	$\Delta t = 1/20$	$\Delta t = 1/40$
.1	1.3/456	1.8/597	2.5/800
$\delta_0 = 1/7$	1.3/410	1.8/528	2.5/706
.2	1.1/357	1.8/463	2.5/616
.4	$\star$	1.3/319	2.2/428

Table 4.2. (cd / N)-values for the fourth order $EP-BD$ method; the total number of arrays equals $1/\Delta t + 4$.

δ	$\Delta t = 1/10$	$\Delta t = 1/20$	$\Delta t = 1/40$
.01	2.1/697	3.2/895	4.3/1192
$\delta_0 = 1/31$	1.9/543	3.2/698	4.3/936
.1	1.6/395	3.1/511	4.3/686
.2	.7/304	1.6/392	$\star$

Table 4.3. (cd / N)-values for the sixth order $EP-BD$ method; the total number of arrays equals $1/\Delta t + 4$.

δ	$\Delta t = 1/10$	$\Delta t = 1/20$	$\Delta t = 1/40$
.005	2.4/725	4.6/933	6.1/1244
$\delta_0 = 1/127$	2.3/671	4.6/863	6.1/1150
.02	2.2/559	4.5/719	6.1/963
.04	1.8/476	3.8/613	5.7/819

Furthermore, concerning the value of δ we see that a larger value is allowed than indicated by Table 3.1, but the methods gradually lose accuracy as δ increases. This is due to a mild form of instability.

The results obtained by the second order ADI method are listed in Table 4.4. We applied the method

Table 4.4. *cd*-values for the ADI method;
the total number of arrays equals $1/\Delta t + 10$ (including both tridiagonal Jacobian matrices).
The total number of iterations to solve the implicit relations equals $2*NEWT/\Delta t$.

NEWT	$\Delta t = 1/40$	$\Delta t = 1/60$	$\Delta t = 1/80$	$\Delta t = 1/120$
1	*	*	1.8	2.2
2	*	2.2	2.7	3.3
5	*	3.7	4.2	4.5

for several values of NEWT, being the number of Newton iterations performed to "solve" each implicit relation in (4.1). Note that for the EP - BD methods an iteration is simply an f-evaluation; for the ADI method, however, an iteration is of quite a different nature and usually much more expensive (i.e. one evaluation of f and the solution of a tridiagonal system of equations). Moreover, the Jacobian matrices have to be evaluated (in this experiment we updated the Jacobians every step). Hence, a comparison of both methods in terms of efficiency is not feasible.

However, as the ADI method needs a relatively small time step for stability reasons, its storage requirements tend to become excessive, whereas the EP - BD methods can take rather large time steps, thus reducing the number of y-vectors to be kept in store.

4.2. A strongly nonlinear example

To construct our second test problem we employ the porous medium operator

$$\Delta(u(t,x_1,x_2))^m, \quad m \geqslant 2,$$

and specify the analytic solution as

$$u(t,x_1,x_2) = \tfrac{1}{4}(x_1+x_2)^{2/m}[e^{-2(t-1)^2}+e^{-2(t-3)^2}]. \tag{4.4}$$

The initial function $\phi(t,x_1,x_2)$ and the Dirichlet boundary conditions are prescribed by (4.4). By setting $m=5$ and introducing a delay term and an inhomogeneous term $g(t,x_1,x_2)$ we arrive at

$$\begin{cases} \dfrac{\partial}{\partial t}u(t,x_1,x_2) = \left[\dfrac{\partial^2}{\partial x_1^2}+\dfrac{\partial^2}{\partial x_2^2}\right]u^5(t,x_1,x_2) + 4u(t-2,x_1,x_2) \\ \qquad\qquad + 4(1-t)u(t,x_1,x_2) + g(t,x_1,x_2), \quad 0 \leqslant t \leqslant 4, \\ \phi(t,x_1,x_2) = \tfrac{1}{4}(x_1+x_2)^{2/5}[e^{-2(t-1)^2}+e^{-2(t-3)^2}], \quad t \leqslant 0, \end{cases} \tag{4.5}$$

defined on the unit square in the (x_1,x_2)-plane.

For this problem, a safe upperbound for the spectral radius S of $\partial f/\partial y$ is obtained by

$$S(\partial f/\partial y)|_{t=t_{n-1}} \approx 1.1\frac{120}{h^2}\frac{1}{4^4}\max_{t\in[t_{n-1},t_n]}\left[e^{-2(t-1)^2}+e^{-2(t-3)^2}\right]^4.$$

Similar to the previous example we tested the second, fourth and sixth order EP - BD method as well as the ADI method. The results are given in the Tables 4.5 - 4.8.

Table 4.5. $(cd\,/\,N)$-values for the second order $EP-BD$ method; the total number of arrays equals $2\,/\,\Delta t+4$.

δ	$\Delta t=1\,/\,2$	$\Delta t=1\,/\,4$	$\Delta t=1\,/\,8$	$\Delta t=1\,/\,16$
.1	1.6/72	2.2/88	3.4/118	3.5/178
$\delta_0=1\,/\,7$	1.6/64	2.2/82	2.9/112	3.6/156
.2	*	1.3/70	2.8/92	3.8/144

Table 4.6. $(cd\,/\,N)$-values for the fourth order $EP-BD$ method; the total number of arrays equals $2\,/\,\Delta t+4$.

δ	$\Delta t=1\,/\,2$	$\Delta t=1\,/\,4$	$\Delta t=1\,/\,8$	$\Delta t=1\,/\,16$
.01	*	1.3/134	3.7/176	5.1/258
$\delta_0=1\,/\,31$	*	1.6/106	4.0/138	4.9/210
.05	*	1.9/96	3.9/128	4.9/188
.1	*	2.4/76	3.8/114	4.9/156
.2	*	*	3.3/88	4.7/126

Table 4.7. $(cd\,/\,N)$-values for the sixth order $EP-BD$ method; the total number of arrays equals $2\,/\,\Delta t+4$.

δ	$\Delta t=1\,/\,2$	$\Delta t=1\,/\,4$	$\Delta t=1\,/\,8$	$\Delta t=1\,/\,16$
.005	*	1.8/136	3.7/182	6.0/272
$\delta_0=1\,/\,127$	*	1.4/128	3.9/176	5.8/252
.02	*	1.4/108	3.6/148	5.6/214
.04	*	1.8/96	3.7/128	5.6/186
.08	*	1.9/76	3.1/110	*

Table 4.8. cd-values for the ADI method; the total number of arrays equals $2\,/\,\Delta t+10$; the total number of Newton iterations equals $8*NEWT\,/\,\Delta t$.

NEWT	$\Delta t=1\,/\,2$	$\Delta t=1\,/\,4$	$\Delta t=1\,/\,8$	$\Delta t=1\,/\,16$	$\Delta t=1\,/\,32$
1	*	1.5	1.9	2.2	2.4
2	*	*	2.1	2.9	3.6
5	*	*	3.7	4.9	5.4

Again, an * denotes an unstable behaviour of the integration process and the quantities cd, N and NEWT have the same meaning as defined in Section 4.1. The results of this example give rise to conclusions similar to those of the previous example: to obtain a stable result, the ADI method needs a smaller time step Δt than the GPC method does.

5. Conclusion

We have indicated how a class of methods for certain parabolic equations with delay can be derived by extending the GPC methods for semi-discretized parabolic equations. The resulting methods have the following properties:

(i) The GPC method consists of an (explicit) linear multistep predictor, an (implicit) linear multistep

corrector and an (unconventional) iteration scheme.

(ii) In order to relax the stability conditions for this method the predictor should be based on extrapolation of preceding y_n-values and the corrector should be $A(\alpha)$-stable where α is allowed to be relatively small.

(iii) The integration step may be freely choosen because the number of iterations is automatically adapted to ensure stability; therefore, the integration step is determined only by accuracy considerations and not limited by stability.

(iv) By choosing a high order corrector the method can take large integration steps thereby limiting the number of backvalues which should be stored to compute the delay term; the reduction in storage is considerable when compared with conventional methods.

(v) For two test examples we compared the GPC methods with the ADI method. The storage reduction factors are roughly 5 and 2 for these problems (in favour of the GPC methods). Moreover, in terms of CP seconds (measured on a CDC 750 computer) the GPC methods are at least competitive.

(vi) Finally, because of its explicit character, the GPC method can also be applied to non-5-point-space-discretizations which allows us to integrate problems with mixed derivatives, or to employ high order space-molecules; in the latter case, the magnitude of the spatial meshes can be increased resulting in a smaller spectral radius $S(\partial f / \partial y)$ and as a consequence a smaller number of stages per step.

References

[1] Artola, M., *Equations paraboliques à retardement,* C.R. Acad. Sci. Paris 264 (1967), 668-671.

[2] Baker, C.T.H., *The numerical treatment of integral equations,* Clarendon Press, Oxford (1977).

[3] Barwell, V.K., *Special stability problems for functional differential equations,* BIT 15 (1975), 130-135.

[4] Bellman, R. & K.L. Cooke, *Differential-Difference Equations,* Academic Press, New York, (1963).

[5] Chosky, N.H., *Time-lag controls: a bibliography,* IRE Transactions on Automatic Control $AC-5$ (1966), 66-70.

[6] Cryer, C.W., *Highly stable multistep methods for retarded differential equations,* SIAM J. Numer. Anal. 11 (1974), 788-797.

[7] El'sgol'ts, L.E. & S.B. Norkin, *Introduction to the theory and application of differential equations with deviating arguments.* [Translation by J.L. Casti], Academic Press, New York (1973).

[8] Houwen, P.J. van der & B.P. Sommeijer, *Improved absolute stability of predictor-corrector methods for retarded differential equations.* In: *Differential-difference equations* ISNM 62, Birkhauser-Verlag (1983), 137-148.

[9] -----&-----, *Predictor-corrector methods with improved absolute stability regions,* IMA J. Numer. Anal. 3 (1983), 417-437.

[10] Lambert, J.D., *Computational methods in ordinary differential equations,* John Wiley, London (1973).

[11] Stetter, H.J., *Improved absolute stability of predictor-corrector schemes,* Computing 3 (1968), 286-296.

[12] Tavernini, L., *Linear multistep methods for the numerical solution of Volterra functional differential equations,* J. Applicable Anal. 3 (1973), 169-185.

[13] Travis, C.C. & G.F. Webb, *Existence and stability for partial functional differential equations,* Transactions Amer. Math. Soc. 200 (1974), 395-418.

[14] Wang, P.K.C., *Asymptotic stability of a time-delayed diffusion system,* Transactions Amer. Math. Soc. (Ser. E) 30 (1963), 500-504.

[15] Wang, P.K.C., *Optimal control of parabolic systems with boundary conditions involving time delays.* SIAM J. Control 13 (1975), 274-293.

[16] Weiss, R., *Transportation lag— an annotated bibliography,* IRE Transactions on Automatic Control, $AC-4$ (1959), 56-64.

[17] Wiederholt, L.F., *Stability of multistep methods for delay differential equations,* Math. of Comp. 30 (1976), 283-290.