DOI: 10.14464/gammas.v6i1.629 GAMMAS 2024, Vol. 6, No. 1, p. 1–13

International Association
of Applied Mathematics and Mechanics
– Archive for Students –

Symplectic exponential Runge-Kutta methods for solving large nonlinear Hamiltonian systems

Till Peters [a,⋆]

[a] Institute of Numerical Mathematics, Braunschweig University of Technology, Braunschweig, Germany

received 27.09.2023, accepted 02.04.2024, published 17.04.2024

⋆ corresponding author: till.peters@tu-braunschweig.de
supervisor: Heike Faßbender, Institute of Numerical Mathematics, Braunschweig University of Technology, Braunschweig, Germany

Abstract: *We study exponential Runge-Kutta methods for solving large dimensional stiff Hamiltonian systems. Due to the Hamiltonian structure we would like to preserve this property of the system. So we describe a way to approximate the matrix exponential terms via Krylov subspace techniques.*

Keywords: Hamiltonian systems, model order reduction, Runge-Kutta methods, exponential integrators

1 Introduction

We study stiff Hamiltonian systems of the form

$$y'(t) = My(t) + f(y(t)), \quad y(t_0) = y_0, \quad t \in [t_0, T] \tag{1}$$

with a Hamiltonian matrix $M \in \mathbb{R}^{2d\times 2d}$ and a suitable nonlinear function $f : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$. Here M is called Hamiltonian if M fulfills $(JM)^T = JM$ with the skew-symmetric orthogonal matrix $J = J_{2d} = \begin{bmatrix} 0_d & I_d \\ -I_d & 0_d \end{bmatrix}$. Furthermore, M forms the main part of the stiffness in system (1) by having only eigenvalues on the imaginary axis with large magnitude. More precisely, if we denote the spectral radius of a matrix M by $\rho(M)$, we assume $\rho(M) \gg \rho\left(\frac{\partial f}{\partial y}(y)\right)$ for all $y \in \mathbb{R}^{2d}$ (see also [13, p.574]).

Exponential integrators are known to be useful for integrating stiff systems as in (1). These integrate the linear part of the differential equation exactly, which can help to loose typical stability issues in the numerical solution of stiff ordinary differential equations. Symplectic methods are suitable for integrating Hamiltonian systems because they preserve the symplectic flow of the system [6]. Combining these two aspects, Mei and Wu show in [13] starting from symplectic Runge-Kutta-Methods how to generate symplectic exponential Runge-Kutta-Methods. For large dimensions d, the evaluation of the action of a matrix exponential on a vector is costly. A standard approach to reduce these costs is the use of Krylov subspace methods to approximate the matrix-vector-products of the form $e^M v$ with $M \in \mathbb{R}^{2d\times 2d}$ and $v \in \mathbb{R}^{2d}$. This idea is proposed in [13, p.577]. In [9] or [11], the authors investigate systems like (1), without M having to be Hamiltonian, and approximate the exponential terms by the standard Arnoldi method [5, pp.499-500] and the Lanczos method [5, pp.503-504]. Now, we will also test other Krylov methods that focus on preservation of the Hamiltonian structure. We investigate different Krylov subspace methods such as the standard Arnoldi method [5, pp.499-500] or different symplectic methods like the symplectic Lanczos method [1]. Fur-

thermore, we compare the resulting approximative integrators in terms of accuracy in the matrix exponential approximation, accuracy in solving the system and preservation of the Hamiltonian structure.

This paper is structured as follows: In the first section some aspects of Hamiltonian systems and its properties are treated. Then we briefly talk about the different methods for solving (1) such as Runge-Kutta methods, exponential Runge-Kutta methods and symplectic methods. In the next section, we present different Krylov subspace methods to approximate the exponential terms appearing in the exponential Runge-Kutta methods. Afterwards, we test these approximation methods with respect to approximating the matrix exponential term itself. Then the exponential Runge-Kutta methods, approximated by different Krylov methods, are compared in terms of accuracy of computing the solution of the system and also in terms of preservation of the Hamiltonian.

2 Basics

2.1 Hamiltonian systems

Let us first take a closer look at (1). Setting $M = J^{-1}Q$ with a symmetric matrix Q and $f(y(t)) = J^{-1}\nabla U(y(t))$ with a smooth potential function U yields

$$y'(t) = J^{-1}Qy(t) + J^{-1}\nabla U(y(t)) = J^{-1}\nabla H(y(t)) \tag{2}$$

which is the classical structure of a Hamiltonian system with the Hamiltonian $H : \mathbb{R}^{2d} \to \mathbb{R}, H(y) = \frac{1}{2}y^T Qy + U(y)$ [13, p.568-569]. In the following we refer to a Hamiltonian system as

$$y'(t) = J^{-1}\nabla H(y(t)), \quad y(t_0) = y_0 \tag{3}$$

and to an autonomous differential system of order one as

$$y'(t) = g(y(t)), \quad y(t_0) = y_0. \tag{4}$$

Definition 1 (Flow of a Hamiltonian system)**.** [6, p.184] For $t \in [t_0, \infty)$ the map $\varphi_t : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ which maps the initial state of the system (3) to the solution at time t is called *flow*:

$$\varphi_t(y_0) = y(t).$$

For Hamiltonian systems the flow is always symplectic which implies for the Jacobian of the flow

$$\left(\frac{\partial \varphi_t}{\partial y_0}\right)^T J \left(\frac{\partial \varphi_t}{\partial y_0}\right) = J. \tag{5}$$

Moreover, the symplecticity of the flow is locally equivalent to the system (4) being Hamiltonian, i.e. it exists a Hamiltonian function $H : \mathbb{R}^{2d} \to \mathbb{R}$ such that $y'(t) = J^{-1}\nabla H(y(t))$. It can be shown that the Hamiltonian function H is preserved along solutions, i.e.: $\frac{d}{dt}H(y(t)) = 0$ [6, Section IV.1, Example 1.2, p.98]. The Hamiltonian generally describes the energy of the system which is an important quantity for such systems. So, we would like to attach importance to this property of preserving the Hamiltonian while solving the system. Therefore, we will use symplectic methods to solve the system which should approximate the flow of the system and preserve the Hamiltonian [6, Section VI.2, pp.182-187].

Definition 2 (Symplectic/Hamiltonian matrix)**.** [6, Definition 2.1, p.183], [7, p.381]
The matrix $S \in \mathbb{R}^{2d\times 2d}$ is called *symplectic* if $S^T JS = J$. Similarly $H \in \mathbb{R}^{2d\times 2d}$ is called a *Hamiltonian matrix* if $H^T J = -JH$ or $(JH)^T = JH$ respectively.

Lemma 3. *Let $H \in \mathbb{R}^{2d\times 2d}$ Hamiltonian and $\lambda \in \mathbb{C}$ an eigenvalue of H. Then, $-\lambda, \bar{\lambda}$ and $-\bar{\lambda}$ are eigenvalues of H too.*

Proof. Let $x \in \mathbb{C}^{2d}$ be an eigenvector of H to the eigenvalue λ such that $Hx = \lambda x$. Therefore it holds

$$H^T Jx = -JHx = -J\lambda x = -\lambda Jx.$$

Because H and H^T possess the same eigenvalues, $-\lambda$ is also an eigenvalue of H. Hence, the statement follows as the complex eigenvalues for real-valued matrices appear always in pairs λ and $\bar{\lambda}$. □

To ensure the stability of the solution to the system it is wise to restrict ourselves to matrices with eigenvalues with non-positive real part. Taking into account Lemma 3, we assume the system matrix M to have only purely imaginary eigenvalues as already stated in the introduction.

With these definitions we can easily see that $M = J^{-1}Q$ is Hamiltonian with Q symmetric as mentioned in the introduction. A symplectic matrix is always nonsingular with inverse $S^{-1} = J^T S^T J$. Moreover, we see that a similarity transformation of a Hamiltonian matrix with a symplectic matrix yields a Hamiltonian matrix again.

Lemma 4. *[2, p.2]*
Let $S \in \mathbb{R}^{2d\times 2d}$ be symplectic and $H \in \mathbb{R}^{2d\times 2d}$ Hamiltonian. Then $S^{-1}HS \in \mathbb{R}^{2d\times 2d}$ is a Hamiltonian matrix.

This aspect can be transferred to non-quadratic matrices with so called symplectic by columns matrices.

Definition 5 (Symplectic by columns matrix)**.** [2, p.3]
We call a matrix $S^{2d,2m} \in \mathbb{R}^{2d\times 2m}, m \le d$ *symplectic by columns* if $(S^{2d,2m})^T J_{2d} S^{2d,2m} = J_{2m}$.

Lemma 6. *[2, p.3]*
Let $S^{2d,2m} \in \mathbb{R}^{2d\times 2m}$, $m \leq d$ be symplectic by columns and $H \in \mathbb{R}^{2d\times 2d}$ Hamiltonian. Then it holds

- $J_{2m}^T (S^{2d,2m})^T J_{2d} \in \mathbb{R}^{2m\times 2d}$ *is the left inverse of* $S^{2d,2m}$.
- $(J_{2m}^T (S^{2d,2m})^T J_{2d}) H S^{2d,2m} \in \mathbb{R}^{2m\times 2m}$ *is Hamiltonian.*

Thus the projection with symplectic by columns matrices onto smaller dimensional spaces preserves the property of being Hamiltonian.

2.2 Runge-Kutta methods

For solving the system (1) we will investigate some special Runge-Kutta-methods which are made for the architecture of our problem.

Definition 7 (Runge-Kutta method (RK))**.** [13, p.574]
For the initial value problem

$$y'(t) = g(t, y(t)), \qquad y(t_0) = y_0 \tag{6}$$

we call the one-step method

$$Y_{n,i} = y_n + h \sum_{j=1}^{s} a_{ij} g(t_n + c_j h, Y_{n,j}), \qquad i = 1, \dots, s, \tag{7a}$$

$$y_{n+1} = y_n + h \sum_{i=1}^{s} b_i g(t_n + c_i h, Y_{n,i}) \tag{7b}$$

with $n = 0, 1, 2, \dots$ *Runge-Kutta method of order s*. Here, y_n is an approximation to $y(t_n)$ for equidistantly chosen time points $t_{n+1} = t_n + h$ with a fixed time step $h \in \mathbb{R}$.

Considering the stiffness of the system, we better use some other solving method developed for the designed structure with the linear part. This help us to solve the system more efficiently, e.g. with larger stepsizes, as we will see later.

Definition 8 (Exponential Runge-Kutta method (ERK))**.** [13, p.568-569]
For the initial value problem

$$y'(t) = My(t) + f(y(t)), \qquad y(t_0) = y_0$$

we call the one-step method

$$Y_{n,i} = e^{c_i hM} y_n + h \sum_{j=1}^{s} \bar{a}_{ij}(hM) f(Y_{n,j}), \qquad i = 1, \dots, s, \tag{8a}$$

$$y_{n+1} = e^{hM} y_n + h \sum_{i=1}^{s} \bar{b}_i(hM) f(Y_{n,i}) \tag{8b}$$

with $n = 0, 1, 2, \dots$ *exponential Runge-Kutta method of order s*. Here, $\bar{a}_{ij}(hM)$ and $\bar{b}_i(hM)$ are certain matrix exponential terms. Moreover, y_n is an approximation to $y(t_n)$ for equidistantly chosen time points $t_{n+1} = t_n + h$ with a fixed time step $h \in \mathbb{R}$.

In the case $M \to 0$ the exponential Runge-Kutta method reduces to a standard Runge-Kutta method which then is called corresponding RK. Starting from system (1), we can choose either the standard Runge-Kutta method with $g(t, y(t)) = My(t) + f(y(t))$ or the exponential Runge-Kutta method to solve the system. The ERK integrates the linear part of the differential equation with the matrix M exactly and uses an approximation for the nonlinear part. In contrast to that, the RK approximates the whole nonlinear function g directly. In the case of implicit methods, equation (8a) describes implicit equations which have to be solved e.g. with fixed-point iterations or Newton's Method. After that, equation (8b) can be simply evaluated. We choose to use here the fixed-point iterations. Furthermore, we have assumed in the introduction that the main part of the stiffness of the system should be included in the linear part with the Hamiltonian matrix, i.e. $\rho(M) \gg \rho\left(\frac{\partial f}{\partial y}(y)\right)$ for all $y \in \mathbb{R}^{2d}$. Mei and Wu have used this in [13] to show the following. In some cases where the fixed-point iterations for solving the inner equations of RK do not converge due to too large time steps, the ones for the ERK do [13, pp.574-576]. By choosing ERK, we can therefore benefit from larger time steps for integrating the system. Additionally, if the fixed-point iterations for solving the inner equations converge for RK and ERK, the ones for ERK converge much faster [13, pp.574-576]. In summary, ERK methods are well suited for integrating stiff systems.

The matrix exponential terms can be chosen in a simple way as described in the next definition.

Definition 9 (Integrating Factor Runge-Kutta Method (IFRK))**.** [3, p.A594]
The special exponential Runge-Kutta method with the coefficients

$$\bar{a}_{ij}(hM) = a_{ij} e^{(c_i - c_j)hM}, \quad \bar{b}_i(hM) = b_i e^{(1-c_i)hM}$$

for $i, j = 1, \dots, s$ is called *Integrating Factor Runge-Kutta Method.*

For the next theorem, we need to introduce the local truncation error and the order of a one-step method to compare the local accuracy of the methods.

Definition 10 (Scaled local truncation error)**.** [8, p.25]
For an one-step method of the form

$$y_{n+1} = y_n + h\phi(t_n, y_n, y_{n+1}, h)$$

we define the *scaled local truncation error* δ as

$$\delta(t_{n+1}, y(t_{n+1}), h) = \frac{1}{h}\big(y(t_{n+1}) - y(t_n) - h\phi(t_n, y(t_n), y_{n+1}, h)\big).$$

Definition 11 (Consistency order/Order)**.** [8, p.26]
If it holds

$$\delta(t_{n+1}, y(t_{n+1}), h) = \mathcal{O}(h^p)$$

for largest possible $p \in \mathbb{N}$, then the one-step method has the *consistency order* p.

For the consistency order of Runge-Kutta methods (7) there exist order conditions to the coefficients a_{ij}, b_i, c_j. If the order conditions from Table 1 are fulfilled up to order $p \in \mathbb{N}, p \leq 4$, then the RK has the order p.

Table 1 – Order conditions up to order 4 [8, p.34]

No.	Order	Order condition
1	1	$\sum_{i=1}^{s} b_i = 1$
2	2	$\sum_{i=1}^{s} b_i c_i = \frac{1}{2}$
3	3	$\sum_{i=1}^{s} b_i c_i^2 = \frac{1}{3}$
4	3	$\sum_{i,j=1}^{s} b_i a_{ij} c_j = \frac{1}{6}$
5	4	$\sum_{i=1}^{s} b_i c_i^3 = \frac{1}{4}$
6	4	$\sum_{i,j=1}^{s} b_i c_i a_{ij} c_j = \frac{1}{8}$
7	4	$\sum_{i,j=1}^{s} b_i a_{ij} c_j^2 = \frac{1}{12}$
8	4	$\sum_{i,j,k=1}^{s} b_i a_{ij} a_{jk} c_k = \frac{1}{24}$

We assume an additional condition for Runge-Kutta methods which appear later in Section 2.3:

$$c_i = \sum_{j=1}^{s} a_{ij}, \quad i = 1, \ldots, s. \tag{9}$$

Now we state a short theorem about the consistency order of the IFRK.

Theorem 12. *We assume a Runge-Kutta method with coefficients a_{ij}, b_i, c_j for $i, j = 1, \ldots s$ to be of order $p \in \mathbb{N}$. Then the associated IFRK, determined by*

$$\bar{a}_{ij}(hM) = a_{ij} e^{(c_i - c_j)hM}, \quad \bar{b}_i(hM) = b_i e^{(1-c_i)hM}$$

for $i, j = 1, \ldots, s$ is also of order p.

Proof. With setting $z(t) = e^{-tM} y(t)$ we obtain out of

$$y'(t) = My(t) + f(y(t)), \qquad y(t_0) = y_0$$

the system

$$z'(t) = e^{-tM} f(e^{tM} z(t)), \qquad z(t_0) = e^{-t_0 M} y_0,$$

where the applied Runge-Kutta method

$$Z_{n,i} = z_n + h \sum_{j=1}^{s} a_{ij} e^{-(t_n + c_j h)M} f(e^{(t_n + c_j h)M} Z_{n,j}),$$
$$z_{n+1} = z_n + h \sum_{i=1}^{s} b_i e^{-(t_n + c_i h)M} f(e^{(t_n + c_i h)M} Z_{n,i})$$

is of order p. This means for the scaled local discretization error δ_z of the Runge-Kutta method

$$\begin{aligned}\delta_z = &\frac{1}{h}\Big[z(t_{n+1}) - z(t_n) \\ &- h \sum_{i=1}^{s} b_i e^{-(t_n + c_i h)M} f(e^{(t_n + c_i h)M} Z_{n,i})\Big] \\ = &\mathcal{O}(h^p).\end{aligned}$$

Hence, it holds for the discretization error δ_y of the exponential Runge-Kutta method by backtransformation

$$\begin{aligned}\delta_y = &\frac{1}{h}\left[y(t_{n+1}) - \left(e^{hM} y(t_n) + h \sum_{i=1}^{s} \bar{b}_i(hM) f(Y_{n,i})\right)\right] \\ = &\frac{1}{h}\left[e^{t_{n+1}M} z(t_{n+1}) \right. \\ &\left. - \left(e^{hM} e^{t_n M} z(t_n) + h \sum_{i=1}^{s} b_i e^{(1-c_i)hM} f(Y_{n,i})\right)\right] \\ = &\frac{1}{h} e^{t_{n+1}M}\Big[z(t_{n+1}) - z(t_n) \\ &- h \sum_{i=1}^{s} b_i e^{-(t_n + c_i h)M} f(Y_{n,i})\Big] \\ = &e^{t_{n+1}M} \delta_z = \mathcal{O}(h^p).\end{aligned}$$

□

So the order of accuracy is transferred from standard RK to the associated IFRK which simplifies investigations on the order of these special ERK methods. But the order is not the only property which can be transferred. As shortly mentioned in Section 2.1, we need symplectic methods for solving Hamiltonian system.

Definition 13 (Symplectic method)**.** [6, p.187]
An one-step method is called *symplectic* if the iteration map ϕ_h, given by $y_{n+1} = \phi_h(y_n)$, is symplectic applied to a smooth Hamiltonian system. The symplecticity of ϕ_h is meant in the sense of (5).

Theorem 14. *[6, Theorem 4.3, p.192]*
A Runge-Kutta method, applied to the system (1) or (6) with $g(y) = My + f(y) = J^{-1}\nabla H(y)$, is symplectic if (10) holds for all $i, j = 1, \ldots, s$.

$$b_i b_j = b_i a_{ij} + b_j a_{ji}. \tag{10}$$

These simple conditions for the Runge-Kutta methods are replaced by some more complicated conditions for the exponential Runge-Kutta methods.

Theorem 15. *[13, Theorem 2.1, p.569]*
Let an exponential Runge-Kutta method with coefficients $\bar{a}_{ij} = \bar{a}_{ij}(hM), \bar{b}_i = \bar{b}_i(hM)$ and c_i for $i,j = 1,\dots,s$ be given. Furthermore, let $S = e^{hM}$ and $S_i = e^{c_i hM}$ for $i = 1,\dots,s$. If it holds that

$$\begin{aligned} \bar{b}_i^T J S S_i^{-1} &= S_i^{-T} S^T J \bar{b}_i = \gamma_i J, \\ \bar{b}_i^T J \bar{b}_j &= \bar{b}_i^T J S S_i^{-1} \bar{a}_{ij} + \bar{a}_{ji}^T S_j^{-T} S^T J \bar{b}_j, \\ \gamma_i &\in \mathbb{R} \text{ arbitrarily}, \quad i,j = 1,\dots,s, \end{aligned}$$

then the exponential Runge-Kutta method is symplectic.

Theorem 16. *[13, Theorem 3.2, p.573]*
Starting from a symplectic RK the associated IFRK is symplectic.

So, we see that the order of a RK and the property of being symplectic is directly transferred to the corresponding IFRK. In this work, we focus on implicit symplectic methods because (10) cannot be fulfilled for explicit methods.

2.3 Convergence of exponential Runge-Kutta methods

In addition to consistency the concept of convergence of one-step methods is important too. An one-step method is called convergent of order $p \in \mathbb{N}$ if for the approximated values $y_n \approx y(t_n)$ at time points $t_n = n \cdot h$, $n = 1,2,\dots,\lfloor T/h \rfloor$ it holds that

$$\left\| y_n - y(t_n) \right\| \leq C h^p. \tag{11}$$

$C > 0$ should be independent of the iteration number n or the time step h, but it is allowed to depend on the length T of the interval. For standard systems (6), this can be deduced from the consistency of an one-step method and a Lipschitz condition on the nonlinear function g. But the occurring Lipschitz constant can become very large for stiff systems, so the constant C can become even larger. If we want to keep the error $\left\| y_n - y(t_n) \right\|$ small, we have to choose very small time steps h which is not useful in practice. Another approach is to use the structure of our starting equation (1). We assume that the main part of the stiffness is contained in the matrix M, so we can use smaller Lipschitz constants or bounds at f and its derivatives. So we seek for an expression (11) for exponential Runge-Kutta methods. This section is oriented to [10] and [13], but in [10] only explicit methods are investigated.

As assumed in the beginning, M shall possess only eigenvalues on the imaginary axis. Then e^{tM} is bounded for $0 \leq t \leq T$ with a constant which is independent of h. Next, we define for an ERK the expressions $\varphi_j(hM), \psi_{i,j}(hM)$ and $\psi_j(hM)$ by

$$\begin{aligned} \varphi_j(hM) &= \frac{1}{h^j} \int_0^h e^{(h-\tau)M} \frac{\tau^{j-1}}{(j-1)!} d\tau, \\ \psi_{j,i}(hM) &= \varphi_j(c_i hM) c_i^j - \sum_{k=1}^{s} \bar{a}_{ik}(hM) \frac{c_k^{j-1}}{(j-1)!}, \\ \psi_j(hM) &= \varphi_j(hM) - \sum_{k=1}^{s} \bar{b}_k(hM) \frac{c_k^{j-1}}{(j-1)!} \end{aligned}$$

with $i,j = 1,\dots,s$. These terms are bounded because $\bar{a}_{ik}(hM)$ and $\bar{b}_k(hM)$ are built from exponential terms according to Definition 8.

Under certain conditions on the derivatives of f we can find order conditions on the exponential terms for an ERK to be convergent. Hence we define Table 2 where J and K should represent bounded operators. The proofs of the following theorems are omitted from this paper due to space limitations but can be found in [14, Kapitel 4].

Theorem 17. *Let there be a sufficiently smooth solution to (1) and let all occurring derivatives be bounded. We apply an ERK to solve (1) and assume that its coefficients fulfill the order conditions from Table 2 up to order $p \in \mathbb{N}, p \leq 4$. Then it holds*

$$\left\| y_n - y(t_n) \right\| \leq C h^p$$

with $0 \leq nh \leq T$ for the solution y_n computed by ERK. $C > 0$ is a suitable constant which does not depend on n or h, only on T.

If we focus now on IFRK, these large amount of conditions can be simplified and traced back to the order conditions of their corresponding RK.

Theorem 18. *Let a given Runge-Kutta method fulfill the order conditions from Table 1 up to order $p \in \mathbb{N}, p \leq 4$ and additionally the condition (9). Then the order conditions for the corresponding IFRK from Table 2 are fulfilled up to order p too.*

The following corollary adds these two results together to receive the desired equation (11).

Corollary 19. *Let there be a sufficiently smooth solution to (1) and let all occurring derivatives be bounded. We apply an IFRK to solve (1) and assume that the coefficients of its corresponding RK fulfill the order conditions*

Table 2 – Stiff order conditions up to order 4

No.	Order	Order condition
1	1	$\psi_1(hM) = \mathcal{O}(h^p)$
2	2	$\psi_2(hM) = \mathcal{O}(h^{p-1})$
3	2	$\sum_{i=1}^s \bar{b}_i(hM)J\psi_{1,i}(hM) = \mathcal{O}(h^{p-1})$
4	3	$\psi_3(hM) = \mathcal{O}(h^{p-2})$
5	3	$\sum_{i=1}^s \bar{b}_i(hM)J\psi_{2,i}(hM) = \mathcal{O}(h^{p-2})$
6	3	$\sum_{i=1}^s \bar{b}_i(hM)J\sum_{j=1}^s \bar{a}_{ij}(hM)J\psi_{1,j}(hM) = \mathcal{O}(h^{p-2})$
7	3	$\sum_{i=1}^s \bar{b}_i(hM)c_iK\psi_{1,i}(hM) = \mathcal{O}(h^{p-2})$
8	4	$\psi_4(hM) = \mathcal{O}(h^{p-3})$
9	4	$\sum_{i=1}^s \bar{b}_i(hM)J\psi_{3,i}(hM) = \mathcal{O}(h^{p-3})$
10	4	$\sum_{i=1}^s \bar{b}_i(hM)J\sum_{j=1}^s \bar{a}_{ij}(hM)J\psi_{2,j}(hM) = \mathcal{O}(h^{p-3})$
11	4	$\sum_{i=1}^s \bar{b}_i(hM)c_iK\psi_{2,i}(hM) = \mathcal{O}(h^{p-3})$
12	4	$\sum_{i=1}^s \bar{b}_i(hM)J\sum_{j=1}^s \bar{a}_{ij}(hM)J\sum_{k=1}^s \bar{a}_{jk}(hM)J\psi_{1,k}(hM) = \mathcal{O}(h^{p-3})$
13	4	$\sum_{i=1}^s \bar{b}_i(hM)J\sum_{j=1}^s \bar{a}_{ij}(hM)c_jK\psi_{1,j}(hM) = \mathcal{O}(h^{p-3})$
14	4	$\sum_{i=1}^s \bar{b}_i(hM)c_iK\sum_{j=1}^s \bar{a}_{ij}(hM)J\psi_{1,j}(hM) = \mathcal{O}(h^{p-3})$

from Table 1 up to order $p \in \mathbb{N}, p \leq 4$. Additionally, the condition (9) should be satisfied. Then it holds

$$\left\| y_n - y(t_n) \right\| \leq Ch^p$$

with $0 \leq nh \leq T$ for the solution y_n computed by IFRK. $C > 0$ is a suitable constant which does not depend on n or h, only on T.

3 Approximation and Krylov subspace methods

In this section, which is based on [2, pp.1-2], we talk about the approximation of $e^M v$ via different Krylov subspace methods. For large dimensions, the exact computation of matrix exponentials is very costly and takes a lot of time, so we would like to approximate the expression $e^M v$ with the aim to evaluate the matrix exponential on a matrix with much smaller dimension.

We start with a matrix $M \in \mathbb{R}^{p\times p}$ and $v \in \mathbb{R}^p$ and seek for an approximation of $f(M)v$ with a matrix function $f : \mathbb{R}^{p\times p} \to \mathbb{R}^{p\times p}$. We assume M to be sparse and $f(M) = e^{\tau M}, \tau \in \mathbb{R}$. Now we want to perform a dimension reduction to a smaller dimension $k \ll p$ using

$$V, W \in \mathbb{R}^{p\times k}, W^T V = I_k, \tilde{M} = W^T MV \in \mathbb{R}^{k\times k} \tag{12}$$

such that

$$f(M)v \approx Vf(\tilde{M})W^T v. \tag{13}$$

Additionally, if V is constructed by the choice of $Ve_1 = v/\|v\|$ for the first canonical unit vector $e_1 \in \mathbb{R}^p$, (13) can be simplified to

$$\begin{aligned} f(M)v &\approx Vf(\tilde{M})W^T v \\ &= \|v\| Vf(\tilde{M})W^T Ve_1 \\ &= \|v\| Vf(\tilde{M})e_1 \end{aligned}$$

with no need to compute W explicitly.

M is Hamiltonian in our starting equation, so we want to preserve this property for the reduced model. But in general, for chosen matrices V and W after equation (12) the projected matrix $\tilde{M}$ is not Hamiltonian again. In the following, set $p = 2d, k = 2m$. Now, we use the symplectic by columns matrices from Definition 5. Let $S^{2d,2m} \in \mathbb{R}^{2d\times 2m}$ be a symplectic by columns matrix and $M \in \mathbb{R}^{2d\times 2d}$ a Hamiltonian matrix. With the help of Lemma 6, we see that the reduced matrix $\tilde{M}$ is Hamiltonian using $V = S^{2d,2m}, W^T = J_{2m}^T (S^{2d,2m})^T J_{2d}$:

$$\tilde{M} = W^T MV = J_{2m}^T (S^{2d,2m})^T J_{2d} M S^{2d,2m} \in \mathbb{R}^{2m\times 2m}.$$

For the construction of these projection matrices, we use different Krylov subspace methods. These methods are well known for constructing subspaces with large and sparse matrices because they only require cheap matrix-vector products. But first, we give a short definition of a Krylov subspace.

Definition 20. Let $M \in \mathbb{R}^{p\times p}, v \in \mathbb{R}^p$ and $k \in \mathbb{N}$. Then the *Krylov subspace* $\mathcal{K}_k(M, v)$ is defined by

$$\mathcal{K}_k(M, v) := \text{span}\{v, Mv, M^2 v, \ldots, M^{k-1} v\}.$$

This is a list of Krylov subspace methods used here:

- Arnoldi Method(A) [5, pp.499-500]:
  This standard well known Krylov subspace method produces a sequence like

$$MV_k = V_k H_k + h_{k+1,k} v_{k+1} e_k^T,$$

  where the columns of $V_k \in \mathbb{R}^{p\times k}$, $V_k^T V_k = I_k$ form the orthonormal basis of the Krylov subspace $\mathcal{K}_k(M,v)$ and $H_k \in \mathbb{R}^{k\times k}$ is an upper Hessenberg matrix. In this case, the matrices introduced in (12) are $V = V_k$ and $\tilde{M} = H_k$.
- Symplectic Lanczos Method (SL) [1]:
  This method yields a sequence

$$MS^{2d,2m} = S^{2d,2m}\tilde{M}^{2m,2m} + \xi_{m+1} v_{m+1} e_{2m}^T$$

  with $\xi_{m+1} \in \mathbb{R}$ and a symplectic by columns matrix $S^{2d,2m} \in \mathbb{R}^{2d\times 2m}$ whose columns span the Krylov subspace $\mathcal{K}_{2m}(M,v)$. Here, the matrices from (12) are given as $V = S^{2d,2m}, W^T = J_{2m}^T (S^{2d,2m})^T J_{2d}$ and $\tilde{M} = \tilde{M}^{2m,2m}$.
- Symplectic Arnoldi Method (SA) [4, p.65]:
  In this method the orthogonal vectors from the standard Arnoldi method are reorthogonalized with respect to the standard scalar product and additionally with respect to the skew-symmetric bilinear form induced by J.
- Isotropic Arnoldi Method (IA) [4, pp.65-66]:
  In comparison to the Symplectic Arnoldi method, here the vectors are directly reorthogonalized in the Arnoldi method with respect to the standard scalar product and the skew-symmetric bilinear form induced by J.
- Hamiltonian Extended Krylov Subspace Method (HEKS) [2]:
  In this algorithm we look at the Krylov subspaces $\mathcal{K}_{2r}(M,v)$ and $\mathcal{K}_{2s}(M^{-1},M^{-1}v)$. It yields a symplectic by columns matrix $S_{r+s} \in \mathbb{R}^{2n\times 2(r+s)}$ such that its columns span the subspace $\mathcal{K}_{2r}(M,v) + \mathcal{K}_{2s}(M^{-1},M^{-1}v)$.
- Block-J-orthogonal Method (BJ) [12]:
  This method constructs an orthonormal and symplectic by columns block matrix $\left[\begin{smallmatrix} W & 0 \\ 0 & W \end{smallmatrix}\right]$ from the standard Arnoldi method applied to $\mathcal{K}_k(M,v)$.

It needs to be mentioned here that the Arnoldi method is the only method which yields a standard basis for the Krylov subspace whereas the other methods use different approaches to construct the symplecticity of the basis vectors. In the symplectic Lanczos method, the symplectic Arnoldi method, the isotropic Arnoldi method and the HEKS breakdowns can occur. To be more precise, in some cases these algorithms lead to a division by nearly zero. This requires to stop the algorithms even if the desired accuracy is not already reached. Due to these breakdowns we will see that we cannot use all of the methods in all cases.

For the Arnoldi method there exist some approximation theorems which give an upper bound on the approximation error. The following one is appropriate for Hamiltonian matrices treated here.

Theorem 21. *[9, Theorem 4, p.1918]*
Let $M \in \mathbb{C}^{p\times p}$, $v \in \mathbb{C}^p$ with $\sigma(M) \subseteq i[\alpha-2\rho, \alpha+2\rho]$ where $\alpha,\rho \in \mathbb{R}$ and $\sigma(M)$ denotes the spectrum of the matrix M. Furthermore, let V_k and H_k be generated by the Arnoldi method with $V_k e_1 = \frac{v}{\|v\|}$. Then it holds for the approximation error ε_m:

$$\begin{aligned}\varepsilon_m &= \left\| e^{\tau M} v - \|v\| V_k e^{\tau H_k} e_1 \right\| \\ &= \|v\| \left\| e^{\tau M} \frac{v}{\|v\|} - V_k e^{\tau H_k} e_1 \right\| \\ &\leq 12 \|v\| e^{-(\rho\tau)^2/k} \left(\frac{e\rho\tau}{k}\right)^k, \qquad k \geq 2\rho\tau.\end{aligned}$$

Because we have assumed M to have purely imaginary eigenvalues, the eigenvalues lie in a suitable interval with $\alpha = 0$. An important aspect is that there is no decrease in the error for $k < \rho\tau$ in general, but for $k \geq 2\rho\tau$ a rapid decay can be observed. So this theorem is strongly dependent on the spectral radius of the matrix. Our numerical experiments will demonstrate this later.

4 Numerical Experiments

In this section we compare the different Krylov methods in terms of accuracy in approximating the matrix exponential via (13) and then also their effects on the accuracy of the exponential Runge-Kutta integrators. We also investigate the preservation of the Hamiltonian regarding the different methods.

We test here mainly the Gauß-Legendre method of order 4 (RK4) which is an implicit and symplectic method with the following coefficients:

$$A = \begin{bmatrix} 1/4 & 1/4 - \sqrt{3}/6 \\ 1/4 + \sqrt{3}/6 & 1/4 \end{bmatrix},$$
$$b = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}, \qquad c = \begin{bmatrix} 1/2 - \sqrt{3}/6 \\ 1/2 + \sqrt{3}/6 \end{bmatrix}.$$

The associated IFRK of order 4 (ERK4) is then given by

$$\begin{aligned}\bar{A}(hM) &= \begin{bmatrix} \bar{a}_{11}(hM) & \bar{a}_{12}(hM) \\ \bar{a}_{21}(hM) & \bar{a}_{22}(hM) \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{4}I & \left(\frac{1}{4}-\frac{\sqrt{3}}{6}\right)e^{-\frac{\sqrt{3}}{3}hM} \\ \left(\frac{1}{4}+\frac{\sqrt{3}}{6}\right)e^{\frac{\sqrt{3}}{3}hM} & \frac{1}{4}I \end{bmatrix}, \\ \bar{b}(hM) &= \begin{bmatrix} \frac{1}{2}e^{\left(\frac{1}{2}+\frac{\sqrt{3}}{6}\right)hM} \\ \frac{1}{2}e^{\left(\frac{1}{2}-\frac{\sqrt{3}}{6}\right)hM} \end{bmatrix}, \; c = \begin{bmatrix} 1/2-\sqrt{3}/6 \\ 1/2+\sqrt{3}/6 \end{bmatrix}.\end{aligned}$$

For some treated examples there exist no analytic solutions, so we use the solution by the associated IFRK (ERK6) of the Gauß- Legendre method of order 6 (RK6) as a reference solution. Its coefficients can be found in [6, p.34]. Furthermore, we use fixed-point iterations for solving the inner equations of the methods in (7) and (8). Then we approximate the ERK4 by the mentioned Krylov subspace methods which is abbreviated by "ERK4 with A/SL/SA/IA/HEKS/BJ". The experiments are performed on an Intel(R) Core(TM) i7-1255U CPU @ 1.70 GHz with 16GB RAM. We work here with MATLAB 2022b and its exponential method `expm` for computing the matrix exponential.

4.1 Examples

We investigate three different examples for the system (1) and split the state as $y(t) = \begin{bmatrix} y^{(1)}(t) \\ y^{(2)}(t) \end{bmatrix}$ assuming an even state dimension $2N$. Their components are indexed as follows:

$$y^{(1)} = \begin{bmatrix} y^{(1)}_{(1)} & \dots & y^{(1)}_{(N)} \end{bmatrix}^T, y^{(2)} = \begin{bmatrix} y^{(2)}_{(1)} & \dots & y^{(2)}_{(N)} \end{bmatrix}^T.$$

The occurring functions and operations are meant to be evaluated componentwise and $1:N$ should abbreviate the vector $\begin{bmatrix} 1 & 2 & \dots & N \end{bmatrix}^T \in \mathbb{R}^N$, similarly $\mathbb{1} := \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}^T \in \mathbb{R}^N$. Also, the system matrix in all examples is given by

$$M = \begin{bmatrix} 0 & D \\ I & 0 \end{bmatrix}$$

where $D \in \mathbb{R}^{N\times N}$ is the sparse differentiation matrix for the second derivative from [4, p.68]:

$$D = \frac{1}{(\Delta x)^2}\begin{bmatrix} -2 & 1 & & & 1 \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ 1 & & & 1 & -2 \end{bmatrix}.$$

Furthermore, the time integration over $[0,T]$ is dependent on the parameters h and T, so in later discussions we will compute solutions y_n at a certain time step t_n, so that $y_n \approx y(t_n), t_n = n\cdot h, n = 1,\dots,T/h$.

1. Sine-Gordon equation (SG)[13, p.580]:
A discretization of the Sine-Gordon equation leads us to the system with

$$\begin{aligned} f(y(t)) &= \begin{bmatrix} -\sin(y^{(2)}(t)) \\ 0 \end{bmatrix}, \\ y_0 &= \begin{bmatrix} \sqrt{N}\left(10^{-2}+\sin\left(\frac{2\pi(1:N)}{N}\right)\right) \\ \pi\cdot\mathbb{1} \end{bmatrix}. \end{aligned}$$

Moreover, we use $N = 2^8, \Delta x = 10/N$ and for the time integration $h = 1/40, T = 100$. The Hamiltonian for this system is given by

$$\begin{aligned} H(y) =& \frac{1}{2}(y^{(1)})^T y^{(1)} - \frac{1}{2}(y^{(2)})^T D y^{(2)} \\ & - \sum_{i=1}^{N}\cos(y^{(2)}_{(i)}). \end{aligned}$$

2. Nonlinear Klein-Gordon equation (NLKG) [13, p.580]:
Similarly to the first example, we obtain a system with

$$\begin{aligned} f(y(t)) &= \begin{bmatrix} -y^{(2)}(t) - (y^{(2)}(t))^3 \\ 0 \end{bmatrix}, \\ y_0 &= \begin{bmatrix} 0 \\ 20\left(1+\cos\left(\frac{2\pi(1:N)}{N}\right)\right) \end{bmatrix}. \end{aligned}$$

Furthermore, we use $N = 2^9, \Delta x = 1.28/N, h = 1/100$,
$T = 10$. The Hamiltonian is given by

$$\begin{aligned} H(y) =& \frac{1}{2}(y^{(1)})^T y^{(1)} - \frac{1}{2}(y^{(2)})^T D y^{(2)} \\ & + \sum_{i=1}^{N}\frac{1}{2}(y^{(2)}_{(i)})^2 + \frac{1}{4}(y^{(2)}_{(i)})^4. \end{aligned}$$

3. Linear wave equation (LW) [4, pp.68-69]:
Here the system is described by

$$\begin{aligned} f(y(t)) &= \begin{bmatrix} c \\ 0 \end{bmatrix}, \\ y_0 &= \begin{bmatrix} 0 \\ \left(1+\sin^2\left(\frac{\pi L(1:N)}{N+1}\right)\right)^{-1} - 1 \end{bmatrix} \end{aligned}$$

with a time-independent vector $c \in \mathbb{R}^N$. We use $N = 400, \Delta x = 2/(N+1), h = 2.5\cdot 10^{-2}, T = 50$. The Hamiltonian is given by

$$\begin{aligned} H(y) =& \frac{1}{2}(y^{(1)})^T y^{(1)} - \frac{1}{2}(y^{(2)})^T D y^{(2)} \\ & - (y^{(2)})^T c. \end{aligned}$$

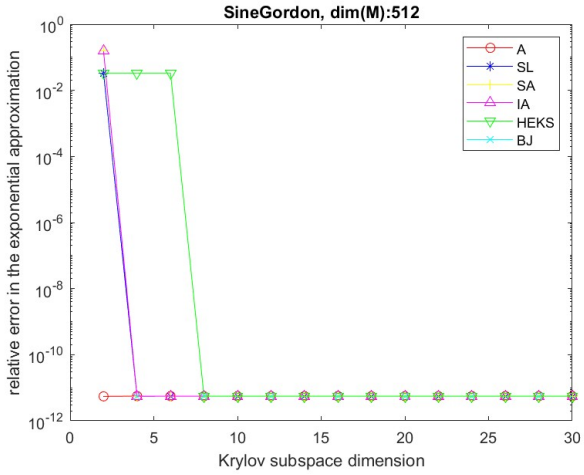

Figure 1 – Approximation of the matrix exponential for the Sine-Gordon equation

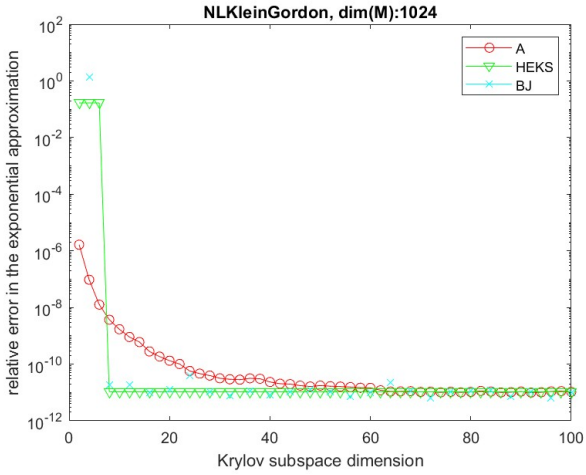

Figure 2 – Approximation of the matrix exponential for the nonlinear Klein-Gordon equation

4.2 Approximation of the matrix exponential

In this section, we investigate the error in the approximation of the matrix exponential $e^M y_0$ for the three different examples. Referring to (13) we use the relative error

$$err_{rel} = \frac{\left\| e^M y_0 - V e^{\tilde{M}} W^T y_0 \right\|}{\left\| e^M y_0 \right\|}$$

and compare the errors for the different methods for each example with increasing Krylov subspace dimension. This chosen subspace dimension is even because of the comparability between the standard Arnoldi and the other methods from Section 3.

For the first example (see Figure 1), all of the considered methods work very well. With a Krylov subspace dimension of no more than 8, a relative error of 10^{-11} is reached. The spectral radius of M is here approximately 51.2 so we can choose $2\rho = 51.2$ in Theorem 21. Then, this theorem for the approximation via the Arnoldi algorithm yields only an assertion for $k \geq 51.2$ which is not useful for our small choices of subspace dimensions.

In contrast to this, in the second example (see Figure 2) not all of the considered methods can be used to approximate the exponential term because of serious breakdowns in some of the algorithms. But for the three remaining methods, the Arnoldi, the HEKS and the BJ method we can see a decrease of the relative error with increasing subspace dimension. From a Krylov subspace dimension 22 all errors are approximately 10^{-11}. For this example, the spectral radius of M is approximately 200, so Theorem 21 is not useful, only for $k \geq 200$.

The last example (see Figure 3) shows a different behavior of the relative errors with increasing subspace dimension. Except for the HEKS method, no method

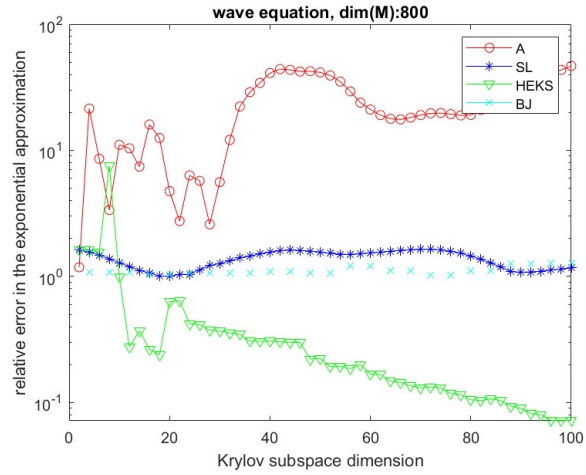

Figure 3 – Approximation of the matrix exponential for the linear wave equation

produces decreasing errors for larger subspace dimensions. Again, not all Krylov methods can be applied to this problem. The Symplectic Arnoldi and the Isotropic Arnoldi method cause some serious breakdowns. This behavior of not decreasing errors can be reasoned by the even larger spectral radius of M compared to the examples before. The eigenvalues with largest magnitude are here approximately ± 401i, so Theorem 21 yields only a statement for subspace dimensions larger than 401.

In conclusion to this section, we have seen that not all of the methods are applicable due to some breakdowns in the algorithms themselves. The spectral radius has a big influence on the course of the graph, so that due to the larger spectral radius, a decrease of the relative errors with increasing subspace dimension for the dimensions we have chosen can not be guaranteed. Theorem 21 only yields a helpful statement for larger subspace dimensions. As the authors note in [9, p.1918],

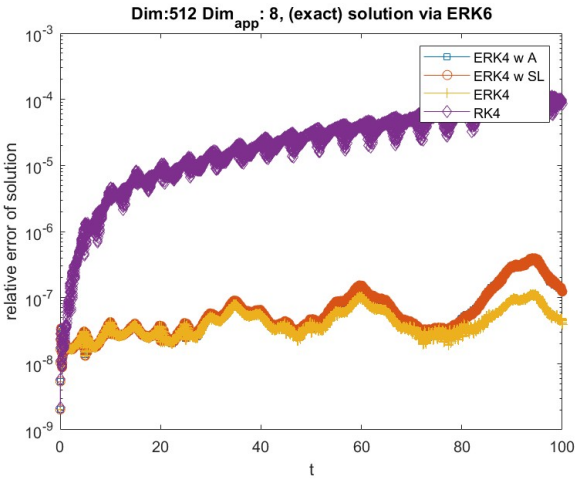

Figure 4 – Approximation of the reference solution for the Sine-Gordon equation

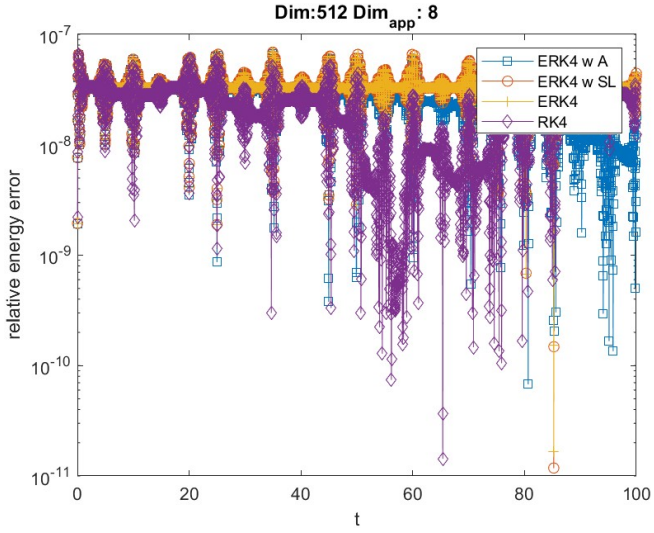

Figure 5 – Preservation of energy for the Sine-Gordon equation

in some cases we do not see a significant reduction in errors for smaller dimensions here. But the required dimensions imply many more computations, so that the time benefit of using the Krylov methods would no longer be useful. Even with smaller dimensions, the approximations often lead to satisfactory results. All in all, we stick to our choice of the small subspaces to benefit from the time advantage of the approximations by the Krylov methods.

4.3 Testing the approximated integrators on the systems

In this section, we analyze the effects on the computations of solutions of the systems if we approximate the exponential terms appearing in ERK4 with the different Krylov subspace methods. We will compute $y_n \approx y(t_n), t_n = n \cdot h, n = 1, \ldots, T/h$. So for each example, we plot the relative error in approximating the true solution of the ODE as

$$err_{rel,sol}(t_n) = \frac{\|y_n - y(t_n)\|}{\|y(t_n)\|}$$

and also the relative error in preserving the Hamiltonian

$$err_{rel,energy}(t_n) = \frac{|H(y_n) - H(y_0))|}{|H(y_0)|}$$

over time. Here we have selected fixed subspace dimensions for each example and the supremum norm to compute the errors.

First, we take a look at the first example (see example 1) where we have reduced the system dimension from 512 to 8. An exact solution is not given so we compute the relative solution error to the reference solution

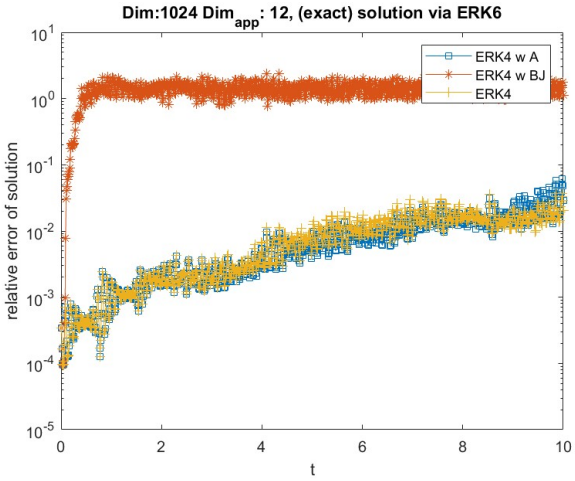

Figure 6 – Approximation of the reference solution for the nonlinear Klein-Gordon equation

computed by ERK6 with the exact exponential. The symplectic and the isotropic Arnoldi, the HEKS and the BJ methods are not applicable due to breakdowns again. We can see in Figure 4 that the ERK4 method, approximated by the Arnoldi and the symplectic Lanczos, perform the best with respect to the solution error. The non-exponential RK4 is slightly worse. Regarding the relative energy error, the methods produce similar results with an approximate error of about 10^{-7} to 10^{-8} (see Figure 5). Furthermore, we see in Table 3 that RK4 is the fastest method. Its exponential variant, the ERK4 without approximations, is much slower. In comparison between the two applicable approximation methods, the symplectic Lanczos method requires slightly more time.

Next, we see for the second example, that only the

Table 3 – Times (in sec) for the experiments of Section 4.3 for each example

Method	Sine-Gordon equation	Nonlinear Klein-Gordon equation	Linear wave equation
ERK4	90.8441	254.6947	139.6264
ERK4 with A	4.2468	5.8492	3.9903
ERK4 with SL	16.0963	-	13.3427
ERK4 with HEKS	-	-	55.8899
ERK4 with BJ	-	7.1846	4.337
RK4	1.1647	-	-

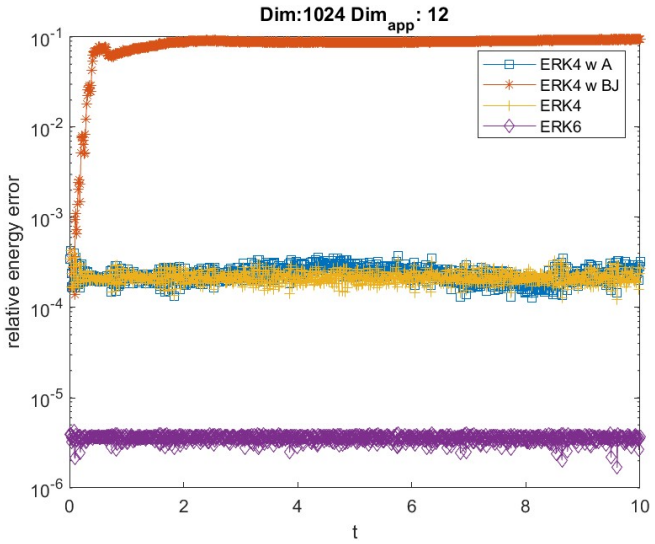

Figure 7 – Preservation of energy for the nonlinear Klein-Gordon equation

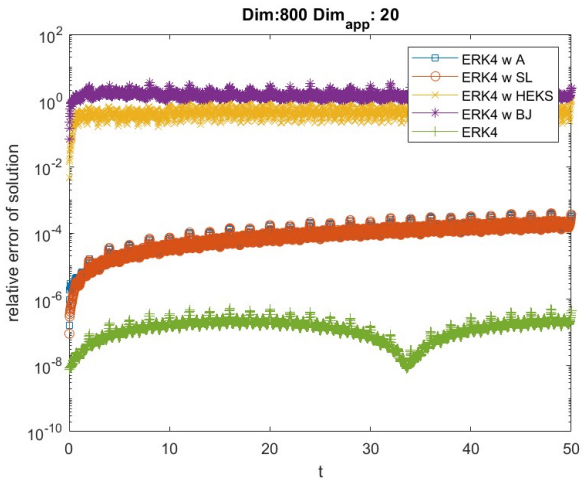

Figure 8 – Approximation of the reference solution for the linear wave equation

Arnoldi and the BJ method can be used to approximate ERK4. The reference solution is computed by ERK6 again and the RK4 method does not work either because the fixed-point iterations for the inner equations do not converge. In Figure 6 the approximation with BJ is inaccurate while the Arnoldi approximation works reliably. This can be transferred to the preservation of the Hamiltonian in Figure 7. It needs to be mentioned that the errors for this example are a bit worse than for the first which can be justified by the bigger dimension and also the different architecture of the example, e.g. the zero entries in y_0. If we look at the computation times in Table 3, we again observe the large acceleration of the ERK4 method due to the approximations using the Krylov methods.

In the last example, we reduce the dimension from 800 to 20. Here the RK4 and the ERK4 with symplectic Arnoldi and the isotropic Arnoldi do not work due to breakdowns. An analytic solution does not exist either but we can compute the solution of the system via the exponential Euler rule which is exact for this linear problem (see [4, p.60]). Regarding the relative solution error in Figure 8 the not-approximated version of ERK4 performs the best while its approximations with the Arnoldi and the symplectic Lanczos method are similar and about two orders of magnitude worse. The HEKS and BJ approximations of ERK4 are not useful with respect to approximating reference solution and the preservation of the Hamiltonian, portrayed in Figure 9. In this figure, it can be seen that the ERK4 works best again, but here the ERK4 with the symplectic Lanczos method is slightly better than its Arnoldi variant. For the comparison of the computation times of the different methods the results of example 3 are collected in Table 3. The time benefit by the approximation is clearly visible where the Arnoldi approximation is the fastest. This can be justified by the additional effort used in the other methods to create the symplecticity. For larger column dimensions of the basis matrix it is often useful to re-J-orthogonalize/reorthogonalize to maintain the symplecticity/orthogonality of the columns which unfortunately costs additionally.

As a short conclusion, we have observed that the RK4 does not work in all examples because the fixed-point iterations do not converge. In contrast to that, these

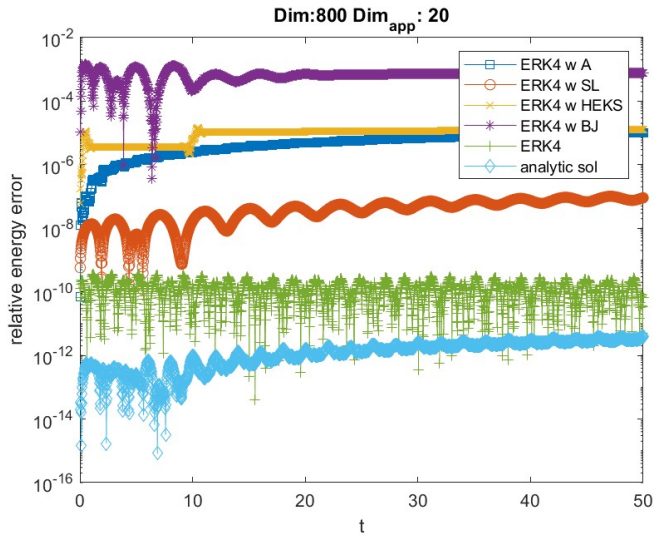

Figure 9 – Preservation of energy for the linear wave equation

problems do not occur with the ERK4, but the associated computations take considerably more time. This can be a problem with even larger examples. Among the different Krylov methods for approximating the ERK4 to speed up the computations, the Arnoldi and the symplectic Lanczos method are suitable. The Arnoldi method can be used in all treated examples and is a bit faster than the symplectic Lanczos method. But if we would like to attach more importance to the preservation of the Hamiltonian, we may choose the symplectic Lanczos method as we have seen in Figure 9. Regarding the accuracy of the computation of the solution of the system, they are approximately on the same level.

5 Summary

In this article, we approximated symplectic exponential Runge-Kutta methods by different Krylov subspace methods and then applied the resulting methods to three test examples. First, we treated some basic properties of Hamiltonian systems in the first section and introduced Hamiltonian and symplectic matrices. Next, we spoke about different variants of standard and exponential Runge-Kutta methods for solving (1). We focus on the IFRK methods because their properties can be easily deduced by their corresponding RK. After that, Krylov subspace methods were presented and how to use them for approximating the exponential terms in the exponential Runge-Kutta methods. In the last section, we tested these methods on three test examples.

A standard symplectic Runge-Kutta method is not usable in all presented examples. This is because the fixed-point iterations do not converge due to the large stiffness of the system. So, exponential Runge-Kutta methods are more suitable as they integrate the main part of the system stiffness exactly. But for large system dimensions it is very costly to evaluate the matrix exponential terms. Therefore we sought for an approximation of the appearing exponential terms in the ERK method to gain a time benefit.

We saw that the accuracy of approximation of the matrix exponential differs strongly from example to example which can be reasoned by the different spectral radii of the system matrices. Among all treated Krylov subspace methods, the Arnoldi and the symplectic Lanczos methods provide the most reliable and best approximations. The symplectic or isotropic Arnoldi method lead often to breakdowns while the approximation of the exponential integrators with the HEKS or BJ method yield bad results.This can be seen on the one hand with respect to the accuracy of the computation of the solution itself and on the other hand regarding the preservation of the Hamiltonian. In comparison between the two remaining Krylov subspace methods, the Arnoldi method is a bit faster than the symplectic Lanczos method. It can also be applied in all examples whereas the symplectic Lanczos method fails in one case. Regarding the accuracy of computing the solution of the system they perform both similarly well. But with respect to the preservation of the Hamiltonian the symplectic Lanczos has small advantages. So if additional importance to the preservation of the Hamiltonian is required, we suggest to choose the symplectic Lanczos method. In these cases where this method cannot be used, a slight increase of the subspace dimension in the Arnoldi approximation leads to satisfying results too.

Code Availability: Source Code for the experiments and methods in this paper is available from

DOI: 10.5281/zenodo.8379737.

Acknowledgements: I would like to thank Professor Heike Faßbender for her advice during my work and my colleague Michel-Niklas Senn for providing some of the codes for the Krylov subspace methods.

References

[1] Peter Benner, Heike Faßbender, and Martin Stoll. A Hamiltonian Krylov-Schur-type method based on the symplectic Lanczos process. *Linear Algebra and its Applications*, 435(3): 578–600, 2010. ISSN 0024-3795. Special Issue: Dedication to Pete Stewart on the occasion of his 70th birthday.

[2] Peter Benner, Heike Faßbender, and Michel-Niklas Senn. The Hamiltonian Extended Krylov Subspace Method. *ArXiv*, abs/2202.12640, 2022.

[3] Ashish Bhatt and Brian E. Moore. Structure-Preserving Exponential Runge-Kutta Methods. *SIAM Journal on Scientific Computing*, 39(2):A593–A612, 2017.

[4] Timo Eirola and Antti Koskela. Krylov integrators for Hamiltonian systems. *BIT Numerical Mathematics*, 59:58–76, 2019.

[5] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, third edition, 1996.

[6] Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric Numerical Integration*. Springer-Verlag Berlin Heidelberg New York, second edition, 2006.

[7] William F. Harris and João Ribeiro Cardoso. The exponential-mean-log-transference as a possible representation of the optical character of an average eye. *Ophthalmic & Physiological Optics : the Journal of the British College of Ophthalmic Opticians (Optometrists)*, 26:380–383, 08 2006.

[8] Martin Hermann. *Numerik gewöhnlicher Differentialgleichungen Band 1: Anfangswertprobleme und lineare Randwertprobleme*. De Gruyter, Berlin, Boston, 2 edition, 2017. ISBN 9783110498882.

[9] Marlis Hochbruck and Christian Lubich. On Krylov Subspace Approximations to the Matrix Exponential Operator. *SIAM Journal on Numerical Analysis*, 34(5):1911–1925, 1997.

[10] Marlis Hochbruck and Alexander Ostermann. Explicit Exponential Runge-Kutta Methods for Semilinear Parabolic Problems. *SIAM Journal on Numerical Analysis*, 43(3):1069–1090, 2005.

[11] Marlis Hochbruck, Christian Lubich, and Hubert Selhofer. Exponential integrators for large systems of differential equations. *SIAM Journal on Scientific Computing*, 19(5):1552–1574, 1998.

[12] Lu Li and Elena Celledoni. Krylov projection methods for linear hamiltonian systems. *Numerical Algorithms 81(4)*, pages 1361–1378, 2019.

[13] Lijie Mei and Xinyuan Wu. Symplectic exponential Runge-Kutta-methods for solving nonlinear Hamiltonian systems. *Journal of Computational Physics*, 338:567–584, 2017. ISSN 0021-9991.

[14] Till Peters. Symplektische exponentielle Runge-Kutta-Methoden zur Lösung großer nichtlinearer Hamiltonscher Systeme. 2024.