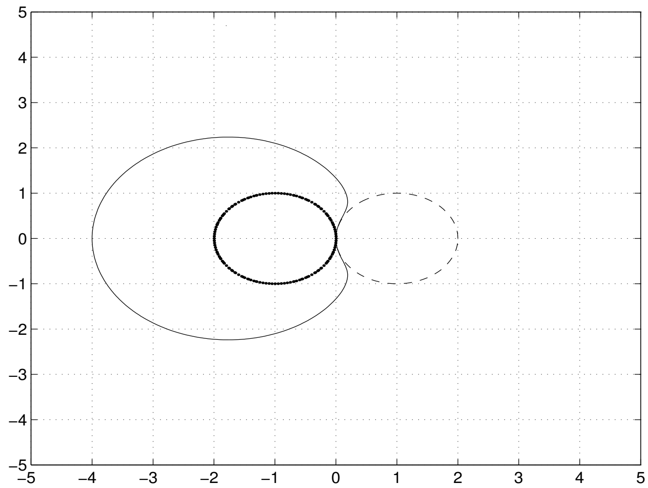


Universiteit Utrecht

Department of Mathematics

Low-dimensional Krylov subspace iterations for enhancing stability of time-step integration schemes

by

Mikhail A. Botchev, Gerard L.G. Sleijpen, and Henk A. van der Vorst

Preprint

nr. 1004

March, 1997

Low-dimensional Krylov subspace iterations for enhancing stability of time-step integration schemes*

Mikhail A. Botchev†‡ Gerard L. G. Sleijpen†§ Henk A. van der Vorst†§

March 3, 1997

Abstract

In a conventional integration scheme of the Predictor–Corrector (PC) type, solution on the next time layer is obtained by solving the Corrector scheme equation with few (usually one) iterative steps of Richardson's method where the initial guess is taken from the Predictor scheme.

Aiming to enhance stability of such a scheme by performing a few optimal Krylov subspace iterations (e.g., k steps of GMRES, $k \leqslant 5$) instead of Richardson's method steps, we get a family of Minimal Residual PC (MR-PC) time step integration schemes. The optimality (residual reduction) property of iterative schemes like GMRES leads to a scheme which is closest in the residual sense to the implicit Corrector scheme.

Two particular MR-PC schemes are investigated here: Forward Euler Predictor – Backward Euler Corrector (of the first order) and Adams(2) Predictor – BDF2 Corrector (of the second order). Practical aspects of using MR-PC scheme including adaptive step size control strategy will be discussed.

1 Introduction

For the numerical integration of a system of stiff ODEs

$$\frac{d\boldsymbol{y}}{dt} = f(t, \boldsymbol{y}), \qquad \boldsymbol{y}\big|_{t=0} = \boldsymbol{y}^0 \in \mathbb{R}^n, \tag{1}$$

one usually considers implicit schemes, for instance, BDF, and Newton's method can be used to solve the nonlinear system of equations in $\boldsymbol{y}^{m+1}$ (vector of the solution on the next time layer). When the system is large and has a sparse Jacobian, an iterative method may be attractive as a linear solver in each Newton's iteration [11, 4, 10]. This can be naturally combined in the framework of inexact Newton method [2, 3, 7, 5].

Robust iterative Krylov subspace methods like GMRES [19], BiCGSTAB [23, 21] or QMR [9] are good candidates here. In the case of using GMRES (perhaps, the most typical situation), the ODE context gives many opportunities to tune the performance of the iteration process:

*Preprint nr. 1004, Dept. of Mathematics, Utrecht University; This is corrected version (March 21, 1997)

†Mathematical Institute, Utrecht University, P.O.Box 80.010, 3508 TA Utrecht, the Netherlands. E-mail: botchev@math.ruu.nl, sleijpen@math.ruu.nl, vorst@math.ruu.nl

‡Work of this author was supported by the Netherlands organization for scientific research NWO, project 95MPR04

§Work of these authors was supported in part by the Netherlands organization for scientific research NWO, project 95MPR04

truncated and (or) nested (flexible) Krylov subspace iterative processes [11, 24, 18], use the same Krylov basis several times [11], reusing Krylov basis with sophisticated injecting technique [20], and application of Krylov-like projections directly on the nonlinear level [22, 6, 8], are only some possibilities. These techniques are mainly aimed to capture the information, delivered by Krylov methods during previous iterations, with a modest Krylov subspace dimension.

Another possibility offered by Krylov iterative methods is building up an integration time-stepping scheme for the simulation of the action of the exponential operator $e^{\tau A}$ (A is the current Jacobian of f). For example, in a first order method investigated in [25], a fixed number of Chebyshev iterations, tuned to approximate $e^{\tau A}\boldsymbol{y}^m$, is done at each time step. This method is more accurate and less computation-consuming than the standard Backward Euler scheme with regular Chebyshev iterations for solving $(I-\tau A)\boldsymbol{y}^{m+1}=\boldsymbol{y}^m$ (assuming f to be linear) [25]. In [26] some modifications (also of second order) for this method are described and, for a model problem, an explanation of the observed faster convergence of Chebyshev iteration for $e^{\tau A}\boldsymbol{y}^m$, rather than for $(I-\tau A)^{-1}\boldsymbol{y}^m$, is given. The results from [10] show that for Krylov methods faster (superlinear) convergence can be expected for Krylov dimension $k \gg \|\tau A\|$. This, however, is too crude a bound for the explanation of the actually observed performance. Theory developed in [15] sheds more light on these phenomena.

The aim of this paper is to investigate the usage of low-dimensional Krylov subspace methods for the time-step integration of stiff ODEs. Roughly speaking, we take an implicit scheme and we try to solve the associated (linearized) scheme (i.e., the system of equations for $\boldsymbol{y}^{m+1}$) inexactly, by GMRES each time step. This leads to a class of schemes which can be regarded as generalized Predictor–Corrector (GPC) schemes, in contrast to the standard Predictor–Corrector (PC) schemes that are obtained by using some Richardson's iterations per time step [12, 13].

The motivation for this research is the solution of very large stiff ODE systems obtained from spatial discretization of 3D magneto-fluid dynamics simulations. These problems are currently solved by explicit schemes, due to computer memory storage limitations. The question arises whether the stability restriction can be relaxed by a Krylov subspace method. For such kind of problems one can not afford high Krylov subspace dimensions and, for the allowed modest dimensions of Krylov subspace, the superlinear convergence for $e^{\tau A}\boldsymbol{y}^m$ can hardly be expected [10, 15]. That is why the attractive exponential technique of [15] has not been considered for our purpose.

Although our framework may be adapted for any Predictor–Corrector scheme (for any implicit scheme where the system in $\boldsymbol{y}^{m+1}$ is solved inexactly), in this paper we restrict our analysis to the linear autonomous case $f(t,\boldsymbol{y}) = A\boldsymbol{y}$. In numerical tests, however, the GPC approach turned out to be successful for nonlinear problems as well. With the usual assumption that A is diagonalizable, we will analyze stability for two GPC schemes: Forward Euler Predictor – Backward Euler Corrector and Adams (2) Predictor – BDF2 Corrector. In both cases our results are based on the standard A-stability theory (see, e.g. [12, 13]) and some typical features of the specific iterative solver, for example, the fast convergence with respect to the eigenvalues which are largest in modulus.

Aside from this feature of the linear solver, one has still some freedom in the choice of one. In this paper we suggest to take k steps of GMRES. Then, the minimal residual property of GMRES [19] gives an inexact implicit scheme (which, for small k, may be regarded as an explicit one) which is closest in the residual sense to the exact implicit scheme. This fact is also exploited in our analysis.

2 Generalized and Minimal Residual PC schemes

For the solution of the linear system of ODEs

$$\frac{d\boldsymbol{y}}{dt} = A\boldsymbol{y} + \boldsymbol{g}(t)\,, \qquad \boldsymbol{y}\big|_{t=0} = \boldsymbol{y}^0 \in \mathbb{R}^n\,, \tag{2}$$

for which the real matrix A has eigenvalues with negative real part, we consider the following time-stepping process. Assume that we have an appropriate implicit scheme for the solution $\boldsymbol{y}_{\rm im}^{m+1}$ on the next time layer t_{m+1}, then one has to solve a linear system

$$\tilde{A}\boldsymbol{y}_{\rm im}^{m+1} = \boldsymbol{b}^m, \qquad \tilde{A} = I - \hat{\beta}\tau A\,, \tag{3}$$

where $\boldsymbol{b}^m$ depends on $\boldsymbol{g}(t)$, some known $\boldsymbol{y}^m$, $\boldsymbol{y}^{m-1}$, ... and possibly also on A, and $\hat{\beta}$ is a scheme-dependent parameter. We want to get an inexpensive approximation for the solution (3), namely, by a few iterations $\boldsymbol{y}_0, \boldsymbol{y}_1, \ldots$ of a Krylov subspace method. In that case, we have to choose the initial guess $\boldsymbol{y}_0$ properly. For consistency, one should take

$$\boldsymbol{y}_0 = \boldsymbol{y}_{\rm ex}^{m+1}\,, \tag{4}$$

where $\boldsymbol{y}_{\rm ex}^{m+1}$ is a numerical solution (at the moment $t = t_{m+1}$) of (2), obtained by some explicit scheme which we will call the Predictor. With k iterations for (3), we get an approximate solution on the next time layer

$$\boldsymbol{y}^{m+1} = \boldsymbol{y}_{\rm ex}^{m+1} + \hat{P}_{k-1}(\tilde{A})\boldsymbol{r}_0 \qquad (\approx \boldsymbol{y}_{\rm im}^{m+1})\,, \tag{5}$$

where $\boldsymbol{r}_0 = \boldsymbol{b}^m - \tilde{A}\boldsymbol{y}_0$ is the initial residual and $\hat{P}_{k-1}(\tilde{A}) \approx \tilde{A}^{-1}$ is a polynomial of order $\leqslant k-1$, implicitly generated by the iteration scheme. The above equality can be interpreted as a generalized Corrector step for the implicit scheme (3).

The adoption of the Predictor – Corrector terminology is justified by the observation that with one Richardson's iteration for (3), we obtain the standard Predictor – Corrector (PC) scheme. Indeed, for a conventional PC scheme we have that [12, 13]

$$\boldsymbol{y}^{m+1} - \hat{\beta}\tau A\boldsymbol{y}_{\rm ex}^{m+1} = \boldsymbol{b}^m\,, \tag{6}$$

and this corresponds to the Richardson's splitting of matrix in (3):

$$\tilde{A} = I - (I - \tilde{A}) = I - \hat{\beta}\tau A\,. \tag{7}$$

With the initial guess (4) we obtain

$$\boldsymbol{y}_1 \equiv \boldsymbol{y}^{m+1} = (I - \tilde{A})\boldsymbol{y}_0 + \boldsymbol{b}^m = \boldsymbol{y}_{\rm ex}^{m+1} + \boldsymbol{b}^m - \tilde{A}\boldsymbol{y}_{\rm ex}^{m+1} = \boldsymbol{y}_{\rm ex}^{m+1} + \boldsymbol{r}_0\,.$$

The last coincides with (5) for $\hat{P}_0 \equiv 1$, so that the PC scheme can be viewed as a special case of (5).

If q-step linear integration methods are taken for the Predictor and Corrector, then k iterations for the solver in the Corrector, as in (5), gives an explicit generalized Corrector formula which is again q-step but not linear anymore (coefficients of the characteristic equation depend on μ^j, $j \leqslant k+1$, $\mu = \lambda\tau$).

It is easy to verify that regardless of how accurate we solve (3), the above described generalized PC (GPC) scheme has a local error of order $p = \min\{p_1, p_2\}$ where p_1 and p_2 are the orders of the Predictor and the Corrector, respectively. To see this, assume that all $\boldsymbol{y}^m$, $\boldsymbol{y}^{m-1}$, ..., involved in the scheme, are exact, and that $\boldsymbol{\epsilon}_{p_1} = \mathcal{O}(\tau^{p_1+1})$ represents the local errors of the Predictor and, likewise, $\boldsymbol{\epsilon}_{p_2} = \mathcal{O}(\tau^{p_2+1})$ is the local error for the Corrector:

$$\boldsymbol{y}_{\rm ex}^{m+1} = \boldsymbol{y}_*^{m+1} + \boldsymbol{\epsilon}_{p_1}, \qquad \boldsymbol{y}_{\rm im}^{m+1} = \boldsymbol{y}_*^{m+1} + \boldsymbol{\epsilon}_{p_2}\,,$$

where $\boldsymbol{y}_*^{m+1}$ is the exact solution of (2) for $t = t_{m+1}$. Now denoting the residual polynomial of the iterative method used in (5) by $P_k(\tilde{\lambda})$, i.e.,

$$P_k(\tilde{\lambda}) = 1 - \tilde{\lambda}\hat{P}_{k-1}(\tilde{\lambda}), \qquad \boldsymbol{r}_k = P_k(\tilde{A})\boldsymbol{r}_0\,, \tag{8}$$

we note that

$$\boldsymbol{y}^{m+1} = \boldsymbol{y}_{\text{im}}^{m+1} - \tilde{A}^{-1}\boldsymbol{r}_k \qquad \text{and}$$
$$\tilde{A}^{-1}\boldsymbol{r}_k = \tilde{A}^{-1}P_k(\tilde{A})\tilde{A}(\boldsymbol{y}_{\text{im}}^{m+1} - \boldsymbol{y}_{\text{ex}}^{m+1}) = P_k(\tilde{A})(\boldsymbol{\epsilon}_{p_2} - \boldsymbol{\epsilon}_{p_1})\,.$$

Then the local error of the GPC scheme is of the form

$$\boldsymbol{y}^{m+1} - \boldsymbol{y}_*^{m+1} = \boldsymbol{\epsilon}_{p_2} - \tilde{A}^{-1}\boldsymbol{r}_k \tag{9a}$$
$$= \boldsymbol{\epsilon}_{p_2} - P_k(\tilde{A})(\boldsymbol{\epsilon}_{p_2} - \boldsymbol{\epsilon}_{p_1})\,. \tag{9b}$$

Condition (9b) illuminates our claim that the local error order of the GPC is at least the minimum of those of the Predictor and the Corrector, and (9a) shows that the order p_2 may be achieved by reduction of the residual. We note also that in the GPC scheme a different P_k may be used for each time step, so that the error can be differently reduced in each time step.

Of course, the explicit GPC schemes are not unconditionally stable, but the question arises, do we gain enough in stability to outbalance the additional computational work in comparison with conventional explicit formulas, e.g., PC schemes?

Suppose that GMRES [19] is used as a linear solver for the Corrector step. Because of the minimal residual property of GMRES [19], we obtain attractive properties for our GPC scheme. Apart from the desirable property to converge faster with respect to the outermost eigenmodes, we see that, when both Predictor and Corrector are linear q-step formulas, GMRES(k) gives a $(k+1)$-order[1] q-step explicit formula which is closest (in the residual sense) to the implicit Corrector scheme. Indeed, with k steps of GMRES, the residual in (3) is minimized by the choice of the polynomial P_k in (8) (i.e., $\hat{P}_{k-1}$ in (5)):

$$\begin{aligned}\text{Residual norm of (3)} = \|\boldsymbol{r}_k\| &= \left\|\boldsymbol{b}^m - \tilde{A}\boldsymbol{y}^{m+1}\right\| \\ &= \left\|\left(I - \tilde{A}\hat{P}_{k-1}(\tilde{A})\right)(\boldsymbol{b}^m - \tilde{A}\boldsymbol{y}_{\text{ex}}^{m+1})\right\| \\ &= \left\|P_k(\tilde{A})\boldsymbol{r}_0\right\| = \min_{R_k\in\Pi_k,\ R_k(0)=1}\left\|R_k(\tilde{A})\boldsymbol{r}_0\right\|\,.\end{aligned}$$

Here Π_k denotes the set of all real polynomials of degree at most k.

We will call GPC scheme with the minimal residual iterative solver by MR-PC scheme (Minimal Residual PC scheme). Let us agree also that by MR-PC(k) we will denote a MR-PC scheme with k Corrector iterations performed each time step.

3 MR-PC scheme with Forward Euler – Backward Euler

We consider the particular case where Forward Euler is the Predictor and Backward Euler is the Corrector. Then, in the notation of (3–5), we have that

$$\begin{aligned}&\boldsymbol{y}_{\text{ex}}^{m+1} = (I + \tau A)\boldsymbol{y}^m\,,\\ &\tilde{A} = I - \tau A, \qquad \hat{\beta} = 1\,,\\ &\boldsymbol{b}^m = \boldsymbol{y}^m\,,\\ &\boldsymbol{r}_0 = \boldsymbol{b}^m - \tilde{A}\boldsymbol{y}_0 = \boldsymbol{y}^m - \tilde{A}\boldsymbol{y}_{\text{ex}}^{m+1} = \tau^2A^2\boldsymbol{y}^m\,.\end{aligned} \tag{9}$$

[1]Here we mean the order of the formula, i.e., only the fact that the characteristic equation possibly contains μ^{k+1} $(\mu = \tau\lambda)$

With (5), the resulting explicit time-stepping scheme becomes

$$\boldsymbol{y}^{m+1} = (I + \tau A + \hat{P}_{k-1}(\tilde{A})\tau^2 A^2)\boldsymbol{y}^m \,. \tag{10}$$

Note that the order of the scheme is generally 1.

Let A be diagonalizable with eigenvalues $|\lambda_1| \geqslant \ldots \geqslant |\lambda_n|$ and let $\tilde{\lambda}_j = 1 - \hat{\beta}\tau\lambda_j$. Since diagonalization does not affect the stability analysis (the standard way: one just substitutes $\boldsymbol{y} \leftarrow P\boldsymbol{y}$ and multiplies the equation (2) by P; here matrix P gives the diagonal matrix PAP^{-1}), without loss of generality, we may assume that

$$A = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)\,.$$

Another assumption that can be done for the sake of simplicity is the homogeneity of (2). Indeed, let the time-stepping scheme

$$\boldsymbol{y}^{m+1} = S\boldsymbol{y}^m + \tau \boldsymbol{g}^m\,, \tag{11}$$

applied for the integration of (2), appear to be stable for the case $\boldsymbol{g} \equiv 0$, so that

$$\|\boldsymbol{y}^{m+1}\| \leqslant \|\boldsymbol{y}^m\|, \qquad m \geqslant 0\,. \tag{12}$$

Then the scheme (11) is also stable for any $\boldsymbol{g}$, since, as it can be shown,

$$\|\boldsymbol{y}^m\| \leqslant \|\boldsymbol{y}^0\| + \hat{C} \max_{0 \leqslant l \leqslant m-1} \|\boldsymbol{g}^l\|, \quad \hat{C} = m\tau, \quad m \geqslant 0\,.$$

Nominally, these considerations are not valid for the MR-PC schemes where operator S depends adaptively on both $\boldsymbol{y}^m$ and $\boldsymbol{g}^m$. However, in our stability analysis this adaptiveness is frozen out, so that (12) holds at each time step m. In other words, in reality MR-PC schemes may do even better than it is predicted in the theory.

3.1 Forward Euler – Backward Euler MR-PC(1) scheme

To simplify analysis, let us perform only one iteration step of GMRES ($k = 1$), so that we have the MR-PC(1) scheme. The polynomial $\hat{P}_{k-1}$ reduces to a constant $\hat{P}_{k-1} = \alpha$ and (10) is replaced by

$$\boldsymbol{y}^{m+1} = (I + \tau A + \alpha\tau^2 A^2)\boldsymbol{y}^m\,. \tag{13}$$

We note that the scheme would be of second order iff $\alpha = 1/2$, but this is generally not the case, since α is chosen adaptively in each time step in order to minimize the residual l_2-norm $\|\boldsymbol{y}^m - \tilde{A}\boldsymbol{y}^{m+1}\|$. This expresses that we have an explicit scheme closest (in the residual sense) to the Backward Euler among all other formulas of the form (13).

The conventional stability analysis with the characteristic equation (see, e.g. [12, 13]) leads to

$$\xi(\mu_j) = 1 + \mu_j + \alpha\mu_j^2, \qquad \mu_j = \tau\lambda_j\,, \tag{14}$$

and we are looking for a stability region $S \ni \mu$ (the region for which $|\xi(\mu)| \leqslant 1$) [12, 13]. It gives the whole stability picture (restriction for τ) for the one-dimensional Dahlquist equation $y' = \lambda y$ and, hence, for the diagonalizable matrix A one has to satisfy condition

$$\tau\lambda = \mu \in S, \qquad \lambda \in \{\lambda_j\}_{j=1}^n\,.$$

In our case the situation is further complicated by the fact that equation (14) is not homogeneous with respect to λ_j, due to the presence of the parameter α.

It is easy to verify that for GMRES(1), α satisfies the condition (assuming real values):

$$\alpha = \frac{\boldsymbol{r}_0^T \tilde{A} \boldsymbol{r}_0}{\|\tilde{A}\boldsymbol{r}_0\|^2}\,. \tag{15}$$

This means that α is in the numerical range of the symmetric part of $\tilde{A}^{-1}$. Moreover, from the GMRES(1) residual polynomial $P_1(\tilde{\lambda}) = 1 - \tilde{\lambda}\alpha$, we conclude that α is the reciprocal of the root of P_1, which is the harmonic Ritz value [16] of $\tilde{A}$. Denoting this by $\tilde{\theta}$, we get

$$\alpha = \frac{1}{\tilde{\theta}} = \frac{1}{1 - \hat{\beta}\tau\theta}\,, \tag{16}$$

where θ is the corresponding harmonic Ritz value[2] for A. With

$$\beta_j = \frac{\theta}{\lambda_j}\,, \tag{17}$$

and $\mu_j = \tau\lambda_j$, this leads to

$$\alpha = \alpha(\mu_j) = \frac{1}{1 - \hat{\beta}\beta_j\mu_j}\,. \tag{18}$$

For a given θ, relation (18) substituted in (14) provides the whole stability picture, because we easily get stability restrictions $|\xi(\mu_j)| \leqslant 1$ for all desired $\mu_j = \tau\lambda_j$ (the complex parameter β_j governs the positions of all eigenvalues λ_j with respect to the harmonic Ritz value θ). Indeed, (14) with (18) leads to

$$\xi(\mu_j) = \frac{1 + \mu_j(1 + \mu_j)(1 - \beta_j)}{1 - \beta_j\mu_j}\,. \tag{19}$$

If, for some j, the harmonic Ritz value coincides with an eigenvalue λ_j then $\beta_j = 1$ and

$$\xi(\mu_j) = \frac{1}{1 - \mu_j}\,.$$

In other words, with respect to the eigenvalue λ_j the scheme reduces to Backward Euler. Note also that for the scalar Dahlquist equation this would always be the case since GMRES(1) would give the exact solution with $\alpha = \tilde{\lambda}^{-1}$.

Proposition 1 (Time-step restriction) *Let the eigenvalues of A be negative (real) values λ and $\theta \in [\lambda_{\min}; \lambda_{\max}]$. Then the Forward Euler – Backward Euler MR-PC(1) scheme is stable iff the time step is restricted as*

$$\begin{aligned} &\tau \leqslant \tau^*(\theta, \lambda), \qquad \forall \lambda \in \operatorname{spectrum} A\,, \\ &\tau^*(\theta,\lambda) = \begin{cases} \dfrac{1}{\theta - \lambda} & \text{for } \lambda < \theta\,, \\ \dfrac{2\theta - \lambda - \sqrt{4\theta(\theta+\lambda) - 7\lambda^2}}{2\lambda(\lambda - \theta)} & \text{for } \lambda > \theta\,. \end{cases} \end{aligned} \tag{20}$$

Note The condition $\theta \in [\lambda_{\min}; \lambda_{\max}]$ holds for the MR-PC(1) scheme since harmonic Ritz values approximate the spectrum from inside [16].

[2] Actually, θ is the harmonic Ritz value of A with respect to the shift $-(\hat{\beta}\tau)^{-1}$.

Proof For each λ define value μ^* in the following way:

$$|\mu^*| = \max\left\{|\mu| = |\lambda|\tau \;\middle|\; |\xi(\mu)| \leqslant 1 \text{ (stability condition)}\right\}, \qquad \mu^* < 0\,.$$

This implies that the value μ^*/λ gives the maximal time step allowed with respect to this λ. Analysis of (19), with $\beta = \theta/\lambda > 0$, leads to the stability restrictions

$$\mu \geqslant \mu^* = \begin{cases} \dfrac{1}{1-\beta} & \text{for } \beta > 1\,, \\ \dfrac{2\beta - 1 - \sqrt{4\beta(\beta+1) - 7}}{2(1-\beta)} & \text{for } \beta < 1\,. \end{cases}$$

With $\mu = \tau\lambda$, this immediately implies (20). ■

We will consider some particular cases.

First assume that the spectrum of A is negative. Then we may scale the time in such a way that $\lambda_1 = -1$: the spectrum of A is in $[-1; 0)$. Now for some fixed $\theta \in [-1; 0)$, we can study the stability for any possible eigenvalue $\lambda \in [-1; 0)$, by varying $\theta/\lambda \equiv \beta \in (0; \infty)$. As it can be seen from the numerical experiments, θ usually occurs to be very close to the largest in magnitude eigenvalue, in this particular case $-1 \lesssim \theta$. Assume for simplicity that $\theta = -1$.

In Table 1 we give the maximal time steps obtained with (20) for some sample eigenvalues in the spectrum (all values are rounded to two decimal places). In the second column of Table 1 we present the auxiliary value μ^*. In the last column of the Table we give listed a *gain factor*:

$$\varkappa = \frac{\tau_{\mathrm{MR-PC}}}{(k+2)\tau_{\mathrm{FE}}}\,, \tag{21}$$

where $\tau_{\mathrm{MR-PC}}$ and τ_{FE} are the maximum time steps allowed for the MR-PC with GMRES(k) and the Forward Euler schemes, respectively. This gain factor $\varkappa$ depends on the solver for the Corrector scheme equation. The factor $1/(k+2)$ indicates that for the MR-PC with GMRES(k), at each time step, one has to perform $k+2$ matrix–vector multiplications instead of 1 for Forward Euler.

In Figure 1 we have plotted the function $\tau^*(\theta, \lambda)$ according to (20). Apart from $\theta = -1$, here we have also selected $\theta = -0.9$ and $\theta = -0.5$. The minimum of each curve presented corresponds to the actual time-step restriction. As we see, the case $\theta = -1.00$ is not optimal, since $\theta = -0.9$ allows a larger time step.

Observing these data, we note the following interesting property of our schemes: if θ coincides with the largest (in modulus) eigenvalue ($\theta = -1$) then the time step is effectively restricted not by large (in modulus) eigenvalues, but primarily by some eigenvalues of order $\approx 0.6|\lambda_1|$. Indeed, with respect to the largest eigenvalue we have an absolutely stable Backward Euler scheme, and, for some other eigenvalues which are close to the largest one, the scheme is still "close" to the Backward Euler. For other components the scheme behaves like an explicit one (see (5), (13)), which means that the small eigenvalues are not restrictive.

To clarify the situation for other large eigenvalues, we have plotted on Figure 2 the stability regions S for $\theta = -1.00$ and some β close to 1 (these values of β correspond to eigenvalues of order 1 in modulus). We should note that $|\beta_1| = |\theta|/|\lambda_1| \leqslant 1$ because harmonic Ritz values approximate the spectrum from inside [16].

The stability region for the case where λ_j are uniformly distributed in $[-1; 0)$ may be obtained as the intersection of stability regions plotted for all eigenvalues affecting the τ, i.e., for $\lambda_j \lesssim -0.55$. It is represented in Figure 3 and compared with those of the Forward and Backward Euler Schemes.

To indicate weak points of MR-PC(1) scheme, let us consider a very bad case for the scheme, namely, one where the largest (in modulus) eigenvalue λ_1 is complex. Evidently, since our original matrix is real we have a complex-conjugate pair, $\lambda_1 = \bar{\lambda}_2$. Assume that $|\lambda_1| \leqslant 1$. In Table 2 the time step restrictions are presented for some $\lambda_{1,2}$ ($|\lambda_{1,2}| \lesssim 1$) and θ. In Table 2, all data have been rounded to two decimal places. By $\theta_\bullet$ we denote the actual harmonic Ritz value derived from GMRES(1) applied to the 2×2 linear system with a real matrix with $\lambda_{1,2}$ as its eigenvalues, random right hand side and zero initial guess. This harmonic Ritz value which is the reciprocal of α in (15) is a good guess for the actual value that the integration scheme would select. In the last column of Table 2 we present for a given λ the time step restriction of the Forward Euler scheme $2|\mathcal{R}e\,\lambda|/|\lambda|^2$. As we see from the presented data, the position of the harmonic Ritz value has a dramatic effect on the stability of the scheme. Nevertheless, even for this unfavorable situation, the scheme is essentially more stable than Forward Euler with maximal time step $\tau_{\mathrm{FE}} = 2$. Although the gain factor $\varkappa$ calculated according to (21), may indicate superiority of the Forward Euler scheme, we usually observe in practice $\varkappa > 1$ (see section 6, page 20).

Table 1: Maximal time steps τ for the Forward Euler – Backward Euler MR-PC(1) scheme, $\theta = -1.00$, $\lambda \in [-1; 0)$

β	$\|\mu\| \leqslant \mu^* =$	$\lambda = \theta/\beta = -\beta^{-1}$	$\tau \leqslant$	$\varkappa$
1.05	23.70	-0.95	24.89	4.15
1.10	13.45	-0.91	14.74	2.46
1.30	6.35	-0.77	8.25	1.38
1.50	4.82	-0.67	7.19	1.20
1.75	3.99	-0.57	7.00	1.17
2.00	3.56	-0.50	7.12	1.19
5.00	2.45	-0.20	12.25	2.04
10.00	2.21	-0.10	22.10	3.68
100.00	2.02	-0.01	202.00	33.67

Table 2: Maximal time steps τ for the Forward Euler – Backward Euler MR-PC(1) scheme compared to those of the Forward Euler (FE) scheme for some "unfavorable" λ's

θ	$\lambda_{1,2}$	$\|\mu\| \leqslant \mu^* =$	MR-PC: $\tau \leqslant$	FE: $\tau \leqslant$	$\varkappa$
$\theta = -1.00$	$-0.9 \pm 0.1i$	9.30	10.27	2.20	1.56
$\theta = -1.00$	$-0.8 \pm 0.5i$	3.56	3.77	1.80	0.62
$\theta = -1.00$	$-0.7 \pm 0.7i$	2.82	2.84	1.43	0.66
$\theta_\bullet = -1.40$	—"—	2.70	2.72	–"–	0.63
$\theta = -1.00$	$-0.3 \pm 0.9i$	1.61	1.70	0.67	0.85
$\theta_\bullet = -3.00$	—"—	1.12	1.18	–"–	0.59

3.2 Forward Euler – Backward Euler MR-PC(k) scheme

In this section we give stability analysis for the Forward Euler – Backward Euler MR-PC(k) scheme for two particular "ideal" cases. For each case, we make a particular assumption on the polynomials taking part in the scheme. It is done in attempts to predict the real-life stability properties of MR-PC schemes with GMRES which delivers new adaptive polynomial each time step. The assumption of the first case is very artificial, it leads to the time step restriction

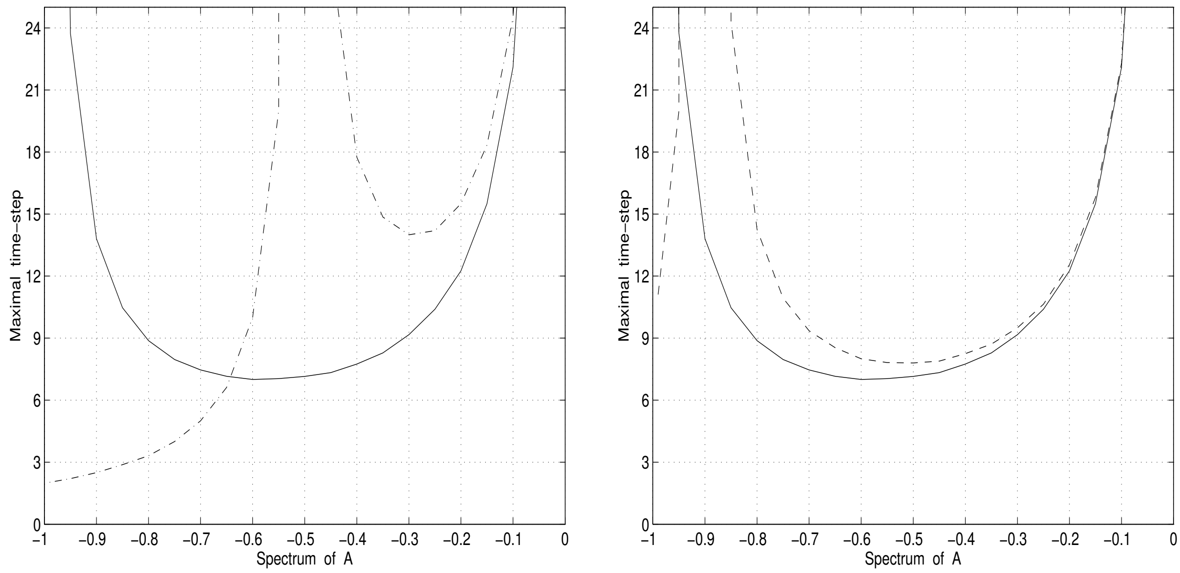


Figure 1: Forward Euler – Backward Euler MR-PC(1) scheme. Maximal time step τ versus $\lambda \in \{-0.99, -0.95 : 0.05 : -0.05, -0.01\}$ affected by θ: $\theta = -0.5$ (dashdotted line), $\theta = -0.9$ (dashed line), $\theta = -1.00$ (solid line)

which occurs to be too pessimistic (too safe) in reality whereas the assumption for the second case proves to be rather close to the reality with an adequate time step restriction. The first case is interesting in the respect that it shows superiority of the adaptive minimal residual iteration in comparison with a stationary iterative scheme (namely, the iteration where the residual polynomial is chosen to be having its roots in the largest part of the spectrum of $\tilde{A}$).

For the Forward Euler – Backward Euler MR-PC(k) scheme (10) we have that

$$\begin{aligned} (\boldsymbol{r}_k =) \quad P_k(\tilde{A})\boldsymbol{r}_0 &= \boldsymbol{y}^m - \tilde{A}\boldsymbol{y}_k\,, \\ \boldsymbol{y}_k &= \tilde{A}^{-1}\big[\boldsymbol{y}^m - P_k(\tilde{A})\boldsymbol{r}_0\big]\,, \end{aligned} \tag{22}$$

where $P_k(\tilde{\lambda}) = 1 - \tilde{\lambda}\hat{P}_{k-1}(\tilde{\lambda})$ is the residual polynomial of the iterative process.

With (9) and $\boldsymbol{y}^{m+1} = \boldsymbol{y}_k$ it follows from (22) that

$$\begin{aligned} \boldsymbol{y}^{m+1} &= S\boldsymbol{y}^m\,, \\ S &= \tilde{A}^{-1}\big[I - P_k(\tilde{A})(\tau A)^2\big] \end{aligned} \tag{23}$$

and the scheme (10) is spectrally stable (i.e., spectral radius $\rho(S) \leqslant 1$) iff

$$\left|\frac{1 - P_{kj}\cdot(\tau\lambda_j)^2}{1 - \tau\lambda_j}\right| \leqslant 1, \quad j = 1, \ldots, n\,, \tag{24}$$

where λ_j are eigenvalues of A and $P_{kj} = P_k(\tilde{\lambda}_j)$.

Furthermore, assuming that $\boldsymbol{y}^m = \sum_j \alpha_j^m \boldsymbol{w}_j$ is an expansion in $\boldsymbol{w}_j$ eigenvectors of A, with (23) we arrive at

$$\alpha_j^{m+1} = \frac{1 - P_{kj}^{m+1}\cdot(\tau\lambda_j)^2}{1 - \tau\lambda_j}\alpha_j^m\,, \quad m \geqslant 0\,, \tag{25}$$

where we added the superscript $\cdot^m$ to P_{kj} to emphasize that for GMRES (and other similar robust Krylov iterations) $P_k(\tilde{\lambda})$ depends on time. As it may be seen from (25), after performing few time

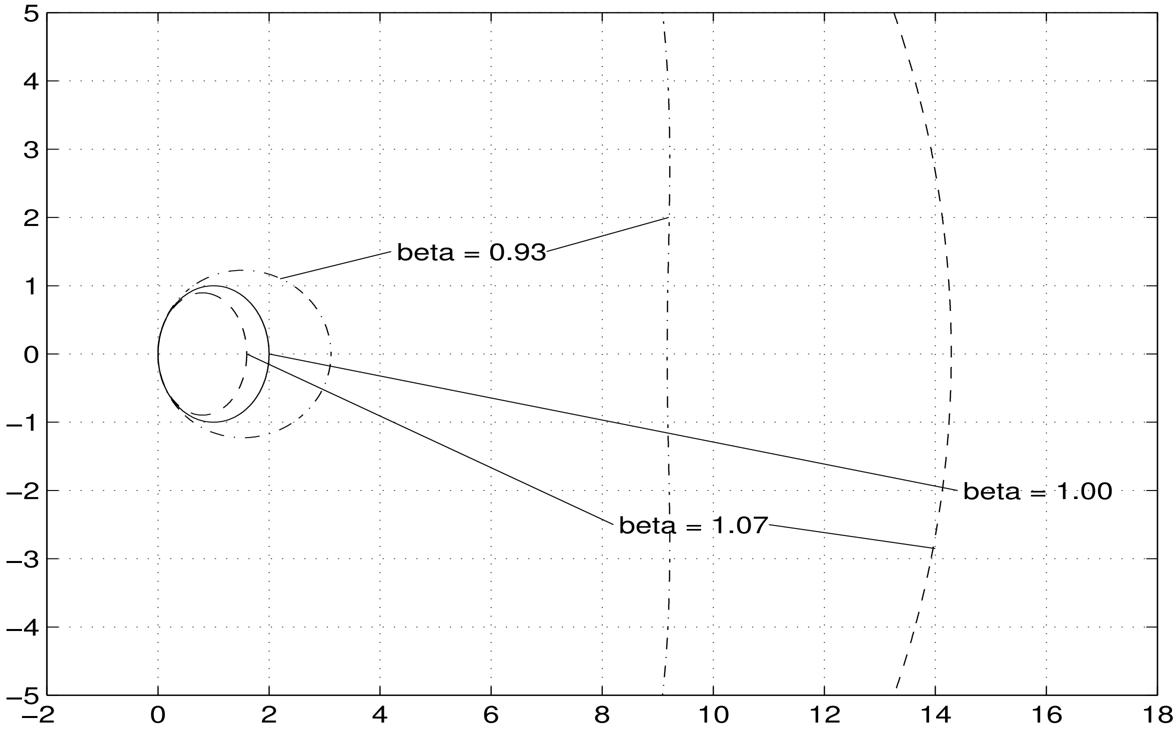


Figure 2: Stability regions $S \subset \mathbb{C}$ of the Forward Euler – Backward Euler MR-PC(1) scheme for $\theta = -1.00$ and $\beta \approx 1$: $\beta = 0.93$ — dashdotted line; $\beta = 1.00$ (Backward Euler scheme) — solid line; $\beta = 1.07$ — dashed line

steps with polynomials $P_k^m(\tilde{\lambda})$ having roots in some points of the spectrum, the corresponding k coefficients α_j^m (namely, those k coefficients for which $\lambda_j \approx \theta_j$) do not dominate anymore as they have had $\approx (1-\tau\lambda_j)^{-1}$ as an amplification factor. Hence for GPC schemes with robust Krylov subspace iterative schemes like GMRES(k), we may expect the feedback phenomena: roots of the $P_k^m(\tilde{\lambda})$ may move to some new positions as soon as $\alpha_1^m, \ldots, \alpha_k^m$ have been getting small enough. Numerical tests (section 6, page 20) confirm these expectations. They show that such a successive "cleaning-off" may be nonmonotone for GMRES(k), so that harmonic Ritz values may fluctuate inside spectrum $I - \tau A$ in accordance with what α_j^m have become undesirably large.

With $\lambda_j \leqslant 0$, requirement (24) leads to the restriction

$$\begin{aligned} \tau &\leqslant \min\{\tau', \tau''\}\,, \\ \tau' &= \min_{j:\, P_{kj}>0} \frac{1+\sqrt{1+8P_{kj}}}{2P_{kj}|\lambda_j|}\,, \\ \tau'' &= \min_{j:\, P_{kj}<0} \frac{1}{|P_{kj}||\lambda_j|}\,, \end{aligned} \tag{26}$$

or, both inequalities hold if

$$\tau \leqslant \min_{j:\, P_{kj}\neq 0} \frac{1}{|P_{kj}||\lambda_j|}\,. \tag{27}$$

Case 1 Rather complicated upper bounds (26),(27) for τ can be interpreted as follow.

Recall that the roots of the residual polynomial $P_k(\tilde{\lambda})$ are of the form $1 - \tau\theta_j$ where θ_j are approximations to some of A eigenvalues λ_j. At the same time we observe from the iterations

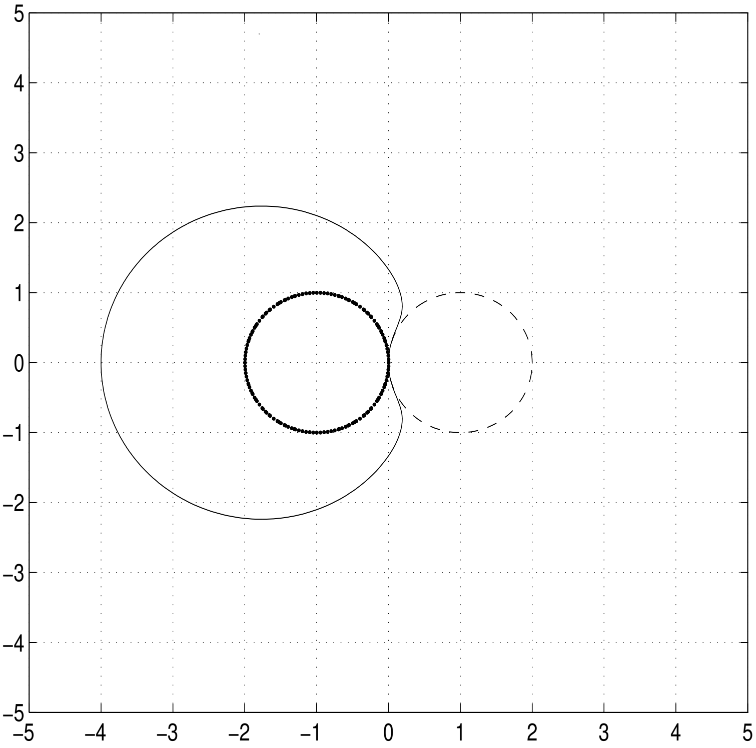


Figure 3: Stability regions $S \subset \mathbb{C}$ for the Forward Euler – Backward Euler MR-PC(1) scheme, $\theta = -1.00$ (inside the solid curve), Forward Euler (inside the dotted circle), and Backward Euler (outside the dashed circle)

that exterior eigenvalues are approximated better than interior ones. Assume that $\theta_j \approx \lambda_j$, $j = 1, \ldots, l \leqslant k$. In such an extreme (and not very likely) case $P_{kj} \approx 0$ for $j = 1, \ldots, l$ and the largest l eigenvalues of A do not play the main role in the restriction (24) anymore[3]. Moreover, it is possible to show that in the case where $\theta_j = \lambda_j$, $j = 1, \ldots, k$ (i.e., if *all* θ's coincide with largest (in modulus) λ's), the remaining eigenvalues $|\lambda_{k+1}| \geqslant \ldots \geqslant |\lambda_n|$ put the same restriction on the time step as for the explicit Forward Euler scheme, i.e., of the form $\tau \leqslant 2/|\lambda_{k+1}|$:

Proposition 2 (Time-step restriction, case 1) *Let A be symmetric negative semidefinite matrix and the residual polynomial $P_k(\tilde{\lambda})$ have its roots in the form $1 - \tau\lambda_j$, $j = 1, \ldots, k \leqslant n$ ($|\lambda_1| \geqslant \ldots \geqslant |\lambda_k| > |\lambda_{k+1}| \geqslant \ldots \geqslant |\lambda_n|$ are eigenvalues of A). Then the Forward Euler – Backward Euler GPC(k) scheme is stable if*

$$\tau \leqslant \frac{2}{\max_{j>k} |\lambda_j|} = \frac{2}{|\lambda_{k+1}|}\,. \tag{28}$$

Proof Since $P_k(\tilde{\lambda})$ is the residual polynomial, $P_k(0) = 1$. Hence, taking into account that there is no root of P_k in $(0\,; 1 - \tau\lambda_k)$ we have $0 < P_{kj} < 1$ for $j = k+1, \ldots, n$. Further proof is just verification of the fact that the stability restriction $\tau \leqslant \tau'$ from (26) is a necessary implication of (28). Indeed, noting that function $\phi(p) = (1 + \sqrt{1 + 8p})/2p$ is monotone decreasing for $p \in (0; 1]$

[3]In fact, GMRES can not do worse because the adaptive selection of θ_j in GMRES leads to the minimal possible residual.

and attains its minimal value 2 at $p = 1$ we get

$$\tau \leqslant \frac{2}{\max_{j>k}|\lambda_j|} = \frac{\min_{p\in(0;1]}\phi(p)}{\max_{j>k}|\lambda_j|} \leqslant$$
$$\leqslant \frac{\min_{j>k}\phi(P_{kj})}{\max_{j>k}|\lambda_j|} \leqslant \min_{j>k}\frac{1+\sqrt{1+8P_{kj}}}{2P_{kj}|\lambda_j|} = \tau' .$$

■

Thus, in the extreme case when the largest k eigenvalues λ_j deliver the polynomial roots $1-\tau\lambda_j$, the time step stability restriction is not worse than for the explicit Forward Euler scheme applied to the problem with absent k largest eigenvalues.

Now let us see whether it is possible to use somehow Proposition 2 for the step size control. Suppose that the extreme assumption of the Proposition on the strict coincidence of the residual polynomial roots with the largest eigenvalues is satisfied. The value λ_{k+1} taking part in the restriction (28) is unknown in practice, therefore we may require instead that

$$\begin{aligned} \tau \leqslant \frac{2}{\theta_k}, \qquad \theta_k = \max_{i=1,\dots k}\theta_i, \\ \text{or} \qquad \eta_k \geqslant -2, \qquad \eta_k = \tau\theta_k . \end{aligned} \tag{29}$$

The last condition is stricter than (28), whereas the value of $\theta_k = \min_{i=1,\dots k}|\theta_i|$ may be easily got from the iteration scheme. As it occurs, applied in practice for MR-PC schemes with GMRES (assume we do neglect all assumptions of Proposition 2 and just impose (29)), restriction (29) is too safe. It gives too small step size and therefore our expectations, that the extreme assumptions of Proposition 2 give us the polynomial $P_k(\tilde{\lambda})$ which is far from the best possible one, are confirmed.

Case 2 Another interpretation of the condition (24) may be given if we make some reasonable assumptions on the polynomial $P_k(\tilde{\lambda})(\tau\lambda)^2$ taking part in that expression. Remind that

$$\begin{aligned} \|\boldsymbol{r}_k\| &= \left\|P_k(\tilde{A})\boldsymbol{r}_0\right\| = \left\|P_k(I-\tau A)(\tau A)^2\boldsymbol{y}^m\right\| \\ &= \min_{R_k\in\Pi_k,\ R_k(0)=1}\left\|R_k(I-\tau A)(\tau A)^2\boldsymbol{y}^m\right\| \leqslant \mathcal{C}_E\left|\max_{\mu=\tau\lambda_j}P_k(I-\mu)\mu^2\right|\|\boldsymbol{y}^m\|, \end{aligned} \tag{30}$$

where $\mathcal{C}_E$ is the condition number of matrix of eigenvectors of τA and Π_k is the set of all real polynomials of degree at most k. Of course, in our implicit assumption that A has been diagonalized we may omit the factor $\mathcal{C}_E$. It is left here only to illuminates the fact that the minimal residual iterations (i.e. polynomial P_k) do depend on the basis in which we consider A.

Let us denote

$$Q(\mu) = P_k(1-\mu)\mu^2 , \tag{31}$$

and now we make the following assumption on $Q(\mu)$ which makes sense as we want to have an optimal reduction of $\|\boldsymbol{r}_k\|$ in (30). Suppose that $Q(\mu)$ is the minimally declining polynomial on the set

$$\mathcal{M} = \{\mu_1 \leqslant \ldots \leqslant \mu_n\} , \qquad \mu_n < 0$$

with at least $k+1$ different negative elements. More precisely, suppose that $Q(\mu)$ solves the following minimax problem:

$$Q(\mu) \text{ is the solution of } \quad \min_{R\in\Pi^*_{k+2}}\max_{\mu\in\mathcal{M}}|R(\mu)| , \tag{32}$$

where Π^*_{k+2} is the set of all polynomials $R(\mu)$ of degree at most $k+2$ such that $R(0) = R'(0) = 0$, $R(1) = 1$. This assumption on $Q(\mu)$ is especially likely to be close to the truth if $\|\boldsymbol{y}^m\|$ has components of approximately the same magnitude.

Notice that condition (32) means that $Q(\mu)$ is exactly of degree $k+2$ and has the alternance property on $[\mu_1; \mu_n]$, so that there are k roots $\eta_1 < \eta_2 < \cdots < \eta_k$ of $Q(\mu)$ in $[\mu_1; \mu_n]$ (this can be proved in a standard way, see, e.g., [17]).

Proposition 3 (Time-step restriction, case 2) *Let A be symmetric negative semidefinite matrix with at least $k+1$ different eigenvalues $\lambda_1 \leqslant \ldots \leqslant \lambda_n$. Assume that the MR-PC residual polynomial $P_k(\tilde{\lambda}) = P_k(1-\mu)$ delivers in (31) polynomial satisfying (32) for $\mathcal{M} = \{\mu_j = \tau\lambda_j\}_{j=1}^n$. Then the condition*

$$\eta_k \geqslant -7 \tag{33}$$

is sufficient for the Forward Euler – Backward Euler MR-PC(k) scheme to be stable. The last corresponds to the time step restriction

$$\tau \leqslant \frac{7}{|\theta_k|}, \tag{34}$$

where θ_k is the largest element of the set $\{\theta_i \mid P_k(1-\tau\theta_i) = 0\}$.
Note For MR-PC with GMRES(k) θ_k are harmonic Ritz values of τA delivered by the iterations.

Proof The stability condition (24) is equivalent to the inequality

$$\mu \leqslant Q(\mu) \leqslant 2 - \mu\,, \qquad \mu = \tau\lambda_j\,, \qquad j = 1, \ldots, n\,.$$

In assumptions of the Proposition the last is fulfilled if the largest nonzero root η_k of the polynomial $Q(\mu)$ satisfies (33) (see Appendix 1). ■

Remark For $k > 1$, following the same approach like in the proof of Proposition 3, it is possible to revise (make more precise) numerically the restriction (33):

$$\begin{array}{cccc} k=2 & k=3 & k=4 & k=5 \\ \eta_k \geqslant -9.69... & \eta_k \geqslant -11.06... & \eta_k \geqslant -11.84... & \eta_k \geqslant -12.32... \end{array} \tag{35}$$

Now let us make some conclusions on the time step restrictions obtained for MR-PC(1) scheme in this section. According to the Proposition 1 (see (20) and Figure 1) the largest value of τ for the case where A has its eigenvalues to be uniformly distributed in $[-1; 0)$ is $\tau \approx 7.5$. This value of τ is achieved for the harmonic Ritz value $\theta = \theta_1 \approx -0.9$ (dashed curve on Figure 1). This is in a good agreement with Proposition 3 (see restriction (34)).

4 MR-PC scheme with Adams(2) – BDF2

Here, all assumptions of the previous section are in force.

As an example of the GPC approach to construct a second order accuracy formula, we will consider a MR-PC scheme with Adams (2) Predictor and BDF2 Corrector. Since the Predictor and the Corrector are of second order, the local error of the MR-PC scheme is also $\mathcal{O}(\tau^3)$.

With the same notation as in (3)–(4), we obtain

$$
\begin{aligned}
\boldsymbol{y}_{\text{ex}}^{m+1} &= \boldsymbol{y}^m + \tau\left(\frac{3}{2}A\boldsymbol{y}^m - \frac{1}{2}A\boldsymbol{y}^{m-1}\right),\\
\tilde{A} &= I - \frac{2\tau}{3}A, \qquad \hat{\beta} = \frac{2}{3},\\
\boldsymbol{b}^m &= \frac{4}{3}\boldsymbol{y}^m - \frac{1}{3}\boldsymbol{y}^{m-1},\\
\boldsymbol{r}_0 &= \boldsymbol{b}^m - \tilde{A}\boldsymbol{y}_{\text{ex}}^{m+1}\\
&= \left(\frac{1}{3}I - \frac{5\tau}{6}A + \tau^2A^2\right)\boldsymbol{y}^m\\
&+ \left(-\frac{1}{3}I + \frac{\tau}{2}A - \frac{\tau^2}{3}A^2\right)\boldsymbol{y}^{m-1},
\end{aligned}
\tag{36}
$$

where the expression for $\boldsymbol{y}_{\text{ex}}^{m+1}$ is the Adams(2) Predictor step and $\tilde{A}$ and $\boldsymbol{b}^m$ delivers in (3) BDF2 formula.

By rewriting (5) explicitly for the case (36), we arrive at the following time-stepping GPC(k) scheme:

$$
\begin{aligned}
\boldsymbol{y}^{m+1} = &\left(\frac{3I + \hat{P}_{k-1}(\tilde{A})}{3} + \tau A\frac{9I - 5\hat{P}_{k-1}(\tilde{A})}{6} + \tau^2A^2\hat{P}_{k-1}(\tilde{A})\right)\boldsymbol{y}^m\\
+ &\left(-\frac{\hat{P}_{k-1}(\tilde{A})}{3} + \tau A\frac{\hat{P}_{k-1}(\tilde{A}) - I}{2} - \tau^2A^2\frac{\hat{P}_{k-1}(\tilde{A})}{3}\right)\boldsymbol{y}^{m-1}.
\end{aligned}
\tag{37}
$$

The polynomial $\hat{P}_{k-1}$ comes from the iterative process used to solve the BDF2 system.

Again, like for the Forward Euler – Backward Euler GPC schemes, we consider the simplest case, the MR-PC(1) scheme, in more detail. Substitution of $\hat{P}_{k-1} = \alpha$ in (37), leads to the MR-PC(1) scheme

$$
\begin{aligned}
\boldsymbol{y}^{m+1} = &\left(\frac{3+\alpha}{3}I + \frac{9-5\alpha}{6}\tau A + \alpha\tau^2A^2\right)\boldsymbol{y}^m\\
+ &\left(-\frac{\alpha}{3}I + \frac{\alpha-1}{2}\tau A - \frac{\alpha}{3}\tau^2A^2\right)\boldsymbol{y}^{m-1},
\end{aligned}
\tag{38}
$$

with the characteristic equation

$$
\xi^2 + \left(-\frac{3+\alpha}{3} - \frac{9-5\alpha}{6} - \alpha\mu^2\right)\xi + \left(\frac{\alpha}{3} - \frac{\alpha-1}{2}\mu + \frac{\alpha}{3}\mu^2\right) = 0, \quad \mu = \mu_j = \tau\lambda_j\,.
\tag{39}
$$

Following the same argumentation and notation as for the Forward Euler – Backward Euler MR-PC scheme (see (16)–(18)), we get the roots of the characteristic equation (39) in the form

$$
\begin{aligned}
\xi_{1,2} &= -\frac{4 + (1-\beta)(3\mu^2 + 2\mu) \pm \sqrt{D}}{2(2\beta\mu - 3)},\\
D &= 4 + 16\mu + 16\mu^2 + 12\mu^3 + 9\mu^4\\
&\quad -(8\mu + 20\mu^2 + 16\mu^3 + 18\mu^4)\beta\\
&\qquad +(4\mu^2 + 4\mu^3 + 9\mu^4)\beta^2\,.
\end{aligned}
\tag{40}
$$

One can verify that if the harmonic Ritz value θ coincides with some eigenvalue λ_j then (39), (40) give for μ_j the BDF2 characteristic equation

$$\begin{aligned} &(2\mu_j - 3)\xi^2 + 4\xi - 1 = 0\,, \\ &\xi_{1,2}(\mu_j) = \frac{-2 \pm \sqrt{1+2\mu_j}}{2\mu_j - 3}, \qquad \mu_j = \tau\lambda_j\,. \end{aligned} \tag{41}$$

This is similar to the Forward Euler – Backward Euler MR-PC scheme, where for $\theta = \lambda_j$ the scheme reduces to Backward Euler (see (19) and below). Again, if θ coincides with an eigenvalue then the MR-PC scheme reduces (for this eigenvalue) to the Corrector scheme.

We consider now some particular cases for (40).

First, assume that the spectrum of A is real and $\lambda_1 = -1$. With relations (16)–(18), we may set θ to some appropriate value (for example, let $\theta = -1$, since in the reality it usually holds $-1 \lesssim \theta$) and, by varying β in (40), we get a stability picture for any A eigenvalue $\lambda = \theta/\beta_j \in [-1; 0)$. The maximal values for the time step, obtained in such a way for some sample eigenvalues and $\theta = -1.00$, are presented in Table 3. As we see, again, for $\theta = -1$ the most critical eigenvalues for the time step are those close to $\approx 0.6|\lambda_1|$. In the last column of the Table 3, the gain factor $\varkappa$, with respect to the conventional Adams(2) – BDF2 PC scheme, is given:

$$\varkappa = \frac{2\tau_{\mathrm{MR-PC}}}{(k+2)\tau_{\mathrm{PC}}}\,. \tag{42}$$

Formula (42) corresponds to the GMRES(k) linear solver and the factor $2/(k+2)$ is due to the fact that we have to perform two matrix–vector multiplications for the PC step (assuming that we keep $A\boldsymbol{y}^m$ in memory for Adams(2), so that both Predictor and Corrector require one multiplication each) and $k+2$ matrix–vector multiplications for the MR-PC step. Note that for this particular case $\tau_{\mathrm{PC}} = 1$ (see, e.g., [12, 13]).

The picture corresponding to Table 3 is given in Figure 4.

In Figure 5 we have plotted the stability regions for $\theta = -1$ and some β close to 1. It gives an impression of what is happening with respect to the largest part of the spectrum (as we noted already, for $\beta = 1$ (or, equivalently, $\theta = \lambda$) we have the BDF2 scheme).

The intersection of all stability regions for $-1 \leqslant \lambda \lesssim -0.55$ (those λ's are putting a restriction on τ), gives the stability region for the scheme applied to the whole system. It is presented in Figure 6, together with those for Adams(2), the Adams(2) – BDF2 PC, and BDF2 schemes.

Finally, for the bad situation where $\lambda_1 \in \mathbb{C}$, $\lambda_1 = \bar{\lambda}_2$, we present maximal values of τ for some $|\lambda_{1,2}| \approx 1$ in Table 4. The values of $\theta_\bullet$ are actual values taken from GMRES(1) in the same manner as for Table 2 (p. 8). The gain factor $\varkappa$ is calculated according to the (42). From the data in Table 4 we note that for this unfavorable case MR-PC(1) does not give much improvement (or, even loses) in comparison with conventional PC scheme. With the MR-PC(2) scheme (with two iterations of GMRES) we may hope to get a better situation because with two harmonic Ritz eigenvalues we may capture the conjugate eigenpair.

5 Some notes on the nonlinear case

For a nonlinear autonomous function f in (1), we have to solve at each time step the equation

$$\boldsymbol{y}^{m+1} - \hat{\beta}\tau f(\boldsymbol{y}^{m+1}) = \boldsymbol{b}^m\,, \tag{43}$$

which is the nonlinear analogue of (3). A standard approach here is the use of (inexact) Newton's method in the following way. Let $\boldsymbol{y}_{(0)}$ be some initial guess for $\boldsymbol{y}^{m+1}$, for instance, $\boldsymbol{y}_{(0)}$ is obtained

Table 3: Maximal time steps τ for the Adams(2) – BDF2 MR-PC scheme, $\theta = -1.00$, $\lambda \in [-1; 0)$

β	$\lvert\mu\rvert \leqslant \mu^* =$	$\lambda = \theta/\beta = -\beta^{-1}$	$\tau \leqslant$	$\varkappa$
1.05	13.34	-0.95	14.04	9.36
1.10	8.38	-0.91	9.28	6.20
1.30	4.22	-0.77	5.51	3.68
1.50	3.23	-0.67	4.86	3.24
1.75	2.66	-0.57	4.66	3.10
2.00	2.35	-0.50	4.70	3.14
5.00	1.46	-0.20	7.30	4.86
10.00	1.23	-0.10	12.30	8.20
100.00	1.02	-0.01	102.00	68.00

Table 4: Maximal time steps τ for the Adams(2) – BDF2 MR-PC scheme compared to those of the Adams(2) – BDF2 PC scheme for some "unfavorable" λ's

θ	$\lambda_{1,2}$	$\lvert\mu\rvert \leqslant \mu^* =$	MR-PC: $\tau \leqslant$	PC: $\tau \leqslant$	$\varkappa$
$\theta = -1.00$	$-0.9 \pm 0.1i$	6.00	6.63	1.09	4.06
$\theta = -1.00$	$-0.8 \pm 0.5i$	2.57	2.72	1.09	1.66
$\theta = -1.00$	$-0.7 \pm 0.7i$	2.15	2.17	1.15	1.26
$\theta_\bullet = -1.40$	—"—	2.02	2.04	–"–	1.18
$\theta = -1.00$	$-0.3 \pm 0.9i$	1.55	1.63	1.47	0.74
$\theta_\bullet = -3.00$	—"—	1.15	1.21	–"–	0.54

with an explicit (Predictor) scheme. The Newton process $\boldsymbol{y}_{(s)} \to \boldsymbol{y}^{m+1}$ may be sketched as

$$
\begin{aligned}
&\texttt{repeat} \\
&\quad \text{solve linear system } \tilde{A}\boldsymbol{x}_{(s)} = \boldsymbol{b}^m_{(s)}\,, \\
&\qquad \tilde{A} = I - \hat{\beta}\tau J_{(s)}\,, \\
&\qquad \boldsymbol{b}^m_{(s)} = \boldsymbol{b}^m - \boldsymbol{y}_{(s)} + \hat{\beta}\tau f(\boldsymbol{y}_{(s)})\,; \\
&\quad \boldsymbol{y}_{(s+1)} = \boldsymbol{y}_{(s)} + \boldsymbol{x}_{(s)}\,; \\
&\texttt{until} \text{ convergence in } \boldsymbol{y}_{(s)} \to \boldsymbol{y}^{m+1}\,.
\end{aligned}
\tag{44}
$$

The matrix $J_{(s)}$ denotes the Jacobian of f in $\boldsymbol{y}_{(s)}$. The popular linearized form of (43), namely,

$$(I - \hat{\beta}\tau J_m)\boldsymbol{y}^{m+1} = \boldsymbol{b}^m + \hat{\beta}\tau(f(\boldsymbol{y}^m) - J_m\boldsymbol{y}^m)\,,$$

is a particular case of (44), where only one Newton iteration is performed with initial guess is $\boldsymbol{y}_{(0)} = \boldsymbol{y}^m$.

With (44), there are many possibilities to reduce the computational work (see, e.g., [11, 4, 10, 5]). To mention just a few ideas, (1) J may be evaluated only once per few iterations; (2) the linear system in (44) may be solved with a modest accuracy; (3) $J_{(s)}$ does not need to be available explicitly in a matrix form, to get $J_{(s)}\boldsymbol{y}$ for a given $\boldsymbol{y}$, we compute it approximately as

$$J_{(s)}\boldsymbol{y} \approx \frac{f(\boldsymbol{y}_{(s)} + \varepsilon\boldsymbol{y}) - f(\boldsymbol{y}_{(s)})}{\varepsilon}\,, \tag{45}$$

where ε is an appropriate small constant.

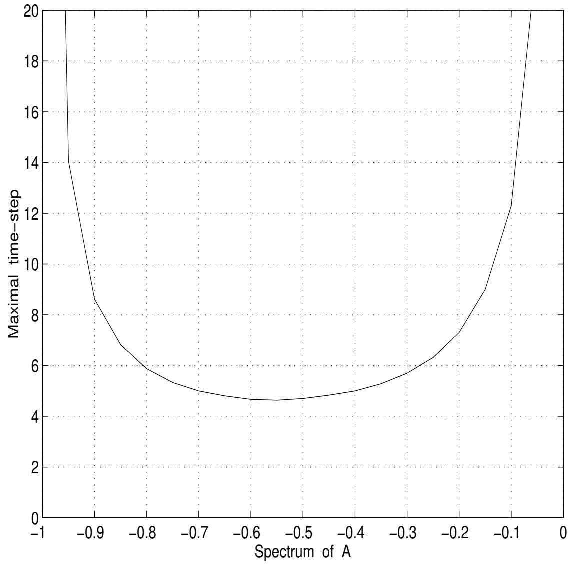


Figure 4: Maximal time step τ depending on $\lambda \in \{-0.99, -0.95 : 0.05 : -0.05, -0.01\}$ for the Adams (2) – BDF2 MR-PC scheme, $\theta = -1.00$

By doing only few linear iterations (to solve system in $\boldsymbol{x}_{(s)}$ in (44) each Newton step), we are in the GPC (or MR-PC) framework. The discussion on how this affects the convergence of the inexact Newton process is out of our scope here. In our limited experience obtained with MR-PC schemes it is better to perform also only a few inexact Newton iterations per time step. In many cases one inexact Newton iteration per time step seems to be sufficient.

Note also that the use of (45) for calculating Jacobian (the so-called "matrix–free approach") may lead to problems in Krylov subspace methods. The reason is that the Krylov subspace basis is generated for a perturbed operator J. The effects of different perturbations may sum up, so that the eventual Krylov subspace has a little to do with the exact J (for more details see [14, Chapter 4]). These problems are limited for the MR-PC schemes if k is taken small.

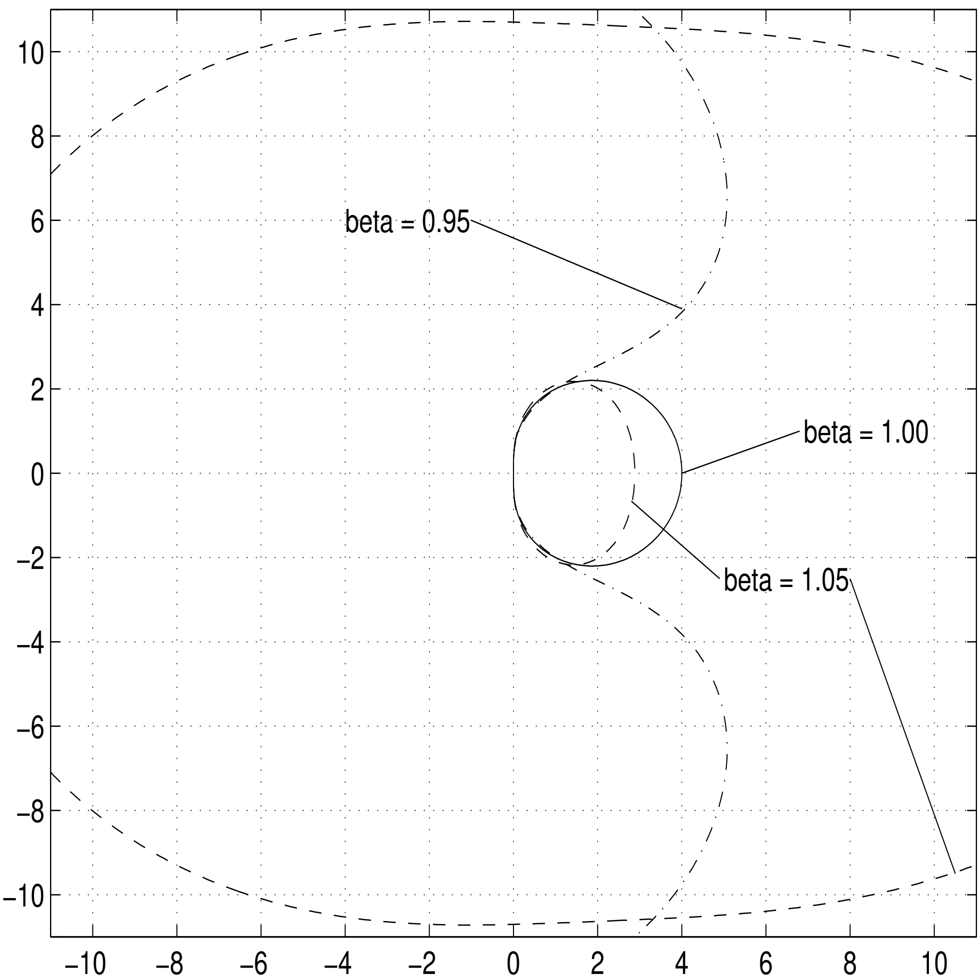


Figure 5: Stability regions $S \subset \mathbb{C}$ of the Adams (2) – BDF2 MR-PC scheme for $\theta = -1.00$ and $\beta \approx 1$: $\beta = 0.95$ — dashdotted line; $\beta = 1.00$ (BDF2 scheme) — solid line; $\beta = 1.05$ — dashed line

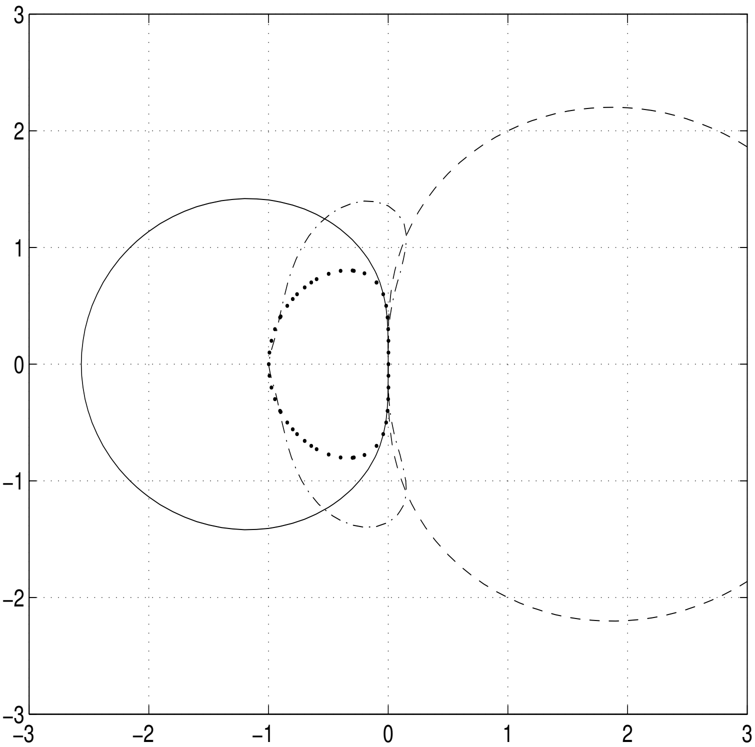


Figure 6: Stability regions $S \subset \mathbb{C}$ for the Adams (2) – BDF2 MR-PC scheme, $\theta = -1.00$ (inside the solid curve), Adams (2) (inside the dotted curve), the conventional Adams (2) Predictor – BDF2 Corrector Scheme (inside the dashdotted curve), and BDF2 (outside the dashed curve)

6 Numerical experiments

6.1 Simple test problems

We will present numerical results obtained with the Forward Euler – Backward Euler MR-PC (FE – BE) and Adams(2) – BDF2 (Ad2 – BDF2) MR-PC scheme for some simple test problems.

In order to have all desired information readily available we have taken the matrix A in (2) to be diagonal:

$$A = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)\,,$$

Ex. 1 $\quad A = \mathrm{Diag}\Big(-1 : h : -0.01\Big)\,,\ h = 0.99/(n-1), \qquad n = 100{,}200, 500\,,$

Ex. 1a $\quad A = \mathrm{Diag}\Big(-1 : h : -0.55, \ldots, -0.55\Big)\,,\ h = 0.99/(n-1), \qquad n = 100{,}200, 500\,,$

Ex. 2 $\quad A = \mathrm{Diag}\Big(-40, -9.5, -9, -8.5, -8 : h : -1\Big)\,, h = 7/(n-5)\,, \qquad n = 200\,,$

Ex. 3 $\quad A = \mathrm{Diag}\Big(-50, -19.5, -19, -18.5, -18 : h : -11\Big)\,, h = 7/(n-5)\,, \qquad n = 200\,,$

Ex. 4 $\quad A = \mathrm{Diag}\Big(-10, -0.9, -0.8, -0.7, -0.6 : h : -0.001\Big)\,,\ h = 0.599/(n-5)\,, \qquad n = 200\,.$

The initial condition for all these examples was $\boldsymbol{y}^0 = (1, \ldots, 1)^T \in \mathbb{R}^n$. In our MATLAB runs the problems have been solved for $t \in [0; 500]$. We have used slightly revised (to obtain harmonic Ritz values) MATLAB implementation of GMRES from [1].

The first example serves to illustrate our theoretical stability results. In Example 3, the matrix A is taken the same as in Example 2, but shifted by -10.

$n \backslash k$	FE	1	2	3	4	5
100	2.0 / 1.15 2.0 / 1.15	7.03 / 0.69 [1.3 / 3.75]	15.7 / 0.44 [2.0 / 1.63]	24.9 / 0.35 [2.5 / 1.63]	35.5 / 0.31 [7.5 / 0.65]	48.5 / 0.28 [12.0 / 0.48]
200	2.0 / 2.3	6.93 / 1.40	15.7 / 0.86	25.0 / 0.65	35.5 / 0.59	48.5 / 0.52
500	2.0 /5.75	6.87 / 3.50	15.7 / 2.15	25.0 / 1.60	36.0 / 1.32	48.5 / 1.24

Table 5: $\max\tau$ / number of megaflops, for FE – BE MR-PC(k) schemes (Example 1)

$n \backslash k$	Adams(2)	1	2	3	4	5
100	1.0 / 2.40 1.0 / 2.40	6.1 / 0.80 [0.3 / 9.02]	14.0 / 0.49 [1.5 / 2.34]	26.0 / 0.33 [2.7 / 1.60]	40.5 / 0.25 [5.0 / 1.01]	58 / 0.21 [10.5 / 0.56]
200	1.0 / 4.79	6.0 / 1.60	14.4 / 0.95	26.0 / 0.64	40.5 / 0.48	57.5 / 0.39
500	1.0 / 11.98	5.95 / 3.99	14.4 / 2.40	26.1 / 1.57	40.5 / 1.17	57.5 / 0.93

Table 6: $\max\tau$ / number of megaflops, for Adams(2) – BDF2 MR-PC(k) schemes (Example 1)

In our numerical tests we regard a scheme to be stable, if for all t_m in the integration segment $[0; T]$, the components y_j^m of the solution $\boldsymbol{y}^m$ are bounded in the following way:

$$|y_j^m| \leqslant \max\{1\,,\, C|(\boldsymbol{y}^m_{\mathrm{exact}})_j|\}\,, \tag{46}$$

where constant C has been set to 1 for these runs. This stability criterion is very crude, it only guarantees that the solution does not blow up. Nevertheless, it appears that (46) allows to

Ex.	$n \backslash k$	FE	1	2	3
1	100	2.0	7.03	15.7	24.9
1a	100	2.0	5.5	24.9	147
1	200	2.0	6.93	15.7	25.0
1a	200	2.0	5.5	24.2	137
1	500	2.0	6.87	15.7	25.0
1a	500	2.0	5.4	23.7	132

Table 7: $\max \tau$ for Forward Euler – Backward Euler MR-PC(k) scheme for Example 1 and Example 1a

Ex.$\backslash k$	FE	1	2	3	4	5	$\kappa(A)$
2	0.05 / 24.1	0.27 / 22.9	0.5 / 11.9	1.4 / 6.5	2.8 / 4.5	4.0 / 3.9	40
3	0.04 / 30.0	0.06 / 58.8	0.4 / 18.1	1.5 / 6.4	6 / 2.2	54 / 0.4	4.55
4	0.2 / 6.4	2.4 / 3.0	3.5 / 2.7	18 / 0.6	32 / 0.49	44 / 0.47	10^4

Table 8: $\max \tau$ / number of megaflops, for FE – BE MR-PC(k) schemes (Examples 2–4)

illustrate sharply the constant $\tau \leqslant 2/\max|\lambda_j|$ for the Forward Euler scheme. We want to check our theoretical restrictions $\tau \leqslant 7$ for the Forward Euler – Backward Euler MR-PC(1) scheme (see Table 1 and Figure 1) and $\tau \leqslant 4.66$ for the Adams2 – BDF2 MR-PC(1) scheme (see Table 3 and Figure 4).

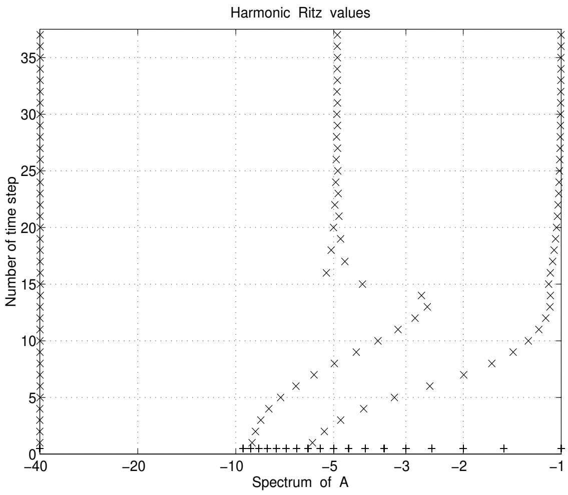
In Table 5 we have listed the maximum time steps (according to (46)) and required number of megaflops for Forward Euler – Backward Euler MR-PC scheme applied to the problem from Example 1. In the second column of Table 5, results for the Forward Euler scheme are shown. Here, we see a good agreement with the theoretically obtained upper bounds on τ for the MR-PC(1) scheme, namely $\tau \lesssim 7$. Note that we can not expect complete agreement because (i) the the actual values of θ are unknown a priori, and, (ii) stability condition (46) is appropriate for Forward Euler but may occur to be too strict for the MR-PC schemes where the temporary growth of the error does not necessarily mean unstability (see also section 7.1). The data given in the third row of Table 5, in square brackets, are obtained with the GPC(k) scheme with Chebyshev iterations (MATLAB implementation from [1]).

Table 6 contains the similar information as Table 5, but for the Adams2 – BDF2 MR-PC scheme.

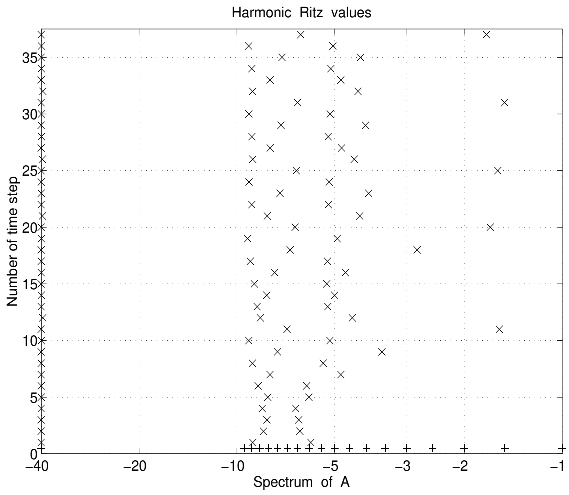
Earlier we noted that for the MR-PC(1) scheme, applied to a linear problem with a uniformly distributed spectrum over $[-1; 0)$, only the part of the spectrum $\lambda \lesssim -0.55$ affected the stability of the scheme. To verify this, we have modified Example 1 by substituting all λ's greater than -0.55 by -0.55. The results for this modified problem (Example 1a) are presented in Table 7. As we see, MR-PC(1) appears to be more sensitive to the spectrum of the problem than it may be expected from our theory: we observed restriction $\tau \lesssim 5.5$ instead of $\tau \lesssim 7$. The results for MR-PC(2) and MR-PC(3) are even better and this is not a surprise, since these schemes are more depth–sensitive with respect to the spectrum.

In Table 8, stability results for the Forward Euler – Backward Euler MR-PC(k) scheme, applied to Examples 2–4, are shown. The second column shows time steps observed for the Forward Euler scheme; they coincide with theoretical upper bound $2/\max|\lambda_j|$. In the last column of Table 8, the condition numbers $\kappa(A) = |\lambda_1|/|\lambda_n|$ are listed.

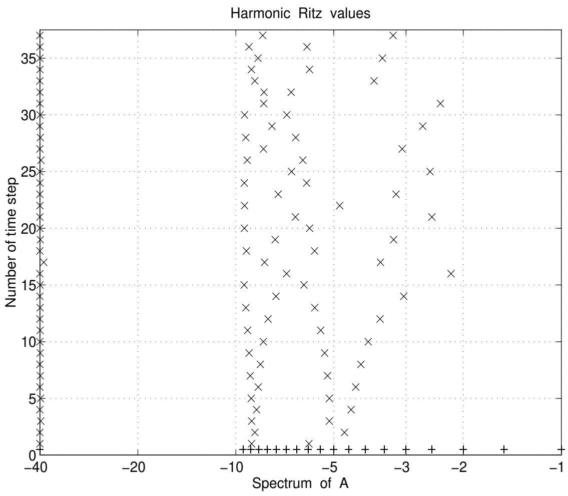
In Figures 7–9, the positions of the harmonic Ritz values are shown for Examples 2, 3, and 4, where n was taken 20. Here, we have used the MR-PC(3) scheme. In the plots, the eigenvalues of A are marked by "+" crosses along x–axis. For each Example, pictures for three values of τ are shown. As we see, for large time steps the θ's take oscillatory positions for successive time steps.



(a) stable $\tau = 0.25$

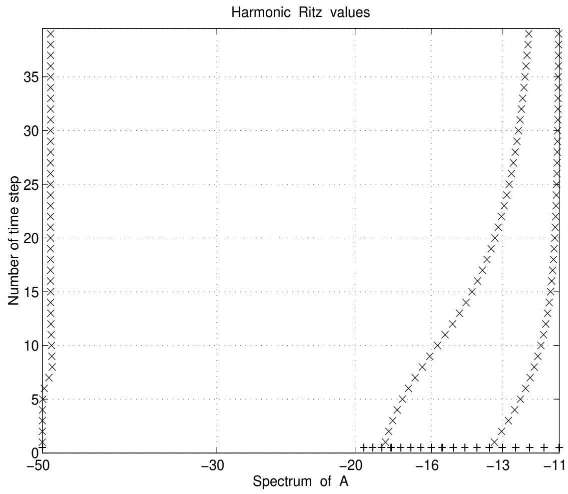


(b) stable $\tau = 0.9$

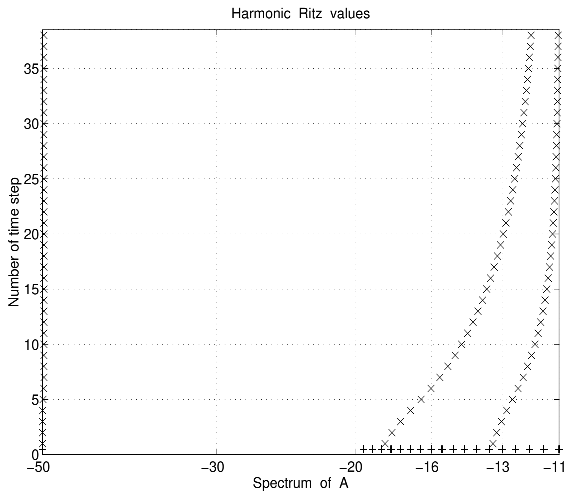


(c) unstable $\tau = 2.0$

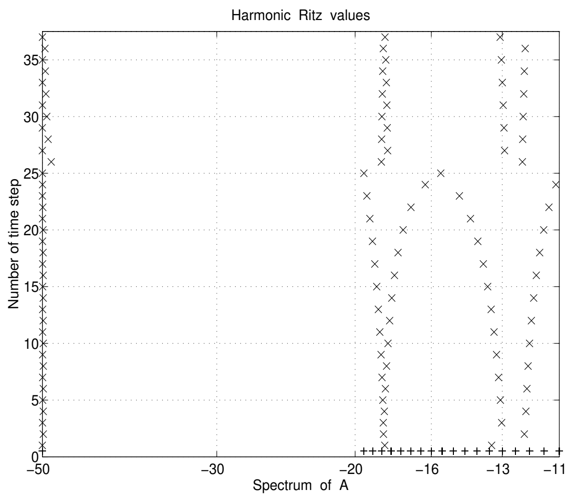
Figure 7: Example 2. FE – BE MR-PC(3) scheme



(a) stable $\tau = 0.1$

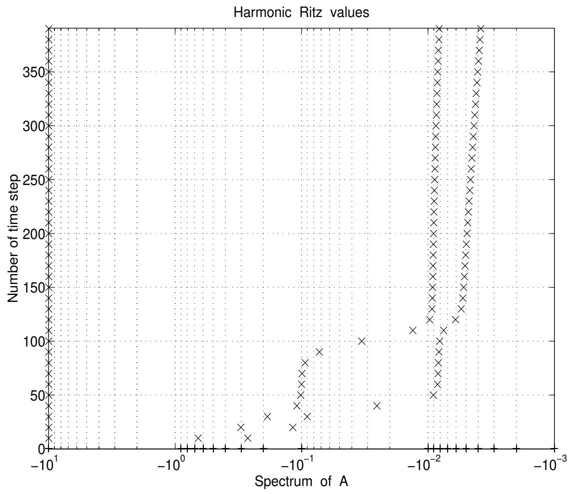


(b) stable $\tau = 0.2$

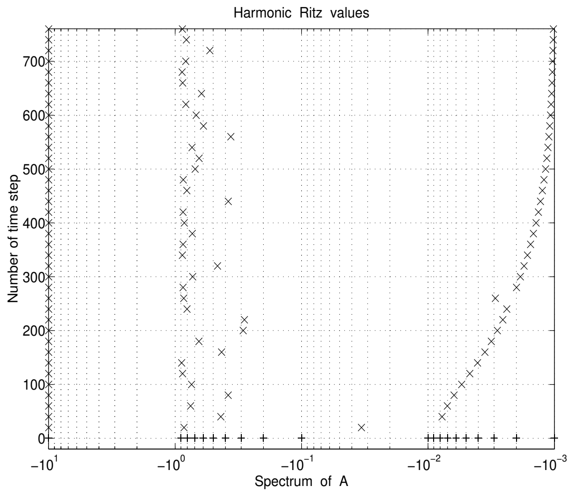


(c) unstable $\tau = 2.0$

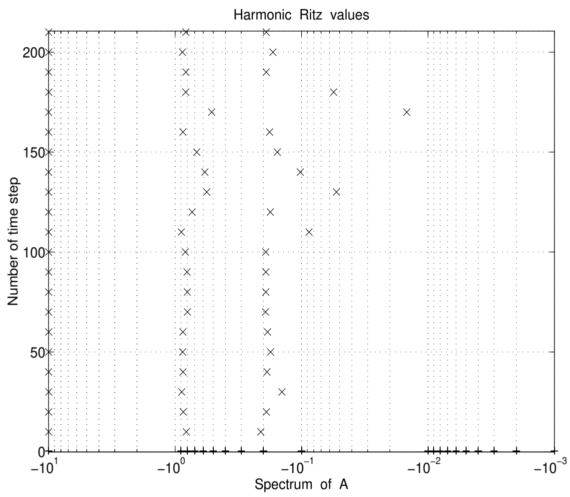
Figure 8: Example 3. FE – BE MR-PC(3) scheme



(a) stable $\tau = 1.0$



(b) stable $\tau = 4.5$



(c) stable $\tau = 7.0$

Figure 9: Example 4. FE – BE MR-PC(3) scheme

6.2 A more difficult test problem

We now consider the following autonomous system, adapted from [11]:

$$\begin{aligned} \frac{d\boldsymbol{y}}{dt} &= A\boldsymbol{y} + \gamma V\left[(V^*\boldsymbol{y}) \circ (V^*\boldsymbol{y})\right], \\ \boldsymbol{y}(0) &= \boldsymbol{y}^0, \end{aligned} \tag{47}$$

where $\circ$ means the elementwise (Hadamard) product and V is a matrix specially chosen in such a way that after the change of variables $\boldsymbol{z} = V^*\boldsymbol{y}$ equation in (47) becomes (componentwise):

$$\frac{dz_j}{dt} = \lambda_j z_j + \gamma z_j^2 . \tag{47$'$}$$

The system (47$'$) has an analytical solution

$$z_j(t) = \frac{-\lambda_j}{\gamma + C_j e^{-\lambda_j t}}, \qquad C_j = C_j(z_j(0)), \tag{48}$$

so we can easily compare our computed solution with the exact one. In [11], the matrix V was taken as

$$V = V^{-1} = I - \frac{2}{v^T u} u v^T ,$$

where u and v are given vectors, and A in (47) is taken $A = V\mathrm{Diag}(\lambda_j)V$. We proceed slightly differently. We take the matrix A as a block-diagonal one, with 2×2 blocks along the diagonal:

$$A = \mathrm{Diag}(A_1, \ldots, A_{n/2}), \qquad A_j = \begin{pmatrix} a_j & -b_j \\ b_j & a_j \end{pmatrix}, \tag{49}$$

and we assume n to be even. The unitary matrix V is taken as

$$V = \mathrm{Diag}(V_1, \ldots, V_{n/2}), \qquad V_j = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}. \tag{50}$$

V has the eigenvectors of A as its columns. The eigenvalues of A are

$$\lambda_{1,2} = a_1 \pm ib_1, \quad \lambda_{3,4} = a_2 \pm ib_2, \quad \ldots,$$

which can be used right away in (48). Note that A is normal, regardless of a_j and b_j. By varying γ, we control the degree of nonlinearity of the problem (47).

To give an impression of the solution of problem (47), we present in Figure 10 solutions for $n = 50$, $\boldsymbol{y}^0 = (1, \ldots, 1)^T$ and for a real spectrum $(-1, \ldots, -0.01)$ as well as for complex one $(-1 \pm i, \ldots, -0.01 \pm i)$.

In our further numerical tests, we have chosen $n = 500$, $\boldsymbol{y}^0 = (1, \ldots, 1)^T$, a_j have been 250 equidistant numbers from $[-1; -0.01]$ and b_j have been taken as

$$\begin{aligned} &(1) \quad b_j = 0 \quad \text{(real spectrum)}, \\ &(2) \quad b_j = 0.5 + 0.5\sin(12a_j) \quad \text{(complex spectrum)}. \end{aligned}$$

(The description of matrix V in (50) remains to be valid for the case $b_j = 0$.)

For these model problems we examine several time-stepping schemes with the crude stability criterion (46), where $C = 2$ and $\boldsymbol{y}^m_{\text{exact}} = V\boldsymbol{z}^m$ ($\boldsymbol{z}^m$ is the vector with components $z_j(t_m)$ cf. (48)). The integration segment was taken $t \in [0; 100]$.

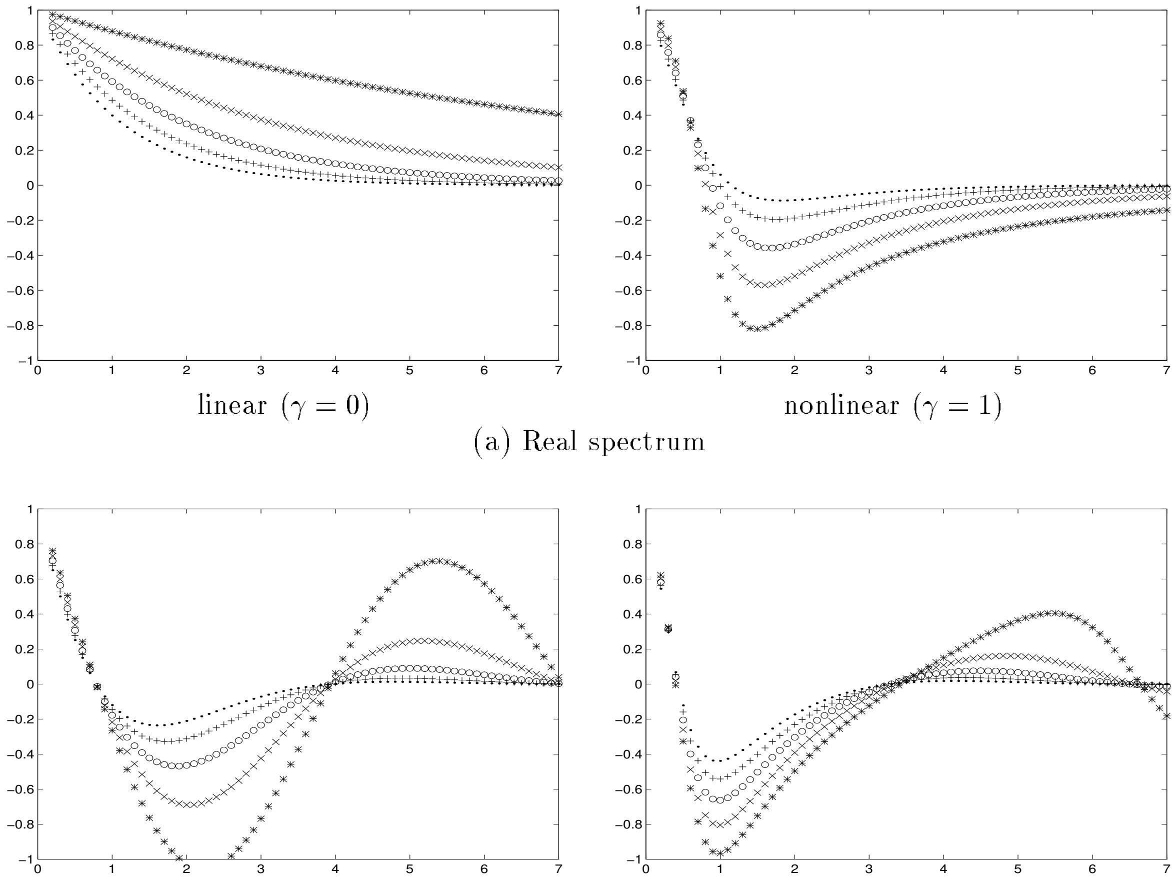
In Tables 9 and 10, we give the maximum τ (to satisfy (46)) for four schemes: Adams(2), the conventional PC scheme with Adams(2) – BDF2, the Adams2 – BDF2 MR-PC scheme, and the

BDF2 scheme with Newton's iterations. For the MR-PC scheme, we used an implementation with inexact Newton process as described in section 5. By the BDF2 scheme with Newton's iteration (the BDF2 – Newton scheme), we understand a scheme exactly the same as the MR-PC scheme but with the exact linear solver for the Jacobian each Newton iteration.

In fourth columns of Tables 9 and 10 we give a number of (inexact) Newton iterations.

We observe that with inexact k-step solvers in the MR-PC schemes, instead of possibly expensive exact solver in the BDF2 – Newton scheme, we arrive at almost the same timesteps for stability. This may lead to big savings in specific circumstances.

In Figure 11 we have plotted harmonic Ritz values θ, observed for the MR-PC(1) scheme applied to the linear "complex spectrum" problem, with respect to the spectrum of A. As we see, during the time-stepping process, θ's monotonically converge to a specific position.



linear ($\gamma = 0$) nonlinear ($\gamma = 1$)

(a) Real spectrum

linear ($\gamma = 0$) nonlinear ($\gamma = 1$)

(b) Complex spectrum

Figure 10: Solutions of the test problem. Components $\boldsymbol{y}_5$ $(\cdot)$, $\boldsymbol{y}_{15}$ $(+)$, $\boldsymbol{y}_{25}$ $(\circ)$, $\boldsymbol{y}_{35}$ $(\times)$, $\boldsymbol{y}_{45}$ $(*)$ for $n = 50$

Table 9: "Real spectrum" problem: max τ

γ	Adams(2)	PC		MR-PC(k)			BDF2
			Newt	$k=1$	$k=3$	$k=5$	
0.1	1.0	0.95	1	3.8	4.2	4.3	4.3
			2	9.0	8	9.3	9.3
1	0.25	0.45	1	0.8	0.75	0.75	0.75
			2	1.3	2.3	1.6	1.5

Table 10: "Complex spectrum" problem: max τ

γ	Adams(2)	PC		MR-PC(k)			BDF2
			Newt	$k=1$	$k=3$	$k=5$	
0	0.7	0.85	1	1.5	3	4.5	"∞"
0.1	0.7	0.85	1	1.5	2.4	2.5	2.7
			2	2.0	4.2	5.2	4.0
1	0.3	0.45	1	0.45	0.7	0.67	0.67
			2	0.63	1.4	0.8	0.9

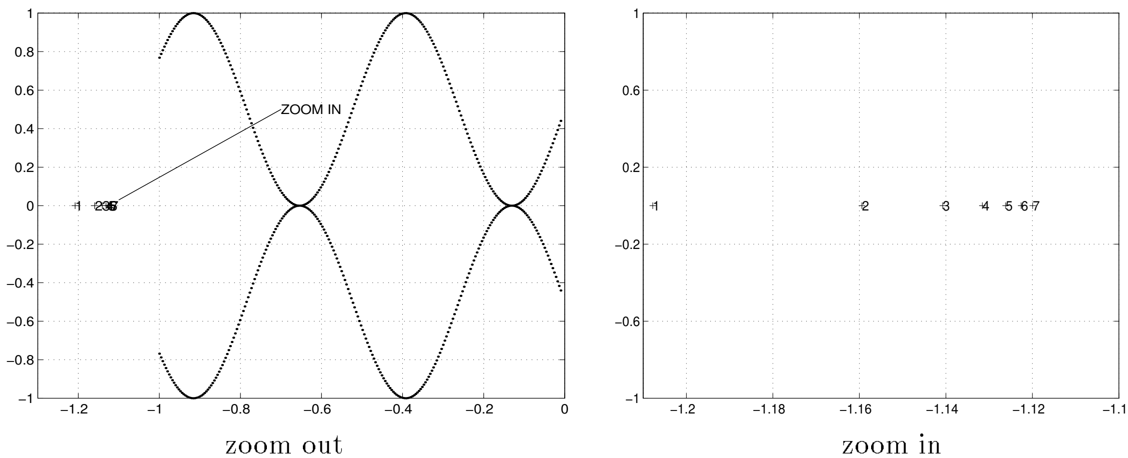


Figure 11: Harmonic Ritz value of the Adams(2) – BDF2 MR-PC(1) scheme w.r.t. λ_j (plotted by "·"). Numbers correspond to the 7 equidistributed moments of time $t \in [0; 100]$. Linear "complex spectrum" problem ($\gamma = 0$)

7 Automatic step size control for the MR-PC schemes

For MR-PC schemes it is possible to ensure stability by automatic step size control based on the spectral information delivered by GMRES. For this purpose we may use time step restriction given by Proposition 3 (condition (33)); as we noted earlier and as it will be confirmed practically in section 7.1, restrictions of Proposition 2 (namely, (28) and (29)) prove to be too pessimistic. We assume here that A has negative spectrum.

We consider automatic step size control only for Forward Euler – Backward Euler MR-PC scheme; the adaptation of the similar technique for checking stability of higher order MR-PC schemes is rather straightforward and left out of the scope of the paper for simplicity.

We will demonstrate how to implement step size control efficiently. Recall that for the Forward Euler – Backward Euler MR-PC scheme we have to perform a limited number of GMRES steps for the system

$$\tilde{A}\boldsymbol{y} = \boldsymbol{y}^m, \qquad \tilde{A} = I - \tau A\,. \tag{51}$$

With k steps of GMRES, an $n \times (k+1)$-matrix V_{k+1} with orthonormal columns $\boldsymbol{v}_i$ is computed and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ form basis for the Krylov subspace:

$$\text{span}\,(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k) = \mathcal{K}_k(\tilde{A}, \boldsymbol{r}_0) \equiv \text{span}\left(\boldsymbol{r}_0, \tilde{A}\boldsymbol{r}_0, \ldots, \tilde{A}^{k-1}\boldsymbol{r}_0\right)\,.$$

Simultaneously, a $(k{+}1)\times k$ upper–Hessenberg matrix $\underline{\tilde{H}}_k$ is delivered by GMRES and $V_{k+1}^*\tilde{A}V_k = \underline{\tilde{H}}_k$. Then, the iterative approximation to the solution $\boldsymbol{y}_{\text{im}}^{m+1}$ of (51) is

$$\boldsymbol{y}_k = \boldsymbol{y}_0 + V_k u_k\,,$$

where u_k is the least squares solution of the projected system

$$\underline{\tilde{H}}_k u_k = \|\boldsymbol{r}_0\| e_1, \qquad e_1 = (1, 0, \ldots, 0)^T \in \mathbb{R}^{k+1}\,. \tag{52}$$

For the Forward Euler – Backward Euler MR-PC scheme, one has that

$$\boldsymbol{y}_0 = \boldsymbol{y}_{\text{ex}}^{m+1} = (I + \tau A)\boldsymbol{y}^m, \qquad \boldsymbol{r}_0 = \tau^2 A^2 \boldsymbol{y}^m\,,$$

and it is easy to verify that by construction V_{k+1} does not depend on τ and

$$\underline{\tilde{H}}_k = \underline{I}_k - \tau \underline{H}_k, \qquad \underline{H}_k = V_{k+1}^* A V_k\,, \tag{53}$$

where $\underline{I}_k$ is $(k+1)\times k$ matrix with ones on its main diagonal and zeros elsewhere. Furthermore, the harmonic Ritz values $\tilde{\theta}_i$ of matrix $\tilde{A}$ are eigenvalues of the matrix [16]

$$\tilde{H}_k^{-*}\underline{\tilde{H}}_k^*\underline{\tilde{H}}_k\,, \tag{54}$$

where $\tilde{H}_k$ is a square matrix containing the first k rows of $\underline{\tilde{H}}_k$. We control the values

$$\eta_i \equiv 1 - \tilde{\theta}_i\,.$$

According to the Proposition 3, the MR-PC scheme is stable if $\eta_k = \max_i \eta_i \geqslant -7$.

Notice that η_k grows in modulus with the growth of τ. As we want τ to be not too small and not too large, we have to check whether

$$\eta_k \in [b_L, b_R]\,, \tag{55}$$

where left bound b_L and right bound b_R are some reasonable values, for example, we may take $b_L = -7$ and $b_R = -5.5$.

$\boldsymbol{y}^0$ is given
set b_L and b_R
`for` $m = 1, 2, \ldots$ `do`
    $\hat{\boldsymbol{r}}_0 = A^2 \boldsymbol{y}^m$
    NewStep = true
    GoodTau = false
    `repeat`
        `if` NewStep
            compute V_{k+1} and $\underline{H}_k$ such that
                $V_{k+1}^* A V_k = \underline{H}_k, \quad \operatorname{span} V_k = \mathcal{K}_k(A, \hat{\boldsymbol{r}}_0)$
            NewStep = false
        $\underline{\tilde{H}}_k = I_k - \tau \underline{H}_k$
        find eigenvalues $\tilde{\theta}_i$ of $\tilde{H}_k^{-*} \underline{\tilde{H}}_k^* \underline{\tilde{H}}_k$
        $\eta_k = \max_{i=1,\ldots,k}(1 - \tilde{\theta}_i)$
        `if` $(\eta_k \geqslant b_L)$`and`$(\eta_k \leqslant b_R)$
            GoodTau = true
        `if` $\eta_k > b_R$
            $c = b_R/\eta_k, \quad \tau = \tau c$ (increase)
        `if` $\eta_k < b_L$
            $c = \eta_k/b_L, \quad \tau = \tau/c$ (decrease)
    `until` GoodTau
    $\boldsymbol{y}_{\text{ex}}^{m+1} = (I + \tau A)\boldsymbol{y}^m$
    $\boldsymbol{r}_0 = \tau^2 \hat{\boldsymbol{r}}_0$
    find least-square solution u_k of (52)
    $\boldsymbol{y}^{m+1} = \boldsymbol{y}_{\text{ex}}^{m+1} + V_k u_k$

Figure 12: An example of Forward Euler – Backward Euler MR-PC scheme with automatic step size control for linear homogeneous problem (2) with negative spectrum

Assume that η_k fails to satisfy (55), namely, let $\eta_k < b_L$, so that τ has to be decreased. Then, the only thing we have to recalculate for the new step size is the matrix $\underline{\tilde{H}}_k$. We do not need to recalculate neither $\boldsymbol{r}_0$ (because only the direction of $\boldsymbol{r}_0$ is important) nor $\boldsymbol{y}_{\text{ex}}^{m+1}$ (it may be calculated thereafter when the choice of τ is done). For the new matrix $\underline{\tilde{H}}_k$, we again calculate the harmonic Ritz values and check condition (55). All this work has to be (re)done in k-dimensional space.

Since harmonic Ritz values depend on τ continuously, it is always possible to satisfy (55) for any appropriate b_L and b_R by varying τ. There are a lot of possibilities to implement the outlined above step size control algorithm. On Figure 12 we present a simple variant of the one which has been used in our numerical tests. We form Krylov basis with respect to A and recalculate $\underline{\tilde{H}}_k$ according to (53). In the algorithm we vary τ in such a way as the dependence η_k on τ were linear. Therefore robustness of this algorithm depends on b_L and b_R; as it occurs, it is important that b_L and b_R are not too close.

7.1 Numerical experiments with automatic step size control

$$
\begin{aligned}
& A = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)\,, \\
\text{Ex. 1} \quad & A = \mathrm{Diag}\Big(-1 : h : -0.01\Big)\,,\ h = 0.99/(n-1), \qquad n = 500\,, \\
\text{Ex. 2} \quad & A = \mathrm{Diag}\Big(-1 : h_1 : -0.9, -0.1 : h_2 : -0.01\Big)\,,\ h_1 = 0.1/(n-11), h_2 = 0.01,\ n = 500\,.
\end{aligned}
$$

The first Example is the same as one in section 6. The second example was chosen like hopefully difficult one: there is a gap in the spectrum, so that the ten smallest in modulus eigenvalues inhibit convergence of the harmonic Ritz values to the more important larger eigenvalues.

The time stepping process was performed for $t \in [0; 500]$. Each time step, apart from the described above step size control (Figure 12) done a priori, we verify, at each time step, the stability condition (46) (where C has been set to 1) a posteriori. In the experiments, parameters of the step size control algorithm b_L and b_R were set to -7 and -5.5 respectively; for this choice of bounds condition (46) was always satisfied. Nevertheless, (46) may occur to be too rough a criterion. (More specifically, with $b_L = -7$ and $b_R = -5.5$ we get value of τ very close to the maximal possible, so that the error does not increase but is not dumped either.) Therefore in such cases it is wise to decrease (in modulus) initial values of bounds, e.g., set $b_L = -5$, $b_R = -3.5$.

Results of the experiments with MR-PC(1) and MR-PC(3) are presented in Figures 13 and 14. Looking at the two first plots in Figure 13 we may compare values of τ delivered by the step size algorithm (namely, $\tau \approx 6.5$ for MR-PC(1) and $\tau \approx 22$ for MR-PC(3)) with maximal possible ones, according to the table 5 (p. 20) they are 6.87 for MR-PC(1) and 25.0 for MR-PC(3). The difference may be explained by some roughness of the used implementation of the step size algorithm.

Finally, we show results of some experiments made to understand how sharp the restriction (34) of Proposition 3 is actually. For this aim, we used modification of the step size control algorithm 12, where the bisection method was used to find such a value value of τ that η_k satisfies the condition (55) for any prescribed values of b_L and b_R. (To be more precise, the bisection algorithm is applied to the function

$$\eta_k = \eta_k(\tau) = \max\left\{1 - \tilde{\theta}_i \;\middle|\; \tilde{\theta}_i \text{ are harm. Ritz values of } \tilde{A}\right\}\,,$$

where θ_i are calculated according to the (54).) In Figure 15 the maximal (in absolute value) component of the numerical solution, namely $\max|y_i^m|$, is plotted versus respective moments of time t_m. Here, we present the plots for MR-PC(1) and MR-PC(3) schemes for different segments $[b_L; b_R] \ni \eta_k$. As we see, the bound $\eta_k \geqslant -7$ occurs to be sharp for MR-PC(1) but not for MR-PC(3). It is in a good agreement with the theory (see the Remark below the Proposition 3 on page 13).

8 Conclusions

To get an efficient compromise in time-step integration between many inexpensive explicit steps and a few expensive implicit steps, in this paper we are looking for possibilities provided by residually optimal Krylov subspace iterative methods. The key difference of our approach from the conventional implicit time-stepping combined with some iterative solution method is that, at each integration step, we perform a fixed (small) number of the iterations, so that we get something between explicit and implicit schemes, namely Generalized PC (GPC) schemes. Among all GPC schemes, we consider the case of use minimal residual iterations (like GMRES) and this leads to the MR-PC (Minimal Residual PC) schemes.

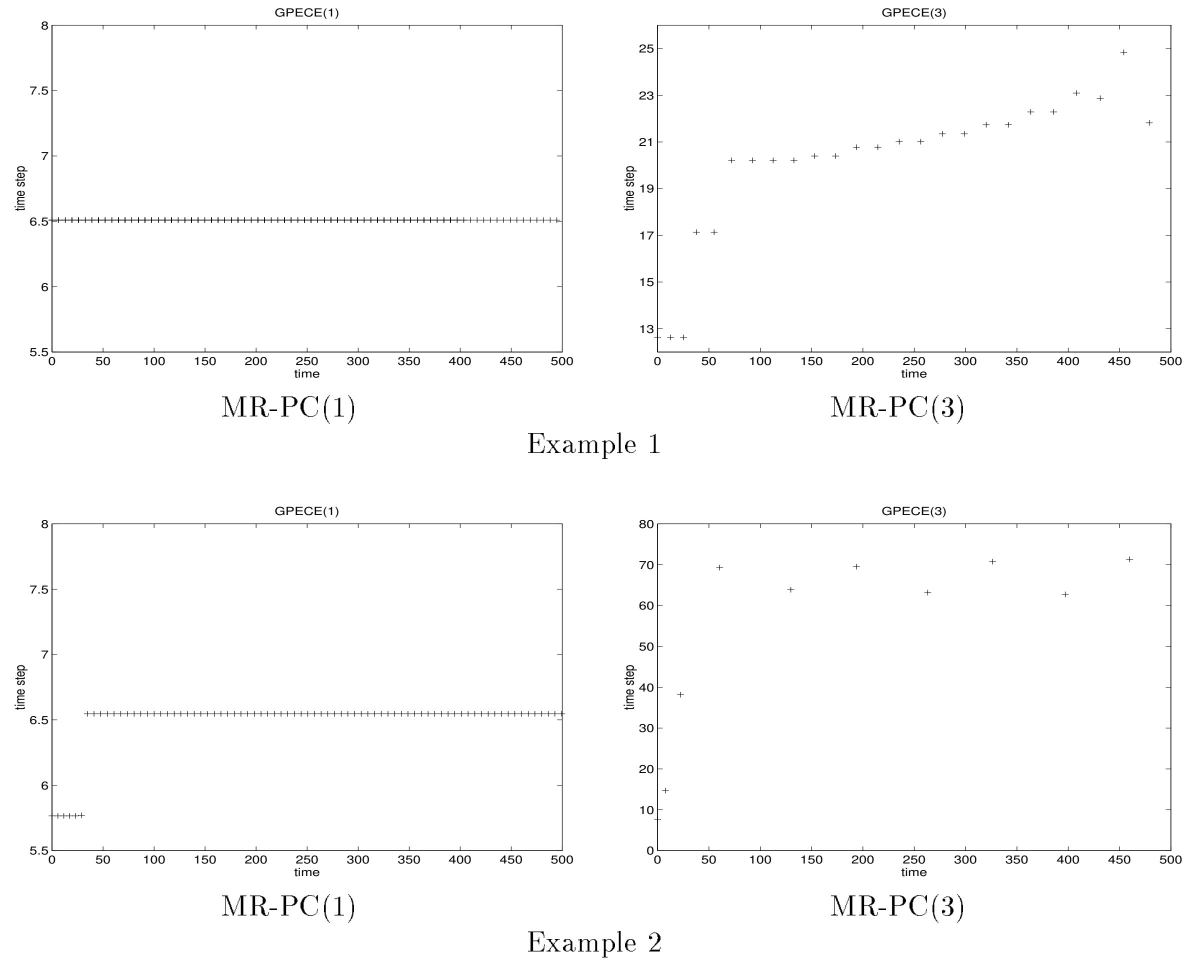
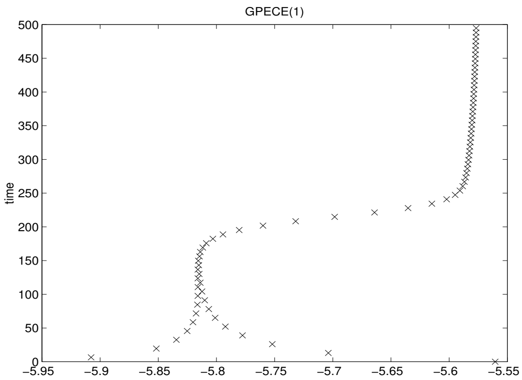


Figure 13: Time step size during the integration

The optimal residual reduction property of the iterations gives a significant gain in stability even for a very small number of iterative steps. The crucial point here is the adaptive nature of the iterations which makes residually optimal MR-PC scheme superior to GPC scheme with other type of Krylov iterations (like, for example, Chebyshev methods).

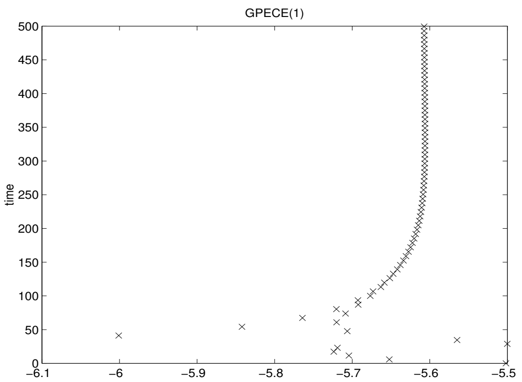
The attractive feature of Krylov subspaces methods in the MR-PC framework to provide with spectral information is exploited here to get an algorithm for automatic step size control with respect to the stability.

There are several open questions remained. The first one is about applicability of the MR-PC framework for nonlinear problems. Apparently, some of results of this paper may be generalized for the case of the dissipative system of ODEs (1). On other hand, some problems arise with the implementation of the step size control technique for nonlinear case. The first straightforward idea is to make the described above step size control like a preprocessing part of only the first Newton iteration each time step, but the validness of such an approach remains to be seen. Furthermore, the question arises of proper adaptation of the step size control strategy for the case of the complex spectrum of the Jacobian.



MR-PC(1) MR-PC(3)

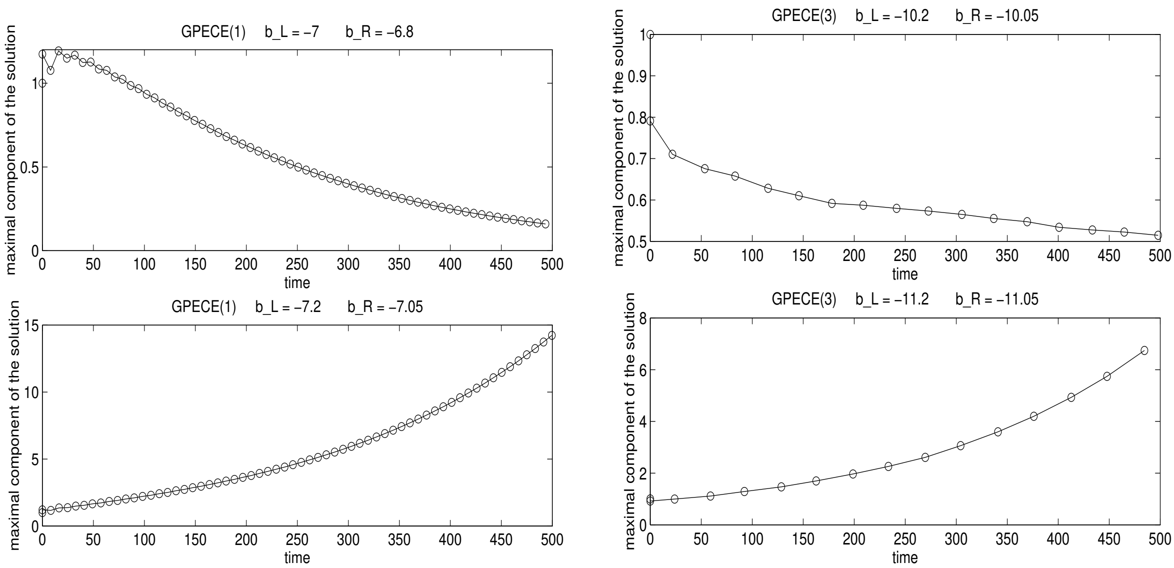
Example 1



MR-PC(1) MR-PC(3)

Example 2

Figure 14: The control value η_k during the integration



MR-PC(1) MR-PC(3)

Figure 15: The maximal component of the numerical solution during the integration in dependence on $\eta_k \in [b_L; b_R]$. (Growth means unstability)

Acknowledgments The first author is grateful to Jos van Dorsselaer for helpful discussions.

References

[1] R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. A. van der Vorst. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods.* SIAM, Philadelphia, PA, 1994.

[2] P. N. Brown. A local convergence theory for combined inexact–Newton / finite difference projections methods. *SIAM J. Numer. Anal.*, 24(2):407–434, 1987.

[3] P. N. Brown and Y. Saad. Convergence theory of nonlinear Newton–Krylov algorithms. *SIAM J. Optimization*, 4(2):297–330, 1994.

[4] T. F. Chan and K. R. Jackson. The use of iterative linear-equation solvers in codes for large systems of stiff IVPs for ODEs. *SIAM J. Sci. Stat. Comput.*, 7(2):378–417, 1986.

[5] R. Choquet and J. Erhel. Newton–GMRES algorithm applied to compressible flows. *Int. J. Numer. Meth. Fluids*, 23:177–190, 1996.

[6] A. T. Chronopoulos. Nonlinear CG-like iterative methods. *J. Comput. Appl. Math.*, 40:73–89, 1992.

[7] S. C. Eisenstat and H. F. Walker. Globally convergent inexact Newton methods. *SIAM J. Optimization*, 4(2):393–422, 1994.

[8] D. R. Fokkema, G. L. G. Sleijpen, and H. A. van der Vorst. Accelerated inexact Newton schemes for large systems of nonlinear equations. Technical Report 918, Department of Mathematics, Utrecht University, July 1995. To appear in SIAM J. Sci. Comput.

[9] R. W. Freund and N. M. Nachtigal. QMR: a quasi–minimal residual method for non–Hermitian linear systems. *Numer. Math.*, 60:315–339, 1991.

[10] E. Gallopoulos and Y. Saad. Efficient solution of parabolic equations by Krylov approximation methods. *SIAM J. Sci. Stat. Comput.*, 13(5):1236–1264, 1992.

[11] C. W. Gear and Y. Saad. Iterative solution of linear equations in ODE codes. *SIAM J. Sci. Stat. Comput.*, 4(4):583–601, 1983.

[12] E. Hairer, S. P. Nørsett, and G. Wanner. *Solving Odinary Differential Equations I. Nonstiff Problems.* Springer Series in Computational Mathematics 8. Springer–Verlag, 1987.

[13] E. Hairer and G. Wanner. *Solving Odinary Differential Equations II. Stiff and Differential–Algebraic Problems.* Springer Series in Computational Mathematics 14. Springer–Verlag, 1991.

[14] M. Heyouni. *Methode de Hessenberg generalisee et applications.* PhD thesis, Universite de Lille, 1996.

[15] M. Hochbruck and C. Lubich. On Krylov subspace approximations to the matrix exponential operator. Technical report, Department of Mathematics, University of Tübingen, January 1995. (To appear in SIAM J. Numer. Anal., 34, 1997).

[16] C. C. Paige, B. N. Parlett, and H. A. van der Vorst. Approximate solutions and eigenvalue bounds from Krylov subspaces. *Numer. Lin. Alg. Appl.*, 2(2):115–133, 1995.

[17] M. J. D. Powell. *Approximation theory and methods.* Cambridge U.P., 1981.

[18] Y. Saad. A flexible inner–outer preconditioned GMRES algorithm. *SIAM J. Sci. Comput.*, 14:461–469, 1993.

[19] Y. Saad and M. H. Schultz. GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.*, 7(3):856–869, 1986.

[20] B. A. Schmitt and R. Weiner. Krylov–W–methods for large stiff equations. In J. G. Verwer and J. M. Sanz-Serna, editors, *Workshop Innovative Time Integrators*, number Note NM-N9602, pages 70–72, Amsterdam, July 1996. CWI, Dept. of Numer. Math.

[21] G. L. G. Sleijpen, D. R. Fokkema, and H. A. van der Vorst. BiCGStab(ℓ) and other hybrid BiCG methods. *Numerical Algorithms*, 7:75–105, 1994.

[22] H. A. van der Vorst. An iterative solution method for solving $f(A)x = b$, using using Krylov subspace information obtained for the symmetric positive definite matrix A. *J. Comput. Appl. Math.*, 18:249–263, 1987.

[23] H. A. van der Vorst. BiCGSTAB: a fast and smoothly converging variant of BiCG for the solution of nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.*, 13(2):631–644, 1992.

[24] H. A. van der Vorst and C. Vuik. GMRESR: a family of nested GMRES methods. *Numer. Lin. Alg. Appl.*, 1:369–386, 1994.

[25] V. T. Zhukov. Numerical experiments solving the heat equation by the local iteration method. Technical Report 97, Keldysh Institute for Applied Mathematics, USSR Acad. Sci., 1984. In Russian.

[26] V. T. Zhukov. *Explicitly Iterative Difference Schemes for Parabolic Equations*, volume 4 of *Mathematical Modelling of Physical Processes*, pages 40–46. VNIIEF, 1993. In Russian.

Appendix 1

We give the remaining part of the proof of Proposition 2. We prove that restriction (33) guarantees that

$$\mu \leqslant Q(\mu) \leqslant 2-\mu\,, \qquad \mu \in [\mu_1;\mu_n]\,, \tag{56}$$

where $Q(\mu)$ is from (31),(32). Evidently, according to the (32), $Q(\mu)$ has to be of the form

$$Q(\mu) = \mu^2 \Pi_{i=1}^k \frac{\mu-\eta_i}{1-\eta_i}\,. \tag{57}$$

Consider first the right inequality in (56). Since $Q(\mu)$ attains its extremal values at the ends of the alternance segment $[\mu_1;\mu_n]$, and is positive in $(\eta_k;0)$, we may estimate

$$Q(\mu) \leqslant \max_{\mu\in[\mu_1;\mu_n]} |Q(\mu)| = Q(\mu_n) \leqslant \max_{\mu\in[\eta_k;0]} Q(\mu)\,.$$

Then,

$$0 < \left|\frac{\mu-\eta_i}{1-\eta_i}\right| = \frac{\mu-\eta_i}{1-\eta_i} < 1, \qquad i=1,\dots,k-1, \qquad \mu\in[\eta_k;0]\,,$$

and hence

$$Q(\mu) \leqslant s(\mu) = \frac{\mu-\eta_k}{1-\eta_k}\mu^2, \qquad \mu\in[\eta_k;0]\,. \tag{58}$$

Thus, inequality

$$\frac{\mu-\eta_k}{1-\eta_k}\mu^2 \leqslant 2-\mu, \qquad \mu\in[\eta_k;0]$$

holds iff

$$\eta_k \geqslant -7\,, \tag{59}$$

and it is sufficient for the right inequality in (56) to be true. It remains to prove the left inequality in (56) under restriction (59). With $s(\mu)$ from (58), it is possible to estimate

$$\max_{\mu\in[\mu_1;\mu_n]} |Q(\mu)| = Q(\mu_n) \leqslant s(\mu_n) \leqslant \max_{\mu\in[\eta_k;0]} s(\mu) = s\left(\frac{2}{3}\eta_k\right) \leqslant \eta_k\,. \tag{60}$$

The rightmost inequality in (60) is equivalent to

$$4\eta_k^2 + 27\eta_k - 27 \leqslant 0 \quad (\eta_k<0) \qquad \Rightarrow$$

$$\eta_k \geqslant \frac{-27-\sqrt{1161}}{8} \approx -7.63\dots\,. \tag{61}$$

On the other hand, (60) implies that

$$Q(\mu) \geqslant -\max_{\mu\in[\mu_1;\mu_n]} |Q(\mu)| \geqslant \eta_k, \qquad \mu\in[\mu_1;\eta_k]\,,$$

$$\text{or} \qquad Q(\mu) \geqslant \mu, \qquad \mu\in[\mu_1;\eta_k]\,.$$

The last inequality is fulfilled for the whole segment $\mu\in[\mu_1;\mu_n]$ because $Q(\mu)>0$ for $\mu\in(\eta_k;\mu_n]$.

Thus, condition (61) guarantees of the left inequality in (56), whereas the stricter condition (59) guarantees the whole (56).