mathematics

MDPI

Article

Integrating Semilinear Wave Problems with Time-Dependent Boundary Values Using Arbitrarily High-Order Splitting Methods

Isaías Alonso-Mallo [1,†,‡] and Ana M. Portillo [2,*,†,‡]

[1] Departamento de Matemática Aplicada, Facultad de Ciencias, Universidad de Valladolid, Paseo de Belén 7, 47011 Valladolid, Spain; isaias@mac.uva.es

[2] Departamento de Matemática Aplicada, Escuela de Ingenierías Industriales, Universidad de Valladolid, Paseo del Cauce 59, 47011 Valladolid, Spain

[*] Correspondence: ana.portillo@uva.es

[†] Instituto de Investigación en Matemáticas de la Universidad de Valladolid (IMUVa), 47011 Valladolid, Spain.

[‡] These authors contributed equally to this work.

Abstract: The initial boundary-value problem associated to a semilinear wave equation with time-dependent boundary values was approximated by using the method of lines. Time integration is achieved by means of an explicit time method obtained from an arbitrarily high-order splitting scheme. We propose a technique to incorporate the boundary values that is more accurate than the one obtained in the standard way, which is clearly seen in the numerical experiments. We prove the consistency and convergence, with the same order of the splitting method, of the full discretization carried out with this technique. Although we performed mathematical analysis under the hypothesis that the source term was Lipschitz-continuous, numerical experiments show that this technique works in more general cases.

Keywords: splitting methods; method of lines; initial boundary-value problem; consistency; convergence

MSC: 65M12; 65M20; 65M22

check for updates

Citation: Alonso-Mallo, I.; Portillo, A.M. Integrating Semilinear Wave Problems with Time-Dependent Boundary Values Using Arbitrarily High-Order Splitting Methods. *Mathematics* **2021**, *9*, 1113. https://doi.org/10.3390/math9101113

Academic Editor: Mechthild Thalhammer

Received: 9 April 2021
Accepted: 12 May 2021
Published: 14 May 2021

1. Introduction

We consider full discretizations by means of the method of lines of a semilinear second order in time-evolutionary problems with time-dependent boundary values. For this, we first discretize in space by means of finite differences, obtaining a system of ordinary differential equations.

For time integration, we rewrite the semidiscrete system as a first-order system in time and apply an arbitrary splitting scheme. Useful descriptions of splitting methods can be found in review articles [1–3]. Splitting schemes are especially useful in the field of geometric integration. In fact, splitting integrators preserve the structural properties of the original problem flow for as long as the flow of intermediate problems does. The good performance of geometric integrators in the long-term integration of systems of Hamiltonian ODEs is well-demonstrated in [4,5].

In this paper, we thus obtain a time integrator that is explicit and has the advantage of being cheap to implement, but with the disadvantage that its stability interval is finite. However, for these second-order in-time evolutionary problems, the stability condition is acceptable, and the step size in time and space may be taken to be of a similar size. It is also possible to use implicit methods, (see, for example, [6]), where Gautschi methods are studied avoiding the order-reduction phenomenon that appears with these methods.

The way in which a splitting method works requires three steps [3]: first, by choosing how to split the problem into several simpler intermediate problems, integrating each intermediate problem either exactly or approximately, and lastly composing the solution of the intermediate problems to obtain an approximation of a certain order of the original problem.

Denoting by h the step size in space and by k the step size in time, and separating the problem, we integrate into two intermediate problems with exact flows given by $\Phi_{h,t}^{[1]}$ and $\Phi_{h,t}^{[2]}$. We consider a general splitting integrator with m stages and coefficients a_j, b_j, $j = 1, \ldots, m$,

$$\Psi_{h,k} \quad = \quad \Phi_{h,b_m k}^{[2]} \Phi_{h,a_m k}^{[1]} \ldots \Phi_{h,a_2 k}^{[1]} \Phi_{h,b_1 k}^{[2]} \Phi_{h,a_1 k}^{[1]}. \tag{1}$$

In the case of the standard method of lines, each must be integrated in time (either exactly or numerically using a sufficiently accurate quadrature formula); however, for this, the term due to the space discretization of the nonvanishing boundary values must be addressed in the same way as the source term is. Since each stage of the splitting method applied to the spatial discretization is exactly integrated, we deduced that the optimal order was achieved, at least for a fixed thickness of the spatial discretization.

We propose in this paper a technique to cheaply and effectively incorporate the boundary values to the time integration carried out by the splitting method. This technique is consistent with these values being used to approximate the spatial differential operator on the boundary. We prove consistency and convergence with optimal order of the full discretization obtained with our technique under the hypothesis that the nonlinear term of the original second order in time problem was Lipschitz-continuous; however, numerical experiments showed that this hypothesis is not necessary in practice. Moreover, numerical experiments clarified the superiority of this technique compared to the use of the standard line method, with minor errors refining the discretization in both space and time.

Throughout this paper, several constants that are independent of the time step of the time integration could be likewise denoted (usually with the letter C, and possibly with some subscript).

The paper is organized as follows. The studied problem and spatial discretization are introduced in Section 2. Section 3 is devoted to the standard method of lines, time discretization being performed with a splitting method. In Section 4, we explain the alternative method that we propose to incorporate boundary values when implementing the splitting method. Numerical experiments that clearly show the better accuracy of the proposed method versus the standard method of lines are carried out in Section 5. Mathematical analysis of the convergence is developed in Section 6, where consistency is proved, and in Section 7, where convergence is stated along a brief review of the needed linear stability.

2. Preliminaries

2.1. Partial Differential Equation

Let X and Y be Hilbert spaces, $D(A) \subset X$ a dense subspace, and $A : D(A) \subset X \to X$, $B : D(A) \subset X \to Y$ two closed linear operators. We consider the abstract second-order in time semilinear equation given by

$$\begin{array}{rcl} u''(t) & = & Au(t) + f(t, u(t)),\ t \in [0, T], \\ u(0) & = & u_0, \\ u'(0) & = & v_0, \\ Bu(t) & = & g(t), \end{array} \tag{2}$$

where source term $f : [0, T] \times D(A) \to X$ is a smooth function that is generally nonlinear. In practice, data f and g, solution u, and operators A and B are defined on a domain $\Omega \subset \mathbb{R}^n$, and they could depend on spatial variables. We did not make this dependence explicit in the abstract formulation (2) in order to simplify the notation.

We make the following hypotheses on operators A and B:

(A1) Operator B is onto.

(A2) $\mathrm{Ker}(B) = D(A_0)$ is a dense subspace of X and $A_0 = A|_{D(A_0)}$ is a negative definite self-adjoint operator. We denote $S_0 = (-A_0)^{1/2}$.

(A3) Steady-state problem

$$\begin{aligned} Ax &= 0, \\ Bx &= v \in Y, \end{aligned}$$

possesses a unique solution denoted by $x = K(0)v$, and there exists a constant C, such that linear operator $K(0) : Y \to D(A)$ satisfies

$$\|K(0)v\|_X \leq C\|v\|_Y.$$

(A4) Solution u in (2) satisfies $u(t) \in D(A)$ for $t \in [0, T]$ and is smooth enough in time.

(A5) Source term $f(t, u)$ is a Lipschitz-continuous function with respect to variable u.

Remark 1. *Because of Hypotheses (A2) and (A3), we deduce that linear problem*

$$\begin{aligned} u''(t) &= Au(t) + f(t),\ t \in [0, T], \\ u(0) &= u_0, \\ u'(0) &= v_0, \\ Bu(t) &= g(t), \end{aligned}$$

is well-posed; see [7,8]. Moreover, Hypothesis (A2) may be generalized to the case of A being a cosine operator [9,10].

2.2. Spatial Discretization

Our first step to discretize (2) by means of the method of lines is spatial discretization. Let $h \in (0, h_0]$ be a parameter that is used to measure the thickness of spatial discretization. We assumed that X_h is a family of finite-dimensional spaces that approximate X. The discrete norm in X_h is denoted by $\|\cdot\|_h$. Moreover, there is a subspace $X_{h,0} \subset X_h$ where the elements of $D(A_0)$ are well-approximated by using $P_h : D(A) \subset X \to X_{h,0}$, that is, we assumed that $P_h u$ is the best approximation when $u \in D(A_0)$. The boundary values are discretized by means of the linear operator $Q_h : Y \to X_{h,b}$, where $X_h = X_{h,0} \oplus X_{h,b}$.

Operator A is approximated by using operators $A_h : X_h \to X_{h,0}$ and $A_{h,0} = A_h|_{X_{h,0}}$. When $u \in D(A)$,

$$A_h(P_h u + Q_h Bu) = A_{h,0}P_h u + A_h Q_h Bu \approx P_h Au. \tag{3}$$

In practice, if we look for solution $u \in D(A)$ of steady-state problem

$$\begin{aligned} Au &= F, \\ Bu &= g, \end{aligned}$$

where $F \in X$ and $g \in Y$, we cannot obtain $P_h u$. Instead, we can compute $R_h u$ satisfying

$$A_h(R_h u + Q_h g) = A_{h,0}R_h u + A_h Q_h g = P_h F. \tag{4}$$

In order to discretize the source term, we suppose that function f can be defined in space $X_{h,0}$. That is, we can consider $f : [0, T] \times X_{h,0} \to X_{h,0}$, and

$$P_h f(t, u) = f(t, P_h u),$$

for each $u \in X$ and $t \in [0, T]$.

With this spatial discretization, we obtain semidiscrete ordinary differential system

$$\begin{aligned} u_h''(t) &= A_{h,0}u_h(t) + A_hQ_hg(t) + f(t,u_h(t)), \\ u_h(0) &= P_hu_0, \\ u_h'(0) &= P_hv_0. \end{aligned} \tag{5}$$

We make the following hypotheses:

(H1) There exists a constant C independent of h, such that, for $u \in D(A)$ and small enough h,

$$\|P_hu\|_h \leq C\|u\|.$$

(H2) Operator $A_{h,0}$ is symmetric and negative definite. Let $S_{h,0}$ be the symmetric and positive definite operator, such that $S_{h,0}^2 = -A_{h,0}$. We also assumed that $A_{h,0}$ and $S_{h,0}$ were invertible and their inverses were uniformly bounded on h.

(H3) There exists a subspace $Z \subset D(A)$ with norm $\|\cdot\|_Z$, such that

$$\|A_{h,0}(R_h - P_h)u\|_h \leq \varepsilon_h\|u\|_Z,$$

for each $u \in Z$, where we suppose that $\varepsilon_h \to 0$ when $h \to 0$.

(H4) $f : [0,T] \times X_{h,0} \to X_{h,0}$ is Lipschitz-continuous.

With these hypotheses, we can prove that the solution of (5) is a good approximation of the one of (2).

Theorem 1. *We assumed Hypotheses (A1–A5) and (H1–H4), that $g \in C^1([0,T],Y)$, $f \in C([0,T] \times X_{h,0}, X_{h,0})$ and $u \in C([0,T],Z)$. Then, spatial error $e_h(t) = P_hu(t) - u_h(t)$ satisfies*

$$\begin{aligned} \|e_h(t)\|_{E_h} &= \|[P_hu(t) - u_h(t), P_hu'(t) - u_h'(t)]^T\|_{E_h} \\ &= \Big(\|S_{h,0}(P_hu(t) - u_h(t))\|_h^2 + \|P_hu'(t) - u_h'(t)\|_h^2\Big)^{1/2} \leq C\varepsilon_h, \end{aligned}$$

where C only depends on T, u, u' and Lipschitz constant L.

Proof. Applying P_h to (2), considering (4), and making the difference with (5),

$$\left.\begin{aligned} P_hu''(t) - u_h''(t) &= A_{h,0}(P_hu(t) - u_h(t)) + f(t,P_hu(t)) - f(t,u_h(t)) \\ &\quad -A_{h,0}(P_hu(t) - R_hu(t)), \\ P_hu(0) - u_h(0) &= 0, \\ P_hu'(0) - u_h'(0) &= 0. \end{aligned}\right\} \tag{6}$$

Rewriting (6) as a first-order in-time problem,

$$\begin{aligned} \begin{bmatrix} P_hu(t) - u_h(t) \\ P_hu'(t) - u_h'(t) \end{bmatrix}' &= \begin{bmatrix} 0 & I_h \\ A_{h,0} & 0 \end{bmatrix}\begin{bmatrix} P_hu(t) - u_h(t) \\ P_hu'(t) - u_h'(t) \end{bmatrix} \\ &\quad + \begin{bmatrix} 0 \\ f(t,P_hu(t)) - f(t,u_h(t)) - A_{h,0}(P_hu(t) - R_hu(t)) \end{bmatrix}, \\ \begin{bmatrix} P_hu(0) - u_h(0) \\ P_hu'(0) - u_h'(0) \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \end{aligned}$$

Using that

$$\exp\left(t\begin{bmatrix} 0 & I_h \\ A_{h,0} & 0 \end{bmatrix}\right) = \begin{bmatrix} \cos(tS_{h,0}) & S_{h,0}^{-1}\sin(tS_{h,0}) \\ -\sin(tS_{h,0}) & \cos(tS_{h,0}) \end{bmatrix},$$

the variation of the constant formula and the vanishing initial conditions, we deduce that

$$\begin{bmatrix} P_h u(t) - u_h(t) \\ P_h u'(t) - u_h'(t) \end{bmatrix} =$$
$$\int_0^t \begin{bmatrix} S_{h,0}^{-1} \sin((t-s)S_{h,0})(f(s, P_h u(s)) - f(s, u_h(s)) - A_{h,0}(P_h u(s) - R_h u(s))) \\ \cos((t-s)S_{h,0})(f(s, P_h u(s)) - f(s, u_h(s)) - A_{h,0}(P_h u(s) - R_h u(s))) \end{bmatrix} ds.$$

Now, using (H2–H4), and that $u \in C([0,T], Z)$, we have

$$\begin{aligned} \|P_h u(t) - u_h(t)\|_{E_h} &= \|S_{h,0}(P_h u(t) - u_h(t))\|_h^2 + \|P_h u'(t) - u_h'(t)\|_h^2 \\ &\leq 2L \int_0^t \|P_h u(s) - u_h(s)\|_h ds + 2T\varepsilon_h \max_{t\in[0,T]} \|u(t)\|_Z. \end{aligned}$$

Applying the Gronwall lemma, we obtain for $t \in [0,T]$,

$$\|P_h u(t) - u_h(t)\|_{E_h} \leq \max_{t\in[0,T]} \|u(t)\|_Z 2Te^{2LT}\varepsilon_h.$$

□

Therefore, we can obtain a good approximation of the solution of (2) by considering the solution of (5) with a small enough value of h. Next, we use a time integrator to achieve full discretization. For this, we consider a splitting scheme. In the next two sections, we study two different ways of incorporating the boundary values with the full discretization.

3. Full Discretization: Standard Method of Lines

First, we rewrite (5) as a first-order differential problem. We denote $\mathbf{u}_h(t) = [u_{1,h}(t), u_{2,h}(t)] = [u_h(t), u_h'(t)]$ and obtain system

$$\begin{aligned} \begin{bmatrix} u_{1,h} \\ u_{2,h} \end{bmatrix}' &= \begin{bmatrix} 0 & I_h \\ A_{h,0} & 0 \end{bmatrix} \begin{bmatrix} u_{1,h} \\ u_{2,h} \end{bmatrix} + \begin{bmatrix} 0 \\ A_h Q_h g(t) \end{bmatrix} + \begin{bmatrix} 0 \\ f(t, u_{1,h}) \end{bmatrix}, \\ \begin{bmatrix} u_{1,h}(0) \\ u_{2,h}(0) \end{bmatrix} &= \begin{bmatrix} P_h u_0 \\ P_h v_0 \end{bmatrix}, \end{aligned} \tag{7}$$

of which the exact flow is denoted by $\mathbf{u}_h(t) = \Phi_{h,t}\mathbf{u}_h(0)$.

Second, we apply to (7) a splitting scheme in the usual way. We then choose a split of (7) in two intermediate problems. The first is

$$\begin{bmatrix} z_{1,h} \\ z_{2,h} \end{bmatrix}' = \begin{bmatrix} 0 & I_h \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_{1,h} \\ z_{2,h} \end{bmatrix}, \tag{8}$$

of which the exact flow is denoted by $\mathbf{z}_h(t) = \Phi_{h,t}^{I_h}\mathbf{z}_h(0)$, and the second is

$$\begin{bmatrix} \tilde{z}_{1,h} \\ \tilde{z}_{2,h} \end{bmatrix}' = \begin{bmatrix} 0 & 0 \\ A_{h,0} & 0 \end{bmatrix} \begin{bmatrix} \tilde{z}_{1,h} \\ \tilde{z}_{2,h} \end{bmatrix} + \begin{bmatrix} 0 \\ A_h Q_h g(t) \end{bmatrix} + \begin{bmatrix} 0 \\ f(t, \tilde{z}_{1,h}) \end{bmatrix}, \tag{9}$$

of which the exact flow is denoted by $\tilde{\mathbf{z}}_h(t) = \Phi_{h,t}^{A_{h,0}+g+f}\tilde{\mathbf{z}}_h(0)$.

Let $k > 0$ be a time step; we consider $t_n = nk$, $n \geq 1$, and $\mathbf{u}_{h,n} = [u_{1,h,n}, u_{2,h,n}]^T \approx [u_{1,h}(t_n), u_{2,h}(t_n)] = \mathbf{u}_h(t_n)$. We now consider a general splitting integrator with m stages and coefficients $a_j, b_j, j = 1, \dots, m$. This splitting scheme composes the flows of intermediate problems to obtain order p as follows:

$$\begin{aligned} \widetilde{\mathbf{z}}_{0,h,n} &= \mathbf{u}_{h,n}, \\ \widetilde{\mathbf{z}}_{j,h,n} &= \Phi^{A_{h,0}+g+f}_{h,b_jk} \Phi^{I_h}_{h,a_jk} \widetilde{\mathbf{z}}_{j-1,h,n}, \quad j = 1, \dots, m, \\ \mathbf{u}_{h,n+1} &= \widetilde{\mathbf{z}}_{m,h,n}, \end{aligned} \tag{10}$$

where $\widetilde{\mathbf{z}}_{j,h,n} = [\widetilde{z}_{1,j,h,n}, \widetilde{z}_{2,j,h,n}]^T$.

To more explicitly write the j-stage in (10), we first consider initial-value problem

$$\begin{aligned} z'_{1,j,h,n}(s) &= z_{2,j-1,h,n}(s), \quad s \in [0, a_jk], \\ z'_{2,j,h,n}(s) &= 0, \\ z_{1,j,h,n}(0) &= \tilde{z}_{1,j-1,h,n}(b_{j-1}k), \\ z_{2,j,h,n}(0) &= \tilde{z}_{2,j-1,h,n}(b_{j-1}k), \end{aligned}$$

and, by advancing a step a_jk in time,

$$\begin{aligned} z_{1,j,h,n}(a_jk) &= \tilde{z}_{1,j-1,h,n}(b_{j-1}k) + a_jk\tilde{z}_{2,j-1,h,n}(b_{j-1}k), \\ z_{2,j,h,n}(a_jk) &= \tilde{z}_{2,j-1,h,n}(b_{j-1}k). \end{aligned}$$

Second, we consider

$$\begin{aligned} \widetilde{z}'_{1,j,h,n}(s) &= 0, \quad s \in [0, b_jk], \\ \widetilde{z}'_{2,j,h,n}(s) &= A_{h,0}\widetilde{z}_{1,j,h,n}(s) + A_hQ_hg(t_n + k\sum_{i=1}^{j-1} b_i + s) + f(t_n + k\sum_{i=1}^{j-1} b_i + s, \widetilde{z}_{1,j,h,n}(s)), \\ \widetilde{z}_{1,j,h,n}(0) &= z_{1,j,h,n}(a_jk), \\ \widetilde{z}_{2,j,h,n}(0) &= z_{2,j,h,n}(a_jk), \end{aligned}$$

and we advance a step b_jk in time,

$$\begin{aligned} \widetilde{z}_{1,j,h,n}(b_jk) &= z_{1,j,h,n}(a_jk), \\ \widetilde{z}_{2,j,h,n}(b_jk) &= z_{2,j,h,n}(a_jk) + b_jkA_{h,0}z_{1,j,h,n}(a_jk) + A_hQ_h \int_0^{b_jk} g(t_n + k\sum_{i=1}^{j-1} b_i + \tau)d\tau \\ &\quad + \int_0^{b_jk} f(t_n + k\sum_{i=1}^{j-1} b_i + \tau, z_{1,j,h,n}(a_jk))d\tau. \end{aligned} \tag{11}$$

With this full discretization, we obtain order p in time since this is the order of the splitting method that we use, and the flows of the intermediate problems are exactly calculated when both integrals must be exactly calculated. In any case, we can always use a quadrature rule with the same accuracy as that of the splitting method. Convergence must be obtained in the discrete energy norm, and a suitable stability hypothesis is needed, similarly to the case of the discretization studied in the next section.

4. Full Discretization: An Alternative Way to Incorporate Boundary Values

The standard method of lines studied in Section 3 seems to be optimal, especially when integrals in (11) are exactly calculated. However, the two integrals had very different origins, since one of them came from the source term and the other one from the boundary value. Furthermore, the integral that came from the discretization of the boundary values in (11) can be arbitrarily large when spatial discretization is refined. This is because operator A_h arises from the approximation at the boundary of differential operator A, which is unbounded.

We now introduce full discretization that more suitably incorporates the boundary values, as we show in the numerical experiments in Section 5. For this purpose, we need a more consistent notation with the fact that we discretize an initial boundary value problem (cf. [11]).

Suppose that $x_h \in X_{h,0}$, $v_h \in X_{h,b}$ and we want to calculate $A_{h,0}x_h + A_h v_h$. Then, we denote

$$B_h x_h = v_h,$$

and

$$A_h x_h = A_{h,0}x_h + A_h B_h x_h = A_{h,0}x_h + A_h v_h.$$

In this way, semidiscrete Problem (5) can be rewritten as

$$\begin{aligned} u_h'' &= A_h u_h + f(t, u_h), \\ u_h(0) &= P_h u_0, \\ u_h'(0) &= P_h v_0, \\ B_h u_h(t) &= Q_h g(t), \end{aligned} \tag{12}$$

which is more similar to the original problem.

With this notation, we rewrite Problem (12) as first-order differential system

$$\begin{aligned} \begin{bmatrix} u_{1,h} \\ u_{2,h} \end{bmatrix}' &= \begin{bmatrix} 0 & I_h \\ A_h & 0 \end{bmatrix} \begin{bmatrix} u_{1,h} \\ u_{2,h} \end{bmatrix} + \begin{bmatrix} 0 \\ f(t, u_{1,h}) \end{bmatrix}, \\ \begin{bmatrix} u_{1,h}(0) \\ u_{2,h}(0) \end{bmatrix} &= \begin{bmatrix} P_h u_0 \\ P_h v_0 \end{bmatrix}, \\ B_h u_{1,h}(t) &= Q_h g(t), \end{aligned} \tag{13}$$

of which the exact flow is given by $\mathbf{u}_h(t) = \Phi_{h,t}\mathbf{u}_h(0)$, as in Section 3.

For time discretization, we consider the same splitting scheme as that in Section 3, and the challenge is to obtain a suitable way with which to incorporate the boundary values. For this, we split Problem (13) into two intermediate problems; the first is

$$\begin{bmatrix} v_{1,h} \\ v_{2,h} \end{bmatrix}' = \begin{bmatrix} 0 & I_h \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_{1,h} \\ v_{2,h} \end{bmatrix}, \tag{14}$$

of which the exact flow is given by $\mathbf{v}_h(t) = \Phi_{h,t}^{I_h}\mathbf{v}_h(0)$, and the second is

$$\begin{bmatrix} w_{1,h} \\ w_{2,h} \end{bmatrix}' = \begin{bmatrix} 0 & 0 \\ A_h & 0 \end{bmatrix} \begin{bmatrix} w_{1,h} \\ w_{2,h} \end{bmatrix} + \begin{bmatrix} 0 \\ f(t, w_{1,h}) \end{bmatrix}, \tag{15}$$

of which the exact flow is $\mathbf{w}_h(t) = \Phi_{h,t}^{A_h+f}\mathbf{w}_h(0)$, where the value of $B_h w_{1,h}(0)$, necessary to compute $A_h w_{1,h}(0) = A_{h,0}w_{1,h}(0) + B_h w_{1,h}(0)$, must be chosen every time that this second intermediate problem is used in the final scheme.

We supposed that we computed $\mathbf{u}_{h,n} = [u_{1,h,n}, u_{2,h,n}]^T \approx [u_{1,h}(t_n), u_{2,h}(t_n)] = \mathbf{u}_h(t_n)$. As in Section 3, we use a general splitting method with m stages, with coefficients a_j, b_j, $j = 1, \ldots, m$. Therefore,

$$\begin{aligned} \mathbf{w}_{0,h,n} &= \mathbf{u}_{h,n}, \\ \mathbf{w}_{j,h,n} &= \Phi_{h,b_j k}^{A_h+f}\Phi_{h,a_j k}^{I_h}\mathbf{w}_{j-1,h,n}, \quad j = 1, \ldots, m, \\ \mathbf{u}_{h,n+1} &= \mathbf{w}_{m,h,n}, \end{aligned} \tag{16}$$

where $\mathbf{w}_{j,h,n} = [w_{1,j,h,n}, w_{2,j,h,n}]^T$.

To more explicitly write the j-stage in (16), we first consider initial-value problem

$$\begin{aligned}
&v'_{1,j,h,n}(s) = v_{2,j-1,h,n}(s), \quad s \in [0, a_j k],\\
&v'_{2,j,h,n}(s) = 0,\\
&v_{1,j,h,n}(0) = w_{1,j-1,h,n}(b_{j-1}k),\\
&v_{2,j,h,n}(0) = w_{2,j-1,h,n}(b_{j-1}k),
\end{aligned}$$

of which the solution at $s = a_j k$ is

$$\begin{aligned}
v_{1,j,h,n}(a_j k) &= w_{1,j-1,h,n}(b_{j-1}k) + a_j k w_{2,j-1,h,n}(b_{j-1}k),\\
v_{2,j,h,n}(a_j k) &= w_{2,j-1,h,n}(b_{j-1}k),
\end{aligned} \tag{17}$$

and we assign boundary value

$$B_h v_{1,j,h,n}(a_j k) = Q_h g(t_n + k \sum_{i=1}^{j} a_i). \tag{18}$$

Then,

$$\begin{aligned}
&w'_{1,j,h,n}(s) = 0, \quad s \in [0, b_j k],\\
&w'_{2,j,h,n}(s) = A_h w_{1,j,h,n}(s) + f(t_n + k \sum_{i=1}^{j-1} b_i + s, w_{1,j,h,n}(s)),\\
&w_{1,j,h,n}(0) = v_{1,j,h,n}(a_j k),\\
&w_{2,j,h,n}(0) = v_{2,j,h,n}(a_j k),
\end{aligned}$$

where using (18), we can calculate

$$\begin{aligned}
A_h w_{1,j,h,n}(s) &= A_{h,0} w_{1,j,h,n}(s) + A_h B_h w_{1,j,h,n}(s)\\
&= A_{h,0} w_{1,j,h,n}(s) + A_h B_h v_{1,j,h,n}(a_j k)\\
&= A_{h,0} w_{1,j,h,n}(s) + A_h Q_h g(t_n + k \sum_{i=1}^{j} a_i),
\end{aligned}$$

and we deduce that its solution at $s = b_j k$ is

$$\begin{aligned}
w_{1,j,h,n}(b_j k) &= v_{1,j,h,n}(a_j k),\\
w_{2,j,h,n}(b_j k) &= v_{2,j,h,n}(a_j k) + b_j k A_{h,0} v_{1,j,h,n}(a_j k) + b_j k A_h Q_h g(t_n + k \sum_{i=1}^{j} a_i)\\
&\quad + \int_0^{b_j k} f(t_n + k \sum_{i=1}^{j-1} b_i + \tau, v_{1,j,h,n}(a_j k)) d\tau.
\end{aligned} \tag{19}$$

Remark 2. *If we compare Formulas (11) and (19), the only difference is the treatment of the boundary values. The two ways of dealing with the boundary are*

$$\begin{aligned}
EX(g) &= \int_0^{b_j k} g(t_n + k \sum_{i=1}^{j-1} b_i + \tau) d\tau,\\
B(g) &= b_j k \, g(t_n + k \sum_{i=1}^{j} a_i),
\end{aligned}$$

for the standard method of lines and for the one that we propose, respectively.

Obviously, the second option is much simpler since the first option may even require to numerically evaluate the integral. We see in the numerical experiments in the next section that the second option also allows for obtaining much more precise results.

We prove in Sections 6 and 7 that, if the full discretization described in this section is used to approximate Problem (12), then the method is convergent, and the optimal order p of the splitting method is achieved.

5. Numerical Experiments

For the numerical experiments in this section, we consider splitting integrators of m stages, with coefficients a_j, b_j, $j = 1, \dots, m$, and $b_m = 0$, which are particular cases of the symmetric ones [3]. More specifically, we use the Strang method with order $p = 2$, and coefficients $a_1 = a_2 = 0.5$ and $b_1 = 1$ and two other methods that are particular cases of symmetric-splitting methods, whose coefficients are given in the following way: Let $l \in N$ be an even number; then, we consider a method with $m = l + 1$ stages satisfying $a_{l/2+1} = 1 - 2(a_1 + \dots + a_{l/2})$, $b_{l/2} = 1/2 - (b_1 + \dots + a_{l/2-1})$, $a_{l+2-i} = a_i$ and $b_{l+1-i} = b_i$, for $i = 1, \dots, l/2$. (Note that our parameter l is called m in [3]).

For $l = 6$, we consider method Ψ_{S4} with order $p = 4$ obtained from

$$\begin{aligned} a_1 &= 0.0792036964311957, & b_1 &= 0.209515106613362, \\ a_2 &= 0.353172906049774, & b_2 &= -0.143851773179818, \\ a_3 &= -0.0420650803577195, \end{aligned}$$

and, for $l = 10$, method Ψ_{S6} with order $p = 6$, given by coefficients

$$\begin{aligned} a_1 &= 0.0502627644003922, & b_1 &= 0.148816447901042, \\ a_2 &= 0.413514300428344, & b_2 &= -0.132385865767784, \\ a_3 &= 0.0450798897943977, & b_3 &= 0.067307604692185, \\ a_4 &= -0.188054853819569, & b_4 &= 0.432666402578175, \\ a_5 &= 0.541960678450780. \end{aligned}$$

5.1. Numerical Experiment: Test 1

We consider test problem

$$\begin{aligned} u_{tt}(x,t) &= u_{xx}(x,t) - \sin(u) + e^{-\mu t}(25(x^2+1) - 2) + \sin(e^{-\mu t}(x^2+1)), \\ u(x,0) &= x^2 + 1, \\ u_t(x,0) &= -\mu(x^2+1), \\ u(0,t) &= e^{-\mu t}, \\ u(1,t) &= 2e^{-\mu t}, \end{aligned} \tag{20}$$

where $x \in [0,1]$, $t \in [0,T]$, of which the exact solution is $u(x,t) = e^{-\mu t}(x^2+1)$.

This problem can be written in abstract format (2) by taking $X = L^2(0,1)$, $D(A) = H^2(0,1)$, A is the second-order derivative in the spatial variable, B is the Dirichlet trace operator, and $Y = \mathbb{R}^2$. Then, $D(A_0) = H^2(0,1) \cap H_0^1(0,1)$ and operator $A_0 = A|_{D(A_0)}$ is self-adjoint and definite negative.

We consider grid $x_j = jh$, $0 \le j \le J+1$, of interval $[0,1]$. We look for approximations $u_{j,n} \approx u(x_j, t_n)$, $0 \le j \le J+1$, $0 \le n \le N$. We take $X_h = \mathbb{R}^{J+2}$, and elements $u_h \in X_h$ are denoted by $u_h = [u_0, u_1, \dots, u_J, u_{J+1}]^T$. In this way, $X_{h,0} = \{u_h \in X_h \text{ such that } u_0 = u_{J+1} = 0\}$ but, for the sake of simplicity, we use $X_{h,0} = \mathbb{R}^J$ and their elements are denoted by $u_h = [u_1, \dots, u_J]^T$. Moreover, $X_{h,b} = \{u_h \in X_h \text{ such that } u_1 = \dots = u_J = 0\}$.

Operator $P_h : D(A) \to X_{h,0}$ is given by $P_h u = [u(x_1), \dots, u(x_J)]^T$, and operator $Q_h : Y \to X_{h,b}$ is given by $Q_h(u_0, u_{J+1}) = (u_0, 0, \dots, 0, u_{J+1})$. Operator A_h arises from the approximation of the second-order derivative in space by using central finite differences, that is,

$$A_h = \frac{1}{h^2}\begin{bmatrix} 1 & -2 & 1 & & & & \\ & 1 & -2 & 1 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & 1 & -2 & 1 & \\ & & & & 1 & -2 & 1 \end{bmatrix}.$$

Now, by using the previous notation, we can write

$$A_{h,0} = \frac{1}{h^2}\begin{bmatrix} -2 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 1 & -2 \end{bmatrix}, \tag{21}$$

which is a symmetric and definite negative operator on $X_{h,0}$.

Hypothesis (H3) can be verified in the standard way by using Taylor series for the local truncation error, with $\varepsilon_h = O(h^2)$ and $Z = C^2(0,1)$. Taking into account that the exact solution of (20) is a second-order polynomial in variable x, there is no spatial error, and we can focus on the error due to time discretization.

In this problem, for $f(t,u) = -\sin(u) + e^{-\mu t}(\mu^2(x^2+1) - 2) + \sin(e^{-\mu t}(x^2+1))$, we exactly compute $\int_0^k f(t_n + \tau, \tilde{z}_{1,h}(\frac{k}{2}))d\tau$

$$\begin{aligned} EX(f) \;\; &= \;\; -k\sin(\tilde{z}_{1,h}(\frac{k}{2})) + (e^{-\mu t_n} - e^{-\mu(t_n+k)})(\mu^2(x_j^2+1) - 2)/\mu \\ &\qquad +(\text{sinint}(e^{-\mu t_n}(x_j^2+1)) - \text{sinint}(e^{-\mu(t_n+k)}(x_j^2+1)))/\mu, \end{aligned}$$

where $1 \le j \le J$.

For the three symmetric-splitting methods, we compare the errors in the energy norm for the choice of boundary $B(g)$ and for the $EX(g)$ option in Remark 2, for values 1, 3, and 5 of μ. As Figure 1 shows, in all cases, option $B(g)$ (solid line) obtained smaller errors than $EX(g)$ did (dashed line). Moreover, the difference was more noticeable for Ψ_{S4} and Ψ_{S6}. Dependence on the size of boundary function g was also observed; for example, the difference between errors of $B(g)$ and $EX(g)$ was larger for $\mu = 1$ than that for values $\mu = 3$ and $\mu = 5$, where the values taken by g were smaller.

In addition, as shown in Figure 1 the slopes of the lines were 2 (Strang), 4 (Ψ_{S4}) and 6 (Ψ_{S6}), which coincides with the expected optimal order of the three methods.

5.2. *Numerical Experiment: Test 2*

We consider test problem

$$\begin{aligned} u_{tt}(x,t) &= u_{xx}(x,t) - \sin(u) + e^{-t^2}((-2+4t^2)(x^2+1) - 2) + \sin(e^{-t^2}(x^2+1)), \\ u(x,0) &= x^2+1, \\ u_t(x,0) &= 0, \\ u(0,t) &= e^{-t^2}, \\ u(1,t) &= 2e^{-t^2}, \end{aligned} \tag{22}$$

where $x \in [0,1]$, $t \in [0,T]$, of which the exact solution is $u(x,t) = e^{-t^2}(x^2+1)$. In this problem, a primitive of function $f(t,u) = -\sin(u) + e^{-t^2}((-2+4t^2)(x^2+1) - 2) + \sin(e^{-t^2}(x^2+1))$ cannot be exactly expressed using elementary functions, so we used a quadrature formula of appropriate order. In the calculations to obtain the data in Figure 2 we used the 3-point Gaussian quadrature of order 6, denoted by $G3(g)$. Errors are compared for Strang's method (black), Ψ_{S4} (red), and Ψ_{S6} (blue), and for options $B(g)$ (solid line) and $G3(g)$ (dashed line). The slopes of the lines indicate order 2 for Strang's method, order 4 for Ψ_{S4}, and order 6 for Ψ_{S6}. Errors are always smaller for option $B(g)$ than those for option $G3(g)$. Moreover, this difference was more pronounced as the order of the method increased.

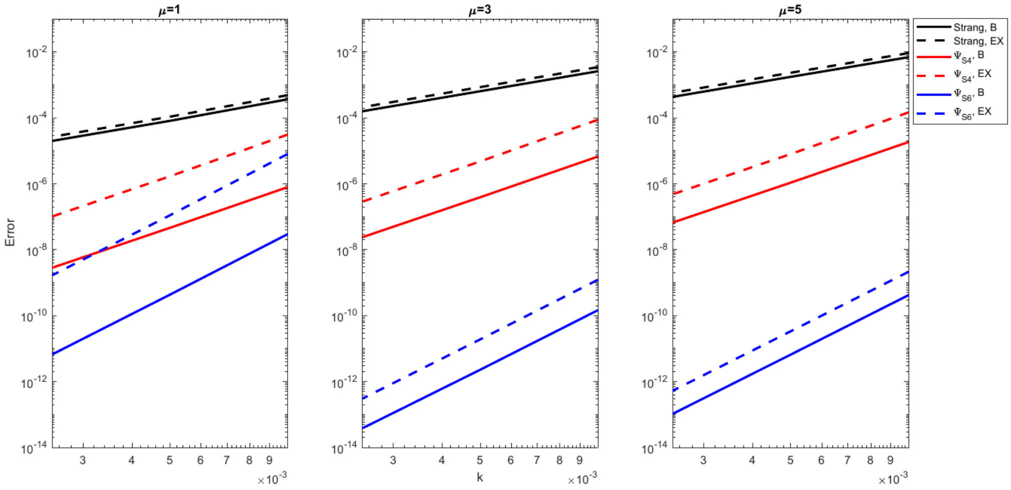

Figure 1. Error, in logarithmic scale, for the energy norm of the solution of Test 1, with $\mu = 1$, $\mu = 3$ and $\mu = 5$, $N = 100$, final time $T = 1$ and $EX(f)$, for Strang's method (black), Ψ_{S4} (red) and Ψ_{S6} (blue).

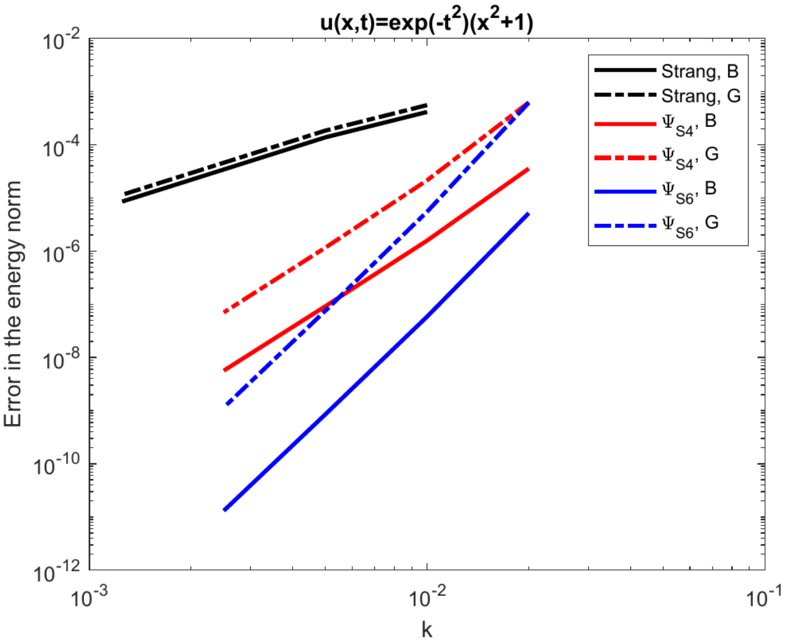

Figure 2. Error in logarithmic scale of energy norm for Test 2, for several options of symmetric-splitting methods, with $N = 100$ for final time $T = 1$ and G3(f).

5.3. *Numerical Experiment: Test 3*

Now, we study the behavior of the error of both methods when spatial discretization is refined. The source term of semidiscrete Problem (5) grows when $h \to 0$ due to the boundary. However, it is expected that this growth has no influence when it is treated as part of the discretization of operator A at the boundary, as in the method we propose in Section 4.

For this, we consider test problem

$$\begin{aligned} u_{tt}(x,t) &= u_{xx}(x,t) - \sin(u) + \sin(e^{-t+x}), \\ u(x,0) &= e^{x}, \\ u_t(x,0) &= -e^{x}, \\ u(0,t) &= e^{-t}, \\ u(1,t) &= e^{-t+1}, \end{aligned} \tag{23}$$

where $x \in [0,1]$, $t \in [0,T]$, of which the exact solution is $u(x,t) = e^{-t+x}$.

In this experiment, spatial error dominates, and we expected to observe order 2 of spatial discretization. Figure 3 shows errors for the splitting using $B(g)$ in continuous line and $EX(g)$ in dashed line, in red for Ψ_{S4} and in blue Ψ_{S6}. The values of $h = k = 1/N$ were used for $N = 25, 50, 100, 200$. In this way, we remained in the stability interval of both methods. In the two methods with $B(g)$, order 2 of spatial discretization was observed; when using the $EX(g)$ option, the errors were larger, and order 2 was lost when h decreased.

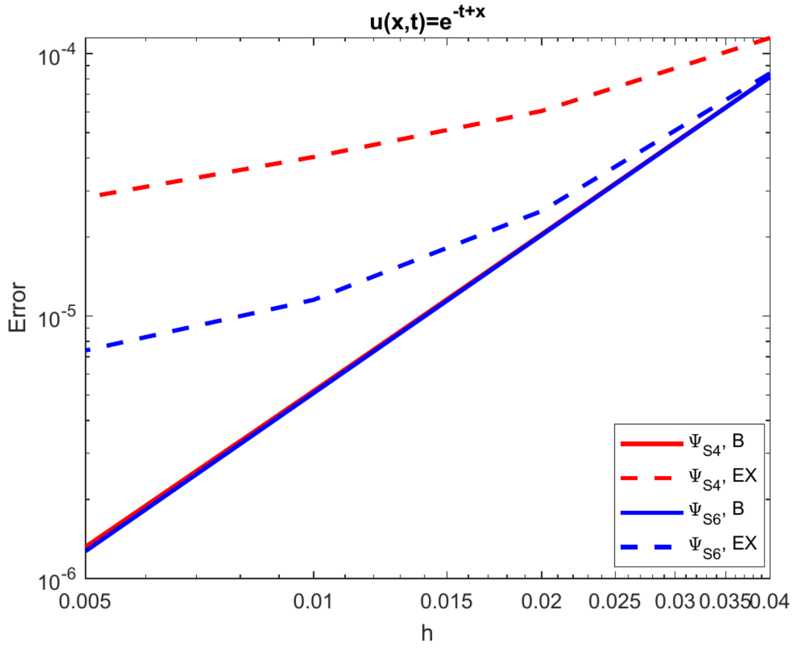

Figure 3. Error in energy norm for Test 3, $N = 25, 50, 100, 200$ and $h = k = 1/N$.

5.4. Numerical Experiment: Test 4

Although theoretical results were shown for the case where f was Lipschitz-continuous, we see in a couple of examples that it also works if f is locally Lipschitz-continuous. We now consider test problem

$$\begin{aligned} u_{tt}(x,t) &= u_{xx}(x,t) + u^2(x,t) + e^{-t}(x^2-1) - e^{-2t}(x^2+1)^2, \\ u(x,0) &= x^2+1, \\ u_t(x,0) &= -(x^2+1), \\ u(0,t) &= e^{-t}, \\ u(1,t) &= 2e^{-t}, \end{aligned}$$

where $x \in [0,1]$, $t \in [0,T]$, of which the exact solution is $u(x,t) = e^{-t}(x^2+1)$.

Table 1 shows the time errors for the three symmetric-splitting methods with $h = 1/100$ for final time $T = 1$. The evolution of $\log_2(\text{error}(k)/\text{error}(k/2))$ is displayed in Table 2. Orders 2, 4, and 6 are shown for Strang's method, Ψ_{S4}, and Ψ_{S6}, respectively. The little loss of order in the lower-right corner was due to the influence of rounding errors.

Table 1. Time error for three symmetric-splitting methods with $h = 1/100$ for final time $T = 1$.

h = 1/100	k = 1/100	k = 1/200	k = 1/400
Strang	4.8394×10^{-4}	1.0727×10^{-4}	2.6416×10^{-5}
Ψ_{S4}	3.0924×10^{-5}	1.6787×10^{-6}	1.0135×10^{-7}
Ψ_{S6}	3.5808×10^{-10}	5.5734×10^{-12}	9.6204×10^{-14}

Table 2. Evolution of $\log_2(\text{error}(k)/\text{error}(k/2))$ for three symmetric-splitting methods with $h = 1/100$.

h = 1/100	k = 1/100	k = 1/200
Strang	2.1737	2.0217
Ψ_{S4}	4.2034	4.0499
Ψ_{S6}	6.0056	5.8563

5.5. Numerical Experiment: Test 5

Lastly, we consider an example in two spatial dimensions, and in the locally Lipschitz = continuous case. The test problem is

$$\begin{aligned}
u_{tt}(x,y,t) &= u_{xx}(x,y,t) + u_{yy}(x,y,t) + u^2(x,y,t) + e^{-t}(x^2+y^2-3) - e^{-2t}(x^2+y^2+1)^2,\\
u(x,y,0) &= x^2+y^2+1,\\
u_t(x,y,0) &= -(x^2+y^2+1),\\
u(0,y,t) &= e^{-t}(y^2+1),\\
u(1,y,t) &= e^{-t}(y^2+2),\\
u(x,0,t) &= e^{-t}(x^2+1),\\
u(x,1,t) &= e^{-t}(x^2+2),
\end{aligned}$$

for values $x \in [0,1], y \in [0,1], t \in [0,T]$, of which the exact solution is $u(x,t) = e^{-t}(x^2+y^2+1)$.

This problem can be written in the abstract format (2). For this, we denote $\Omega = (0,1) \times (0,1)$ and Γ is the boundary of Ω. We take $X = L^2(\Omega)$, $D(A) = H^2(\Omega)$; A is the Laplacian operator in the spatial variables x and y, B is the Dirichlet trace operator on Γ and $Y = h^{1/2}(\Gamma)$. Then, $D(A_0) = H^2(\Omega) \cap H_0^1(\Omega)$ and operator $A_0 = A|_{D(A_0)}$ is self-adjoint and definite negative.

We consider grid $(x_j, y_l) = (jh, lh)$, $0 \le j,l \le J+1$, of $\overline{\Omega}$. We look for approximations $u_{j,l,n} \approx u(x_j, y_l, t_n)$, $0 \le j,l \le J+1$, $0 \le n \le N$. We take $X_h = \mathbb{R}^{J+2} \times \mathbb{R}^{J+2}$, and elements $u_h \in X_h$ are denoted by $u_h = [u_{0,1}, u_{0,2}, \dots, u_{J+1,J}, u_{J+1,J+1}]^T$. In this way, $X_{h,0} = \{u_h \in X_h \text{ such that } u_{0,l} = u_{J+1,l} = u_{j,0} = u_{j,J+1} = 0\}$; however , for the sake of simplicity, we use $X_{h,0} = \mathbb{R}^J \times \mathbb{R}^J$, and their elements are denoted by $u_h = [u_{1,1}, u_{1,2}, \dots, u_{J,J-1}, u_{J,J}]^T$. Subspace $X_{h,b}$ is similarly defined.

Operator $P_h : D(A) \to X_{h,0}$ is given by

$$P_h u = [u(x_1,y_1), u(x_1,y_2), \dots, u(x_J,y_{J-1}), u(x_J,y_J)]^T,$$

and A_h arises from the approximation of the Laplacian operator by using central finite differences in each spatial direction, that is, considering second-order spatial discretization

$$A_{h,0} = \frac{1}{h^2}\begin{bmatrix} B_J & I_J & & & \\ I_J & B_J & I_J & & \\ & \ddots & \ddots & \ddots & \\ & & I_J & B_J & I_J \\ & & & I_J & B_J \end{bmatrix}, \tag{24}$$

where I_J is the identity matrix, and

$$B_J = \begin{bmatrix} -4 & 1 & & & \\ 1 & -4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -4 & 1 \\ & & & 1 & -4 \end{bmatrix}.$$

Taking into account that the exact solution is a second-order polynomial in variables x, y, there is no spatial error; then, we can again focus on the error due to time discretization.

Table 3 shows the time errors for the three symmetric-splitting methods, with $h = 1/100$ and for final time $T = 1$. The missing value for the Strang method and $k = 1/100$ was due to the instability of the numerical solution in this case; see Section 7.1. The evolution of $\log_2(\text{error}(k)/\text{error}(k/2))$ is displayed in Table 4. Orders 2, 4, and 6 are shown for Strang's method, Ψ_{S4} and Ψ_{S6}, respectively.

Table 3. Time error for three symmetric-splitting methods with $h = 1/100$ for final time $T = 1$.

h = 1/100	k = 1/100	k = 1/200	k = 1/400
Strang	–	1.3639×10^{-4}	3.3576×10^{-5}
Ψ_{S4}	1.3720×10^{-6}	7.9938×10^{-8}	4.9125×10^{-9}
Ψ_{S6}	5.2613×10^{-8}	7.6468×10^{-10}	1.1734×10^{-11}

Table 4. Evolution of $\log_2(\text{error}(k)/\text{error}(k/2))$ for three symmetric-splitting methods with $h = 1/100$.

h = 1/100	k = 1/100	k = 1/200
Strang	–	2.0222
Ψ_{S4}	4.1012	4.0244
Ψ_{S6}	6.1044	6.0261

6. Consistency Correctly Incorporating Boundary Values

Here, we deduce consistency in the energy norm of the implementation of a splitting method with the boundary values that we chose in Section 4.

As a first step in the proof of consistency, we introduce ordinary differential system

$$\begin{aligned} \begin{bmatrix} u_{1,h} \\ u_{2,h} \\ u_{3,h} \end{bmatrix}' &= \begin{bmatrix} 0 & I_h & 0 \\ A_{h,0} & 0 & A_h \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_{1,h} \\ u_{2,h} \\ u_{3,h} \end{bmatrix} + \begin{bmatrix} 0 \\ f(t, u_{1,h}) \\ Q_h g'(t) \end{bmatrix}, \\ \begin{bmatrix} u_{1,h}(0) \\ u_{2,h}(0) \\ u_{3,h}(0) \end{bmatrix} &= \begin{bmatrix} P_h u_0 \\ P_h v_0 \\ Q_h g(0) \end{bmatrix}. \end{aligned} \tag{25}$$

We have

$$\begin{aligned} u_{3,h}'(t) &= Q_h g'(t), \\ u_{3,h}(0) &= Q_h g(0). \end{aligned}$$

and, therefore,

$$u_{3,h}(t) = Q_h g(t). \tag{26}$$

Then, for the first two components of (25), we deduce that

$$\begin{aligned} u_{1,h}'(t) &= u_{2,h}(t), \\ u_{2,h}'(t) &= A_{h,0} u_{1,h}(t) + A_h Q_h g(t) + f(t, u_{1,h})(t), \end{aligned}$$

which is the same problem as (7) (and as (12) with the notation of Section 4). We also deduce that

$$u_{1,h}(t) = u_h(t), \quad u_{2,h}(t) = u_h'(t), \tag{27}$$

being $u_h(t)$ the solution of (5).

We split (25) into two intermediate problems that are similar to the ones used in Section 4, and we applied to it the same splitting method. The solution of (25) was approximated with order p of the splitting method. This particularly is true for the two first components that match those of the solution of (12). Therefore, to prove consistency, it suffices to see that the obtained approximations for the first two components are the same as those described in Section 4 with the choice of boundary values made in (18).

We choose the following split of Problem (25). The first intermediate problem is

$$\begin{bmatrix} v_{1,h} \\ v_{2,h} \\ v_{3,h} \end{bmatrix}' = \begin{bmatrix} 0 & I_h & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_{1,h} \\ v_{2,h} \\ v_{3,h} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ Q_h g'(t) \end{bmatrix}, \tag{28}$$

of which the exact flow is denoted as $\mathbf{v}_h(t) = \Phi_{h,t}^{[1]} \mathbf{v}_h(0)$, and the second is

$$\begin{bmatrix} w_{1,h} \\ w_{2,h} \\ w_{3,h} \end{bmatrix}' = \begin{bmatrix} 0 & 0 & 0 \\ A_{h,0} & 0 & A_h \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} w_{1,h} \\ w_{2,h} \\ w_{3,h} \end{bmatrix} + \begin{bmatrix} 0 \\ f(t, w_{1,h}) \\ 0 \end{bmatrix}. \tag{29}$$

of which the exact flow is denoted as $\mathbf{w}_h(t) = \Phi_{h,t}^{[2]} \mathbf{w}_h(0)$.

Assuming that approximation $\mathbf{u}_{h,n} = [u_{1,h,n}, u_{2,h,n}, u_{3,h,n}]^T \approx [u_{1,h}(t_n), u_{2,h}(t_n), u_{3,h}(t_n)]^T$ is already calculated and, as in previous sections, we apply a general splitting of order p,

$$\begin{aligned} \mathbf{w}_{0,h,n} &= \mathbf{u}_{h,n}, \\ \mathbf{w}_{j,h,n} &= \Phi_{h,b_j k}^{[2]} \Phi_{h,a_j k}^{[1]} \mathbf{w}_{j-1,h,n}, \quad j = 1, \ldots, m, \\ \mathbf{u}_{h,n+1} &= \mathbf{w}_{m,h,n}, \end{aligned} \tag{30}$$

where $\mathbf{w}_{j,h,n} = [w_{1,j,h,n}, w_{2,j,h,n}, w_{3,j,h,n}]^T$.

The performance of a time step $k > 0$ is as follows. For each $j = 1, \ldots, m$, the first problem to be solved is

$$\begin{aligned} v_{1,j,h,n}'(s) &= v_{2,1,h,n}(s), \quad s \in [0, a_j k], \\ v_{2,j,h,n}'(s) &= 0, \\ v_{3,j,h,n}'(s) &= Q_h g'(t_n + k \sum_{l=1}^{j-1} a_l + s), \\ v_{1,j,h,n}(0) &= w_{1,j-1,h,n}(b_{j-1}k), \\ v_{2,j,h,n}(0) &= w_{2,j-1,h,n}(b_{j-1}k), \\ v_{3,j,h,n}(0) &= w_{3,j-1,h,n}(b_{j-1}k) = Q_h g(t_n + k \sum_{l=1}^{j-1} a_l), \end{aligned}$$

whose solution at $s = a_j k$ is

$$\begin{aligned} &v_{1,j,h,n}(a_j k) = w_{1,j-1,h,n}(b_{j-1}k) + a_j k w_{2,j-1,h,n}(b_{j-1}k), \\ &v_{2,j,h,n}(a_j k) = w_{2,j-1,h,n}(b_{j-1}k), \\ &v_{3,j,h,n}(a_j k) = Q_h g(t_n + k \sum_{l=1}^{j} a_l). \end{aligned} \tag{31}$$

Since the third component of (31) provides the boundary value (18), the first and second components of (31) are the same as (17). Then, the second problem is

$$\begin{aligned}
w'_{1,j,h,n}(s) &= 0, \quad s \in [0, b_j k], \\
w'_{2,j,h,n}(s) &= A_{h,0} w_{1,j,h,n} + A_h w_{3,j,h,n} + f(t_n + k \sum_{l=1}^{j-1} b_l + s, w_{1,j,h,n}(s)), \\
w'_{3,j,h,n}(s) &= 0, \\
w_{1,j,h,n}(0) &= v_{1,j,h,n}(a_j k), \\
w_{2,j,h,n}(0) &= v_{2,j,h,n}(a_j k), \\
w_{3,j,h,n}(0) &= v_{3,j,h,n}(a_j k),
\end{aligned}$$

whose solution at $s = b_j k$ is

$$\begin{aligned}
w_{1,j,h,n}(b_j k) &= v_{1,j,h,n}(a_j k), \\
w_{2,j,h,n}(b_j k) &= v_{2,j,h,n}(a_j k) + b_j k A_{h,0} v_{1,j,h,n}(a_j k) + b_j k A_h Q_h g(t_n + k \sum_{l=1}^{j} a_l) \\
&\quad + \int_0^{b_j k} f(t_n + k \sum_{l=1}^{j-1} b_l \tau, v_{1,j,h,n}(a_j k)) d\tau, \\
w_{3,j,h,n}(b_j k) &= Q_h g(t_n + k \sum_{l=1}^{j} a_l).
\end{aligned} \tag{32}$$

The first two components of (32) are the same as those of (19).

Remark 3. *Although the splitting approximation (30) provides the same approximation as (16) through its first two components, it is not convenient to use it in practice. It is more useful to use the implementation of Section 4, which can be carried out with minimal modifications to the standard method of lines.*

Now, we consider $\mathbf{u}_h(t) = [u_{1,h}(t), u_{2,h}(t), u_{3,h}(t)]^T$, the solution of (25). We can easily deduce that

$$\widetilde{\mathbf{u}}_h(t) = [\widetilde{u}_{1,h}(t), \widetilde{u}_{2,h}(t), \widetilde{u}_{3,h}(t)]^T = [S_{h,0} u_{1,h}(t), u_{2,h}(t), u_{3,h}(t)]^T$$

is the solution of ordinary differential system

$$\begin{aligned}
\begin{bmatrix} \widetilde{u}_{1,h} \\ \widetilde{u}_{2,h} \\ \widetilde{u}_{3,h} \end{bmatrix}' &= \begin{bmatrix} 0 & S_{h,0} & 0 \\ -S_{h,0} & 0 & A_h \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \widetilde{u}_{1,h} \\ \widetilde{u}_{2,h} \\ \widetilde{u}_{3,h} \end{bmatrix} + \begin{bmatrix} 0 \\ f(t, S_{h,0}^{-1} \widetilde{u}_{1,h}) \\ Q_h g'(t) \end{bmatrix}, \\
\begin{bmatrix} \widetilde{u}_{1,h}(0) \\ \widetilde{u}_{2,h}(0) \\ \widetilde{u}_{3,h}(0) \end{bmatrix} &= \begin{bmatrix} S_{h,0} P_h u_0 \\ P_h v_0 \\ Q_h g(0) \end{bmatrix}.
\end{aligned} \tag{33}$$

We deduce that $\widetilde{u}_{3,h}(t) = Q_h g(t)$, and that

$$\begin{bmatrix} \widetilde{u}_{1,h}(t) \\ \widetilde{u}_{2,h}(t) \end{bmatrix}' = \begin{bmatrix} 0 & S_{h,0} \\ -S_{h,0} & 0 \end{bmatrix} \begin{bmatrix} \widetilde{u}_{1,h}(t) \\ \widetilde{u}_{2,h}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ A_h Q_h g(t) + f(t, S_{h,0}^{-1} u_{1,h}(t)), \end{bmatrix}.$$

Therefore, the solution of Problem (33) is the appropriate one to calculate the energy norm of the solution of (25). Now, we see that approximating the solution of Problem (33) by means of a splitting method is equivalent to applying the same change of variables to (30), as is stated in (37).

The first intermediate problem is

$$\begin{bmatrix} \widetilde{v}_{1,h} \\ \widetilde{v}_{2,h} \\ \widetilde{v}_{3,h} \end{bmatrix}' = \begin{bmatrix} 0 & S_{h,0} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \widetilde{v}_{1,h} \\ \widetilde{v}_{2,h} \\ \widetilde{v}_{3,h} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ Q_h g'(t) \end{bmatrix}, \tag{34}$$

of which the exact flow is denoted as $\widetilde{\mathbf{v}}_h(t) = \widetilde{\Phi}_t^{[1]}\widetilde{\mathbf{v}}_h(0)$ and the second intermediate problem

$$\begin{bmatrix} \widetilde{w}_{1,h} \\ \widetilde{w}_{2,h} \\ \widetilde{w}_{3,h} \end{bmatrix}' = \begin{bmatrix} 0 & 0 & 0 \\ -S_{h,0} & 0 & A_h \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \widetilde{w}_{1,h} \\ \widetilde{w}_{2,h} \\ \widetilde{w}_{3,h} \end{bmatrix} + \begin{bmatrix} 0 \\ f(t, S_{h,0}^{-1}\widetilde{w}_{1,h}) \\ 0 \end{bmatrix}. \tag{35}$$

of which the exact flow is denoted as $\widetilde{\mathbf{w}}_h(t) = \widetilde{\Phi}_t^{[2]}\widetilde{\mathbf{w}}_h(0)$.

As in the previous sections, we use a general splitting method with m stages and order p, with coefficients $a_j, b_j, j = 1, \ldots, m$. Therefore,

$$\begin{aligned} \widetilde{\mathbf{w}}_{0,h,n} &= \widetilde{\mathbf{u}}_{h,n}, \\ \widetilde{\mathbf{w}}_{j,h,n} &= \widetilde{\Phi}_{h,b_jk}^{[2]}\widetilde{\Phi}_{h,a_jk}^{[1]}\widetilde{\mathbf{w}}_{j-1,h,n}, \quad j = 1, \ldots, m, \\ \widetilde{\mathbf{u}}_{h,n+1} &= \widetilde{\mathbf{w}}_{m,h,n}, \end{aligned} \tag{36}$$

where $\widetilde{\mathbf{w}}_{j,h,n} = [\widetilde{w}_{1,j,h,n}, \widetilde{w}_{2,j,h,n}, \widetilde{w}_{3,j,h,n}]^T$.

We now prove that

$$\widetilde{\mathbf{u}}_{h,n} = [\widetilde{u}_{1,h,n}, \widetilde{u}_{2,h,n}, \widetilde{u}_{3,h,n}]^T = [S_{h,0}u_{1,h,n}, u_{2,h,n}, u_{3,h,n}]^T, \tag{37}$$

where $\mathbf{u}_{h,n} = [u_{1,h,n}, u_{2,h,n}, u_{3,h,n}]^T$ is the solution of the splitting given by (30).

For $j = 1, \ldots, m$, the following problems are solved.

$$\begin{aligned} &\widetilde{v}'_{1,j,h,n}(s) = S_{h,0}\widetilde{v}_{2,j,h,n}(s), \quad s \in [0, a_jk], \\ &\widetilde{v}'_{2,j,h,n}(s) = 0, \\ &\widetilde{v}'_{3,j,h,n}(s) = Q_h g'(t_n + k\sum_{l=1}^{j-1} a_l + s), \\ &\widetilde{v}_{1,j,h,n}(0) = \widetilde{w}_{1,j-1,h,n}(b_{j-1}k) = w_{1,j-1,h,n}(b_{j-1}k), \\ &\widetilde{v}_{2,j,h,n}(0) = \widetilde{w}_{2,j-1,h,n}(b_{j-1}k) = w_{2,j-1,h,n}(b_{j-1}k), \\ &\widetilde{v}_{3,j,h,n}(0) = \widetilde{w}_{3,j-1,h,n}(b_{j-1}k) = Q_h g(t_n + k\sum_{l=1}^{j-1} a_l) = w_{3,j-1,h,n}(b_{j-1}k), \end{aligned}$$

whose solution at $s = a_jk$ is

$$\begin{aligned} &\widetilde{v}_{1,j,h,n}(a_jk) = S_{h,0}w_{1,j-1,h,n}(b_{j-1}k) + a_jkS_{h,0}w_{2,j-1,h,n}(b_{j-1}k) = S_{h,0}v_{1,j,h,n}(a_jk), \\ &\widetilde{v}_{2,j,h,n}(a_jk) = w_{2,j-1,h,n}(b_{j-1}k) = v_{2,j,h,n}(a_jk), \\ &\widetilde{v}_{3,j,h,n}(a_jk) = Q_h g(t_n + k\sum_{l=1}^{j} a_l) = v_{3,j,h,n}(a_jk). \end{aligned} \tag{38}$$

Then, in $s \in [0, b_j k]$, we solve system

$$\begin{aligned}
&\widetilde{w}'_{1,j,h,n}(s) = 0,\\
&\widetilde{w}'_{2,j,h,n}(s) = -S_{h,0}\widetilde{w}_{1,j,h,n} + A_h\widetilde{w}_{3,j,h,n} + f(t_n + k\sum_{l=1}^{j-1} b_l + s, S_{h,0}^{-1}\widetilde{w}_{1,j,h,n}(s)),\\
&\widetilde{w}'_{3,j,h,n}(s) = 0,\\
&\widetilde{w}_{1,j,h,n}(0) = \widetilde{v}_{1,j,h,n}(a_j k) = S_{h,0}v_{1,j,h,n}(a_j k),\\
&\widetilde{w}_{2,j,h,n}(0) = \widetilde{v}_{2,j,h,n}(a_j k) = v_{2,j,h,n}(a_j k),\\
&\widetilde{w}_{3,j,h,n}(0) = \widetilde{v}_{3,j,h,n}(a_j k) = Q_h g(t_n + k\sum_{l=1}^{j} a_l) = v_{3,j,h,n}(a_j k),
\end{aligned} \tag{39}$$

whose solution at $s = b_j k$ is

$$\begin{aligned}
\widetilde{w}_{1,j,h,n}(b_j k) &= \widetilde{v}_{1,j,h,n}(a_j k) = S_{h,0}v_{1,j,h,n}(a_j k) = S_{h,0}w_{1,j,h,n}(b_j k),\\
\widetilde{w}_{2,j,h,n}(b_j k) &= \widetilde{v}_{2,j,h,n}(a_j k) + b_j k(-S_{h,0}\widetilde{v}_{1,j,h,n}(a_j k) + A_h Q_h g(t_n + k\sum_{l=1}^{j} a_l))\\
&\quad + \int_0^{b_j k} f(t_n + k\sum_{l=1}^{j-1} b_l + \tau, S_{h,0}^{-1}\widetilde{v}_{1,j,h,n}(a_j k))d\tau,\\
&= v_{2,j,h,n}(a_1 k) + b_j k(A_{h,0}v_{1,j,h,n}(a_j k) + A_h Q_h g(t_n + k\sum_{l=1}^{j} a_l))\\
&\quad + \int_0^{b_j k} f(t_n + k\sum_{l=1}^{j-1} b_l + \tau, v_{1,j,h,n}(a_j k))d\tau,\\
&= w_{2,j,h,n}(b_j k)\\
\widetilde{w}_{3,j,h,n}(b_j k) &= Q_h g(t_n + k\sum_{l=1}^{j} a_l) = w_{3,j,h,n}(b_j k),
\end{aligned} \tag{40}$$

and we obtain (37).

Theorem 2. *We assume that the time discretization of (12) is obtained by means of a splitting method of order p, applied as described in Formulas (16)–(17), with the choice of intermediate boundary values (18). Let $\mathbf{u}_h(t_n) = [u_{1,h}(t_n), u_{2,h}(t_n)]^T$ be the value at t_n of the solution of (13), and let $\overline{\mathbf{u}}_{h,n} = [\overline{u}_{1,h,n}, \overline{u}_{2,h,n}]^T$ be its approximation obtained with the splitting method (16) by taking a step of size k starting from the exact value $\mathbf{u}_h(t_{n-1}) = [u_{1,h}(t_{n-1}), u_{2,h}(t_{n-1})]^T$.*

Then, local error $\rho_{h,n} = \mathbf{u}_h(t_n) - \overline{\mathbf{u}}_{h,n}$ satisfies

$$\begin{aligned}
\|\rho_{h,n}\|_{E_h} &= \|\mathbf{u}_h(t_n) - \overline{\mathbf{u}}_{h,n}\|_{E_h}\\
&= \Big(\|S_{h,0}(u_{1,h}(t_n) - \overline{u}_{1,h,n})\|_h^2 + \|u_{2,h}(t_n) - \overline{u}_{2,h,n}\|_h^2\Big)^{1/2} = O(k^{p+1}).
\end{aligned}$$

Proof.

$$\begin{aligned}
\|\rho_{h,n}\|_{E_h} &= \|\mathbf{u}_h(t_n) - \overline{\mathbf{u}}_{h,n}\|_{E_h}\\
&= \Big(\|S_{h,0}(u_{1,h}(t_n) - \overline{u}_{1,h,n})\|_h^2 + \|u_{2,h}(t_n) - u_{2,h,n}\|_h^2\Big)^{1/2}\\
&\leq \Big(\|S_{h,0}(u_{1,h}(t_n) - \overline{u}_{1,h,n})\|_h^2 + \|u_{2,h}(t_n) - \overline{u}_{2,h,n}\|_h^2 + \|u_{3,h}(t_n) - \overline{u}_{3,h,n}\|_h^2\Big)^{1/2}\\
&= \Big(\|\widetilde{u}_{1,h}(t_n) - \overline{\widetilde{u}}_{1,h,n}\|_h^2 + \|\widetilde{u}_{2,h}(t_n) - \overline{\widetilde{u}}_{2,h,n}\|_h^2 + \|\widetilde{u}_{3,h}(t_n) - \overline{\widetilde{u}}_{3,h,n}\|_h^2\Big)^{1/2}\\
&= \|\widetilde{\mathbf{u}}_h(t_n) - \overline{\widetilde{\mathbf{u}}}_{h,n}\|_{E_h} = O(k^{p+1}),
\end{aligned}$$

where $\overline{\widetilde{\mathbf{u}}}_{h,n} = [\overline{\widetilde{u}}_{1,h,n}, \overline{\widetilde{u}}_{2,h,n}, \overline{\widetilde{u}}_{3,h,n}]^T$ is the approximation obtained at $t = t_n$ with the splitting method (36) by taking a step of size k starting from the exact value at t_{n-1} of the solution of

(33) $\widetilde{\mathbf{u}}_h(t_{n-1}) = [\widetilde{u}_{1,h}(t_{n-1}), \widetilde{u}_{2,h}(t_{n-1}), \widetilde{u}_{3,h}(t_{n-1})]^T$. Therefore, the result is a consequence that p is the order of the splitting method. □

7. Stability and Convergence

7.1. Stability

To achieve convergence in energy norm, we need time discretization to be stable. In our case, it is sufficient to have linear stability. That is, it is enough that time discretization with the splitting method is stable for fully homogeneous linear problem

$$\begin{aligned} u_h''(t) &= A_{h,0}u_h(t), \\ u_h(0) &= P_h u_0, \\ u_h'(0) &= P_h v_0, \end{aligned} \tag{41}$$

corresponding to the space discretization of (2) with vanishing boundary values and source term.

To test linear stability, we first apply the splitting method to the harmonic oscillator $y'' + \lambda^2 y = 0$, $\lambda > 0$. We denote $[p,q]^T = [\lambda y, y']^T$ and we consider the standard splitting

$$\begin{bmatrix} p \\ q \end{bmatrix}' = \left\{ \begin{bmatrix} 0 & \lambda \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ -\lambda & 0 \end{bmatrix} \right\} \begin{bmatrix} p \\ q \end{bmatrix} \tag{42}$$

If we now apply a splitting method with a time step $k > 0$, we obtain numerical method

$$\begin{bmatrix} p_{n+1} \\ q_{n+1} \end{bmatrix} = R(\omega) \begin{bmatrix} p_n \\ q_n \end{bmatrix}$$

where $\omega = k\lambda > 0$.

Matrix R, of which the elements are polynomials in the ω variable, is called the stability matrix and is given by

$$R(\omega) = \begin{bmatrix} R_{11}(\omega) & R_{12}(\omega) \\ R_{21}(\omega) & R_{22}(\omega) \end{bmatrix},$$

an it can be computed as

$$R(\omega) = \prod_{j=1}^{m} \begin{bmatrix} 1 & 0 \\ -b_j\omega & 1 \end{bmatrix} \begin{bmatrix} 1 & a_j\omega \\ 0 & 1 \end{bmatrix} = \prod_{j=1}^{m} R_j(\omega), \tag{43}$$

where the product of matrices must be calculated in the correct order.

To obtain stability for the harmonic oscillator, the boundedness of the powers of stability matrix (43) is required. For this, the following definition of "stability interval" is very useful (cf. [12,13]).

Definition 1. *The stability interval of a method with stability matrix $R(\omega)$ is $[0, \omega^*)$ if ω^* is the supremum of values $\omega \geq 0$, such that, for all $\omega \in [0, \omega^*)$,*

$$\rho(R(\omega)) \leq 1,$$

and $R(\omega)$ is simple when $\rho(R(\omega)) = 1$, where $\rho(R(\omega))$ is the spectral radius of $R(\omega)$.

In this way, we have linear stability for test Problem (42) when $k\lambda < \omega_*$. The larger the value of ω_* is, the less restrictive the stability condition is.

Remark 4. *In order to calculate the ω_* value of the stability interval for the methods used in the numerical experiments, Strang, Ψ_{S4} and Ψ_{S6}, we follow the same technique as in [14,15]. Taking into account that* $\det(R(\omega)) = 1$ *from (43) and, for the three methods, $R_{11}(\omega) = R_{22}(\omega)$, the eigenvalues of $R(\omega)$ are the solutions of $\lambda^2 - 2R_{11}(\omega)\lambda + 1 = 0$. Then, to obtain the stability interval, it is enough to study the greatest real value ω_*, such that $R_{11}(\omega)^2 - 1 \leq 0$ for all $\omega \in [0, \omega_*]$. The value of ω_* for the Strang method is 2, for Ψ_{S4} is 6.31 and for Ψ_{S6} is 3.44.*

Regarding Problem (41), linear stability in the energy norm states that the powers of matrix

$$R(kS_{h,0}) = \begin{bmatrix} R_{11}(kS_{h,0}) & R_{12}(kS_{h,0}) \\ R_{21}(kS_{h,0}) & R_{22}(kS_{h,0}) \end{bmatrix},$$

are bounded in the matrix norm induced by the discrete norm in $X_{h,0}$, that is, if $T > 0$ is fixed,

$$\|R^n(kS_{h,0})\|_h \leq C, \tag{44}$$

where C is a constant independent of h, n and k when $nk \leq T$.

Therefore, we need that $k|\lambda_{h,0}| < \omega_*$ for all $\lambda_{h,0} \in \rho(S_{h,0})$. Taking into account that $S_{h,0}$ is symmetric and positive definite, any of its eigenvalues are positive; therefore, it is enough that

$$k\lambda^*_{h,0} < \omega_* \tag{45}$$

is satisfied, where $\lambda^*_{h,0}$ is the largest eigenvalue of $S_{h,0}$.

Remark 5. *We can now deduce the ratio between parameters k and h that must be satisfied to have stability in the energy norm for the numerical experiments in Section 5.*

The eigenvalues of the tridiagonal matrix $diag(1, -2, 1)$, which appears in the matrix (21) used in tests from 1 to 4, are given by $-2 + 2\cos(j\pi/(J+1))$, $j = 1, 2, \ldots, J$, and they all belong to the interval $(-4, 0)$ (see for example [16,17]). We conclude that the largest eigenvalue of $S_{h,0}$ satisfies $\lambda^*_{h,0} < \dfrac{2}{h}$ and, to achieve stability, it suffices that

$$\frac{2k}{h} < \omega^* \Leftrightarrow \frac{k}{h} < \frac{\omega^*}{2}. \tag{46}$$

Similarly, for the two-dimensional problem in Test 5*, the eigenvalues of the block tridiagonal matrix appearing in (24) are included in interval $(-8, 0)$, which means that, in this case, the largest eigenvalue of $S_{h,0}$ satisfies $\lambda^*_{h,0} < \dfrac{2\sqrt{2}}{h}$ and, to achieve stability, it suffices that*

$$\frac{2\sqrt{2}k}{h} < \omega^* \Leftrightarrow \frac{k}{h} < \frac{\omega^*}{2\sqrt{2}}. \tag{47}$$

Stability conditions (46) and (47) for the three splitting methods considered in the numerical experiments of Section 5 are given in Table 5. The missing value for the Strang method for $h = 1/100$ and $k = 1/100$ in Table 3 is due to the instability of the numerical solution because the stability condition was not fulfilled.

Table 5. Stability ratios (46) and (47) for the three splitting methods considered in the numerical experiments.

	Strang	**Ψ_{S4}**	**Ψ_{S6}**
1D	1	3.16	1.72
2D	0.71	2.23	1.22

7.2. Convergence

We now study the convergence of the full discretization of Section 4.

Theorem 3. *Assume that the time discretization of (12) is obtained by means of a splitting method of order p, applied as described in Formulas (16)–(17), with the choice of intermediate boundary values (18). Let* $\mathbf{u}_h(t_n) = [u_{1,h}(t_n), u_{2,h}(t_n)]^T$ *be the value at* t_n *of solution of (13), and let* $\mathbf{u}_{h,n} = [u_{1,h,n}, u_{2,h,n}]^T$ *be its approximation obtained with the splitting method (16). Assume also that (45) and the linear stability condition (44) are satisfied. Then, global error* $\mathbf{e}_{h,n} = \mathbf{u}_h(t_n) - \mathbf{u}_{h,n}$ *satisfies*

$$\begin{aligned} \|\mathbf{e}_{h,n}\|_{E_h} &= \|\mathbf{u}_h(t_n) - \mathbf{u}_{h,n}\|_{E_h} \\ &= \Big(\|S_{h,0}(u_{1,h}(t_n) - u_{1,h,n})\|_h^2 + \|u_{2,h}(t_n) - u_{2,h,n}\|_h^2 \Big)^{1/2} = O(k^p), \end{aligned}$$

Proof. Let $\tilde{\mathbf{u}}_h(t) = [\tilde{u}_{1,h}(t), \tilde{u}_{2,h}(t)]^T$ be the first two components of (33). We showed in Section 6 that $\tilde{\mathbf{u}}_h(t) = [S_{h,0}u_{1,h}(t), u_{2,h}(t)]^T$.

On the other hand, let $\tilde{\mathbf{u}}_{h,n} = [\tilde{u}_{1,h,n}, \tilde{u}_{2,h,n}]^T$ be the first two components of (36). We also showed in Section 6 that $\tilde{\mathbf{u}}_{h,n} = [S_{h,0}u_{1,h,n}, u_{2,h,n}]^T$. Therefore, we can obtain $\tilde{\mathbf{u}}_{h,n}$ from Section 4, using Equations (17) and (19). For this, we make $\tilde{\mathbf{v}}_{j,h,n} = [S_{h,0}v_{1,j,h,n}, v_{2,j,h,n}]^T$, $\tilde{\mathbf{w}}_{j,h,n} = [S_{h,0}w_{1,j,h,n}, w_{2,j,h,n}]^T$, $j = 1, \ldots, m$, and we deduce that

$$\tilde{\mathbf{v}}_{j,h,n} = \begin{bmatrix} \tilde{v}_{1,j,h,n} \\ \tilde{v}_{2,j,h,n} \end{bmatrix} = \begin{bmatrix} I_h & a_j k S_{h,0} \\ 0 & I_h \end{bmatrix} \begin{bmatrix} \tilde{w}_{1,j-1,h,n} \\ \tilde{w}_{2,j-1,h,n} \end{bmatrix} = M_j(kS_{h,0})\tilde{\mathbf{w}}_{j-1,h,n}.$$

Then,

$$\begin{aligned} \tilde{\mathbf{w}}_{j,h,n} &= \begin{bmatrix} \tilde{w}_{1,j,h,n} \\ \tilde{w}_{2,j,h,n} \end{bmatrix} \\ &= \begin{bmatrix} I_h & 0 \\ -b_j k S_{h,0} & I_h \end{bmatrix} \begin{bmatrix} \tilde{v}_{1,j,h,n} \\ \tilde{v}_{2,j,h,n} \end{bmatrix} \\ &\quad + \begin{bmatrix} 0 \\ Q_h g(t_n + \sum\limits_{r=1}^{j} a_r k) \end{bmatrix} + \begin{bmatrix} 0 \\ \int_0^{b_j k} f(t_n + \sum\limits_{r=1}^{j-1} b_r k + \tau, S_{h,0}^{-1}\tilde{v}_{1,j,h,n} d\tau) \end{bmatrix} \\ &= \begin{bmatrix} I_h & 0 \\ -b_j k S_{h,0} & I_h \end{bmatrix} \begin{bmatrix} I_h & a_j k S_{h,0} \\ 0 & I_h \end{bmatrix} \begin{bmatrix} \tilde{w}_{1,j-1,h,n} \\ \tilde{w}_{2,j-1,h,n} \end{bmatrix} \\ &\quad + \begin{bmatrix} 0 \\ Q_h g(t_n + \sum\limits_{r=1}^{j} a_r k) \end{bmatrix} + \begin{bmatrix} 0 \\ \int_0^{b_j k} f(t_n + \sum\limits_{r=1}^{j-1} b_r k + \tau, S_{h,0}^{-1}\tilde{v}_{1,j,h,n} d\tau) \end{bmatrix} \\ &= R_j(kS_{h,0})\tilde{\mathbf{w}}_{j,h,n} \\ &\quad + \begin{bmatrix} 0 \\ Q_h g(t_n + \sum\limits_{r=1}^{j} a_r k) \end{bmatrix} + \begin{bmatrix} 0 \\ \int_0^{b_j k} f(t_n + \sum\limits_{r=1}^{j-1} b_r k + \tau, S_{h,0}^{-1}\tilde{v}_{1,j,h,n} d\tau) \end{bmatrix}. \end{aligned}$$

By using recursive reasoning,

$$\begin{aligned}
\widetilde{\mathbf{w}}_{j,h,n} &= \prod_{l=1}^{j} R_l(kS_{h,0})\widetilde{\mathbf{w}}_{0,h,n} \\
&\quad + \sum_{s=1}^{j}\prod_{l=s+1}^{j} R_l(kS_{h,0}) \begin{bmatrix} 0 \\ Q_h g(t_n + \sum_{r=1}^{s} a_r k) \end{bmatrix} \\
&\quad + \sum_{s=1}^{j}\prod_{l=s+1}^{j} R_l(kS_{h,0}) \begin{bmatrix} 0 \\ \int_0^{b_s k} f(t_n + \sum_{r=1}^{s-1} b_r k + \tau, S_{h,0}^{-1}\widetilde{v}_{1,s,h,n}) d\tau \end{bmatrix}.
\end{aligned}$$

We now define $\overline{\widetilde{\mathbf{w}}}_{0,h,n} = \widetilde{\mathbf{u}}_h(t_n)$, and

$$\begin{aligned}
\overline{\widetilde{\mathbf{w}}}_{j,h,n} &= \prod_{l=1}^{j} R_l(kS_{h,0})\overline{\widetilde{\mathbf{w}}}_{0,h,n} \\
&\quad + \sum_{s=1}^{j}\prod_{l=s+1}^{j} R_l(kS_{h,0}) \begin{bmatrix} 0 \\ Q_h g(t_n + \sum_{r=1}^{s} a_r k) \end{bmatrix} \\
&\quad + \sum_{s=1}^{j}\prod_{l=s+1}^{j} R_l(kS_{h,0}) \begin{bmatrix} 0 \\ \int_0^{b_s k} f(t_n + \sum_{r=1}^{s-1} b_r k + \tau, S_{h,0}^{-1}\overline{\widetilde{v}}_{1,s,h,n}) d\tau \end{bmatrix}
\end{aligned}$$

and, subtracting

$$\overline{\widetilde{\mathbf{w}}}_{j,h,n} - \widetilde{\mathbf{w}}_{j,h,n} = \prod_{l=1}^{j} R_l(kS_{h,0})(\overline{\widetilde{\mathbf{w}}}_{0,h,n} - \widetilde{\mathbf{w}}_{0,h,n}) + F_{n,j}$$

where

$$\begin{aligned}
F_{n,j} &= \textstyle\sum_{s=1}^{j}\prod_{l=s+1}^{j} R_l(kS_{h,0}) \begin{bmatrix} 0 \\ \int_0^{b_s k}[f(t_n + \sum_{r=1}^{s-1} b_r k + \tau, S_{h,0}^{-1}\overline{\widetilde{v}}_{1,s,h,n}) - f(t_n + \sum_{r=1}^{s-1} b_r k + \tau, S_{h,0}^{-1}\widetilde{v}_{1,s,h,n})]d\tau \end{bmatrix} \\
&= R_j(kS_{h,0})F_{n,j-1} + \begin{bmatrix} 0 \\ \int_0^{b_j k}[f(t_n + \sum_{r=1}^{j} b_r k + \tau, S_{h,0}^{-1}\overline{\widetilde{v}}_{1,j,h,n}) - f(t_n + \sum_{r=1}^{j} b_r k + \tau, S_{h,0}^{-1}\widetilde{v}_{1,j,h,n})]d\tau \end{bmatrix}.
\end{aligned}$$

To use an inductive reasoning, we first consider

$$F_{n,1} = \begin{bmatrix} 0 \\ \int_0^{b_1 k}[f(t_n + \tau, S_{h,0}^{-1}\overline{\widetilde{v}}_{1,1,h,n}) - f(t_n + \tau, S_{h,0}^{-1}\widetilde{v}_{1,1,h,n})]d\tau \end{bmatrix}$$

and, taking norm,

$$\begin{aligned}
\|F_{n,1}\|_h &\le |b_1|k|L\|S_{h,0}^{-1}\|_h\|M_1(kS_{h,0})\|_h\|\overline{\widetilde{\mathbf{w}}}_{0,h,n} - \overline{\widetilde{\mathbf{w}}}_{0,h,n}\|_h \\
&= kC_{1,1}\|\overline{\widetilde{\mathbf{w}}}_{0,h,n} - \overline{\widetilde{\mathbf{w}}}_{0,h,n}\|_h,
\end{aligned}$$

where, if (45) is satisfied and taking into account Hypothesis (H2), constant $C_{1,1}$ is independent on h, k, and n.

On the other hand,

$$\overline{\widetilde{\mathbf{w}}}_{1,h,n} - \widetilde{\mathbf{w}}_{1,h,n} = R_1(kS_{h,0})(\overline{\widetilde{\mathbf{w}}}_{0,h,n} - \widetilde{\mathbf{w}}_{0,h,n}) + F_{n,1}$$

and, taking again norm

$$\begin{aligned} \|\overline{\widetilde{\mathbf{w}}}_{1,h,n} - \widetilde{\mathbf{w}}_{1,h,n}\|_h &\leq \|R_1(kS_{h,0})\|_h \|\overline{\widetilde{\mathbf{w}}}_{0,h,n} - \widetilde{\mathbf{w}}_{0,h,n}\|_h + \|F_{n,1}\|_h \\ &\leq (\|R_1(kS_{h,0})\|_h + kC_{1,1})\|\|\overline{\widetilde{\mathbf{w}}}_{0,h,n} - \widetilde{\mathbf{w}}_{0,h,n}\|_h \\ &= C_{1,2}\|\overline{\widetilde{\mathbf{w}}}_{0,h,n} - \widetilde{\mathbf{w}}_{0,h,n}\|_h, \end{aligned}$$

where constant $C_{1,2}$ is also independent on k, n, and h because of (45) and Hypothesis (H2).

Now, we use inductive reasoning, assuming that

$$\|F_{n,j-1}\|_h \leq kC_{j-1,1}\|\overline{\widetilde{\mathbf{w}}}_{0,h,n} - \widetilde{\mathbf{w}}_{0,h,n}\|_h,$$

and,

$$\|\overline{\widetilde{\mathbf{w}}}_{j-1,h,n} - \widetilde{\mathbf{w}}_{j-1,h,n}\|_h \leq C_{j-1,2}\|\overline{\widetilde{\mathbf{w}}}_{0,h,n} - \widetilde{\mathbf{w}}_{0,h,n}\|_h$$

where constants $C_{j-1,1}$ and $C_{j-1,2}$ are independent on k, n, and h when (45) is satisfied.

Then,

$$\begin{aligned} \|F_{n,j}\|_h &\leq \|R_j(kS_{h,0})\|_h \|F_{n,j-1}\|_h + |b_j| kL \|S_{h,0}^{-1}\|_h \|M_j(kS_{h,0})\|_h \|\overline{\widetilde{\mathbf{w}}}_{j-1,h,n} - \widetilde{\mathbf{w}}_{j-1,h,n}\|_h \\ &\leq \|R_j(kS_{h,0})\|_h kC_{j-1,1}\|\overline{\widetilde{\mathbf{w}}}_{0,h,n} - \widetilde{\mathbf{w}}_{0,h,n}\|_h \\ &\quad + |b_j| kL \|S_{h,0}^{-1}\|_h \|M_j(kS_{h,0})\|_h C_{j-1,2}\|\overline{\widetilde{\mathbf{w}}}_{0,h,n} - \widetilde{\mathbf{w}}_{0,h,n}\|_h \\ &= kC_{j,1}\|\overline{\widetilde{\mathbf{w}}}_{0,h,n} - \widetilde{\mathbf{w}}_{0,h,n}\|_h \end{aligned}$$

and

$$\begin{aligned} \|\overline{\widetilde{\mathbf{w}}}_{j,h,n} - \widetilde{\mathbf{w}}_{j,h,n}\|_h &\leq \|\prod_{l=1}^{j} R_l(kS_{h,0})\|_h \|\overline{\widetilde{\mathbf{w}}}_{0,h,n} - \widetilde{\mathbf{w}}_{0,h,n}\|_h + \|F_{n,j}\|_h \\ &\leq \|\prod_{l=1}^{j} R_l(kS_{h,0})\|_h \|\overline{\widetilde{\mathbf{w}}}_{0,h,n} - \widetilde{\mathbf{w}}_{0,h,n}\|_h + kC_{j,1}\|\overline{\widetilde{\mathbf{w}}}_{0,h,n} - \widetilde{\mathbf{w}}_{0,h,n}\|_h \\ &= C_{j,2}\|\overline{\widetilde{\mathbf{w}}}_{0,h,n} - \widetilde{\mathbf{w}}_{0,h,n}\|_h. \end{aligned}$$

Lastly, we prove convergence taking into account that $\widetilde{\mathbf{w}}_{0,h,n-1} = \widetilde{\mathbf{u}}_{h,n-1}$, $\overline{\widetilde{\mathbf{w}}}_{0,h,n-1} = \widetilde{\mathbf{u}}_h(t_{n-1})$, $\widetilde{\mathbf{w}}_{m,h,n-1} = \widetilde{\mathbf{u}}_{h,n}$, $\overline{\widetilde{\mathbf{w}}}_{m,h,n-1} = \overline{\widetilde{\mathbf{u}}}_{h,n}$. Moreover, we denote $F_{n,m} = F_n$ and $C_{m,1} = C_1$.

The global error is given by

$$\begin{aligned} \widetilde{\mathbf{e}}_{h,n} &= \begin{bmatrix} \widetilde{e}_{1,h,n} \\ \widetilde{e}_{2,h,n} \end{bmatrix} = \begin{bmatrix} \widetilde{u}_{1,h}(t_n) \\ \widetilde{u}_{2,h}(t_n) \end{bmatrix} - \begin{bmatrix} \widetilde{u}_{1,h,n} \\ \widetilde{u}_{2,h,n} \end{bmatrix} \\ &= \begin{bmatrix} \widetilde{u}_{1,h}(t_n) \\ \widetilde{u}_{2,h}(t_n) \end{bmatrix} - \begin{bmatrix} \overline{\widetilde{u}}_{1,h,n} \\ \overline{\widetilde{u}}_{2,h,n} \end{bmatrix} + \begin{bmatrix} \overline{\widetilde{u}}_{1,h,n} \\ \overline{\widetilde{u}}_{2,h,n} \end{bmatrix} - \begin{bmatrix} \widetilde{u}_{1,h,n} \\ \widetilde{u}_{2,h,n} \end{bmatrix} \\ &= \widetilde{\boldsymbol{\ae}}_{h,n} + R(kS_{h,0})\widetilde{\mathbf{e}}_{h,n-1} + F_n \end{aligned}$$

where $\widetilde{\rho}_{h,n}$ is the semidiscrete local error at $t = t_n$.

Therefore, we deduce that

$$\widetilde{\mathbf{e}}_{h,n} = \sum_{j=1}^{n} R^{n-j}(kS_{h,0})\widetilde{\boldsymbol{\ae}}_{h,j} + \sum_{j=1}^{n} R^{n-j}(kS_{h,0})F_j.$$

Taking norms and using stability Hypothesis (44)

$$\begin{aligned}
\|\widetilde{\mathbf{e}}_{h,n}\|_h &\leq \sum_{j=1}^{n} \|R^{n-j}(kS_{h,0})\|_h \|\widetilde{\textbf{\ae}}_{h,j}\|_h + \sum_{j=1}^{n} \|R^{n-j}(kS_{h,0})\|_h \|F_j\|_h \\
&\leq C \sum_{j=1}^{n} \|\widetilde{\rho}_{h,j}\|_h + C \sum_{j=1}^{n} \|F_j\|_h \\
&\leq C_1 k^p + C_2 \sum_{j=1}^{n} k \|\widetilde{e}_{h,j-1}\|_h \\
&= C_1 k^p + C_2 \sum_{j=0}^{n-1} k \|\widetilde{e}_{h,j}\|_h.
\end{aligned}$$

The proof of convergence is achieved by using the discrete Gronwall lemma. □

Author Contributions: Conceptualization, I.A.-M. and A.M.P.; Formal analysis, I.A.-M. and A.M.P.; Funding acquisition, I.A.-M.; Investigation, I.A.-M. and A.M.P.; Software, A.M.P.; Writing–review and editing, I.A.-M. and A.M.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministerio de Ciencia y Educación, grant number PGC2018-101443-B-I00, and the first author by Consejería de Educación, Junta de Castilla y León y Leónand Feder funds, grant number VA193P20.

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

References

1. Blanes, S.; Casas, F.; Murua, A. Splitting and composition methods in the numerical integration of differential equation. *Bol. Soc. Esp. Mat. Apl.* **2008**, *45*, 89–145.
2. McLachlan, R.I. On the numerical integration of ordinary differential equations by symmetric composition methods. *SIAM J. Sci. Comput.* **1995**, *16*, 151–168. [CrossRef]
3. McLachlan, R.I.; Quispel, G.R. Splitting methods. *Acta Numer.* **2002**, *11*, 341–434. [CrossRef]
4. Hairer, E.; Wanner, G.; Lubich, C. *Geometric Numerical Integration*; Springer: Berlin/Heidelberg, Germany, 2006.
5. Sanz-Serna, J.M.; Calvo, M.P. *Numerical Hamiltonian Problems*; Chapmand and Hall: London, UK, 1994.
6. Cano, B.; Moreta, M.J. A modified Gautschi's method without order reduction when integrating boundary value nonlinear wave problems. *Appl. Math. Comput.* **2020**, *373*, 125022. [CrossRef]
7. Alonso–Mallo, I.; Palencia, C. On the convolutions operators arising in the study of abstract initial boundary value problems. *Proc. Roy. Soc. Edinb. Sect.* **1996**, *126*, 515–539. [CrossRef]
8. Palencia, C.; Alonso-Mallo, I. Abstract initial-boundary value problems. *Proc. R. Soc. Edinb. Sect. A* **1994**, *124*, 879–908. [CrossRef]
9. Arendt, W.; Batty, C.F.K.; Hieber, M.; Neubrander, F. Vector-Valued Laplace Transforms and Cauchy Problems. In *Monographs in Mathematics*; Birkhäuser: Basel, Switzerland, 2001; Volume 96.
10. Fattorini, H.O. *Second Order Linear Differential Equations in Banach Spaces*, 1st ed.; North-Holland Mathematics Studies; Notas de Matemática: Amsterdam, The Netherlands, 2000; Volume 108.
11. Alonso–Mallo, I. Explicit single step methods with optimal order of convergence for partial differential equations. *Appl. Numer. Math.* **1999**, *31*, 117–131. [CrossRef]
12. Alonso–Mallo, I.; Cano, B.; Moreta, M.J. Stability of Runge-Kutta-Nyström methods. *J. Comput. Appl. Math.* **2006**, *189*, 120–131. [CrossRef]
13. Alonso–Mallo, I.; Cano, B.; Moreta, M.J. The stability of rational approximations of cosine functions on Hilbert spaces. *Appl. Numer. Math.* **2009**, *59*, 21–38. [CrossRef]
14. Blanes, S.; Casas, F.; Murua, A. On the linear stability of splitting methods. *Found. Comp. Math.* **2008**, *8*, 357–393. [CrossRef]
15. Portillo, A. M. High order full discretization for anisotropic wave equations. *Appl. Math. Comput.* **2018**, *323*, 1–16. [CrossRef]
16. Elliott, J. F. The Characteristic Roots of Certain Real Symmetric Matrices. Master's Thesis, University of Tennessee, Knoxville, TN, USA, 1953.
17. Gregory, R.T.; Karney, D. *A Collection of Matrices for Testing Computational Algorithm*; Wiley: New York, NY, USA, 1969.