

Article

Implicit A-Stable Peer Triplets for ODE Constrained Optimal Control Problems

Jens Lang ^{1,*}  and Bernhard A. Schmitt ² 
¹ Department of Mathematics, Technical University of Darmstadt, Dolivostraße 15, 64293 Darmstadt, Germany

² Department of Mathematics, Philipps-Universität Marburg, Hans-Meerwein-Straße 6, 35043 Marburg, Germany

* Correspondence: lang@mathematik.tu-darmstadt.de

Abstract: This paper is concerned with the construction and convergence analysis of novel implicit Peer triplets of two-step nature with four stages for nonlinear ODE constrained optimal control problems. We combine the property of superconvergence of some standard Peer method for inner grid points with carefully designed starting and end methods to achieve order four for the state variables and order three for the adjoint variables in a first-discretize-then-optimize approach together with A-stability. The notion triplets emphasize that these three different Peer methods have to satisfy additional matching conditions. Four such Peer triplets of practical interest are constructed. In addition, as a benchmark method, the well-known backward differentiation formula BDF4, which is only $A(73.35^\circ)$ -stable, is extended to a special Peer triplet to supply an adjoint consistent method of higher order and BDF type with equidistant nodes. Within the class of Peer triplets, we found a diagonally implicit $A(84^\circ)$ -stable method with nodes symmetric in $[0, 1]$ to a common center that performs equally well. Numerical tests with four well established optimal control problems confirm the theoretical findings also concerning A-stability.

Keywords: implicit Peer two-step methods; BDF-methods; nonlinear optimal control; first-discretize-then-optimize; discrete adjoints

MSC: 34H05; 49J15; 65L05; 65L06



Citation: Lang, J.; Schmitt, B.A. Implicit A-Stable Peer Triplets for ODE Constrained Optimal Control Problems. *Algorithms* **2022**, *15*, 310. <https://doi.org/10.3390/a15090310>

Academic Editor: Devendra Kumar

Received: 15 August 2022

Accepted: 26 August 2022

Published: 29 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The design of efficient time integrators for the numerical solution of optimal control problems constrained by systems of ordinary differential equations (ODEs) is still an active research field. Such systems typically arise from semi-discretized partial differential equations describing, e.g., the dynamics of heat and mass transfer or fluid flow in complex physical systems. Symplectic one-step Runge–Kutta methods [1,2] exploit the Hamiltonian structure of the first-order optimality system—the necessary conditions to find an optimizer—and automatically yield a consistent approximation of the adjoint equations, which can be used to compute the gradient of the objective function. The first-order symplectic Euler, second-order Störmer–Verlet and higher-order Gauss methods are prominent representatives of this class, which are all implicit for general Hamiltonian systems, see the monograph [3]. Moreover, compositions of basic integrators with different steps sizes and splitting methods have been investigated. Generalized partitioned Runge–Kutta methods which allow one to compute exact gradients with respect to the initial condition are studied in [4]. To avoid the solution of large systems of nonlinear equations, semi-explicit W-methods [5] and stabilized explicit Runge–Kutta–Chebyshev methods [6] have been proposed, too. However, as all one-step methods, also symplectic Runge–Kutta schemes join the structural suffering of order reductions, which may lead to poor results in their application, e.g., to boundary control problems such as external cooling and heating in a manufacturing process; see [7,8] for a detailed study of this behaviour.

In contrast, multistep methods including Peer two-step methods avoid order reductions and allow a simple implementation [9,10]. However, the discrete adjoint schemes of linear multistep methods are in general not consistent or show a significant decrease of their approximation order [11,12]. Recently, we have developed implicit Peer two-step methods [13] with three stages to solve ODE constrained optimal control problems of the form

$$\text{minimize } C(y(T)) \quad (1)$$

$$\text{subject to } y'(t) = f(y(t), u(t)), \quad u(t) \in U_{ad}, \quad t \in (0, T], \quad (2)$$

$$y(0) = y_0, \quad (3)$$

with the state $y(t) \in \mathbb{R}^m$, the control $u(t) \in \mathbb{R}^d$, $f: \mathbb{R}^m \times \mathbb{R}^d \mapsto \mathbb{R}^m$, the objective function $C: \mathbb{R}^m \mapsto \mathbb{R}$, where the set of admissible controls $U_{ad} \subset \mathbb{R}^d$ is closed and convex. Introducing for any $u \in U_{ad}$ the normal cone mapping

$$N_U(u) = \{w \in \mathbb{R}^d : w^T(v - u) \leq 0 \text{ for all } v \in U_{ad}\}, \quad (4)$$

the first-order Karush–Kuhn–Tucker (KKT) optimality system [14,15] reads

$$y'(t) = f(y(t), u(t)), \quad t \in (0, T], \quad y(0) = y_0, \quad (5)$$

$$p'(t) = -\nabla_y f(y(t), u(t))^T p(t), \quad t \in [0, T], \quad p(T) = \nabla_y C(y(T))^T, \quad (6)$$

$$-\nabla_u f(y(t), u(t))^T p(t) \in N_U(u(t)), \quad t \in [0, T]. \quad (7)$$

In this paper, we assume the existence of a unique local solution (y^*, p^*, u^*) of the KKT system with sufficient regularity properties to justify the use of higher order Peer triplets, see, e.g., the smoothness assumption in Section 2 of [14].

Remark 1. The objective function $C(y(T))$ in (1) is specified in the so-called Mayer form using terminal solution values only. Terms given in the Lagrange form

$$C_l(y, u) := \int_0^T l(y(t), u(t)) dt \quad (8)$$

can be equivalently reduced to the Mayer form by introducing an additional state y_{m+1} , the new state vector $\tilde{y} := (y_1, \dots, y_m, y_{m+1})^T$, and an additional differential equation $y'_{m+1}(t) = l(y(t), u(t))$ with initial values $y_{m+1}(0) = 0$. Then, the Lagrange term (8) simply reduces to $y_{m+1}(T)$.

Following a first-discretize-and-then-optimize approach, we apply an s -stage implicit Peer two-step method to (2)–(3) with approximations $Y_{ni} \approx y(t_n + c_i h)$ and $U_{ni} \approx u(t_n + c_i h)$, $i = 1, \dots, s$, on an equi-spaced time grid $\{t_0, \dots, t_{N+1}\} \subset [0, T]$ with step size $h = (T - t_0)/(N + 1)$ and nodes c_1, \dots, c_s , which are fixed for all time steps, to get the discrete constraint nonlinear optimal control problem

$$\text{minimize } C(y_h(T)) \quad (9)$$

$$\text{subject to } A_0 Y_0 = a \otimes y_0 + hb \otimes f(y_0, u_0) + hK_0 F(Y_0, U_0), \quad (10)$$

$$A_n Y_n = B_n Y_{n-1} + hK_n F(Y_n, U_n), \quad n = 1, \dots, N, \quad (11)$$

$$y_h(T) = (w^T \otimes I) Y_N, \quad (12)$$

with long vectors $Y_n = (Y_{ni})_{i=1}^s \in \mathbb{R}^{sm}$, $U_n = (U_{ni})_{i=1}^s \in \mathbb{R}^{sd}$, and $F(Y_n, U_n) = (f(Y_{ni}, U_{ni}))_{i=1}^s$. Further, $y_h(T) \approx y(T)$, $u_0 \approx u(0)$, and $a, b, w \in \mathbb{R}^s$ are additional parameter vectors at both interval ends, $A_n, B_n, K_n \in \mathbb{R}^{s \times s}$, and $I \in \mathbb{R}^{m \times m}$ being the identity matrix. We will use the same symbol for a coefficient matrix like A and its Kronecker product $A \otimes I$ as a mapping from the space \mathbb{R}^{sm} to itself. In contrast to one-step methods, Peer two-step methods com-

pute Y_n from the previous stage vector Y_{n-1} . Hence, also a starting method, given in (10), for the first time interval $[t_0, t_1]$ is required. On each subinterval, Peer methods may be defined by three coefficient matrices A_n, B_n, K_n , where A_n and K_n are assumed to be nonsingular. For practical reasons, this general version will not be used, the coefficients in the inner grid points will belong to a fixed Peer method $(A_n, B_n, K_n) \equiv (A, B, K)$, $n = 1, \dots, N - 1$. The last forward step has the same form as the standard steps but uses exceptional coefficients (A_N, B_N, K_N) to allow for a better approximation of the end conditions.

The KKT conditions (5)–(7) for ODE-constrained optimal control problems on a time interval $[0, T]$ lead to a boundary value problem for a system of two differential equations, see Section 2.1. The first one corresponds to the original *forward* ODE for the state solution $y(t)$ and the second one is a linear, *adjoint* ODE for a Lagrange multiplier $p(t)$. It is well known that numerical methods for such problems may have to satisfy additional order conditions for the adjoint equation [2,5,14,16–18]. While these additional conditions are rather mild for one-step methods they may lead to severe restrictions for other types of methods such as multistep and Peer methods, especially at the right-hand boundary at the end point T . Here, the order for the approximation of the adjoint equation may be limited to one.

For Peer methods, this question was discussed first in [19] and the adjoint boundary condition at T was identified as the most critical point. In a more recent article [13], these bottlenecks could be circumvented by two measures. First, equivalent formulations of the forward method are not equivalent for the adjoint formulation and using a redundant formulation of Peer methods with three coefficient matrices (A, B, K) adds additional free parameters. The second measure is to consider different methods for the first and last time interval. Hence, instead of one single Peer method (which will be called the standard method) we discuss *triplets* of Peer methods consisting of a common standard method (A, B, K) for all subintervals of the grid from the interior of $[0, T]$, plus a starting method (A_0, K_0) for the first subinterval and an end method (A_N, B_N, K_N) for the last one. These two boundary methods may have lower order than the standard method since they are used only once.

The present work extends the results from [13] which considered methods with $s = 3$ stages only, in two ways. We will now concentrate on methods with four stages and better stability properties such as A-stability. The purpose of an accurate solution of the adjoint equation increases the number of conditions on the parameters of the method. Requiring high order of convergence s for the state variable $y(t)$ and order $s - 1$ for the adjoint variable $p(t)$ – which we combine to the pair $(s, s - 1)$ from now on – a variant of the method BDF3 was identified in [13] as the most attractive standard method. However, this method is not A-stable, with a stability angle of $\alpha = 86^\circ$. In order to obtain A-stability, we will reduce the required orders by one. Still, we will show that convergence with the orders $(s, s - 1)$ can be regained by a superconvergence property.

The paper is organized as follows: In Section 2.1, the boundary value problem arising from the KKT system by eliminating the control and its discretization by means of discrete adjoint Peer two-step triplets are formulated. An extensive error analysis concentrating on the superconvergence effect is presented in Section 2.2. The restrictions imposed by the starting and end method on the standard Peer two-step method is studied in Section 2.3. The following Section 2.4 describes the actual construction principles of Peer triplets. Numerical tests are undertaken in Section 3. The paper concludes with a discussion in Section 4.

2. Materials and Methods

2.1. The Boundary Value Problem

Following the usual Lagrangian approach applied in [13], the first order discrete optimality conditions now consist of the forward Equations (10)–(12), the discrete adjoint equations, acting backwards in time,

$$A_N^T P_N = w \otimes p_h(T) + h \nabla_Y F(Y_N, U_N)^T K_N^T P_N, \quad (13)$$

$$A_n^T P_n = B_{n+1}^T P_{n+1} + h \nabla_Y F(Y_n, U_n)^T K_n^T P_n, \quad N-1 \geq n \geq 0, \quad (14)$$

and the control conditions

$$-h \nabla_U F(Y_n, U_n)^T K_n^T P_n \in N_{U^s}(U_n), \quad 0 \leq n \leq N, \quad (15)$$

$$-h \nabla_{u_0} f(y_0, u_0)^T (b^T \otimes I) P_0 \in N_U(u_0). \quad (16)$$

Here, $p_h(T) = \nabla_y C(y_h(T))^T$ and the Jacobians of F are block diagonal matrices $\nabla_Y F(Y_n, U_n) = \text{diag}_i(\nabla_{Y_{ni}} f(Y_{ni}, U_{ni}))$ and $\nabla_U F(Y_n, U_n) = \text{diag}_i(\nabla_{U_{ni}} f(Y_{ni}, U_{ni}))$. The generalized normal cone mapping $N_{U^s}(U_n)$ is defined by

$$N_{U^s}(u) = \left\{ w \in \mathbb{R}^{sd} : w^T(v - u) \leq 0 \text{ for all } v \in U_{ad}^s \subset \mathbb{R}^{sd} \right\}. \quad (17)$$

The discrete KKT conditions (10)–(16) should be good approximations to the continuous ones (5)–(7). In what follows, we assume sufficient smoothness of the optimal control problem such that a local solution (y^*, u^*, p^*) of the KKT system (1)–(3) exists. Furthermore, let the Hamiltonian $H(y, u, p) := p^T f(y, u)$ satisfy a coercivity assumption, which is a strong form of a second-order condition. Then the first-order optimality conditions are also sufficient [14]. If (y, p) is sufficiently close to (y^*, p^*) , the control uniqueness property introduced in [14] yields the existence of a locally unique minimizer $u = u(y, p)$ of the Hamiltonian over all $u \in U_{ad}$. Substituting u in terms of (y, p) in (5)–(6), gives then the two-point boundary value problem

$$y'(t) = g(y(t), p(t)), \quad y(0) = y_0, \quad (18)$$

$$p'(t) = \phi(y(t), p(t)), \quad p(T) = \nabla_y C(y(T))^T, \quad (19)$$

with the source functions defined by

$$g(y, p) := f(y, u(y, p)), \quad \phi(y, p) := -\nabla_y f(y, u(y, p))^T p. \quad (20)$$

The same arguments apply to the discrete first-order optimality system (10)–(16). Substituting the discrete controls $U_n = U_n(Y_n, P_n)$ in terms of (Y_n, P_n) and defining

$$\Phi(Y_n, K_n^T P_n) := \left(\phi(Y_{ni}, (K_n^T P_n)_i) \right)_{i=1}^s, \quad G(Y_n, P_n) := (g(Y_{ni}, P_{ni}))_{i=1}^s, \quad (21)$$

the approximations for the forward and adjoint differential equations read in a compact form

$$A_0 Y_0 = a \otimes y_0 + h b \otimes g(y_0, p_h(0)) + h K_0 G(Y_0, P_0), \quad (22)$$

$$A_n Y_n = B_n Y_{n-1} + h K_n G(Y_n, P_n), \quad 1 \leq n \leq N, \quad (23)$$

$$y_h(T) = (w^T \otimes I) Y_N, \quad (24)$$

$$p_h(0) = (v^T \otimes I) P_0, \quad (25)$$

$$A_n^T P_n = B_{n+1}^T P_{n+1} - h \Phi(Y_n, K_n^T P_n), \quad 0 \leq n \leq N-1, \quad (26)$$

$$A_N^T P_N = w \otimes p_h(T) - h \Phi(Y_N, K_N^T P_N), \quad n = N. \quad (27)$$

Here, the value of $p_h(0)$ is determined by an interpolant $p_h(0) = (v^T \otimes I) P_0 \approx p(0)$ with $v \in \mathbb{R}^s$ of appropriate order. In a next step, these discrete equations are now treated as

a discretization of the two-point boundary value problem (18)–(19). We will derive order conditions and give bounds for the global error.

2.2. Error Analysis

2.2.1. Order Conditions

The local error of the standard Peer method and the starting method is easily analyzed by Taylor expansion of the stage residuals, if the exact ODE solutions are used as stages. Hence, defining $\mathbf{y}_n^{(k)}(h\mathbf{c}) := (y^{(k)}(t_n + hc_i))_{i=1}^s$, $k = 0, 1$, for the forward Peer method, where $\mathbf{c} = (c_1, \dots, c_s)^\top$, local order q_1 means that

$$A_n \mathbf{y}_n(h\mathbf{c}) - B_n \mathbf{y}_{n-1}(h\mathbf{c}) - hK_n \mathbf{y}'(h\mathbf{c}) = O(h^{q_1}). \quad (28)$$

In all steps of the Peer triplet, requiring local order q_1 for the state variable and order q_2 for the adjoint solution leads to the following algebraic conditions from [13]. These conditions depend on the Vandermonde matrix $V_q = (\mathbf{1}, \mathbf{c}, \mathbf{c}^2, \dots, \mathbf{c}^{q-1}) \in \mathbb{R}^{s \times q}$, the Pascal matrix $\mathcal{P}_q = \left(\binom{j-1}{i-1} \right)_{i,j=1}^q$ and the nilpotent matrix $\tilde{E}_q = (i\delta_{i+1,j})_{i,j=1}^q$ which commutes with $\mathcal{P}_q = \exp(\tilde{E}_q)$. For the different steps (22)–(24) and (25)–(27), and in the same succession we write down the order conditions from [13] when K_n is diagonal. The forward conditions are

$$A_0 V_{q_1} = a e_1^\top + b e_2^\top + K_0 V_{q_1} \tilde{E}_{q_1}, \quad n = 0, \quad (29)$$

$$A_n V_{q_1} = B_n V_{q_1} \mathcal{P}_{q_1}^{-1} + K_n V_{q_1} \tilde{E}_{q_1}, \quad 1 \leq n \leq N, \quad (30)$$

$$w^\top V_{q_1} = \mathbf{1}^\top, \quad (31)$$

with the cardinal basis vectors $e_i \in \mathbb{R}^s$, $i = 1, \dots, s$. The conditions for the adjoint methods are given by

$$v^\top V_{q_2} = e_1^\top, \quad (32)$$

$$A_n^\top V_{q_2} = B_{n+1}^\top V_{q_2} \mathcal{P}_{q_2} - K_n V_{q_2} \tilde{E}_{q_2}, \quad 0 \leq n \leq N-1, \quad (33)$$

$$A_N^\top V_{q_2} = w \mathbf{1}^\top - K_N V_{q_2} \tilde{E}_{q_2}, \quad n = N. \quad (34)$$

2.2.2. Bounds for the Global Error

In this section, the errors $\check{Y}_{nj} := y(t_{nj}) - Y_{nj}$, $\check{P}_{nj} := p(t_{nj}) - P_{nj}$, $n = 0, \dots, N$, $j = 1, \dots, s$, are analyzed. According to [13], the equation for the errors $\check{Y}^\top = (\check{Y}_0^\top, \dots, \check{Y}_N^\top)$ and $\check{P}^\top = (\check{P}_0^\top, \dots, \check{P}_N^\top)$ is a linear system of the form

$$\mathbb{M}_h \check{Z} = \tau, \quad \check{Z} = \begin{pmatrix} \check{Y} \\ \check{P} \end{pmatrix}, \quad \tau = \begin{pmatrix} \tau^Y \\ \tau^P \end{pmatrix}, \quad (35)$$

where the matrix \mathbb{M}_h has a 2×2 -block structure and (τ^Y, τ^P) denote the corresponding truncation errors. Deleting all h -depending terms from \mathbb{M}_h , the block structure of the remaining matrix \mathbb{M}_0 is given by

$$\mathbb{M}_0 = \begin{pmatrix} M_{11} \otimes I_m & 0 \\ M_{21} \otimes \nabla_{yy} C_N & M_{22} \otimes I_m \end{pmatrix} \quad (36)$$

with $M_{11}, M_{21}, M_{22} \in \mathbb{R}^{s(N+1) \times s(N+1)}$ and a mean value $\nabla_{yy} C_N \in \mathbb{R}^{m \times m}$ of the symmetric Hessian matrix of C . The index ranges of all three matrices are copied from the numbering of the grid, $0, \dots, N$, for convenience. In fact, $M_{21} = (e_N \otimes \mathbf{1})(e_N \otimes w)^\top$ has rank one only with $e_N = (\delta_{Nj})_{j=0}^N$. The diagonal blocks of \mathbb{M}_0 are nonsingular and its inverse has the form

$$\mathbb{M}_0^{-1} = \begin{pmatrix} M_{11}^{-1} \otimes I_m & 0 \\ -(M_{22}^{-1} M_{21} M_{11}^{-1}) \otimes \nabla_{yy} C_N & M_{22}^{-1} \otimes I_m \end{pmatrix}. \quad (37)$$

The diagonal blocks M_{11}, M_{22} have a bi-diagonal block structure with identity matrices I_s in the diagonal. The individual $s \times s$ -blocks of their inverses are easily computed with M_{11}^{-1} having lower triangular block form and M_{22}^{-1} upper triangular block form, with blocks

$$(M_{11}^{-1})_{nk} = \bar{B}_n \cdots \bar{B}_{k+1}, k \leq n, \quad (M_{22}^{-1})_{nk} = \bar{B}_{n+1}^T \cdots \bar{B}_k^T, k \geq n, \quad (38)$$

with the abbreviations $\bar{B}_n := A_n^{-1}B_n, 1 \leq n \leq N$ and $\bar{B}_{n+1}^T := (B_{n+1}A_n^{-1})^T, 0 \leq n < N$. Empty products for $k = n$ mean the identity I_s .

Defining $\mathbb{U} := h^{-1}(\mathbb{M}_h - \mathbb{M}_0)$ and rewriting (35) in fixed-point form

$$\check{Z} = h\mathbb{M}_0^{-1}\mathbb{U}\check{Z} + \mathbb{M}_0^{-1}\tau, \quad (39)$$

it has been shown in the proof of Theorem 4.1 of [13] for smooth right hand sides f and $h \leq h_0$ that

$$\|\check{Z}\| \leq 2 \max\{\|M_{11}^{-1}\tau^Y\|, \|M_{22}^{-1}\tau^P\|\} \quad (40)$$

in suitable norms, where these norms are discussed in more detail in Section 2.2.3 below. Moreover, due to the lower triangular block structure of \mathbb{M}_0 the estimate for the error in the state variable may be refined (Lemma 4.2 in [13]) to

$$\|\check{Y}\| \leq \|M_{11}^{-1}\tau^Y\| + hL\|\check{Z}\| \quad (41)$$

with some constant L . Without additional conditions, estimates of the terms on the right-hand side of (40) in the form $\|\check{Z}\| = O(h^{-1}\|\tau\|)$ lead to the loss of one order in the global error. However, this loss may be avoided by one additional superconvergence condition on the forward and the adjoint method each, which will be considered next.

In our convergence result Theorem 1 below, we will assume the existence of the discrete solution satisfying (22)–(27), for simplicity. However, at least for quadratic objective functions C , this existence follows quite simply along the same lines used in this paragraph here.

Lemma 1. *Let the right-hand side f of (2) be smooth with bounded second derivatives and let the function C be a polynomial of degree two, at most. Let the coefficients of the standard scheme satisfy*

$$A\mathbb{I} = B\mathbb{I}, \quad \mathbb{I}^T A = \mathbb{I}^T B, \quad |\lambda_2(A^{-1}B)| < 1, \quad (42)$$

where λ_2 denotes the absolutely second largest eigenvalue of the matrix $A^{-1}B$. Then, there exists a unique solution Y, P of the system (22)–(27) for small enough stepsizes h .

Proof. Since $\nabla_{yy}C$ is constant, \mathbb{M}_0^{-1} in (37) is a fixed matrix and similar to (39) the Equations (22)–(27) may be written as a fixed-point problem

$$Z = h\mathbb{M}_0^{-1}U(Z), \quad Z = \begin{pmatrix} Y \\ P \end{pmatrix}, \quad (43)$$

where the function $U(Z)$ is also smooth, having the Jacobian \mathbb{U} considered in (39). Due to (42) there exist norms such that $\|\bar{B}\| \leq 1, \|\bar{B}^T\| \leq 1$ hold. In this case, the proof of Theorem 4.1 in [13] shows the bound $\|\mathbb{M}_0^{-1}\mathbb{U}\| \leq L$ with a constant L depending on the derivatives of g and ϕ . Hence, by the mean value theorem, the right-hand side of (43) has hL as a Lipschitz constant and the map $Z \mapsto h\mathbb{M}_0^{-1}U(Z)$ is a contraction for $h \leq h_0 := 1/(2L)$, establishing the existence of a unique fixed point solution. \square

2.2.3. Superconvergence of the Standard Method

For s -stage Peer methods, global order s may be attained in many cases if other properties of the method have lower priority. For optimal stiff stability properties like

A-stability, however, it may be necessary to sacrifice one order of consistency as in [20,21]. Accordingly, in this paper the order conditions for the standard method are lowered by one compared to the requirements in the recent paper [13] to local orders $(s, s - 1)$, see Table 1. Still, the higher global orders may be preserved to some extent by the concept of superconvergence which prevents the order reduction in the global error by cancellation of the leading error term.

Table 1. Combined order conditions for the Peer triplet, including the compatibility condition (67) and the condition (61) for full matrices K_N .

Steps	forward: $q_1 = s$	adjoint: $q_2 = s - 1$
Start, $n = 0$	(29)	(33),(61), $\beta = 0$
Standard, $1 \leq n < N$	(30)	(33)
Superconvergence	(52)	(53)
Compatibility	(67)	(67)
Last step	(30), $n = N$	(33), $n = N - 1$
End point	(31)	(34),(61), $\beta = N$

Superconvergence is essentially based on the observation that the powers of the forward stability matrix $\bar{B} := A^{-1}B$ may converge to a rank-one matrix which maps the leading error term of τ^Y to zero. This is the case if the eigenvalue 1 of \bar{B} is isolated. Indeed, if the eigenvalues λ_j , $j = 1, \dots, s$, of the stability matrix \bar{B} satisfy

$$1 = \lambda_1 > |\lambda_2| \geq \dots |\lambda_s|, \quad (44)$$

then its powers \bar{B}^n converge to the rank-one matrix $\mathbb{1}\mathbb{1}^T A$ since $\mathbb{1}$ and $\mathbb{1}^T A$ are the right and left eigenvectors having unit inner product $\mathbb{1}^T A \mathbb{1} = 1$ due to the preconsistency conditions $A^{-1}B\mathbb{1} = \mathbb{1}$ and $A^{-T}\bar{B}^T\mathbb{1} = \mathbb{1}$ of the forward and backward standard Peer method, see (30) and (33). The convergence is geometric, i.e.,

$$\|\bar{B}^n - \mathbb{1}\mathbb{1}^T A\| = \|(\bar{B} - \mathbb{1}\mathbb{1}^T A)^n\| = O(\gamma^n) \rightarrow 0, \quad n \rightarrow \infty, \quad (45)$$

for any $\gamma \in (|\lambda_2|, 1)$. Some care has to be taken here since the error estimate (40) depends on the existence of special norms satisfying $\|\bar{B}\|_{X_1} := \|X_1^{-1}\bar{B}X_1\|_\infty = 1$, resp. $\|\bar{B}^T\|_{X_2} = 1$. Concentrating on the forward error $\|M_{11}^{-1}\tau^Y\|$, a first transformation of \bar{B} is considered with the matrix

$$X = \begin{pmatrix} 1 & -\beta^T \\ \mathbb{1}_{s-1} & I_{s-1} - \mathbb{1}_{s-1}\beta^T \end{pmatrix}, \quad X^{-1} = \begin{pmatrix} \beta_1 & \beta^T \\ -\mathbb{1}_{s-1} & I_{s-1} \end{pmatrix}, \quad (46)$$

where $(\beta_1, \beta^T) = \mathbb{1}^T A$. Since $Xe_1 = \mathbb{1}$ and $e_1^T X^{-1} = \mathbb{1}^T A$, the matrix $X^{-1}\bar{B}X$ is block-diagonal with the dominating eigenvalue 1 in the first diagonal entry. Due to (44) there exists an additional nonsingular matrix $\Xi \in \mathbb{R}^{(s-1) \times (s-1)}$ such that the lower diagonal block of $X^{-1}\bar{B}X$ has norm smaller one, i.e.

$$\|\Xi^{-1}(-\mathbb{1}_{s-1}, I_{s-1})\bar{B} \begin{pmatrix} -\beta^T \\ I_{s-1} - \mathbb{1}_{s-1}\beta^T \end{pmatrix} \Xi\|_\infty = \gamma < 1. \quad (47)$$

Hence, with the matrix

$$X_1 := X \begin{pmatrix} 1 & \\ & \Xi \end{pmatrix} \quad (48)$$

the required norm is found, satisfying $\|\bar{B}\|_{X_1} := \|X_1^{-1}\bar{B}X_1\|_\infty = \max\{1, \gamma\} = 1$ and $\|\bar{B} - \mathbb{1}\mathbb{1}^T A\|_{X_1} = \gamma < 1$ in (45). Using this norm in (40) and (41) for $\epsilon^Y := M_{11}^{-1}\tau^Y$, it is seen with (38) that

$$\epsilon_n^Y = \sum_{k=0}^n \mathbb{1}(\mathbb{1}^T A \tau_{n-k}^Y) + \underbrace{\sum_{k=0}^n (\bar{B} - \mathbb{1}\mathbb{1}^T A)^k \tau_{n-k}^Y}_{=O(h^s)}, \quad 0 \leq n < N. \quad (49)$$

Only for ϵ_N^Y a slight modification is required and the factors in the second sum have to be replaced by $(\bar{B}_N - \mathbb{1}\mathbb{1}^T A)(\bar{B} - \mathbb{1}\mathbb{1}^T A)^{k-1}$ for $k > 1$ with norms still of size $O(\gamma^k)$. Now, in all cases the loss of one order in the first sum in (49) may be avoided if the leading $O(h^s)$ -term of τ_{n-k}^Y is canceled in the product with the left eigenvector, i.e., if $\mathbb{1}^T A \tau_{n-k}^Y = O(h^{s+1})$. An analogous argument may be applied to the second term $\|M_{22}^{-1}\tau^P\|$ in (40). The adjoint stability matrix $\bar{B}^T = (BA^{-1})^T$ possesses the same eigenvalues as \bar{B} and its leading eigenvectors are also known: $\bar{B}^T \mathbb{1} = \mathbb{1}$ and $(A\mathbb{1})^T \bar{B}^T = \mathbb{1}^T B^T = (A\mathbb{1})^T$ by preconsistency and an analogous construction applies.

Under the conditions corresponding to local orders $(s, s-1)$ the leading error terms in $\tau_n^Y = \frac{1}{s!}h^s \eta_s \otimes y^{(s)}(t_n) + O(h^{s+1})$ and $\tau_n^P = \frac{1}{(s-1)!}\eta_{s-1}^* \otimes p^{(s-1)}(t_n) + O(h^s)$ are given by

$$\eta_s = \mathbf{c}^s - A^{-1}B(\mathbf{c} - \mathbb{1})^s - sA^{-1}K\mathbf{c}^{s-1}, \quad (50)$$

$$\eta_{s-1}^* = \mathbf{c}^{s-1} - A^{-T}B^T(\mathbf{c} + \mathbb{1})^{s-1} + (s-1)A^{-T}K\mathbf{c}^{s-2}. \quad (51)$$

Considering now $(\mathbb{1}^T A)\eta_s$ in (49) and similarly $(A\mathbb{1})^T \eta_{s-1}^*$ the following result is obtained.

Theorem 1. *Let the Peer triplet with $s > 1$ stages satisfy the order conditions collected in Table 1 and let the solutions satisfy $y \in C^{s+1}[0, T]$, $p \in C^s[0, T]$. Let the coefficients of the standard Peer method satisfy the conditions*

$$\mathbb{1}^T (A\mathbf{c}^s - B(\mathbf{c} - \mathbb{1})^s - sK\mathbf{c}^{s-1}) = 0, \quad (52)$$

$$\mathbb{1}^T (A^T \mathbf{c}^{s-1} - B^T(\mathbf{c} + \mathbb{1})^{s-1} + (s-1)K\mathbf{c}^{s-2}) = 0, \quad (53)$$

and let (44) be satisfied. Assume, that a Peer solution $(Y^T, P^T)^T$ exists and that f and C have bounded second derivatives. Then, for stepsizes $h \leq h_0$ the error of these solutions is bounded by

$$\|Y_{nj} - y(t_{nj})\|_\infty + h\|P_{nj} - p(t_{nj})\|_\infty = O(h^s), \quad (54)$$

$n = 0, \dots, N$, $j = 1, \dots, s$.

Proof. Under condition (52) the representation (49) shows that $\|M_{11}^{-1}\tau^Y\| = O(h^s)$. In the same way follows $\|M_{22}^{-1}\tau^P\| = O(h^{s-1})$ from condition (53). In (40) this leads to a common error $\|\check{Z}\| = O(h^{s-1})$ which may be refined for the state variable with (41) to $\|\check{Y}\| = O(h^s)$. \square

Remark 2. *The estimate (49) shows that superconvergence may be a fragile property and may be impaired if $|\lambda_2|$ is too close to one, leading to very large values in the bound*

$$\sum_{k=0}^n \|(\bar{B} - \mathbb{1}\mathbb{1}^T A)^k\|_{X_1} \leq \frac{1}{1-\gamma}$$

for the second term in (49). In fact, numerical tests showed that the value $\gamma \cong |\lambda_2|$ plays a crucial role. While $|\lambda_2| \doteq 0.9$ was appropriate for the three-stage method AP3o32f which shows order 3 in the tests in Section 3, for two four-stage methods with $|\lambda_2| \doteq 0.9$ superconvergence was not seen in

any of our three test problems below. Reducing $|\lambda_2|$ farther with additional searches we found that a value below $\gamma = 0.8$ may be safe to achieve order 4 in practice. Hence, the value $|\lambda_2|$ will be one of the data listed in the properties of the Peer methods developed below.

Remark 3. By Theorem 1, weakening the order conditions from the local order pair $(s + 1, s)$ to the present pair $(s, s - 1)$ combined with fewer conditions for superconvergence preserves global order h^s for the state variable. However, it also leads to a more complicated structure of the leading error. Extending the Taylor expansion for τ^Y, τ^P and applying the different bounds, a more detailed representation of the state error may be derived,

$$\begin{aligned} \|\check{Y}\| \leq & h^s \left(\frac{\|\eta_s\|}{(1-\gamma)s!} \|y^{(s)}\| + \frac{|\mathbb{1}^T A \eta_{s+1}|}{(s+1)!} \|y^{(s+1)}\| \right) \\ & + h^s (\hat{L} \|p^{(s)}\| + \hat{L} \|p^{(s-1)}\|) + O(h^{s+1}) \end{aligned} \quad (55)$$

with some modified constant \hat{L} . Obviously, the leading error depends on four different derivatives of the solutions. Since the dependence on p is rather indirect, we may concentrate on the first line in (55). Both derivatives there may not be correlated and it may be difficult to choose a reasonable combination of both error constants as objective function in the construction of efficient methods. Still, in the local error the leading term is obviously $\tau^Y \doteq h^s \frac{1}{s!} \eta_s \otimes y^{(s)}$, with η_s defined in (50), and it may be propagated through non-linearity and rounding errors. Hence, in the search for methods,

$$err_s := \frac{1}{s!} \|\eta_s\|_\infty = \frac{1}{s!} \|\mathbf{c}^s - A^{-1}B(\mathbf{c} - \mathbb{1})^s - sA^{-1}K\mathbf{c}^{s-1}\|_\infty \quad (56)$$

is used as the leading error constant.

2.2.4. Adjoint Order Conditions for General Matrices K_n

The number of order conditions for the boundary methods is so large that they may not be fulfilled with the restriction to diagonal coefficient matrices K_0, K_N for $s \geq 3$. Hence, it is convenient to make the step to full matrices K_0, K_N in the boundary methods and the order conditions for the adjoint schemes have to be derived for this case. Unfortunately, for such matrices the adjoint schemes (27) and (26) for $n = 0$ have a rather unfamiliar form. Luckily, the adjoint differential equation $p' = -\nabla_y f(y, u)^T p$ is linear. We abbreviate the initial value problem for this equation as

$$p'(t) = -J(t)p(t), \quad p(T) = p_T, \quad (57)$$

with $J(t) = \nabla_y f(y(t), u(t))^T$ and for some boundary index $\beta \in \{0, N\}$, we consider the matrices $A_\beta = (a_{ij}^{(\beta)})$, $K_\beta = (\kappa_{ij}^{(\beta)})$. Starting the analysis with the simpler end step (27), we have

$$\sum_{j=1}^s a_{ji}^{(N)} P_{Nj} = w_i p_h(T) + hJ(t_{Ni}) \sum_{j=1}^s \kappa_{ji}^{(N)} P_{Nj}, \quad i = 1, \dots, s, \quad (58)$$

which is some kind of half-one-leg form since it evaluates the Jacobian J and the solution p at different time points. This step must be analyzed for the linear equation (57) only. Expressions for the higher derivatives of the solution p follow easily:

$$p'' = (J^2 - J')p, \quad p''' = (-J^3 + 2J'J + JJ' - J'')p. \quad (59)$$

Lemma 2. For a smooth coefficient matrix $J(t)$, the scheme (58) for the linear differential Equation (57) has local order 3 under the conditions

$$A_N^T V_3 + K_N^T V_3 \tilde{E}_3 = w \mathbb{1}^T \quad (60)$$

and with $\beta = N$

$$\sum_{\substack{i=1 \\ i \neq j}}^s (c_i - c_j) \kappa_{ij}^{(\beta)} = 0, \quad j = 1, \dots, s. \quad (61)$$

Proof. Considering the residual of the scheme with the exact solution $p(t)$, Taylor expansion at t_N and the Leibniz rule yield

$$\begin{aligned} \Delta_i &= \sum_{j=1}^s a_{ji}^{(N)} p(t_{Nj}) - w_i p(T) - h J(t_{Ni}) \sum_{j=1}^s \kappa_{ji}^{(N)} p(t_{Nj}) \\ &= \sum_{k=0}^{q-1} \frac{h^k}{k!} \left(\sum_{j=1}^s a_{ji}^{(N)} c_j^k - w_i \right) p^{(k)} - h \underbrace{\sum_{k=0}^{q-2} \frac{h^k}{k!} c_i^k J^{(k)} \sum_{j=1}^s \kappa_{ji}^{(N)} \sum_{\ell=0}^{q-2} \frac{h^\ell}{\ell!} c_j^\ell p^{(\ell)}}_{:=\delta} + O(h^q) \end{aligned}$$

where all derivatives are evaluated at t_N . The second term can be further reformulated as

$$\begin{aligned} \delta &= \sum_{k=1}^{q-1} h^k \sum_{\ell=0}^{k-1} \sum_{j=1}^s \kappa_{ji}^{(N)} \frac{c_i^\ell c_j^{k-\ell-1}}{\ell!(k-\ell-1)!} J^{(\ell)} p^{(k-\ell-1)} \\ &= \sum_{k=1}^{q-1} h^k \sum_{\ell=1}^k \sum_{j=1}^s \kappa_{ji}^{(N)} \frac{c_i^{\ell-1} c_j^{k-\ell}}{(\ell-1)!(k-\ell)!} J^{(\ell-1)} p^{(k-\ell)}. \end{aligned}$$

Looking at the factors of h^0, h^1, h^2 separately leads to the order conditions

$$\begin{aligned} h^0 : 0 &= A_N^T \mathbf{1} - w, \\ h^1 : 0 &= \left(\sum_{j=1}^s a_{ji}^{(N)} c_j - w_i \right) p' - \sum_{j=1}^s \kappa_{ji}^{(N)} J p = \left(- \sum_{j=1}^s a_{ji}^{(N)} c_j + w_i - \sum_{j=1}^s \kappa_{ji}^{(N)} \right) J p, \text{ i.e.} \\ 0 &= A_N^T \mathbf{c} + K_N^T \mathbf{1} - w, \\ h^2 : 0 &= \frac{1}{2} \left(\sum_j a_{ji}^{(N)} c_j^2 - w_i \right) p'' - \sum_{j=1}^s \kappa_{ji}^{(N)} (c_j J p' + c_i J' p) \\ &= \frac{1}{2} \left(\sum_{j=1}^s c_j^2 a_{ji}^{(N)} - w_i + 2 \sum_{j=1}^s c_j \kappa_{ji}^{(N)} \right) J^2 p - \frac{1}{2} \left(\sum_{j=1}^s c_j^2 a_{ji}^{(N)} - w_i + 2 \sum_{j=1}^s \kappa_{ji}^{(N)} c_i \right) J' p. \end{aligned}$$

Cancellation of the factor of J^2 requires the condition $0 = A_N^T \mathbf{c}^2 + 2K_N^T \mathbf{c} - w$, which combines with the h^0, h^1 -conditions to (60). The factor of J' , however, requires $0 = A_N^T \mathbf{c}^2 + 2D_c K_N^T \mathbf{1} - w$ with $D_c = \text{diag}(c_i)$. Subtracting this expression from the factor of J^2 leaves $0 = K_N^T \mathbf{c} - D_c K_N^T \mathbf{1}$, which corresponds to (61). \square

Remark 4. Condition (60) is the standard version for diagonal K_N from [13]. Hence, the half-one-leg form of (58) introduces s additional conditions (61), only, for order 3 while the boundary coefficients $K_\beta, \beta = 0, N$, may now contain $s(s-1)$ additional elements. In fact, a similar analysis for step (26) with $n = \beta = 0$ reveals again (61) as the only condition in addition to (33).

2.3. Existence of Boundary Methods Imposes Restrictions on the Standard Method

In the previous paper [13], the combination of forward and adjoint order conditions for the standard method (A, B, K) into one set of equations relating only A and K already gave insight on some background of these methods such as the advantages of using symmetric nodes. It also simplifies the actual construction of methods leading to shorter expressions

during the elimination process with algebraic software tools. For ease of reference, this singular Sylvester type equation is reproduced here,

$$(V_{q_2} \mathcal{P}_{q_2})^T A (V_{q_1} \mathcal{P}_{q_1}) - V_{q_2}^T A V_{q_1} = (V_{q_2} \mathcal{P}_{q_2})^T K V_{q_1} \tilde{E}_{q_1} + (V_{q_2} \tilde{E}_{q_2})^T K V_{q_1}. \quad (62)$$

A similar combination of the order conditions for the boundary methods, however, reveals crucial restrictions: the triplet of methods (A_0, K_0) , (A, B, K) , (A_N, B_N, K_N) has to be discussed together since boundary methods of an appropriate order may not exist for any standard method (A, B, K) , only for those satisfying certain compatibility conditions required by the boundary methods. Knowing these conditions allows one to design the standard method alone without the ballast of two more methods with many additional unknowns. This decoupled construction also greatly reduces the dimension of the search space if methods are optimized by automated search routines. We start with the discussion of the end method.

2.3.1. Combined Conditions for the End Method

We remind that we now are looking for methods having local order $q_1 = s$ and $q_2 = s - 1$ everywhere, which we abbreviate from now on as $(q_1, q_2) = (s, q)$. In particular this means that $A^T V_q + K^T V_q \tilde{E}_q = B^T V_q \mathcal{P}_q$ for the standard method. Looking for bottlenecks in the design of these methods, we try to identify crucial necessary conditions and consider the three order conditions for the end method (A_N, B_N, K_N) in combination

$$\begin{aligned} A_N V_s - B_N V_s \mathcal{P}_s^{-1} - K_N V_s \tilde{E}_s &= 0, \\ B_N^T V_q \mathcal{P}_q &= A^T V_q + K^T V_q \tilde{E}_q = B^T V_q \mathcal{P}_q, \\ A_N^T V_q + K_N^T V_q \tilde{E}_q &= w \mathbb{1}_q^T. \end{aligned} \quad (63)$$

From these conditions the matrices A_N, B_N may be eliminated, revealing the first restrictions on B . Here, the singular matrix map

$$\mathcal{L}_{q,s} : \mathbb{R}^{q \times s} \rightarrow \mathbb{R}^{q \times s}, \quad X \mapsto \tilde{E}_q^T X + X \tilde{E}_s \quad (64)$$

plays a crucial role.

Lemma 3. A necessary condition for a boundary method (A_N, B_N, K_N) to satisfy (63) is

$$\mathcal{L}_{q,s}(V_q^T K_N V_s) = \mathbb{1}_q \mathbb{1}_s^T - V_q^T B V_s \mathcal{P}_s^{-1}. \quad (65)$$

Proof. The second condition, $V_q^T B_N = V_q^T B$, in (63) leads to the necessary equation

$$V_q^T A_N V_s - V_q^T K_N V_s \tilde{E}_s = V_q^T B V_s \mathcal{P}_s^{-1}$$

due to the first condition. The transposed third condition

$$V_q^T A_N V_s + \tilde{E}_q^T V_q^T K_N V_s = \mathbb{1}_q w^T V_s = \mathbb{1}_q \mathbb{1}_s^T$$

multiplied by the nonsingular matrix V_s may be used to eliminate A_N and leads to

$$\tilde{E}_q^T V_q^T K_N V_s + V_q^T K_N V_s \tilde{E}_s = \mathbb{1}_q \mathbb{1}_s^T - V_q^T B V_s \mathcal{P}_s^{-1}$$

which is the equation (65) from the statement. \square

Unfortunately, this lemma leads to several restrictions on the design of the methods due to the properties of the map $\mathcal{L}_{q,s}$. Firstly, for diagonal matrices K_N the image $\mathcal{L}_{q,s}(V_q^T K_N V_s)$ has a very restricted shape.

Lemma 4. If $K \in \mathbb{R}^{s \times s}$ is a diagonal matrix, then $V_q^T K V_s$ and $\mathcal{L}_{q,s}(V_q^T K V_s)$ are Hankel matrices with constant entries along anti-diagonals.

Proof. With $K = \text{diag}(\kappa_i)$, we have $x_{ij} := e_i^T (V_q^T K V_s) e_j = \sum_{k=1}^s \kappa_k c_k^{i+j-2}$, showing Hankel form of $X = (x_{ij}) =: (\xi_{i+j-1})$ for $i = 1, \dots, s-1, j = 1, \dots, s$. Now, $e_i^T (\tilde{E}_q^T X + X \tilde{E}_s) e_j = (i-1)x_{i-1,j} + x_{i,j-1}(j-1) = (i+j-2)\xi_{i+j-2}$, which shows again Hankel form. \square

This lemma means that an end method with diagonal K_N only exists if also $V_q^T B V_s \mathcal{P}_s^{-1}$ on the right-hand side of (65) has Hankel structure. Unfortunately it was observed that for standard methods with definite K this is the case for $q_2 \leq 2$ only (there exist methods with an explicit stage $\kappa_{33} = 0$).

Remark 5. Trying to overcome this bottleneck with diagonal matrices K_N , one might consider adding additional stages of the end method. However, using general end nodes $(\hat{c}_1, \dots, \hat{c}_{\hat{s}})$ with $\hat{s} \geq s$ does not remove this obstacle. The corresponding matrix $\mathcal{L}_{q,s}(\hat{V}_q^T K_N \hat{V}_s)$ with appropriate Vandermonde matrices \hat{V} still has Hankel form.

However, even with a full end matrix K_N , Lemma 3 and the Fredholm alternative enforce restrictions on the standard method (A, B, K) due to the singularity of $\mathcal{L}_{q,s}$. This is discussed for the present situation with $s = 4, q = 3$, only. The matrix belonging to the map $\mathcal{L}_{q,s}$ is $I_s \otimes \tilde{E}_q^T + \tilde{E}_s^T \otimes I_q$ and its transpose is $I_s \otimes \tilde{E}_q + \tilde{E}_s \otimes I_q$. Hence, the adjoint of the map $\mathcal{L}_{q,s}$ is given by

$$\mathcal{L}_{q,s}^T: \mathbb{R}^{q \times s} \rightarrow \mathbb{R}^{q \times s}, \quad X \mapsto \tilde{E}_q X + X \tilde{E}_s^T.$$

Component-wise the map acts as

$$\mathcal{L}_{q,s}^T: (x_{ij}) \mapsto (ix_{i+1,j} + jx_{i,j+1})$$

with elements having indices $i > q$ or $j > s$ being zero. For $q = 3, s = 4$, one gets

$$\mathcal{L}_{3,4}(X)^T = \begin{pmatrix} x_{12} + x_{21} & 2x_{13} + x_{22} & 3x_{14} + x_{23} & x_{24} \\ x_{22} + 2x_{31} & 2x_{23} + 2x_{32} & 3x_{24} + 2x_{33} & 2x_{34} \\ x_{32} & 2x_{33} & 3x_{34} & 0 \end{pmatrix}.$$

It is seen that the kernel of $\mathcal{L}_{3,4}^T$ has dimension 3 and is given by

$$X = \begin{pmatrix} \xi_1 & \xi_2 & \xi_3 & 0 \\ -\xi_2 & -2\xi_3 & 0 & 0 \\ \xi_3 & 0 & 0 & 0 \end{pmatrix}. \quad (66)$$

In (65), the Fredholm condition leads to restrictions on the matrix B from the standard scheme. However, since the matrix A should have triangular form, it is the more natural variable in the search for good methods and an equivalent reformulation of these conditions for A is of practical interest.

Lemma 5. Assume that the standard method (A, B, K) has local order $(s, q) = (4, 3)$. Then, end methods (A_N, B_N, K_N) of order $(s, q) = (4, 3)$ only exist if the standard method (A, B, K) satisfies the following set of three conditions, either for B or for A ,

$$\left. \begin{aligned} \mathbb{1}^T B \mathbb{1} &= 1 \\ \mathbb{1}^T B \mathbf{c} - \mathbf{c}^T B \mathbb{1} &= 1 \\ \mathbb{1}^T B \mathbf{c}^2 - 2\mathbf{c}^T B \mathbf{c} + (\mathbf{c}^2)^T B \mathbb{1} &= 1 \end{aligned} \right\} \left\{ \begin{aligned} 1 &= \mathbb{1}^T A \mathbb{1} \\ 1 &= \mathbb{1}^T A \mathbf{c} - \mathbf{c}^T A \mathbb{1} \\ 0 &= \mathbb{1}^T A \mathbf{c}^2 - 2\mathbf{c}^T A \mathbf{c} + (\mathbf{c}^2)^T A \mathbb{1}. \end{aligned} \right. \quad (67)$$

Proof. Multiplying equation (65) by \mathcal{P}_s from the right and using $\tilde{E}_s \mathcal{P}_s = \mathcal{P}_s \tilde{E}_s$, an equivalent form is

$$\mathcal{L}_{q,s}(V_q^T K_N V_s \mathcal{P}_s) = \mathbb{1}_q \mathbb{1}_s^T \mathcal{P}_s - V_q^T B V_s =: R,$$

with $\mathbb{1}_s^T \mathcal{P}_s = (1, 2, 4, 8, \dots)$ by the binomial formula. The Fredholm alternative requires that $\text{tr}(X^T R) = 0$ for all X from (66). We now frequently use the identities $\text{tr}((vu^T)R) = u^T R v$ and $V e_i = \mathbf{c}^{i-1}$ with $e_i \in \mathbb{R}^s$. The kernel in (66) is spanned by three basis elements. The first, $X_1 = \bar{e}_1 e_1^T$ (with the convention $\bar{e}_i \in \mathbb{R}^q$) leads to

$$0 \stackrel{!}{=} \text{tr}(X_1^T R) = \bar{e}_1^T \mathbb{1}_q \mathbb{1}_s^T e_1 - \bar{e}_1^T V_q^T B V_s e_1 = 1 - \mathbb{1}_s^T B \mathbb{1}_s.$$

The second basis element is $X_2 = \bar{e}_1 e_2^T - \bar{e}_2 e_1^T$. Here $\text{tr}(X_2^T \mathbb{1}_q \mathbb{1}_s^T \mathcal{P}_s) = \mathbb{1}_s^T \mathcal{P}_s e_2 - \mathbb{1}_s^T \mathcal{P}_s e_1 = 1$ and

$$\text{tr}(X_2^T V_q^T B V_s) = \bar{e}_1^T V_q B V_s e_2 - \bar{e}_2^T V_q B V_s e_1 = \mathbb{1}_s^T B \mathbf{c} - \mathbf{c}^T B \mathbb{1}_s.$$

For the third element, $X_3 = \bar{e}_1 e_3^T - 2\bar{e}_2 e_2^T + \bar{e}_3 e_1^T$, one gets $\text{tr}(X_3^T \mathbb{1}_q \mathbb{1}_s^T \mathcal{P}_s) = \mathbb{1}_s^T \mathcal{P}_s (e_3 - 2e_2 + e_1) = 4 - 4 + 1 = 1$.

The third condition on B is

$$\text{tr}(X_3^T V_q^T B V_s) = \mathbb{1}_s^T B \mathbf{c}^2 - 2\mathbf{c}^T B \mathbf{c} + (\mathbf{c}^2)^T B \mathbb{1}_s.$$

The versions for A follow from the order conditions. Let again $\mathbb{1} := \mathbb{1}_s$. The first columns of (30) and (33) show $B\mathbb{1} = A\mathbb{1}$ and $\mathbb{1}^T B = \mathbb{1}^T A$, which gives $\mathbb{1}^T B\mathbb{1} = \mathbb{1}^T A\mathbb{1} = 1$. The second column of (30) reads $B\mathbf{c} = A\mathbf{c} + A\mathbb{1} - K\mathbb{1}$ and leads to $\mathbb{1}^T A\mathbf{c} = \mathbb{1}^T B\mathbf{c} + \mathbb{1}^T A\mathbb{1} - \mathbb{1}^T K\mathbb{1}$ showing also $\mathbb{1}^T K\mathbb{1} = \mathbb{1}^T A\mathbb{1} = 1$. Hence, the second condition in (67) is equivalent with

$$1 \stackrel{!}{=} \mathbb{1}^T (B\mathbf{c}) - \mathbf{c}^T (B\mathbb{1}) = \mathbb{1}^T (A\mathbf{c} + A\mathbb{1} - K\mathbb{1}) - \mathbf{c}^T A\mathbb{1} = \mathbb{1}^T A\mathbf{c} - \mathbf{c}^T A\mathbb{1}.$$

In order to show the last equivalence in (67), we have to look ahead at the forward condition (30) for order 3, which is $B\mathbf{c}^2 = A(\mathbf{c}^2 + 2\mathbf{c} + \mathbb{1}) - 2K(\mathbf{c} + \mathbb{1})$. This leads to $\mathbb{1}^T A\mathbf{c}^2 = \mathbb{1}^T B\mathbf{c}^2 = \mathbb{1}^T A(\mathbf{c}^2 + 2\mathbf{c} + \mathbb{1}) - 2\mathbb{1}^T K(\mathbf{c} + \mathbb{1})$, which is equivalent to $2(\mathbb{1}^T A\mathbf{c} - \mathbb{1}^T K\mathbf{c}) = 1$. Now, this expression is required in the last reformulation which also uses the second adjoint order condition $B^T \mathbf{c} = A^T \mathbf{c} - A^T \mathbb{1} + K^T \mathbb{1}$, yielding

$$\begin{aligned} 1 &\stackrel{!}{=} (\mathbb{1}^T B)\mathbf{c}^2 + (\mathbf{c}^2)^T (B\mathbb{1}) - 2\mathbf{c}^T (B^T \mathbf{c}) \\ &= \mathbb{1}^T A\mathbf{c}^2 + (\mathbf{c}^2)^T A\mathbb{1} - 2\mathbf{c}^T A^T \mathbf{c} + \underbrace{2(\mathbf{c}^T A^T \mathbb{1} - \mathbf{c}^T K^T \mathbb{1})}_{=1}. \end{aligned}$$

□

Remark 6. It can be shown that only the first condition $\mathbb{1}^T A\mathbb{1} = 1$ in (67) is required if the matrix K is diagonal. This first condition is merely a normalization fixing the free common factor in the class $\{\alpha \cdot (A, B, K) : \alpha \neq 0\}$ of equivalent methods. The other two conditions are consequences of the order conditions on the standard method (A, B, K) with diagonal K . However, the proof is rather lengthy and very technical and is omitted.

The restrictions (67) on the standard method also seem to be sufficient with (61) posing no further restrictions. In fact, with (67) the construction of boundary methods was always possible in Section 2.4.

2.3.2. Combined Conditions for the Starting Method

The starting method has to satisfy only two conditions

$$\begin{aligned} A_0 V_s &= a e_1^T + b e_2^T + K_0 V_s \tilde{E}_s, \\ A_0^T V_q &= B^T V_q \mathcal{P}_q - K_0^T V_q \tilde{E}_q. \end{aligned} \quad (68)$$

The first two columns of the first equation may be solved for $a = A_0 V_s e_1$ and $b = (A_0 V_s - K_0 V_s \tilde{E}_s) e_2$ leading to the reduced conditions

$$(A_0 V_s - K_0 V_s \tilde{E}_s) Q_3 = 0, \quad Q_3 := I_s - e_1 e_1^T - e_2 e_2^T.$$

The presence of the projection Q_3 leads to changes in the condition compared to the end method.

Lemma 6. A necessary condition for the starting method (A_0, K_0) to satisfy (68) is

$$\left(\mathcal{L}_{q,s}(\mathcal{P}_q^{-T} V_q^T K_0 V_s) - V_q^T B V_s \right) Q_3 = 0. \quad (69)$$

Proof. Transposing the second condition from (68) and multiplying with $V_s Q_3$ gives

$$(V_q^T A_0 V_s + \tilde{E}_q^T V_q^T K_0 V_s - \mathcal{P}_q^T V_q^T B V_s) Q_3 = 0,$$

and $V_q^T A_0 V_s Q_3$ may now be eliminated from both equations, yielding

$$(\tilde{E}_q^T V_q^T K_0 V_s + V_q^T K_0 V_s \tilde{E}_s) Q_3 = \mathcal{P}_q^T V_q^T B V_s Q_3.$$

Again, \mathcal{P}_q^T may be moved to the left side and leads to (69) since it commutes with \tilde{E}_q^T . \square

The situation is now similar to the one for the end method (also concerning a diagonal form of K_0) and we consider again the Fredholm condition. The matrix belonging to the matrix product $\mathcal{L}_{q,s}(\cdot) \cdot Q_3$ is $Q_3 \otimes \tilde{E}_q^T + (\tilde{E}_s Q_3)^T \otimes I_q$ and it has the transpose $Q_3 \otimes \tilde{E}_q + (\tilde{E}_s Q_3) \otimes I_q$. This matrix belongs to the map

$$X \mapsto \tilde{E}_q(X Q_3) + (X Q_3) \tilde{E}_s^T = \mathcal{L}_{q,s}^T(X Q_3). \quad (70)$$

For $q = 3, s = 4$, images of this map are given by

$$\mathcal{L}_{3,4}^T(X Q_3) = \begin{pmatrix} 0 & 2x_{13} & x_{23} + 3x_{14} & x_{24} \\ 0 & 2x_{23} & 2x_{33} + 3x_{24} & 2x_{34} \\ 0 & 2x_{33} & 3x_{34} & 0 \end{pmatrix}.$$

Here, the map $\mathcal{L}_{3,4}^T$ alone introduces no new kernel elements, the kernel of (70) coincides with that of the map $X \mapsto X Q_3$ given by matrices of the form $X = \tilde{X}(I - Q_3)$, $\tilde{X} \in \mathbb{R}^{q \times s}$. Since the right-hand side of (69) is $V_q^T B V_s Q_3$, the condition for solvability

$$\text{tr}(X^T V_q^T B V_s Q_3) = \text{tr}(\tilde{X}^T V_q^T B V_s Q_3 (I - Q_3)) = 0$$

is always satisfied since $Q_3(I - Q_3) = 0$. Hence, no additional restrictions on the standard method are introduced by requiring the existence of starting methods.

2.4. Construction of Peer Triplets

The construction of Peer triplets requires the solution of the collected order conditions from Table 1 and additional optimization of stability and error properties. However, it has been observed that some of these conditions may be related in non-obvious ways, see e.g., Remark 6. This means that the accuracy of numerical solutions may be quite poor due to large and unknown rank deficiencies. Instead, all order conditions were solved here exactly by algebraic manipulation with rational coefficients as far as possible.

The construction of the triplets was simplified by the compatibility conditions (67) allowing the isolated construction of the standard method (A, B, K) without the many additional parameters of the boundary methods. Furthermore, an elimination of the matrix B from forward and adjoint conditions derived in [13], see (62), reduces the number of parameters in A, K to $s(s+3)/2$ elements with $s-1$ additional parameters from the nodes. This is so since (A, B, K) is invariant under a common shift of nodes and we chose the

increments $d_1 = c_2 - c_1$, $d_j = c_j - c_2$, $j = 3, \dots, s$ as parameters. Still, due to the mentioned dependencies between conditions, for $s = 4$ a six-parameter family of methods exists which has been derived explicitly (with quite bulky expressions).

However, optimization of stability properties such as $A(\alpha)$ -stability or error constants err_s from (56) was not possible in Maple with six free parameters. Instead, the algebraic expressions were copied to Matlab scripts for some Monte-Carlo-type search routines. The resulting coefficients of the standard method were finally approximated by rational expressions and brought back to Maple for the construction of the two boundary methods.

At first glance, having the full six-parameter family of standard methods at hand may seem to be a good work base. However, the large dimension of the search space may prevent optimal results with reasonable effort. This can be seen below, where the restriction to symmetric nodes or singly-implicit methods yielded methods with smaller err_4 than automated global searches.

$A(\alpha)$ -stability of the method may be checked [13] by considering the eigenvalue problem for the stability matrix $(A - zK)^{-1}B$, i.e.,

$$Bx = \lambda(A - zK)x \iff K^{-1}(A - \lambda^{-1}B)x = zx. \quad (71)$$

as an eigenvalue problem for $z \in \mathbb{C}$ where $|\lambda^{-1}| = 1$ runs along the unit circle. Since we focus on A-stable methods, exact verification of this property would have been preferable, of course, but an algebraic proof of A-stability seemed to be out of reach. It turned out that the algebraic criterion of the second author [22] is rather restrictive (often corresponding to norm estimates, Lemma 2.8 ibd). On the other hand, application of the exact Schur criterion [20,23] is not straight-forward and hardly feasible, since the (rational) coefficients of the stability polynomial are prohibitively large for optimized parameters (dozens of decimal places of the numerators).

2.4.1. Requirements for the Boundary Methods

Since Lemma 5 guarantees the existence of the two boundary methods (A_0, K_0) and (A_N, B_N, K_N) , their construction can follow after that of the standard method. Requirements for these two members of the triplet may also be weakened since they are applied once only. This relaxation applies to the order conditions as shown in Table 1, but also to stability requirements. Still, the number of conditions at the boundaries is so large that the diagonal triangular forms of K_0, K_N and A_0, A_N respectively have to be sacrificed and replaced by some triangular block structure. Compared to the computational effort of the complete boundary value problem, the additional complexity of the two boundary steps should not be an issue. However, for non-diagonal matrices K_0, K_N and $s = 4$, the four additional one-leg-conditions (61) have to be obeyed.

Weakened stability requirements mean that the last forward Peer step (23) for $n = N$ and the two adjoint Peer steps (26) for $n = N - 1$ and $n = 0$ need not be A-stable and only nearly zero stable if the corresponding stability matrices have moderate norms. This argument also applies to the two Runge–Kutta steps without solution output (22) and (27). However, the implementation of these steps should be numerically safe for stiff problems and arbitrary step sizes. These steps require the solution of two linear systems with the matrices $A_0 - hK_0J_0$, $A_N^T - hJ_N^TK_N^T$ or, rather

$$K_0^{-1}A_0 - hJ_0, \quad (K_N^{-1}A_N)^T - hJ_N^T, \quad (72)$$

where J_0, J_N are block diagonal matrices of Jacobians. These Jacobians are expected to have absolutely large eigenvalues in the left complex half-plane. For such eigenvalues, non-singularity of the matrices (72) is assured under the following eigenvalue condition:

$$\mu_0 := \min_j \operatorname{Re} \lambda_j(K_0^{-1}A_0) > 0, \quad \mu_N := \min_j \operatorname{Re} \lambda_j(K_N^{-1}A_N) > 0. \quad (73)$$

In Table 2 the constants μ_0, μ_N are displayed for all designed Peer triplets as well as the spectral radii $\rho(A_N^{-1}B_N)$ and $\rho(BA_0^{-1})$, $\rho(B_NA^{-1})$ for the boundary steps of the Peer triplet. In the search for the boundary methods with their exact algebraic parameterizations, these spectral radii were minimized, and if they were close to one the values $\mu_0, \mu_N > 0$ were maximized.

2.4.2. $A(\alpha)$ -Stable Four-Stage Methods

Although our focus is on A-stable methods, we also shortly consider $A(\alpha)$ -stable methods. We like to consider BDF4 (backward differentiation formulas) as a benchmark, since triplets based on BDF3 were the most efficient ones in [13]. In order to distinguish the different methods, we denote them in the form APso q_1q_2 aaa, where AP stands for Adjoint Peer method followed by the stage number and the smallest forward and adjoint orders q_1 and q_2 in the triplet. The trailing letters are related to properties of the method like diagonal or singly diagonal implicitness.

The Peer triplet AP4o43bdf based on BDF4 has equi-spaced nodes $c_i = i/4$, $i = 1, \dots, 4$, yielding $w = e_4$. The coefficients of the full triplet are given in Appendix A.1. Obviously the method is singly-implicit and its well-known stability angle is $\alpha = 73.35^\circ$. We also monitor the norm of the zero-stability matrix $\|A^{-1}B\|_\infty$, which may be a measure for the propagation of rounding errors. Its value is $\|A^{-1}B\|_\infty \leq 5.80$. Since BDF4 has full global order 4, the error constant from (56) is $err_4 = 0$. Still, the end methods were constructed with the local orders (4, 3) according to Table 1. The matrices of the corresponding starting method have a leading 3×3 block and a separated last stage. We abbreviate this as block structure (3 + 1). All characteristics of the boundary method are given in Table 2.

Motivated by the beneficial properties of Peer methods with symmetric nodes seen in [13,19], the nodes of the next triplet with the diagonally-implicit standard method were chosen symmetric to a common center, i.e., $c_1 + c_4 = c_2 + c_3$. Unfortunately, searches for large stability angles with such nodes in the interval $[0, 1]$ did not find A-stable methods, but the following method AP4o43dif with flip symmetric nodes and $\alpha = 84.00^\circ$, which is an improvement of 10 degrees over BDF4. Its coefficients are given in Appendix A.2. Although there exist A-stable methods with other nodes in $[0, 1]$, this method is of its own interest since its error constant $err_4 \doteq 2.5 \times 10^{-3}$ is surprisingly small. The node vector of AP4o43dif includes $c_4 = 1$, leading again to $w = e_4$. Further properties of the standard method (A, B, K) are $\|A^{-1}B\|_\infty \leq 2.01$ and the damping factor $|\lambda_2| = 0.26$. See Table 2 for the boundary methods.

2.4.3. A-Stable Methods

By extensive searches with the full six-parameter family of diagonally-implicit four-stage methods many A-stable methods were found even with nodes in $[0, 1]$. In fact, regions with A-stable methods exist in at least three of the eight octants in (d_1, d_3, d_4) -space. Surprisingly, however, for none of these methods the last node c_4 was the rightmost one. In addition, it may be unexpected that some of the diagonal elements of A and K have negative signs. This does not cause stability problems if $a_{ii}k_{ii} > 0$, $i = 1, \dots, s$, see also (72). A-stability assured, the search procedure tried to minimize a linear combination $err_s + \delta\|A^{-1}B\|_\infty$ of the error constant and the norm of the stability matrix with small $\delta < 10^{-3}$ to account for the different magnitudes of these data. As mentioned in Remark 2, it was also necessary to include the damping factor $|\lambda_2|$ in the minimization process. One of the best A-stable standard methods with general nodes was named AP4o43dig. Its coefficients are given in Appendix A.3. It has an error constant $err_4 \leq 0.0260$ and damping factor $|\lambda_2| \doteq 0.798$, see Table 3. No block structure was chosen in the boundary methods in order to avoid large norms of the mixed stability matrices BA_0^{-1} , B_NA^{-1} , $A_N^{-1}B_N$, see Table 2.

It was observed that properties of the methods may improve, if the nodes have a wider spread than the standard interval $[0, 1]$. In our setting, the general vector w allows for an end evaluation $y_h(T)$ at any place between the nodes. Since an evaluation roughly

in the middle of all nodes may have good properties, in a further search the nodes were restricted to the interval $[0, 2]$. Indeed, all characteristic data of the method AP4o43die with extended nodes presented in Appendix A.4 have improved. As mentioned before, the standard method is invariant under a common node-shift and a nearly minimal error constant was obtained with the node increments $d_1 = \frac{10}{11}$, $d_3 = -1$ and $d_4 = \frac{2}{3}$. Then, the additional freedom in the choice of c_2 was needed for the boundary methods, since the conditions (73) could only be satisfied in a small interval around $c_2 = \frac{5}{4}$. The full node vector with this choice has alternating node increments since $c_3 < c_1 < c_2 < c_4$. The method is A-stable, its error constant $err_4 \leq 0.0136$ is almost half as large as for the method AP4o43dig, and $\|A^{-1}B\|_\infty \leq 6.1$ and $|\lambda_2| \leq \frac{2}{3}$ are smaller, too. The data of the boundary methods can be found in Table 2.

Table 2. Properties of the boundary methods of Peer triplets.

Triplet	Starting Method			End Method			
	Blocks	μ_0	$\varrho(BA_0^{-1})$	Blocks	μ_N	$\varrho(A_N^{-1}B_N)$	$\varrho(B_NA^{-1})$
AP4o43bdf	3+1	5.47	1	1+3	3.81	1	1.15
AP4o43dif	3+1	6.27	1	1+3	4.40	1	1.03
AP4o43dig	4	0.99	1	4	0.89	1.001	1
AP4o43sil	3+1	1.88	1	4	0.72	1	1.03
AP4o43die	3+1	3.80	1	1+3	0.66	2.6	1.98
AP3o32f	1+1+1	1.50	1.02	1+2	0.94	1	1

Table 3. Properties of the standard methods of Peer triplets.

s	Triplet	Nodes	α	$\ A^{-1}B\ _\infty$	$ \lambda_2 $	err_s	Remarks
4	AP4o43bdf	BDF4	73.35°	5.79	0.099	0	singly-implicit
	AP4o43dif	[0, 1]	84.0°	2.01	0.26	0.0025	diag.-implicit
	AP4o43dig	[0, 1]	90°	24.5	0.798	0.0260	$c_3 = 1$
	AP4o43sil	[0, 1]	90°	32.2	0.60	0.0230	$c_3 = 1$, sing.impl.
	AP4o43die	[0, 2]	90°	6.08	0.66	0.0135	nodes alternate
3	AP3o32f	[0, 1]	90°	15.3	0.91	0.0170	nodes alternate

For medium-sized ODE problems, where direct solvers for the stage equations may be used, an additional helpful feature is diagonal singly-implicitness of the standard method. In our context this means that the triangular matrices $K^{-1}A$ and AK^{-1} have a constant value θ in the main diagonal. Using the ansatz

$$a_{ii} = \theta \kappa_{ij}, \quad i = 1, \dots, s,$$

for $A = (a_{ij})$ and $K = (\kappa_{ij})$, the order conditions from Table 1 lead to a cubic equation for θ with no rational solutions, in general. In order to avoid pollution of the algebraic elimination through superfluous terms caused by rounding errors, numerical solutions for this cubic equations were not used until the very end. This means that also the order conditions for the boundary methods were solved with θ as a free parameter, only the final evaluation of the coefficients in Appendix A.5 was performed with its numerical value. In addition, in the Matlab-search for A-stable methods, the cubic equation was solved numerically and it turned out that the largest positive solution gave the best properties. Hence, this Peer triplet was named AP4o43sil. For a first candidate with nearly minimal $err_s + \delta\|A^{-1}B\|_\infty$, the damping factor $\gamma \doteq 0.89$ was again too close to one to ensure superconvergence in numerical tests, see Remark 2. However, further searches nearby minimizing the damping factor found a better standard method with $|\lambda_2| = 0.60$. Its nodes are $\mathbf{c}^T = \left(\frac{1}{50}, \frac{3}{5}, 1, \frac{41}{85}\right)$, the diagonal parameter $\theta = 3.34552931287687520$ is the largest zero of the cubic equation

$$112673616 \theta^3 + 106686908 \theta^2 - 2102637319 \theta + 1621264295 = 0.$$

Further properties of the standard method are A-stability, nodes in $(0, 1]$ with $c_3 = 1$, norm $\|A^{-1}B\|_\infty = 32.2$, and an error constant $err_4 = 0.0230$. The end method (A_N, B_N, K_N) has full matrices A_N, K_N , see Table 2.

For the sake of completeness, we also present an A-stable diagonally-implicit three-stage method, since in the previous paper [13] we could not find such methods with reasonable parameters. After relaxing the order conditions by using superconvergence, such methods exist. Applying all conditions for forward order $s = 3$ and adjoint order $s - 1 = 2$, there remains a five-parameter family depending on the node differences $d_1 = c_2 - c_1$, $d_3 = c_3 - c_2$ and three elements of A or K . A-stable methods exist in all four corners of the square $[-\frac{1}{2}, \frac{1}{2}]^2$ in the (d_1, d_3) -plane, the smallest errors err_3 were observed in the second quadrant. The method AP3o32f with $(d_1, d_3) = (-\frac{5}{27}, \frac{2}{5})$ has a node vector with $c_3 = 1$. The coefficients can be found in Appendix A.6. The characteristic data are $err_3 \leq 0.017$, $\|A^{-1}B\|_\infty \leq 15.3$, $|\lambda_2| = 0.91$. The starting method is of standard form with lower triangular A_0 and diagonal K_0 .

The main properties of the newly developed Peer triplets are summarized in Table 3 for the standard methods and Table 2 for the boundary methods.

3. Results

We present numerical results for all Peer triplets listed in Table 3 and compare them with those obtained for the third-order four-stage one-step W-method Ros3wo proposed in [5] which is linearly implicit (often called semi-explicit in the literature) and also A-stable. All calculations have been carried out with Matlab-Version R2019a, using the nonlinear solver *fsolve* to approximate the overall coupled scheme (22)–(27) with a tolerance 10^{-14} . To illustrate the rates of convergence, we consider three nonlinear optimal control problems, the Rayleigh problem, the van der Pol oscillator, and a controlled motion problem. A linear wave problem is studied to demonstrate the practical importance of A-stability for larger time steps.

3.1. The Rayleigh Problem

The first problem is taken from [24] and describes the behaviour of a tunnel-diode oscillator. With the electric current $y_1(t)$ and the transformed voltage at the generator $u(t)$, the Rayleigh problem reads

$$\text{Minimize } \int_0^{2.5} (u(t)^2 + y_1(t)^2) dt \quad (74)$$

$$\text{subject to } y_1''(t) - y_1'(1.4 - 0.14y_1'(t)^2) + y_1(t) = 4u(t), \quad t \in (0, 2.5], \quad (75)$$

$$y_1(0) = y_1'(0) = -5. \quad (76)$$

Introducing $y_2(t) = y_1'(t)$ and eliminating the control $u(t)$ yields the following nonlinear boundary value problem (see [5] for more details):

$$y_1'(t) = y_2(t), \quad (77)$$

$$y_2'(t) = -y_1(t) + y_2(t)(1.4 - 0.14y_2(t)^2) - 8p_2(t), \quad (78)$$

$$y_1(0) = -5, y_2(0) = -5, \quad (79)$$

$$p_1'(t) = p_2(t) - 2y_1(t), \quad (80)$$

$$p_2'(t) = -p_1(t) - (1.4 - 0.42y_2(t)^2)p_2(t), \quad (81)$$

$$p_1(2.5) = 0, p_2(2.5) = 0. \quad (82)$$

To study convergence orders of our new methods, we compute a reference solution at the discrete time points $t = t_n$ by applying the classical fourth-order RK4 with

$N = 1280$ steps. Numerical results for the maximum state and adjoint errors are presented in Figure 1 for $N+1 = 20, 40, 80, 160, 320$. AP3o32f and Ros3wo show their expected orders (3,2) and (3,3) for state and adjoint solutions, respectively. Order three for the adjoint solutions is achieved by all new four-stage Peer methods. The smaller error constants of AP4o43bdf, AP4o43dif and Ros3wo are clearly visible. The additional superconvergence order four for the state solutions shows up for AP4o43die and AP4o43sil and nearly for AP4o43dif and AP4o43dig. AP4o43bdf does not reach its full order four here, too.

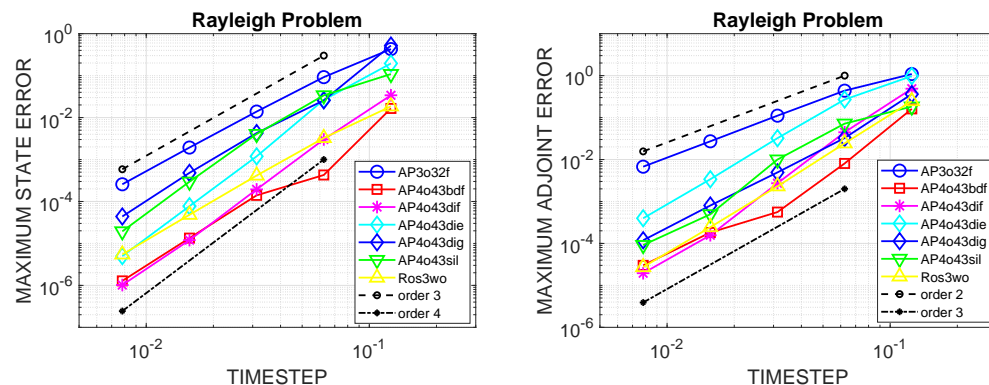


Figure 1. Rayleigh Problem: Convergence of the maximal state errors $\|(w^T \otimes I)Y_n - y(t_{n+1})\|_\infty$ (left) and adjoint errors $\|(v^T \otimes I)P_n - p(t_n)\|_\infty$ (right), $n = 0, \dots, N$.

3.2. The van der Pol Oscillator

Our second problem is the following optimal control problem for the van der Pol oscillator:

$$\text{Minimize } \int_0^2 (u(t)^2 + y(t)^2 + y'(t)^2) dt \quad (83)$$

$$\text{subject to } \varepsilon y''(t) - (1 - y(t)^2)y'(t) + y(t) = u(t), \quad t \in (0, 2], \quad (84)$$

$$y(0) = 0, \quad y'(0) = 2. \quad (85)$$

Introducing Lienhardt's coordinates $y_2(t) = y(t)$, $y_1(t) = \varepsilon y'(t) + y(t)^3/3 - y(t)$, and eliminating the control $u(t)$, we finally get the following nonlinear boundary value problem in $[0, 2]$ (see again [5] for more details):

$$y_1'(t) = -y_2(t) - \frac{p_1(t)}{2}, \quad (86)$$

$$y_2'(t) = \frac{1}{\varepsilon} \left(y_1(t) + y_2(t) - \frac{y_2(t)^3}{3} \right), \quad (87)$$

$$y_1(0) = 2\varepsilon, \quad y_2(0) = 0, \quad (88)$$

$$p_1'(t) = -\frac{1}{\varepsilon} p_2(t) - \frac{1}{\varepsilon^2} \left(y_1(t) + y_2(t) - \frac{y_2(t)^3}{3} \right), \quad (89)$$

$$p_2'(t) = p_1(t) - \frac{1}{\varepsilon} (1 - y_2(t)^2) p_2(t) \quad (90)$$

$$- \frac{2}{\varepsilon^2} \left(y_1(t) + y_2(t) - \frac{y_2(t)^3}{3} \right) (1 - y_2(t)^2) - 2y_2(t), \quad (91)$$

$$p_1(2) = 0, \quad p_2(2) = 0. \quad (92)$$

The van der Pol equation with small positive values of ε gives rise to very steep profiles in $y(t)$, requiring variable step sizes for an efficient numerical approximation. Since the factor ε^{-2} appears in the adjoint equations, the boundary value problem is even harder to solve. In order to apply constant step sizes with $N+1 = 20, 40, 80, 160, 320$, we consider the mildly stiff case with $\varepsilon = 0.1$ for our convergence study.

Numerical results for the maximum state and adjoint errors are presented in Figure 2, where a reference solution is computed with AP4o43bdf for $N = 1279$. Order three for the adjoint solutions is achieved by all new four-stage Peer methods and also by Ros3wo. The three-stage Peer method AP3o32f drops down to order 1.3. For AP4o43dig and AP4o43sil applied with $N = 19$, Matlab's *fsolve* was not able to deliver a solution. The additional superconvergence order four for the state solutions is visible for AP4o43die and nearly for AP4o43dif which performs best and beats AP4o43bdf by a factor five. AP4o43bdf does not reach its full order four here. The methods AP4o43sil and AP4o43dig show order three only, whereas AP3o32f and Ros3wo reach their theoretical order three.

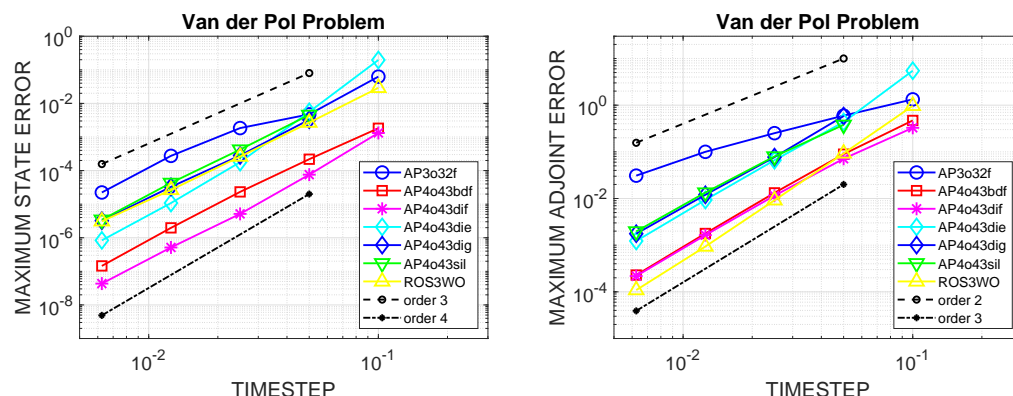


Figure 2. Van der Pol Problem: Convergence of the maximal errors $\|(w^T \otimes I)Y_n - y(t_{n+1})\|_\infty$ (left) and adjoint errors $\|(v^T \otimes I)P_n - p(t_n)\|_\infty$ (right), $n = 0, \dots, N$.

3.3. A Controlled Motion Problem

This problem was studied in [1]. The motion of a damped oscillator is controlled in a double-well potential, where the control $u(t)$ acts only on the velocity $y_2(t)$. The optimal control problem reads

$$\text{Minimize } \frac{\alpha}{2} \|y(6) - y_f\|^2 + \frac{1}{2} \int_0^6 u(t)^2 dt \quad (93)$$

$$\text{subject to } y'_1(t) - y_2(t) = 0, \quad (94)$$

$$y'_2(t) - y_1(t) + y_1(t)^3 + \nu y_2(t) = u(t), \quad t \in (0, 6], \quad (95)$$

$$y_1(0) = -1, y_2(0) = 0, \quad (96)$$

where $\nu > 0$ is the damping parameter and y_f the target final position. As in [1], we set $\nu = 1$, $y_f = (1, 0)^T$, and $\alpha = 10$.

Eliminating the scalar control $u(t)$ yields the following nonlinear boundary value problem:

$$y'_1(t) = y_2(t), \quad (97)$$

$$y'_2(t) = y_1(t) - y_1(t)^3 - \nu y_2(t) - p_2(t), \quad (98)$$

$$y_1(0) = -1, y_2(0) = 0, \quad (99)$$

$$p'_1(t) = (3y_1(t)^2 - 1)p_2(t), \quad (100)$$

$$p'_2(t) = -p_1(t) + \nu p_2(t), \quad (101)$$

$$p_1(6) = \alpha(y_1(6) - 1), p_2(6) = \alpha y_2(6). \quad (102)$$

The optimal control $u^* = -p_2^*$ must accelerate the motion of the particle to follow an optimal path (y_1^*, y_2^*) through the total energy field $E = \frac{1}{2}y_2^2 + \frac{1}{4}y_1^4 - \frac{1}{2}y_1^2$, shown in Figure 3 on the top, in order to reach the final target y_f behind the saddle point. The cost obtained from a reference solution with $N = 1279$ is $\bar{C}(y(6)) = 0.77674$, which is in good agreement with the lower order approximation in [1]. Numerical results for the maximum state and adjoint errors are presented in Figure 3 for $N+1 = 10, 20, 40, 80, 160$. Worthy of

mentioning is the repeated excellent performance of AP4o43bdf and AP4o43dif, but also the convincing results achieved by the third-order method Ros3wo. All theoretical orders are well observable, except for AP4o43dig, which tends to order three for the state solutions. A closer inspection reveals that this is caused by the second state y_2 , while the first one asymptotically converges with fourth order. However, the three methods AP4o43die, AP4o43dig, and AP4o43sil perform quite similar. Observe that AP3o32f has convergence problems for $N=9$.

The stagnation of the state errors for the finest step sizes is due to the limited accuracy of Matlab's *fsolve* – a fact which was already reported in [17].

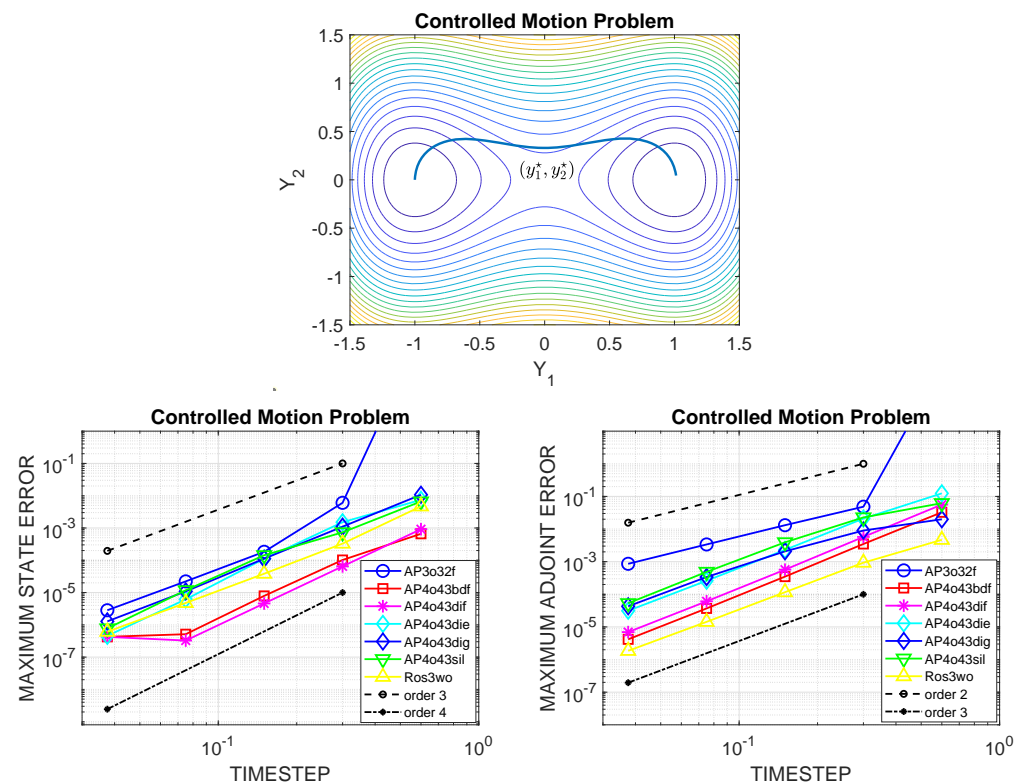


Figure 3. Controlled Motion Problem: optimal path (y_1^*, y_2^*) through the total energy field $E = \frac{1}{2}y_2^2 + \frac{1}{4}y_1^4 - \frac{1}{2}y_1^2$ visualized by isolines and exhibiting a saddle point structure (top). Convergence of the maximal state errors $\|(w^T \otimes I)Y_n - y(t_{n+1})\|_\infty$ (bottom left) and adjoint errors $\|(v^T \otimes I)P_n - p(t_n)\|_\infty$ (bottom right), $n = 0, \dots, N$.

3.4. A Wave Problem

The fourth problem is taken from [25] and demonstrates the practical importance of A-stability. We consider the optimal control problem

$$\text{Minimize } y_1(1) + \frac{1}{2} \int_0^1 u(t)^2 dt \quad (103)$$

$$\text{subject to } y_1''(t) + (2\pi\kappa)^2 y_1(t) = u(t), \quad t \in (0, 1], \quad (104)$$

$$y_1(0) = y_1'(0) = 0, \quad (105)$$

where $\kappa = 16$ is used. Introducing $y_2(t) = y_1'(t)$ and eliminating the control $u(t)$ yields the following linear boundary value problem:

$$y_1'(t) = y_2(t), \quad (106)$$

$$y_2'(t) = -(2\pi\kappa)^2 y_1(t) - p_2(t), \quad (107)$$

$$y_1(0) = 0, y_2(0) = 0, \quad (108)$$

$$p_1'(t) = (2\pi\kappa)^2 p_2(t), \quad (109)$$

$$p_2'(t) = -p_1(t), \quad (110)$$

$$p_1(1) = 1, p_2(1) = 0. \quad (111)$$

The exact solutions are given by

$$y_1^*(t) = \frac{1}{2(2\pi\kappa)^3} \sin(2\pi\kappa t) - \frac{t}{2(2\pi\kappa)^2} \cos(2\pi\kappa t), \quad (112)$$

$$y_2^*(t) = \frac{t}{2(2\pi\kappa)} \sin(2\pi\kappa t), \quad (113)$$

$$p_1^*(t) = \cos(2\pi\kappa t), p_2^*(t) = \frac{1}{2\pi\kappa} \sin(2\pi\kappa t), \quad (114)$$

and the optimal control is $u^*(t) = -p_2^*(t)$. The key observation here is that the eigenvalues of the Jacobian of the right-hand side in (106)–(111) are $\lambda_{1/2} = 2\pi\kappa i$ and $\lambda_{3/4} = -2\pi\kappa i$, which requests appropriate step sizes for only the $A(\alpha)$ -stable methods AP4o43bdf and AP4o34dif due to their stability restrictions along the imaginary axis. Indeed, a closer inspection of the stability region of the (multistep) BDF4 method near the origin reveals that $|\lambda h_{bdf4}| \leq 0.3$ is a minimum requirement to achieve acceptable approximations for problems with imaginary eigenvalues and moderate time horizon. For the four-stage AP4o43bdf with step size $h = 4h_{bdf4}$, this yields $|\lambda h| \leq 1.2$ and hence $h \leq 1/(32\pi) \approx 0.02$ for the wave problem considered. A similar argument applies to AP4o34dif, too. Numerical results for the maximum state and adjoint errors are plotted in Figure 4 for $N+1 = 20, 40, 80, 160, 320$. They clearly show that both methods deliver first feasible results for $h = 1/80$ and below only, but then again outperform the other Peer methods. Once again, Ros3wo performs remarkably well. The orders of convergence for the adjoint solutions are one better than the theoretical values, possibly due to the overall linear structure of the boundary value problem.

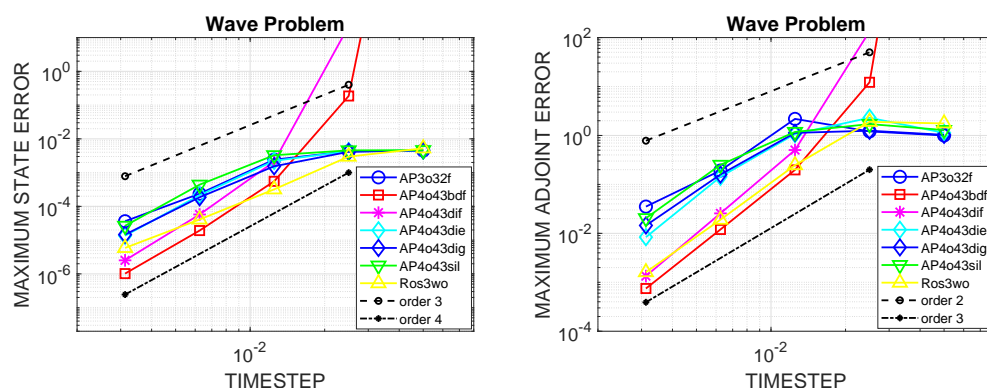


Figure 4. Wave Problem: Convergence of the maximal state errors $\|(w^T \otimes I)Y_n - y(t_{n+1})\|_\infty$ (left) and adjoint errors $\|(v^T \otimes I)P_n - p(t_n)\|_\infty$ (right), $n = 0, \dots, N$.

4. Discussion

We have extended our three-stage adjoint Peer two-step methods constructed in [13] to four stages to not only improve the convergence order of the methods but also their stability. Combining superconvergence of a standard Peer method with a careful design of starting

and end Peer methods with appropriately enhanced structure, discrete adjoint A-stable Peer methods of order (4,3) could be found. Still, new requirements had to be dealt with for the higher order pair (4,3). The property of A-stability comes with larger error constants and a few other minor structural disadvantages. As long as A-stability is not an issue to solve the boundary value problem arising from eliminating the control from the system of KKT conditions, a Peer variant AP4o43bdf of the $A(73.35^\circ)$ -stable BDF4 and the $A(84^\circ)$ -stable AP4o43dif are the most attractive methods, which perform equally well depending on the problem type. The A-stable methods AP4o43dig and AP4o43die with diagonally implicit standard Peer methods are very good alternatives if eigenvalues close or on the imaginary axis are existent. We have also constructed the A-stable method AP4o43sil with a singly-diagonal main part as an additional option if large linear systems can be still solved by a direct solver and hence the property of requesting one LU-decomposition only is highly valuable. In future work, we plan to train our novel methods in a projected gradient approach to also tackle large-scale PDE constrained optimal control problems with semi-discretizations in space. In these applications, Peer triplets may have to satisfy even more severe requirements.

Author Contributions: Conceptualization, J.L. and B.A.S.; investigation, J.L. and B.A.S.; software, J.L. and B.A.S. All authors have read and agreed to the published version of the manuscript.

Funding: The first author is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) within the collaborative research center TRR154 “Mathematical modeling, simulation and optimisation using the example of gas networks” (Project-ID 239904186, TRR154/2-2018, TP B01).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

In what follows, we will give the coefficient matrices which define the Peer triplets discussed above. We have used the symbolic option in Maple as long as possible to avoid any roundoff errors which would pollute the symbolic manipulations by a great number of superfluous terms. If possible, we provide exact rational numbers for the coefficients and give numbers with 16 digits otherwise. It is sufficient to only show pairs (A_n, K_n) and the node vector \mathbf{c} , since all other parameters can be easily computed from the following relations:

$$B_n = (A_n V_s - K_n V_s \tilde{E}_s) P_s V_s^{-1},$$

$$a = A_0 \mathbf{1}, \quad b = A_0 c - K_0 \mathbf{1}, \quad w = V_s^{-T} \mathbf{1}, \quad v = V_s^{-T} e_1, \quad s = 3, 4,$$

with $e_1 = (1, 0, \dots, 0)^T \in \mathbb{R}^s$ and the special matrices

$$V_s = (\mathbf{1}, \mathbf{c}, \mathbf{c}^2, \dots, \mathbf{c}^{s-1}), \quad P_q = \left(\binom{j-1}{i-1} \right)_{i,j=1}^s, \quad \tilde{E}_s = (i \delta_{i+1,j})_{i,j=1}^s.$$

Appendix A.1. Coefficients of AP4o54bdf

$$\mathbf{c}^T = \left(\frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1 \right)$$

$$A_0 = \begin{pmatrix} 2 & \frac{1}{2} & & \\ -\frac{265}{96} & \frac{17}{96} & \frac{11}{288} & \\ \frac{7}{6} & -\frac{47}{24} & \frac{25}{12} & \\ -\frac{21}{32} & \frac{227}{96} & -\frac{1163}{288} & \frac{25}{12} \end{pmatrix}, K_0 = \begin{pmatrix} \frac{1}{2} & & & \\ -\frac{77}{192} & \frac{3}{32} & & \\ \frac{67}{192} & \frac{17}{96} & \frac{155}{576} & \\ -\frac{19}{192} & -\frac{17}{192} & 0 & \frac{1}{4} \end{pmatrix}$$

$$A = \begin{pmatrix} \frac{25}{12} & & & \\ -4 & \frac{25}{12} & & \\ 3 & -4 & \frac{25}{12} & \\ -\frac{4}{3} & 3 & -4 & \frac{25}{12} \end{pmatrix}, K = \begin{pmatrix} \frac{1}{4} & & & \\ & \frac{1}{4} & & \\ & & \frac{1}{4} & \\ & & & \frac{1}{4} \end{pmatrix}$$

$$A_N = \begin{pmatrix} \frac{635}{96} & & & \\ -\frac{1235}{72} & \frac{35}{32} & \frac{67}{96} & -\frac{43}{288} \\ \frac{4475}{288} & -\frac{35}{24} & 0 & \frac{43}{72} \\ -5 & \frac{35}{96} & -\frac{67}{96} & \frac{53}{96} \end{pmatrix}, K_N = \begin{pmatrix} \frac{25}{32} & & & \\ -\frac{5}{3} & \frac{61}{192} & -\frac{1}{192} & \\ \frac{115}{64} & -\frac{13}{48} & \frac{23}{64} & \\ -\frac{185}{288} & \frac{13}{96} & -\frac{1}{192} & \frac{43}{576} \end{pmatrix}$$

Appendix A.2. Coefficients of $AP_{4o43dif}$

$$\mathbf{c}^T = \left(\frac{3}{22}, \frac{53}{132}, \frac{97}{132}, 1 \right)$$

$$A_0 = \begin{pmatrix} 1.1582197171362010 & 0.04624378638947835 & & \\ -0.7020381998219871 & 1.55936795391810600 & 0.02624560436867219 & \\ -0.5084723270832399 & -3.71639374160988500 & 2.21790664582949000 & \\ & 2.30412222276248700 & -2.66055187655540200 & 1.275350214666080 \end{pmatrix}$$

$$K_0 = \begin{pmatrix} 0.10630513138050800 & & & \\ 0.39188176473763570 & 0.18135555683856680 & & \\ -0.05671611789968874 & 0.06216151000641002 & 0.3773797141792473 & \\ -0.08101130220826174 & -0.03462160050989925 & & 0.13000000000000000 \end{pmatrix}$$

$$A = \begin{pmatrix} 2.713996187194519 & & & \\ -5.753019558612675 & 2.063116456071261 & & \\ 5.300000000000000 & -4.392801539381829 & 2.202673482804081 & \\ -2.313267438350870 & 2.523025304770754 & -2.619073109161321 & 1.275350214666080 \end{pmatrix}$$

$$K = \text{diag}(0.2212740342685062, 0.2910929443629617, 0.3576330213685321, 0.13)$$

$$A_N = \begin{pmatrix} 3.321208926131899 & & & \\ -6.825690771130220 & 1.096545539465253 & 0.5537291752348234 & -0.08772005153993665 \\ 5.504481844998321 & -1.589672639210835 & 0.3213889335444567 & 0.44690680951897520 \\ -2.000000000000000 & 0.493127099745582 & -0.8751181087792801 & 0.64081324202096150 \end{pmatrix}$$

$$K_N = \begin{pmatrix} 0.4780945554021703 & & & \\ -0.750000000000000 & 0.3443968810148793 & 0.03064999258349816 & \\ 0.9263584006629529 & -0.2474620802680682 & 0.34743387239297130 & \\ -0.4116869618629235 & 0.1378269814151266 & 0.03853141924782626 & 0.06599889592052376 \end{pmatrix}$$

Appendix A.3. Coefficients of $AP_{4043dig}$

$$\mathbf{c}^T = \left(\frac{139}{1159}, \frac{11}{19}, 1, \frac{1375}{2014} \right)$$

$$A_0 = \begin{pmatrix} -482.1874750642102 & 4.750000000000000 & -5.916666666666667 & -6.500000000000000 \\ 5295.612100386801 & 78.73229010791468 & 60.16394904407432 & 7.222222222222222 \\ 893.8003010294580 & -4.061724320422766 & 9.736228361340601 & 19.67879886925837 \\ -5707.694317901957 & -68.66254446706468 & -58.05689396607474 & -35.61626763467349 \end{pmatrix}$$

$$K_0 = \begin{pmatrix} -49.91295086094522 & 0.525000000000000 & -3.439024390243902 & 2.894736842105263 \\ 405.5730073881453 & 49.31516975831240 & 8.193548387096774 & -3.428571428571429 \\ 53.62032171809015 & 12.67084977396168 & -1.304859285573478 & 4.013292871986014 \\ -414.6382351541371 & -49.08870633833948 & -1.334271117642095 & -15.23643896150272 \end{pmatrix}$$

$$A = \begin{pmatrix} -2.604429828805958 & & & \\ 6.603320924494022 & 11.44234275562775 & & \\ 0.5317173544040980 & -2.710438820206414 & 3.550000000000000 & \\ -5.000000000000000 & 2.026117385005894 & 2.376616772673509 & -15.21524654319290 \end{pmatrix}$$

$$K = \text{diag}(-0.8973222553064913, 3.337407156628221, 1.164566261468968, -2.604651162790698)$$

$$A_N = \begin{pmatrix} -3.754385964912281 & 0.01222493887530562 & 1.014925373134328 & -0.1403508771929825 \\ 11.35280296428295e & 45.64990373363066 & -15.17493010383148 & -20.79910559459065 \\ 0.03205698176794144 & 2.937595714981687 & -3.756123160591242 & 2.666624105937791 \\ -7.630473981138614 & -48.59972438748765 & 18.91612789128839 & 18.27283236584584 \end{pmatrix}$$

$$K_N = \begin{pmatrix} -0.9578456075353955 & -0.3387096774193548 & 0.1194029850746269 & 0.8045112781954887 \\ 11.57142857142857 & -96.60667975350700 & -20.12500000000000 & 102.4846028390543 \\ 3.761888534906390 & -33.44512959236514 & -5.865106921633463 & 34.94704625261853 \\ -15.32045224098695 & 134.2026156894527 & 26.37615509001594 & -141.3196101415549 \end{pmatrix}$$

Appendix A.4. Coefficients of $AP_{4 \times 4 \times 3 \times 3}$

$$\mathbf{c}^T = \left(\frac{15}{44}, \frac{5}{4}, \frac{1}{4}, \frac{23}{12} \right)$$

$$A_0 = \begin{pmatrix} \frac{1573}{27} & & & \\ -\frac{29140496694667}{1728480384000} & \frac{37071572404007}{69715375488000} & \frac{37246788257}{8450348544} & \\ -\frac{98934973036237}{2160600480000} & -\frac{59311823623513}{87144219360000} & -\frac{51770824817}{13203669600} & \\ \frac{7653678714559}{1387052160000} & -\frac{7872573544487}{38730764160000} & \frac{32557703329}{23473190400} & \frac{18014543}{144484600} \end{pmatrix}$$

$$K_0 = \begin{pmatrix} \frac{110}{9} & & & \\ \frac{4}{5} & \frac{168095353644187}{920242956441600} & -\frac{136571614975979}{36809718257664} & \\ -\frac{9857504559041}{1080300240000} & \frac{398472962076949}{11503036955520000} & -\frac{18235836500357}{18404859128832} & \\ -\frac{1423069729157}{1440400320000} & \frac{398472962076949}{7668691303680000} & \frac{136571614975979}{61349530429440} & \frac{2}{61} \end{pmatrix}$$

$$A = \begin{pmatrix} \frac{45808744223}{19505421000} & & & \\ -\frac{279428522}{187552125} & \frac{3}{4} & & \\ \frac{285647}{15004170} & \frac{22704013}{125034750} & -\frac{11}{10} & \\ \frac{1}{4} & -\frac{11824391}{41678250} & \frac{832579}{4167825} & \frac{18014543}{144484600} \end{pmatrix}$$

$$K = \text{diag} \left(\frac{35085281}{25006950}, \frac{2300653}{8335650}, -\frac{1780019}{2500695}, \frac{2}{61} \right)$$

$$A_N = \begin{pmatrix} \frac{46268635184890481}{6231747944448000} & & & \\ -\frac{843924159892681}{239682613248000} & \frac{2095498352743}{2535104563200} & \frac{240331128931}{253510456320} & -\frac{10279031063}{122060590080} \\ & -4 & \frac{3733202770769}{25351045632000} & -\frac{8883867896017}{6337761408000} \\ \frac{39222471881369}{27696657530880} & -\frac{637146509711}{2816782848000} & \frac{191242568381}{352097856000} & \frac{175009899277}{8137372672000} \end{pmatrix}$$

$$K_N = \begin{pmatrix} \frac{95205971609617}{35952391987200} & & & \\ & \frac{28298016708823}{167316901171200} & -\frac{1296366536717}{4182922529280} & \\ -\frac{958020525197}{864240192000} & -\frac{2425439590003}{104573063232000} & -\frac{27877023310129}{41829225292800} & \\ -\frac{958020525197}{14980163328000} & -\frac{2425439590003}{69715375488000} & \frac{1296366536717}{6971537548800} & -\frac{22310489177}{4882423603200} \end{pmatrix}$$

Appendix A.5. Coefficients of $AP4o43sil$

$$\mathbf{c}^T = \left(\frac{1}{50}, \frac{3}{5}, 1, \frac{41}{85} \right)$$

$$A_0 = \begin{pmatrix} -18.6770976012982273 & -1.15212718448036531 & -0.684527356670693701 & \\ 30.2098963703001422 & -19.0677876392318276 & -7.55433120044842482 & \\ -9.81986015015644262 & -2.15227598175855777 & 4.86425591259034856 & \\ -\frac{57}{25} & \frac{695}{72} & 8.28572795643498617 & 9.37534909694128499 \end{pmatrix}$$

$$K_0 = \begin{pmatrix} -11.4061014637853601 & -0.0776818914116313719 & -0.278650826939386227 & \\ \frac{59}{28} & -13.9738118565057040 & 1.13881886074868390 & \\ -\frac{2498819}{583100} & 2.75133184568842863 & 0.477663277652266390 & \\ \frac{161}{25} & \frac{779}{80} & -0.352459713213735910 & 2.80235150260251923 \end{pmatrix}$$

$$A = \begin{pmatrix} -3.40824065799546119 & & & \\ -10.5240959029253065 & -6.21116392867196304 & & \\ 1.24215119761892880 & -3.47742608697035235 & 3.58958480260299081 & \\ 12.1231239821473188 & -3.03082301205062472 & 1.32154050930321039 & 9.37534909694123776 \end{pmatrix}$$

$$K = \text{diag}(-1.01874482010281778, -1.85655642136068493, 1.07294973886097804, 2.80235150260250919)$$

$$A_N = \begin{pmatrix} -3.93487127199009570 & \frac{75}{26} & -\frac{7}{8} & \\ -4.71932036763822580 & 42.7928650670737006 & -3.60582429888170880 & -36.9240613682294166 \\ -\frac{71}{202} & -2.31371287972058488 & 0.191639567731266976 & 1.90723457480523846 \\ 9.00567678814317298 & -43.3637675719685003 & 5.28918473115044182 & 35.0168267934241781 \end{pmatrix}$$

$$K_N = \begin{pmatrix} -0.687420439097535868 & & & \\ \frac{247}{72} & 24.5972883813771674 & -2.02081177860650990 & -24.7095335501766993 \\ -0.427115478996059306 & -\frac{1780}{289} & 0.897376604330415086 & 5.61580307958561348 \\ -3.39815952896990200 & -\frac{356}{17} & 1.56153637437775765 & 22.4504633222483055 \end{pmatrix}$$

Appendix A.6. Coefficients of $AP3o32f$

$$\mathbf{c}^T = \left(\frac{106}{135}, \frac{3}{5}, 1 \right)$$

$$A_0 = \begin{pmatrix} -\frac{13474483}{2809000} & & \\ \frac{2765681}{1404500} & \frac{753641}{273375} & \\ -\frac{48583191}{81461000} & -\frac{1538339}{1093500} & \frac{1783}{580} \end{pmatrix}$$

$$K_0 = \text{diag} \left(-\frac{13474483}{7155000}, \frac{2513302}{1366875}, \frac{11}{10} \right)$$

$$A = \begin{pmatrix} -\frac{11}{2} & & \\ \frac{6493}{2700} & \frac{64}{25} & \\ -\frac{25757}{78300} & -\frac{121}{100} & \frac{1783}{580} \end{pmatrix}$$

$$K = \text{diag} \left(-\frac{93}{50}, \frac{44}{25}, \frac{11}{10} \right)$$

$$A_N = \begin{pmatrix} -3 & & \\ -\frac{559409}{391500} & \frac{5418793}{1458000} & \frac{2257039}{1691280} \\ \frac{1733909}{391500} & -\frac{5418793}{1458000} & -\frac{565759}{1691280} \end{pmatrix}$$

$$K_N = \text{diag} \left(-\frac{1190159}{978750}, \frac{5418793}{3645000}, \frac{2257039}{4228200} \right)$$

References

1. Liu, X.; Frank, J. Symplectic Runge-Kutta discretization of a regularized forward-backward sweep iteration for optimal control problems. *J. Comput. Appl. Math.* **2021**, *383*, 113133.
2. Sanz-Serna, J. Symplectic Runge-Kutta schemes for adjoint equations, automatic differentiation, optimal control, and more. *SIAM Rev.* **2016**, *58*, 3–33.
3. Hairer, E.; Wanner, G.; Lubich, C. *Geometric Numerical Integration, Structure-Preserving Algorithms for Ordinary Differential Equations*; Springer Series in Computational Mathematic; Springer: Berlin/Heidelberg, Germany, 1970; Volume 31.
4. Matsuda, T.; Miyatake, Y. Generalization of partitioned Runge-Kutta methods for adjoint systems. *J. Comput. Appl. Math.* **2021**, *388*, 113308.
5. Lang, J.; Verwer, J. W-Methods in Optimal Control. *Numer. Math.* **2013**, *124*, 337–360.
6. Almuslimani, I.; Vilmart, G. Explicit stabilized integrators for stiff optimal control problems. *SIAM J. Sci. Comput.* **2021**, *43*, A721–A743.
7. Lubich, C.; Ostermann, A. Runge-Kutta approximation of quasi-linear parabolic equations. *Math. Comp.* **1995**, *64*, 601–627.
8. Ostermann, A.; Roche, M. Runge-Kutta methods for partial differential equations and fractional orders of convergence. *Math. Comp.* **1992**, *59*, 403–420.
9. Gerisch, A.; Lang, J.; Podhaisky, H.; Weiner, R. High-order linearly implicit two-step peer—Finite element methods for time-dependent PDEs. *Appl. Numer. Math.* **2009**, *59*, 624–638.
10. Gottermeier, B.; Lang, J. Adaptive Two-Step Peer Methods for Incompressible Navier-Stokes Equations. In Proceedings of the ENUMATH 2009, the 8th European Conference on Numerical Mathematics and Advanced Applications, Uppsala, Sweden, 29 June–3 July 2009; Numerical Mathematics and Advanced Applications; Kreiss, G., Lötstedt, P., Malqvist, A., Neytcheva, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; pp. 387–395.
11. Albi, G.; Herty, M.; Pareschi, L. Linear multistep methods for optimal control problems and applications to hyperbolic relaxation systems. *Appl. Math. Comput.* **2019**, *354*, 460–477.
12. Sandu, A. Reverse automatic differentiation of linear multistep methods. In *Advances in Automatic Differentiation*; Lecture Notes in Computational Science and Engineering; Bischof, C., Bücker, H., Hovland, P., Naumann, U., Utke, J., Eds.; Springer: Berlin, Germany, 2008; Volume 64, pp. 1–12.
13. Lang, J.; Schmitt, B. Discrete adjoint implicit peer methods in optimal control. *J. Comput. Appl. Math.* **2022**, *416*, 114596.
14. Hager, W. Runge-Kutta methods in optimal control and the transformed adjoint system. *Numer. Math.* **2000**, *87*, 247–282.
15. Troutman, J. *Variational Calculus and Optimal Control*; Springer: New York, NY, USA, 1996.
16. Bonnans, F.; Laurent-Varin, J. Computation of order conditions for symplectic partitioned Runge-Kutta schemes with application to optimal control. *Numer. Math.* **2006**, *103*, 1–10.
17. Herty, M.; Pareschi, L.; Steffensen, S. Implicit-Explicit Runge-Kutta schemes for numerical discretization of optimal control problems. *SIAM J. Numer. Anal.* **2013**, *51*, 1875–1899.
18. Sandu, A. On the properties of Runge-Kutta discrete adjoints. *Lect. Notes Comput. Sci.* **2006**, *3394*, 550–557.
19. Schröder, D.; Lang, J.; Weiner, R. Stability and consistency of discrete adjoint implicit peer methods. *J. Comput. Appl. Math.* **2014**, *262*, 73–86.
20. Montijano, J.; Podhaisky, H.; Randez, L.; Calvo, M. A family of L-stable singly implicit Peer methods for solving stiff IVPs. *BIT* **2019**, *59*, 483–502.
21. Schmitt, B.; Weiner, R. Efficient A-stable peer two-step methods. *J. Comput. Appl. Math.* **2017**, *316*, 319–32.
22. Schmitt, B. Algebraic criteria for A-stability of peer two-step methods. Technical Report. *arXiv* **2015**, arXiv:1506.05738.
23. Jackiewicz, Z. *General Linear Methods for Ordinary Differential Equations*; John Wiley&Sons: Hoboken, NJ, USA, 2009.
24. Jacobson, D.; Mayne, D. *Differential Dynamic Programming*; American Elsevier Publishing: New York, NY, USA, 1970.
25. Dontchev, A.; Hager, W.; Veliov, V. Second-order Runge-Kutta approximations in control constrained optimal control. *SIAM J. Numer. Anal.* **2000**, *38*, 202–226.