arXiv:1909.13215v1 [math.NA] 29 Sep 2019

Energy Stability of Explicit Runge–Kutta Methods for Non-autonomous or Nonlinear Problems

Hendrik Ranocha, David I. Ketcheson

September 29, 2019

Many important initial value problems have the property that energy is non-increasing in time. Energy stable methods, also referred to as strongly stable methods, guarantee the same property discretely. We investigate requirements for conditional energy stability of explicit Runge–Kutta methods for nonlinear or non-autonomous problems. We provide both necessary and sufficient conditions for energy stability over these classes of problems. Examples of conditionally energy stable schemes are constructed and an example is given in which unconditional energy stability is obtained with an explicit scheme.

Keywords. Runge–Kutta methods, energy stability, strong stability, monotonicity, semiboundedness, dissipation, conservation

Mathematics Subject Classification (2010). 65L06, 65L20, 65M12, 65M20

1 Introduction

Ever since the construction of numerical methods for ordinary and (time-dependent) partial differential equations (ODEs and PDEs, respectively), their stability has been an important and active topic of research. Monotonicity, meaning that the norm of the solution is bounded by its initial value, is a particularly exacting stability property. For equations, such as parabolic PDEs, that contain a significant amount of dissipation, any reasonable numerical method will typically preserve monotonicity under an appropriate time step restriction. In contrast, for non-dissipative problems such as hyperbolic PDEs (and their slightly dissipative semidiscretizations), common time discretizations may not preserve monotonicity under any finite step size.

The energy method is an effective tool to get stability estimates, e.g. for hyperbolic PDEs [12, 19]. Using summation by parts operators [9, 34], these can be transferred efficiently to the semidiscrete level for many different kinds of schemes [11, 20, 21, 28]. However, applying the same approach in time yields implicit methods [2, 10, 22, 25]. Classical nonlinearly stable methods, such as algebraically stable Runge–Kutta methods, are also implicit. For hyperbolic problems, such implicit methods are usually less efficient than explicit ones. It is possible to obtain conditional energy stability with explicit methods by using modifications that go outside the class of Runge–Kutta methods; e.g. projection methods [5, 6, 13] and relaxation Runge–Kutta schemes [18, 29].

Nevertheless, it is interesting to know what can be achieved within the class of explicit Runge–Kutta methods without modifications. In this setting, results have been obtained for problems that include a certain amount of dissipation [8, 16]. Recently this topic has again attracted the interest of researchers and several results (using the term *strong stability*) for linear, time-independent operators have been discovered [27, 32, 33]. Nonlinear problems have been investigated in [24], where many non-existence results for energy stable and strong stability preserving (SSP) methods of order two and greater have been proved. A first order accurate energy stable SSP method for autonomous problems has also been discovered therein.

This article extends these previous works considerably by studying both time-dependent linear and autonomous nonlinear problems. After introducing the notation and reviewing some basic results in Section 2, the focus lies on time-dependent linear operators in Section 3. The main result, Theorem 3.4, gives necessary conditions for conditional energy stability in this setting. These conditions are not satisfied by any known Runge–Kutta scheme we are aware of. However, an example of a scheme fulfilling these necessary conditions is given, and is proved to be energy stable for a restricted class of relevant problems (Theorem 3.12).

Next, autonomous nonlinear problems are studied in Section 4. The necessary conditions in this setting are, perhaps surprisingly, weaker than in the non-autonomous linear case. These conditions are based on an expansion of the change of energy (26), which is also used to study sufficient conditions for energy stability. Based thereon, we give a procedure for developing energy stable schemes, and give examples of schemes of second and third order (Theorem 4.11 and Corollary 4.15).

While most of the paper is devoted to the guarantee of stability over a whole class of semibounded problems, in Section 5 we ask whether an explicit Runge–Kutta method can be unconditionally energy stable for some specific problem. We show that this is impossible if the problem is linear, but – surprisingly – we give an example of unconditional stability for a nonlinear problem. Finally, in Section 6, the results are summed up, open questions are discussed, and directions of future research are outlined.

2 Energy Evolution by Runge–Kutta Methods

Consider a time-dependent initial value problem

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t} u(t) &= f(t, u(t)), \quad t \in (0, T), \\ u(0) &= u_0, \end{aligned} \tag{1}$$

in a real Hilbert space $\mathcal{H}$ with inner product $\langle \cdot, \cdot \rangle$, inducing the norm $\|\cdot\|$. We refer to $\|\cdot\|^2$ as the energy.

2.1 Energy Stability

For a smooth solution of (1), the time derivative of the energy is

$$\frac{\mathrm{d}}{\mathrm{d}t} \|u(t)\|^2 = 2 \left\langle u(t),\, \tfrac{\mathrm{d}}{\mathrm{d}t} u(t) \right\rangle = 2 \left\langle u(t),\, f(t, u(t)) \right\rangle. \tag{2}$$

Definition 2.1. The right hand side $f\colon [0,T] \times \mathcal{H} \to \mathcal{H}$ of (1) is *semibounded*, if

$$\forall u \in \mathcal{H}, t \in [0,T]\colon \quad \left\langle u,\, f(t,u) \right\rangle \leq 0. \tag{3}$$

If f is semibounded, the ODE (1) will sometimes also be called semibounded. ◃

Remark 2.2. The results in this work extend to complex Hilbert spaces if one assumes that the real part of the inner product $\langle u,\, f(t,u) \rangle$ is non-positive. ◃

Thus, the energy of any smooth solution of (1) is bounded by its initial value if f is semibounded. However, an approximate solution obtained by a numerical method does not necessarily satisfy this inequality. For example, applying one step of the explicit Euler method to (1) yields the new value $u_+ = u_0 + \Delta t\, f(0, u_0)$, satisfying

$$\|u_+\|^2 = \left\|u_0 + \Delta t\, f(0,u_0)\right\|^2 = \|u_0\|^2 + \underbrace{2\Delta t \left\langle u_0,\, f(0,u_0) \right\rangle}_{\leq 0} + \underbrace{\Delta t^2 \left\|f(0,u_0)\right\|^2}_{\geq 0}. \tag{4}$$

Thus, for a general semibounded f, the norm of the numerical solution can increase during one time step, e.g. if $\langle u_0,\, f(0,u_0) \rangle = 0$. In particular, this happens if $f(t,u) = L(t)u$, where $L(t) \neq 0$ is a skew-symmetric operator.

Definition 2.3. A one-step numerical scheme for approximating the solution of (1) is *conditionally energy stable* with respect to a class $\mathcal{F}$ of semibounded problems if for each $f \in \mathcal{F}$ there exists $\Delta t_{\max} > 0$ such that $\|u_+\|^2 \leq \|u_0\|^2$ for all $0 < \Delta t \leq \Delta t_{\max}$. ◁

Here $\Delta t_{\max}$ may depend on f, u_0, and the method itself. In the following we will consider the classes $\mathcal{F}$ of linear non-autonomous and nonlinear autonomous problems. We will often omit the word *conditional* for brevity.

2.2 Runge–Kutta Methods

A general (explicit or implicit) Runge–Kutta method with s stages can be described by its Butcher tableau [4, 14]

$$\begin{array}{c|c} c & A \\ \hline & b \end{array} \tag{5}$$

where $A \in \mathbb{R}^{s\times s}$ and $b, c \in \mathbb{R}^s$. For (1), a step from u_0 to u_+ is given by

$$u_i = u_0 + \Delta t \sum_{j=1}^{s} a_{ij}\, f(c_j \Delta t, u_j), \qquad u_+ = u_0 + \Delta t \sum_{i=1}^{s} b_i\, f(c_i \Delta t, u_i). \tag{6}$$

Here, u_i are the stage values of the Runge–Kutta method. It is also possible to express the method via the slopes $k_i = f(c_i \Delta t, u_i)$.

Using the stage values u_i as in (6), the change in energy is given by [4, equation (357e)]

$$\begin{aligned} \|u_+\|^2 - \|u_0\|^2 = 2\Delta t \sum_{i=1}^{s} b_i \left\langle u_i,\ f(c_i \Delta t, u_i)\right\rangle \\ + (\Delta t)^2 \left[\sum_{i,j=1}^{s} \left(b_i b_j - b_i\, a_{ij} - b_j a_{ji} \right) \left\langle f(c_i \Delta t, u_i),\ f(c_j \Delta t, u_j)\right\rangle \right]. \end{aligned} \tag{7}$$

The first term on the right hand side is consistent with $\int_{t_0}^{t_0+\Delta t} 2 \left\langle u(t),\ f(t, u(t))\right\rangle \mathrm{d}t$, if the Runge–Kutta method is consistent, i.e. $\sum_{i=1}^{s} b_i = 1$. Semiboundedness of f implies that this term is non-positive if all b_i are non-negative.

We recall the classical property of algebraic stability, which guarantees energy stability for semibounded operators; cf. [4, section 357] and references cited therein.

Definition 2.4. A Runge–Kutta method is *algebraically stable* if $b_i \geq 0$ for all i and the matrix with entries $(b_i b_j - b_i\, a_{ij} - b_j a_{ji})_{i,j}$ is negative semidefinite. ◁

Comparison with (7) shows that algebraically stable Runge–Kutta methods are energy stable for any time step size $\Delta t > 0$, cf. [4, section 357] and references cited therein. While there are Runge–Kutta methods with these nice stability properties such as Gauß, Radau IA/IIA or Lobatto IIIC schemes, these are necessarily implicit.

For linear and semibounded problems (1) with constant coefficients, several results concerning the conditional energy stability of explicit Runge–Kutta methods have been achieved [27, 32, 33, 35] (note that the term *conditional energy stability* herein is precisely what is meant by *strong stability* in the latter works). Typically, conditional energy stability can be guaranteed for problems $f(t, u) = Lu$ in this class under a time step restriction of the form $\Delta t \leq C\|L\|^{-1}$, corresponding to a classical CFL criterion for discretizations of hyperbolic conservation laws. Similar results have been obtained for some first order accurate schemes and autonomous semibounded nonlinear problems [24]. In the latter setting, the maximal time step is proportional to the inverse of the Lipschitz constant of the nonlinear right-hand side f of the ODE.

3 Non-autonomous Linear Operators

In this section, the special case of non-autonomous linear operators is studied. Hence,

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t} u(t) &= L(t) u(t), \quad t \in (0, T), \\ u(0) &= u_0, \end{aligned} \tag{8}$$

is considered as special case of (1).

To formulate our results, we need the following definitions regarding the abscissae, or nodes, of a Runge–Kutta method.

Definition 3.1. We say node c_k of a Runge–Kutta method is *unique* if there is no other node $c_j = c_k$ such that $j \neq k$. ◁

Definition 3.2. A Runge–Kutta method is said to be *non-confluent* if each of its nodes is unique. Otherwise, it is called *confluent*. ◁

Definition 3.3. We say the node c_k of a Runge–Kutta method is a *quadrature node* if $b_k \neq 0$. ◁

3.1 Main Result

For linear problems with constant coefficients, some common and practical Runge–Kutta methods have been shown to be conditionally energy stable [27, 32, 33, 35]. For linear problems with varying coefficients, energy stability is more difficult to attain. Given almost any Runge–Kutta method, we can choose a non-autonomous problem (8) that makes the given method behave like Euler's method, leading to energy growth.

Theorem 3.4. *An explicit Runge–Kutta method with a unique quadrature node cannot be energy stable for semibounded linear problems* (8) *under a time step restriction depending only on an upper bound of* $\|L(t)\|$ *and the Lipschitz constant of* $t \mapsto L(t)$.

Proof. Given $\Delta t > 0$, choose $u_0 \neq 0$, $L(t) = \lambda(t)\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ with $\lambda(c_k \Delta t) > 0$ sufficiently small and $\lambda(c_j) = 0$ for all $j \neq k$. The function λ can be continued as a Lipschitz continuous function with arbitrarily small Lipschitz constant depending on $\lambda(c_k \Delta t) > 0$, cf. e.g. Kirszbraun's theorem [30, Theorem 1.31]. Hence, $\|L(t)\|$ and the Lipschitz constant of $t \mapsto L(t)$ can be made arbitrarily small.

Then, the first step of the given explicit Runge–Kutta method yields

$$\begin{aligned} u_+ &= u_0 + \Delta t \sum_{i=1}^{s} b_i L(c_i \Delta t) u_i = u_0 + \Delta t b_k L(c_k \Delta t) u_k \\ &= u_0 + \Delta t b_k L(c_k \Delta t) \left(u_0 + \Delta t \sum_{j=1}^{k-1} a_{kj} L(c_j \Delta t) u_j \right) = u_0 + \Delta t b_k L(c_k \Delta t) u_0. \end{aligned} \tag{9}$$

This is equivalent to an explicit Euler step with time step $b_k \Delta t \neq 0$. Since $L(c_k \Delta t) \neq 0$ is skew-symmetric and injective,

$$\|u_+\|^2 - \|u_0\|^2 = b_k^2 \Delta t^2 \|L(c_k \Delta t) u_0\|^2 > 0. \tag{10}$$

Hence, the explicit Runge–Kutta method is not energy stable. □

Remark 3.5. The function L appearing in the proof can be made arbitrarily smooth by considering classical cut-off functions (Friedrichs mollifier). ◁

Remark 3.6. The proof holds also for linear scalar problems in the complex plane if $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ is replaced by the imaginary unit i. ◁

Remark 3.7. In a non-confluent method, every node is unique, so Theorem 3.4 implies such methods cannot be energy stable. Moreover, it seems that all methods currently used in practice have at least one unique quadrature node. The only schemes not covered by the theorem are those for which *every quadrature node is repeated*. ◁

Remark 3.8. The proof shows additionally that $a_{kk} \neq 0$ (taking k as in the proof) is a necessary condition for an implicit Runge–Kutta scheme to be energy stable. In particular, the Lobatto IIIA and IIIB schemes cannot be energy stable because they have a zero row or column in A and only positive b_i. ◁

Remark 3.9. The technique used in the proof of Theorem 3.4 cannot be extended to arbitrary confluent Runge–Kutta schemes. Indeed, consider the Butcher tableau

$$
\begin{array}{c|ccc}
0 & 0 & & \\
1 & 1 & 0 & \\
0 & 1 & -1 & 0 \\
\hline
 & -1/2 & 0 & 3/2
\end{array}
\tag{11}
$$

and the associated Runge–Kutta method for (8)

$$
\begin{aligned}
u_1 &= u_0, \\
u_2 &= u_0 + \Delta t L(0) u_1, \\
u_3 &= u_0 + \Delta t L(0) u_1 - \Delta t L(\Delta t) u_2, \\
u_+ &= u_0 + \frac{\Delta t}{2} L(0)(-u_1 + 3u_3).
\end{aligned}
\tag{12}
$$

If $L(t)$ is skew-symmetric, $L(0) \neq 0$, and $L(\Delta t) = 0$,

$$
u_+ = u_0 + \frac{\Delta t}{2} L(0)(-u_0 + 3u_0 + 3\Delta t L(0) u_0) = u_0 + \Delta t L(0) u_0 + \frac{3}{2}\Delta t^2 L(0)^2 u_0,
\tag{13}
$$

and

$$
\begin{aligned}
\|u_+\|^2 - \|u_0\|^2 &= 2\Delta t \left\langle u_0,\, L(0)u_0 \right\rangle + 3\Delta t^2 \left\langle u_0,\, L(0)^2 u_0 \right\rangle + \Delta t^2 \left\| L(0) u_0 \right\|^2 \\
&\quad + 3\Delta t^3 \left\langle L(0)u_0,\, L(0)^2 u_0 \right\rangle + \frac{9}{4}\Delta t^4 \left\| L(0)^2 u_0 \right\|^2 \\
&= -2\Delta t^2 \left\| L(0) u_0 \right\|^2 + \frac{9}{4}\Delta t^4 \left\| L(0)^2 u_0 \right\|^2 .
\end{aligned}
\tag{14}
$$

If Δt is sufficiently small, $\|u_+\|^2 - \|u_0\|^2 < 0$.

Similarly, if $L(t)$ is skew-symmetric, $L(0) = 0$, and $L(\Delta t) \neq 0$, $u_+ = u_0$ and the first time step is energy stable.

However, this does not mean that the scheme (11) is energy stable for semibounded operators. Indeed, its stability function

$$
\varphi(z) = \frac{\det(\mathrm{I} - zA + zeb^T)}{\det(\mathrm{I} - zA)} = 1 + z - \frac{3}{2} z^3
\tag{15}
$$

satisfies $\left|\varphi(\mathrm{i}y)\right| > 1$ for $0 \neq y \in \mathbb{R}$. Hence, the scheme is not even stable for linear skew-symmetric operators with constant coefficients. ◁

Remark 3.10. The proof above relies on the fact that if a Runge–Kutta method has a unique node c_k, then the value of the corresponding stage derivative $L(c_k \Delta t)$ can be chosen independently of the values of all other stages. For methods with only duplicated nodes, the values of the stage derivatives become coupled and the problematic construction in the proof is precluded. ◁

The assumption of Theorem 3.4 may be necessary. This conjecture is supported by the following scheme.

Example 3.11. The explicit Runge–Kutta method with Butcher coefficients

$$
\begin{array}{c|cccc}
0 & & & & \\
1 & 1 & & & \\
0 & 1 & -1 & & \\
1 & -1 & 1 & 1 & \\
\hline
 & 1/4 & 1/4 & 1/4 & 1/4
\end{array}
\tag{16}
$$

has four stages, is second order accurate, and does not have a unique node. ◁

The proof of Theorem 3.4 relies on the construction of a suitable operator L such that $L(t_1)L(t_2) = L(t_2)L(t_1)$ for all $t_1, t_2 \in [0, T]$ and $\forall t \in [0, T]$: $L(t) = -L(t)^T$. For such operators, the scheme (16) is energy stable.

Theorem 3.12. *The Runge–Kutta scheme with coefficients* (16) *is conditionally energy stable for the class of linear non-autonomous ODEs* (8) *where the operator L is bounded and satisfies $\forall t_1, t_2 \in [0,T]\colon L(t_1)L(t_2) = L(t_2)L(t_1)$ and $\forall t \in [0,T]\colon L(t) = -L(t)^T$.*

Proof. Using $L_0 := L(0)$ and $L_1 := L(\Delta t)$, the change of the norm can be calculated explicitly for general L as

$$
\begin{aligned}
\|u_+\|^2 - \|u_0\|^2 = \Delta t\Big(&\langle u_0,\, L_0u_0\rangle + \langle u_0,\, L_1u_0\rangle\Big) \qquad (17)\\
+ \frac{1}{2}\Delta t^2\Big(&\langle u_0,\, L_0^2u_0\rangle - \langle u_0,\, L_0L_1u_0\rangle + \langle u_0,\, L_1L_0u_0\rangle + \langle u_0,\, L_1^2u_0\rangle\\
&+ \frac{1}{2}\|L_0u_0\|^2 + \langle L_0u_0,\, L_1u_0\rangle + \frac{1}{2}\|L_1u_0\|^2\Big)\\
+ \frac{1}{2}\Delta t^3\Big(&-\langle u_0,\, L_0L_1L_0u_0\rangle + \langle u_0,\, L_1L_0^2u_0\rangle - \langle u_0,\, L_1L_0L_1u_0\rangle + \langle u_0,\, L_1^2L_0u_0\rangle\\
&+ \frac{1}{2}\langle L_0u_0,\, L_0^2u_0\rangle - \frac{1}{2}\langle L_0u_0,\, L_0L_1u_0\rangle + \frac{1}{2}\langle L_0u_0,\, L_1L_0u_0\rangle + \frac{1}{2}\langle L_0u_0,\, L_1^2u_0\rangle\\
&+ \frac{1}{2}\langle L_0^2u_0,\, L_1u_0\rangle + \frac{1}{2}\langle L_1u_0,\, L_1L_0u_0\rangle - \frac{1}{2}\langle L_1u_0,\, L_0L_1u_0\rangle + \frac{1}{2}\langle L_1u_0,\, L_1^2u_0\rangle\Big)\\
+ \frac{1}{2}\Delta t^4\Big(&-\langle u_0,\, L_1L_0L_1L_0u_0\rangle - \frac{1}{2}\langle L_0u_0,\, L_0L_1L_0u_0\rangle + \frac{1}{2}\langle L_0u_0,\, L_1L_0^2u_0\rangle\\
&- \frac{1}{2}\langle L_0u_0,\, L_1L_0L_1u_0\rangle + \frac{1}{2}\langle L_0u_0,\, L_1^2L_0u_0\rangle + \frac{1}{8}\left\|L_0^2u_0\right\|^2 - \frac{1}{4}\langle L_0^2u_0,\, L_0L_1u_0\rangle\\
&+ \frac{1}{4}\langle L_0^2u_0,\, L_1L_0u_0\rangle + \frac{1}{4}\langle L_0^2u_0,\, L_1^2u_0\rangle + \frac{1}{2}\langle L_1u_0,\, L_1L_0^2u_0\rangle\\
&- \frac{1}{4}\langle L_0L_1u_0,\, L_1L_0u_0\rangle + \frac{1}{8}\|L_1L_0u_0\|^2 + \frac{1}{4}\langle L_1L_0u_0,\, L_1^2u_0\rangle\\
&- \frac{1}{2}\langle L_1u_0,\, L_0L_1L_0u_0\rangle + \frac{1}{2}\langle L_1u_0,\, L_1^2L_0u_0\rangle - \frac{1}{2}\langle L_1u_0,\, L_1L_0L_1u_0\rangle\\
&+ \frac{1}{8}\|L_0L_1u_0\|^2 - \frac{1}{4}\langle L_0L_1u_0,\, L_1^2u_0\rangle + \frac{1}{8}\left\|L_1^2u_0\right\|^2\Big) + \mathcal{O}(\Delta t^5).
\end{aligned}
$$

Inserting the assumptions on L, this equation reduces to

$$
\begin{aligned}
\|u_+\|^2 - \|u_0\|^2 = &-\frac{1}{4}\Delta t^2\|L_0u_0 - L_1u_0\|^2 \qquad (18)\\
&+ \frac{1}{16}\Delta t^4\left(\left\|L_0^2u_0\right\|^2 - 7\|L_0L_1u_0\|^2 + \left\|L_1^2u_0\right\|^2\right) + \mathcal{O}(\Delta t^5).
\end{aligned}
$$

The Δt^2 term is nonpositive. If the Δt^2 term is negative, the energy is non-increasing if the time step Δt is small enough.

Otherwise, L_0u_0 must be equal to L_1u_0 and the Δt^4 term is $-\frac{1}{4}\Delta t^4\left\|L_0^2u_0\right\|^2 \le 0$. If this term vanishes, all products of higher powers of L_0, L_1 and u_0 must vanish, too, i.e. $\|u_+\|^2 = \|u_0\|^2$. □

3.2 Boundedness

Using an additional technical assumption, the construction used to prove Theorem 3.4 can be used to show that the growth of the norm is unbounded.

Theorem 3.13. *Let an explicit Runge–Kutta method be given. Suppose there exists a quadrature node c_k such that for all $j \ne k$ the difference $c_j - c_k$ is not an integer. Then there exists a semibounded problem* (8) *such that the numerical solution given by the method grows monotonically without bound.*

Proof. The construction used in the proof of Theorem 3.4 can be applied to all steps of the given Runge–Kutta method simultaneously, since $\forall l \in \mathbb{Z}, j \ne k\colon c_k \ne c_j + l$. Hence, for every step from $u^{(n)}$ to $u^{(n+1)}$,

$$\left\|u^{(n+1)}\right\|^2 - \left\|u^{(n)}\right\|^2 = b_k^2 \Delta t^2 \left|\lambda(c_k \Delta t)\right|^2 \left\|u^{(n)}\right\|^2 > 0. \tag{19}$$

Therefore, the norms of the numerical solutions grow monotonically and without bounds. □

Example 3.14. Besides SSPRK(10,4) of [17], Theorem 3.13 can be applied to the popular explicit strong stability preserving Runge–Kutta method SSPRK(3,3) of [31]. For this scheme we have found other ODEs that appear to lead to unbounded growth of the energy; e.g. $L(t) = \sin(t)\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and the example in Subsection 3.4. ◁

3.3 Connection to Algebraic Stability

Adapting a result of Burrage and Butcher [3] slightly, energy stability for (8) is equivalent to algebraic stability for some schemes.

Theorem 3.15. *A non-confluent Runge–Kutta method (i.e. with distinct nodes c_i) that is energy stable for semibounded linear problems* (8) *under a time step restriction depending on an upper bound of $\left\|L(t)\right\|$ and the Lipschitz constant of $t \mapsto L(t)$ must be algebraically stable.*

Proof. This proof is more or less a repetition of a result of Burrage and Butcher [3], enhanced by noting that the varying coefficient can be chosen such that $\left\|L(t)\right\|$ and the Lipschitz constant of $t \mapsto L(t)$ are arbitrarily small. For completeness, the proof is given in the following.

Choosing $L(t) = -\lambda(t) \in \mathbb{R}$, $\lambda(c_i \Delta t) = \varepsilon$, and $\lambda(c_j \Delta t) = 0$ for $j \neq i$ with $\varepsilon > 0$ sufficiently small, L can be extended to a smooth mapping with arbitrarily small $\left\|L(t)\right\|$ and Lipschitz constant of $t \mapsto L(t)$. Hence, energy stability and (7) imply

$$\|u_+\|^2 - \|u_0\|^2 = -2\varepsilon \Delta t b_i \|u_i\|^2 + \varepsilon^2 \Delta t^2 (b_i^2 - 2 b_i a_{ii}) \|u_i\|^2 \leq 0. \tag{20}$$

For small $\varepsilon \Delta t > 0$ and $u_0 \neq 0$, $u_i \neq 0$ and the second term is negligible. Hence, $b_i \geq 0$.

Considering $L(t) = \lambda(t)\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ with $\lambda(t) \in \mathbb{R}$, (7) yields

$$\|u_+\|^2 - \|u_0\|^2 = \Delta t^2 \sum_{i,j=1}^{s} (b_i b_j - b_i a_{ij} - b_j a_{ji}) \lambda(c_i \Delta t) \lambda(c_j \Delta t) \left\langle u_i,\, u_j \right\rangle. \tag{21}$$

For $\lambda(c_i \Delta t) = \varepsilon \xi_i$ with arbitrary $\xi \in \mathbb{R}^s$ and sufficiently small $\varepsilon > 0$, L can be extended to a smooth mapping with arbitrarily small $\left\|L(t)\right\|$ and Lipschitz constant of $t \mapsto L(t)$. Since $u_i = u_0 + O(\varepsilon \Delta t)$, energy stability and (21) imply

$$\varepsilon^2 \Delta t^2 \sum_{i,j=1}^{s} (b_i b_j - b_i a_{ij} - b_j a_{ji}) \xi_i \xi_j \|u_0\|^2 + O(\varepsilon^3 \Delta t^3) \leq 0. \tag{22}$$

Hence, the matrix with entries $(b_i b_j - b_i a_{ij} - b_j a_{ji})$ must be negative semidefinite. □

Theorem 3.15 shows that conditional stability for non-autonomous linear problems is equivalent to unconditional stability for general nonlinear problems in the context of non-confluent Runge–Kutta methods and ODEs with semibounded operators.

Remark 3.16. If a Runge–Kutta method satisfying the assumptions of Theorem 3.15 is irreducible in the sense of Dahlquist and Jeltsch, the weights b_i must be strictly positive [8]. ◁

3.4 Numerical Results

The problems constructed in the proofs above are rather special and perhaps not typical of applications. The following example shows that the (poor) behaviors suggested in the above theorems also occur for a more natural problem. Consider the linear advection equation

$$\begin{aligned} \partial_t u(t,x) + \sin(t^2) \partial_x u(t,x) &= 0, && t \in (0,100), x \in [-1,1], \\ u(0,x) &= \sin(\pi x), && x \in [-1,1], \end{aligned} \tag{23}$$

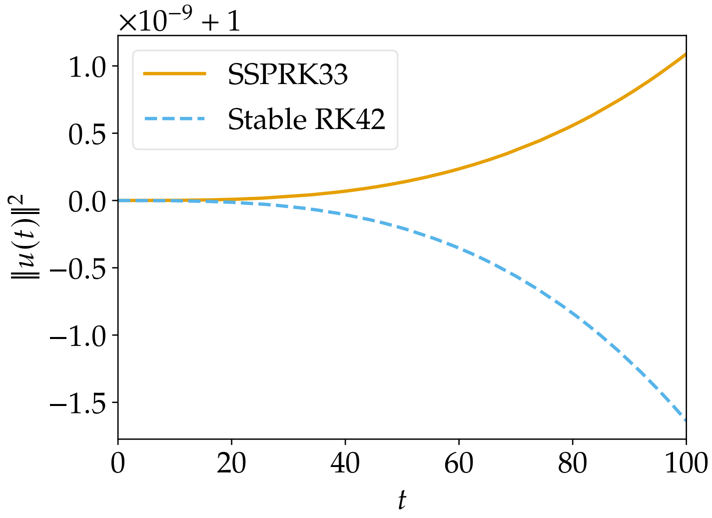


Figure 1: Discrete energies of the numerical solution of (23) using a skew-symmetric finite difference operato with $\Delta t = 10^{-5}$.

with periodic boundary conditions. A finite difference semidiscretization using the classical second order accurate central stencil and 50 grid points results in a skew-symmetric ODE (1). Applying SSPRK(3,3) [31], the time step $\Delta t = 10^{-5}$ is approximately three orders of magnitude smaller than required for energy stability of a corresponding constant coefficient problem [27, 35]. However, the energy $\left\|u(t)\right\|^2 = \Delta x \sum_i u_i^2$ increases exponentially, as can be seen in Figure 1. In contrast, the energy of the numerical approximation using the scheme (16) is decreasing, in accordance with Theorem 3.12.

The methods are implemented using double precision numbers `Float64` in the package `DifferentialEquations.jl` [23] in Julia [1]. The source code for these numerical experiments is available at [26].

4 Time-Independent Nonlinear Operators

In this section, the special case of time-independent but possibly nonlinear operators is studied. Hence,

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t} u(t) &= f(u(t)), \quad t \in (0,T), \\ u(0) &= u_0, \end{aligned} \tag{24}$$

is considered as special case of (1).

The technique used in the stability proof of [24, Theorem 7.1] can be described as follows.

1. Expand $\|u_+\|^2 - \|u_0\|^2$ as in (7).
2. Use the Lipschitz continuity of the right-hand side of (24) in order to expand the Δt^2 term in a power series in Δt.
3. Use the coefficients of the scheme and signs of the dominant terms in the power series to determine conditions for energy stability.

The second step can be generalised to higher order schemes by considering analytic right-hand sides f and the expansion [4, equation (313b)]

$$f(u_i) = \sum_{|t| \le n} \Delta t^{|t|-1} \frac{1}{\sigma(t)} (\Phi_i D)(t)\, F(t)(u_0) + O(\Delta t^{n+1}). \tag{25}$$

This is a sum over all trees t of order $|t| \le n$, $\sigma(t)$ is the symmetry of t, $(\Phi_i D)(t)$ is the i-th derivative weight of t, and $F(t)(u_0)$ the elementary differential associated with the tree t and the right hand side f, evaluated at u_0.

Inserting (25) into (7) yields

$$\|u_+\|^2 - \|u_0\|^2 = 2\Delta t \sum_{i=1}^{s} b_i \left\langle u_i,\, f(u_i)\right\rangle$$
$$+ \sum_{|t_1|,|t_2|\le n} \Delta t^{|t_1|+|t_2|} \sum_{i,j=1}^{s} (b_i b_j - b_i\, a_{ij} - b_j a_{ji}) \frac{(\Phi_i D)(t_1)\,(\Phi_j D)(t_2)}{\sigma(t_1)\sigma(t_2)} \left\langle F(t_1)(u_0),\, F(t_2)(u_0)\right\rangle$$
$$+ \mathcal{O}(\Delta t^{n+3}). \quad (26)$$

Example 4.1. For the three stage, third order method SSPRK(3,3) of [31] given by the Butcher coefficients

$$\begin{array}{c|ccc} 0 & & & \\ 1 & 1 & & \\ 1/2 & 1/4 & 1/4 & \\ \hline & 1/6 & 1/6 & 2/3 \end{array} \quad (27)$$

the first terms of the sum over trees in (26) are

$$\Delta t^4 \left[\frac{1}{6} \left\langle f,\, f'f'f\right\rangle - \frac{1}{12} \left\langle f,\, f''(f,f)\right\rangle + \frac{1}{12} \left\|f'f\right\|^2 \right] + \mathcal{O}(\Delta t^5). \quad (28)$$

By a suitable choice of a non-dissipative right hand side f such that $\forall u\colon \left\langle u,\, f(u)\right\rangle = 0$, the leading order term can be made positive. Hence, the energy increases for every sufficiently small time step Δt.

Of course, there are also problems for which SSPRK(3,3) is conditionally energy stable. For example, the right hand side $f\colon \mathbb{R}^2 \to \mathbb{R}^2$, $f(u) = \|u\|^2 (-u_2, u_1)^T$, yields $-\frac{7}{12}\|u\|^{10} + \mathcal{O}(\Delta t^5)$ as first terms of the expansion (26). ◁

Example 4.2. It is appealing to choose a Runge–Kutta method such that the coefficients in front of scalar products of elementary differentials whose sign cannot be controlled vanish and such that the coefficients multiplying non-negative terms such as $\left\|f'f\right\|^2$ are negative. However, it seems difficult to do this in a way that guarantees energy stability for all semibounded f. For example,

$$\begin{array}{c|ccc} 0 & & & \\ 3/8 & 3/8 & & \\ 1 & -1 & 2 & \\ \hline & 1/22 & 8/11 & 5/22 \end{array} \quad (29)$$

represents a three stage, second order scheme. The corresponding first terms of the sum over trees in (26) are

$$\begin{aligned}
&\Delta t^4 \left[-\frac{1}{11} \left\|f'f\right\|^2 \right] + \Delta t^5 \left[-\frac{15}{176} \left\langle f'f,\, f'f'f\right\rangle - \frac{49}{704} \left\langle f'f,\, f''(f,f)\right\rangle \right] \qquad (30)\\
&+ \Delta t^6 \Bigg[-\frac{45}{2816} \left\langle f'f,\, f'f''(f,f)\right\rangle - \frac{15}{176} \left\langle f'f,\, f''(f,f'f)\right\rangle - \frac{347}{16896} \left\langle f'f,\, f'''(f,f,f)\right\rangle \\
&\qquad + \frac{225}{7744} \left\|f'f'f\right\|^2 + \frac{255}{30976} \left\langle f'f'f,\, f''(f,f)\right\rangle - \frac{149}{30976} \left\|f''(f,f)\right\|^2 \Bigg] \\
&+ \Delta t^7 \Bigg[-\frac{45}{22528} \left\langle f'f,\, f''(f'f'f,f)\right\rangle - \frac{45}{2816} \left\langle f'f,\, f''(f''(f,f),f)\right\rangle \\
&\qquad - \frac{45}{1408} \left\langle f'f,\, f''(f'f,f'f)\right\rangle - \frac{15}{352} \left\langle f'f,\, f'''(f'f,f,f)\right\rangle \\
&\qquad - \frac{2641}{540672} \left\langle f'f,\, f''''(f,f,f,f)\right\rangle + \frac{675}{61952} \left\langle f'f'f,\, f'f''(f,f)\right\rangle \\
&\qquad + \frac{225}{3872} \left\langle f'f'f,\, f''(f'f,f)\right\rangle + \frac{205}{22528} \left\langle f'f'f,\, f'''(f,f,f)\right\rangle \\
&\qquad + \frac{765}{495616} \left\langle f''(f,f),\, f'f''(f,f)\right\rangle + \frac{255}{30976} \left\langle f''(f,f),\, f''(f'f,f)\right\rangle
\end{aligned}$$

$$
\begin{aligned}
&\quad - \frac{1}{16896}\left\langle f''(f,f),\, f'''(f,f,f)\right\rangle \Bigg] \\
&+ \Delta t^8 \Bigg[-\frac{135}{720896}\left\langle f'f,\, f'f''''(f,f,f,f)\right\rangle - \frac{45}{22528}\left\langle f'f,\, f''(f'''(f,f,f),f)\right\rangle \\
&\quad - \frac{135}{11264}\left\langle f'f,\, f''(f''(f,f),f'f)\right\rangle - \frac{45}{5632}\left\langle f'f,\, f'''(f''(f,f),f,f)\right\rangle \\
&\quad - \frac{45}{1408}\left\langle f'f,\, f'''(f'f,f'f,f)\right\rangle - \frac{5}{352}\left\langle f'f,\, f''''(f'f,f,f,f)\right\rangle \\
&\quad - \frac{-20723}{21626880}\left\langle f'f,\, f'''''(f,f,f,f,f)\right\rangle + \frac{675}{495616}\left\langle f'f'f,\, f'f'''(f,f,f)\right\rangle \\
&\quad + \frac{675}{61952}\left\langle f'f'f,\, f''(f''(f,f),f)\right\rangle + \frac{675}{30976}\left\langle f'f'f,\, f''(f'f,f'f)\right\rangle \\
&\quad + \frac{255}{7744}\left\langle f'f'f,\, f'''(f'f,f,f)\right\rangle + \frac{22765}{7929856}\left\langle f'f'f,\, f''''(f,f,f,f)\right\rangle \\
&\quad + \frac{765}{3964928}\left\langle f''(f,f),\, f'f'''(f,f,f)\right\rangle + \frac{765}{495616}\left\langle f''(f,f),\, f''(f''(f,f),f)\right\rangle \\
&\quad + \frac{765}{247808}\left\langle f''(f,f),\, f''(f'f,f'f)\right\rangle + \frac{265}{61952}\left\langle f''(f,f),\, f'''(f'f,f,f)\right\rangle \\
&\quad + \frac{1667}{5947392}\left\langle f''(f,f),\, f''''(f,f,f,f)\right\rangle + \frac{2025}{1982464}\left\| f'f''(f,f)\right\|^2 \\
&\quad + \frac{675}{61952}\left\langle f'f''(f,f),\, f''(f'f,f)\right\rangle + \frac{615}{360448}\left\langle f'f''(f,f),\, f'''(f,f,f)\right\rangle \\
&\quad + \frac{225}{7744}\left\| f''(f'f,f)\right\|^2 + \frac{205}{22528}\left\langle f''(f'f,f),\, f'''(f,f,f)\right\rangle \\
&\quad + \frac{1019}{1622016}\left\| f'''(f,f,f)\right\|^2 \Bigg] + O(\Delta t^9).
\end{aligned}
$$

If $\left\|f'f\right\| \neq 0$, the scheme is dissipative for sufficiently small time step $\Delta t > 0$. Similarly, if $\left\|f'f\right\| = 0$ and $\left\|f''(f,f)\right\| \neq 0$, the norm of the numerical solutions cannot increase for sufficiently small $\Delta t > 0$. If $f'f = 0$ and $f''(f,f) = 0$, the Δt^7 terms vanish additionally. However, the energy of the numerical solutions increases for arbitrarily small $\Delta t > 0$ if $f'f = 0$, $f''(f,f) = 0$ and $\left\|f'''(f,f,f)\right\| \neq 0$, since there is a positive term $\Delta t^8 \frac{1019}{1622016}\left\|f'''(f,f,f)\right\|^2 + O(\Delta t^9)$ and all other terms vanish. ◁

4.1 Necessary Conditions for Energy Stability

The Examples 4.1 and 4.2 demonstrate the importance of the bushy trees $\ldots$ with corresponding derivative weights c_i, c_i^2, c_i^3, $\ldots$ and elementary differentials $f'f$, $f''(f,f)$, $f'''(f,f,f)$ etc. While it is known that the elementary differentials are linearly independent for general right hand sides f [4, Section 314], it is of interest to study the independence of these terms for semibounded f and in particular for non-dissipative f.

In order to do that, the setting will be changed. Instead of considering a Hilbert space, $\mathcal{H}$ is a real semi inner product space. The semi inner product is still written as $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ denotes the induced seminorm (i.e. nearly a norm but not necessarily definite). The same definitions of energy stability etc. as for inner product spaces are used. Semi inner products will be considered only in this subsection.

Theorem 4.3. *For each $k \in \mathbb{N}$, there is an autonomous ODE with semibounded right hand side f in a semi inner product space such that the elementary differentials evaluated at u_0 satisfy $f^{(k)}(f,\dots,f) \neq 0$ and $f^{(l)}(f,\dots,f) = 0$ for all $l \neq k$. Additionally, $\left\langle f,\, f^{(k)}\right\rangle = 0$ and $\forall u\colon\ \left\langle u,\, f(u)\right\rangle = 0$.*

Proof. Consider the space $\mathbb{R}^3$ equipped with the semi inner product induced by the matrix

$P = \operatorname{diag}(0, 1, 1)$, i.e. $\langle u, v \rangle_P = u^T P v$. Consider the ODE

$$u'(t) = f(u(t)), \quad u(0) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad f(u) := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + u_1^k \begin{pmatrix} 0 \\ -u_3 \\ u_2 \end{pmatrix}. \tag{31}$$

The right hand side f satisfies $\forall u \in \mathbb{R}^3\colon \big\langle u, f(u) \big\rangle_P = 0$ and $f(u_0) = (1, 0, 0)^T \neq 0$. Hence, for $l \in \mathbb{N}$, the i-th component of the elementary differential $f^{(l)}(f, \dots, f)$ evaluated at u_0 is [4, Definition 310A]

$$\Big[f^{(l)}(f, \dots, f) \Big]^i = f^i_{j_1, \dots, j_l} f^{j_1} \dots f^{j_l} = f^i_{1, \dots, 1} \implies f^{(l)}(f, \dots, f) = \delta_{k,l} k! \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{32}$$

Here, summation over repeated indices is implied, upper indices denote components, lower indices denote derivatives, and $\delta_{k,l}$ is the Kronecker delta. □

Remark 4.4. If semi inner products are allowed, problems depending explicitly on time have to be considered again. Indeed, the test problem constructed in the proof of Theorem 4.3 is of this form since u_1 can be interpreted as time t and the numerical approximation of u_1 is equal to t if the classical condition $c_i = \sum_{j=1}^{s} a_{ij}$ is satisfied.

Hence, the results of section 3 can be applied, showing that there is nothing to gain if there is a c_i distinct from the others with $b_i \neq 0$. Thus, it is interesting whether the independence of the elementary differentials for energy conservative right hand sides holds also in inner product spaces. Since this problem seems to be intractable with the current methods, it is left for future investigations. ◁

Theorem 4.3 shows that the choice of elementary differentials made in Example 4.2 is possible. Hence, the method mentioned there is not energy stable for general autonomous and semibounded problems. The basic argument used there can be formulated as follows.

Theorem 4.5. *Consider a Runge–Kutta method with order of accuracy at least two. If there is a $k \in \mathbb{N}$ such that the Butcher coefficients satisfy*

$$\sum_{i,j=1}^{s} (b_i b_j - b_i a_{ij} - b_j a_{ji}) c_i^k c_j^k > 0, \tag{33}$$

then the method is not energy stable for general autonomous and semibounded ODEs in semi inner product spaces.

Proof. Since the method is at least second order accurate, the lowest order term in the sum involving trees in (26) vanishes since

$$\begin{aligned} \sum_{i,j=1}^{s} (b_i b_j - b_i a_{ij} - b_j a_{ji}) (\Phi_i D)(\bullet)\, (\Phi_j D)(\bullet) &= \sum_{i,j=1}^{s} (b_i b_j - b_i a_{ij} - b_j a_{ji}) \\ &= \left(\sum_{i=1}^{s} b_i \right)^2 - 2 \sum_{i=1}^{s} b_i c_i = 1 - 2 \cdot \frac{1}{2} = 0. \end{aligned} \tag{34}$$

Because of Theorem 4.3, it is possible to choose f such that the remaining terms all vanish except the one corresponding to the bushy tree with k leaves. While this is not formulated directly there, a close inspection of the proof reveals that this is indeed true. For example, one can choose f so that $f' = 0$ and $\big\langle f, f' f''(f, f) \big\rangle = 0$ while $f \neq 0$ and $f''(f, f) \neq 0$.

By choosing such an f all terms of order up to Δt^{2k+1} vanish and (26) takes the form

$$\|u_+\|^2 - \|u_0\|^2 = \Delta t^{2k+2} \underbrace{\sum_{i,j=1}^{s} (b_i b_j - b_i a_{ij} - b_j a_{ji}) c_i^k c_j^k \Big\| f^{(k)}(f, \dots, f) \Big\|^2}_{>0} + O(\Delta t^{2k+3}). \tag{35}$$

Hence, $\|u_+\|^2 > \|u_0\|^2$ for arbitrarily small $\Delta t > 0$. □

Theorem 4.5 implies in particular that $\sum_{i,j=1}^{s}(b_i b_j - b_i a_{ij} - b_j a_{ji}) c_i^k c_j^k$ must be non-positive for B-stable methods. This is related to algebraically stable schemes, where the matrix with entries $(b_i b_j - b_i a_{ij} - b_j a_{ji})$ is negative semidefinite. It can be proved that B-stable schemes are indeed algebraically stable if certain additional (technical) assumptions are satisfied, e.g. if the method is non-confluent [15, Corollary IV.12.14] or irreducible [15, Theorem IV.12.18].

Example 4.6. For the B-stable but not algebraically stable reducible schemes with $A = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$, $b = (2, -1)^T$ [7, page 80] and $A = \begin{pmatrix} 1/2 & 0 \\ 1/4 & 1/4 \end{pmatrix}$, $b = (1/2, 1/2)^T$ [15, Table IV.12.2], the terms $\sum_{i,j=1}^{s}(b_i b_j - b_i a_{ij} - b_j a_{ji})(\Phi_i D)(t_1)\,(\Phi_j D)(t_2)$ vanish, exactly as for the implicit midpoint method to which these schemes can be reduced. ◁

While explicit Runge–Kutta methods cannot be algebraically stable, i.e. unconditionally stable for all semibounded problems, it is interesting to study whether they can be stable for these problems under a suitable time step restriction. Theorem 7.1 of [24] proves that there are indeed conditionally energy stable schemes of first order. For higher order schemes, it remains to check the condition given in Theorem 4.5. While this can be done for every scheme given explicitly, there are some results for general classes of schemes.

Theorem 4.7. *Consider an explicit Runge–Kutta method. Assume that there is a unique $i_{\max}$ such that $|c_{i_{\max}}| = \max_i |c_i|$ and that $b_{i_{\max}} \neq 0$. Then there is a $k \in \mathbb{N}$ such that (33) is satisfied. Hence, if the method is at least second order accurate, it is not energy stable for general autonomous and semibounded ODEs in semi inner product spaces.*

Proof. The expression on the left hand side of (33) can be written as

$$\left(b^T c^k\right)^2 - 2\left(c^k\right)^T \operatorname{diag}(b) A c^k, \tag{36}$$

where the exponentiation is performed componentwise. Using the given assumptions,

$$\left(\frac{c}{c_{i_{\max}}}\right)^k \to e_{i_{\max}}, \quad k \to \infty, \tag{37}$$

where $e_{i_{\max}}$ is the standard unit vector with components $(e_{i_{\max}})_j = \delta_{j,i_{\max}}$. Since the Runge–Kutta scheme is explicit, A is a strictly lower triangular matrix and $e_{i_{\max}}^T \operatorname{diag}(b) A e_{i_{\max}} = 0$. Because of $b^T e_{i_{\max}} = b_{i_{\max}} \neq 0$,

$$\left(\frac{b^T c^k}{c_{i_{\max}}}\right)^2 - 2\frac{(c^k)^T}{c_{i_{\max}}^k} \operatorname{diag}(b) A \frac{c^k}{c_{i_{\max}}^k} \to \left(b^T e_{i_{\max}}\right)^2 - 2 e_{i_{\max}}^T \operatorname{diag}(b) A e_{i_{\max}}^k = b_{i_{\max}}^2 \neq 0, \tag{38}$$

for $k \to \infty$. Hence, there is a $k \in \mathbb{N}$ such that (33) is satisfied. □

Remark 4.8. Theorem 4.7 can also be applied to many confluent methods such as the ten-stage, fourth order, explicit strong stability preserving method SSPRK(10,4) of [17]. Indeed, $i_{\max} = 10$, $c_{i_{\max}} = 1$, and $b_{i_{\max}} = \frac{1}{10}$ in that case. That this scheme is not energy stable for autonomous and semibounded problems has also been proved using some specific counterexamples in [24, Sections 4.3 and 6]. ◁

Remark 4.9. The argument used to prove Theorem 4.7 can also be applied to implicit Runge–Kutta methods with $a_{i_{\max}, i_{\max}} = 0$. ◁

4.2 Sufficient Conditions for Energy Stability

Theorem 4.7 does not imply that all explicit Runge–Kutta methods of order two or greater cannot be energy stable.

Example 4.10. The Runge–Kutta method with Butcher coefficients

$$\begin{array}{c|cccc} 0 & & & & \\ 1 & 1 & & & \\ 0 & 1 & -1 & & \\ 1 & -1 & 1 & 1 & \\ \hline & 1/4 & 1/4 & 1/4 & 1/4 \end{array} \tag{39}$$

has four stages and is second order accurate. Since $c_2 = 1 = c_4$, it does not satisfy the assumptions of Theorem 4.7. Because $c_i \in \{0,1\}$, we have $c^k = c$ for $k \in \mathbb{N}$. Hence, it suffices to consider $k = 1$ in (33), i.e.

$$\sum_{i,j=1}^{s} (b_i b_j - b_i\, a_{ij} - b_j\, a_{ji}) c_i c_j = -\frac{1}{4} < 0. \tag{40}$$

Thus, the sum in (33) is negative for all $k \in \mathbb{N}$. ◁

Theorem 4.11. *The Runge–Kutta method given in Example 4.10 is conditionally energy stable with respect to autonomous ODEs* (24) *with analytical and semibounded right hand side.*

Proof. Since the right hand side f is analytical, the energy difference after one time step can be expanded as in (26). Because of the semiboundedness of f, the term proportional to Δt is non-positive.

There are no inner products between $f = F(\bullet)$ and higher order elementary differentials in the remaining terms because

$$\sum_{i,j=1}^{s} (b_i b_j - b_i a_{ij} - b_j a_{ji}) \frac{(\Phi_i D)(\bullet)\,(\Phi_j D)(t_2)}{\sigma(\bullet)\,\sigma(t_2)} = \sum_{j=1}^{s} \underbrace{\left(b_j - \sum_{i=1}^{s} b_i a_{ij} - b_j c_j \right)}_{=0} \frac{(\Phi_j D)(t_2)}{\sigma(t_2)}, \tag{41}$$

which can be computed explicitly. The first remaining terms are of the form

$$\begin{aligned} &-\frac{1}{4}\Delta t^4 \|f'f\|^2 + \Delta t^5 \left[-\frac{1}{2}\langle f'f,\, f'f'f\rangle - \frac{1}{4}\langle f'f,\, f''(f,f)\rangle \right] \\ &+ \Delta t^6 \Bigg[\frac{1}{4}\langle f'f,\, f'f'f'f\rangle - \frac{1}{4}\langle f'f,\, f'f''(f,f)\rangle - \frac{1}{4}\langle f'f,\, f''(f'f,f)\rangle \\ &\qquad - \frac{1}{12}\langle f'f,\, f'''(f,f,f)\rangle + \frac{1}{2}\|f'f'f\|^2 - \frac{1}{4}\langle f'f'f,\, f''(f,f)\rangle - \frac{1}{16}\|f''(f,f)\|^2 \Bigg]. \end{aligned} \tag{42}$$

If $f'f \neq 0$, the Δt^4 term dominates the other ones for sufficiently small $\Delta t > 0$ and the scheme is energy stable. If $f'f = 0$ and $f''(f,f) \neq 0$, the Δt^4, Δt^5, and most of the Δt^6 terms vanish. Only $-\frac{1}{16}\Delta t^6 \|f''(f,f)\| < 0$ remains and dominates higher order terms, resulting in a stable scheme for small $\Delta t > 0$.

This argument can be applied similarly to all other terms. Suppose that $f'f = \cdots = f^{(k-1)}(f,\dots,f) = 0$ and $f^{(k)}(f,\dots,f) \neq 0$. The terms up to (and including) $O(\Delta t^{2k+1})$ vanish, since (as described at the beginning of the proof) for this method, the series (26) does not include any terms involving an inner product of f and other elementary differentials. Most of the Δt^{2k+2} terms vanish too, except the one proportional to $\left\| f^{(k)}(f,\dots,f) \right\|$. Because $\sum_{i,j=1}^{s} (b_i b_j - b_i a_{ij} - b_j a_{ji}) c_i^k c_j^k < 0$, this term is negative and dominates higher order terms. Hence, the scheme is energy stable for sufficiently small $\Delta t > 0$. □

The key ingredients of the proof of Theorem 4.11 are distilled in

Proposition 4.12. *Consider a second or third order accurate explicit Runge–Kutta method satisfying*

- $\forall i \in \{1, \dots, s\}\colon b_i \geq 0,$
- $\forall j \in \{1, \dots, s\}\colon b_j - \sum_{i=1}^{s} b_i a_{ij} - b_j c_j = 0,$
- $\forall i \in \{1, \dots, s\}\colon c_i = \sum_{j=1}^{s} a_{ij},$
- $\forall k \in \mathbb{N}\colon$ *The sum in* (33) *is negative.*

Such a scheme is conditionally energy stable for any autonomous ODE (24) *with right hand side that is analytical and semibounded with respect to an inner product.*

The proof is basically the same as that of Theorem 4.11. The restriction to second and third order accurate schemes is explained in Section 4.3 below.

Remark 4.13. Since infinitely many constraints have to be satisfied to apply Proposition 4.12, it is useful to consider additional simplifying assumptions/constraints in order to find feasible solutions. The Runge–Kutta method given in Example 4.10 and other schemes have been constructed using the following additional steps.

- Because of Theorem 4.7, the node with biggest absolute value should appear at least twice. In order to facilitate the search for a solution, it has been useful to choose the nodes c_i manually. Here, the biggest node is chosen as 1 (twice). In numerical experiments, it seemed to be useful/necessary to specify also the node 0 twice.
- The sum in (33) with $k = 1$ has to be negative. Additionally, the same sum involving only the nodes 0 and 1 should be negative. In numerical experiments, this additional condition has often be sufficient to guarantee that the sum in (33) is negative for general $k \in \mathbb{N}$.

◁

Example 4.14. The third order accurate Runge–Kutta method with Butcher coefficients

$$
\begin{array}{c|ccccc}
0 & & & & & \\
1/2 & 1/2 & & & & \\
1 & 1 & & & & \\
0 & 1 & 0 & -1 & & \\
1 & -3 & 2 & 1 & 1 & \\
\hline
 & 0 & 2/3 & 0 & 1/6 & 1/6
\end{array}
\tag{43}
$$

has been constructed using the approach just described. ◁

Corollary 4.15. *The Runge–Kutta method of Example 4.14 is energy stable for autonomous ODEs* (24) *with analytical and semibounded right hand side in inner product spaces if the time step* $\Delta t > 0$ *is sufficiently small.*

Proof. Since

$$
\sum_{i,j=1}^{s} (b_i b_j - b_i\, a_{ij} - b_j\, a_{ji}) c_i^k c_j^k = -\frac{11}{36} - \frac{4}{9} 2^{-k} \left(1 - 2^{-k}\right) < 0,
\tag{44}
$$

the sum in (33) is negative for general $k \in \mathbb{N}$ and Proposition 4.12 can be applied. □

4.3 Limitations for Higher Order Schemes

The conditions listed in Proposition 4.12 are not sufficient to create energy stable fourth order methods, since the coefficient of $\|f'f\|^2$ in the expansion (26) vanishes because of accuracy constraints. However, terms like $\langle f'f,\, f'f'f \rangle$ etc. appear later, which cannot be controlled in general. Hence, one has to impose additionally that all scalar products of $f'f$ with higher order differentials vanish. Because of

$$
\sum_{i,j=1}^{s} (b_i b_j - b_i a_{ij} - b_j a_{ji}) \frac{(\Phi_i D)(\begin{smallmatrix}\bullet\\\bullet\end{smallmatrix})\, (\Phi_j D)(t_2)}{\sigma(\begin{smallmatrix}\bullet\\\bullet\end{smallmatrix})\, \sigma(t_2)}
$$

$$= \sum_{j=1}^{s}\left(b_j\sum_{i=1}^{s}c_i - \sum_{i=1}^{s}b_ic_ia_{ij} - b_j\sum_{i=1}^{s}a_{ji}c_i\right)\frac{(\Phi_jD)(t_2)}{\sigma(t_2)}, \quad (45)$$

this additional constraint is

$$b_j\sum_{i=1}^{s}c_i - \sum_{i=1}^{s}b_ic_ia_{ij} - b_j\sum_{i=1}^{s}a_{ji}c_i = 0. \quad (46)$$

By summing over j and using the order conditions for order 3, it becomes clear that some nodes c_i must be negative if this constraint should be satisfied.

Nevertheless, even this additional constraint does not suffice to guarantee energy stability. Indeed, terms involving higher order differentials of the form $\left\|f'f'f\right\|^2$ appear and cannot be controlled by the previous terms with negative coefficients.

5 Unconditional Stability of Explicit Runge–Kutta Discretizations

In this section we investigate the possibility of obtaining unconditional stability with explicit Runge–Kutta methods. It is usually true in numerical analysis that explicit methods can be only conditionally stable. The following (unsurprising) theorem confirms this view, in the context of the linear autonomous initial-value problem:

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t}u(t) &= Lu(t), \quad t \in (0,T), \\ u(0) &= u_0, \end{aligned} \quad (47)$$

Theorem 5.1. *Let an at least first order accurate explicit Runge-Kutta method and constant (possibly complex) matrix $L \neq 0$ be given. Then, there exist initial values u_0 and a step size $\Delta t_*(A,b,L)$ such that the numerical solution u^n of* (47) *blows up as $n \to \infty$ for any $\Delta t > \Delta t_*$.*

Proof. An s-stage explicit Runge–Kutta method applied to (47) gives the solution $u^{n+1} = R(\Delta tL)u^n$, where $R(z) = \sum_{j=0}^{d}\alpha_jz^j$ is a polynomial with degree $d \leq s$ and $\alpha_0 = \alpha_1 = 1$. Two cases can occur.

1. L has an eigenvalue $\lambda \neq 0$ with eigenvector u_0: Then, $R(\Delta tL)u_0 = (I + \alpha_1\Delta tL + \cdots + \alpha_d\Delta t^dL^d)u_0 = (I + \alpha_1\Delta t\lambda + \ldots + \alpha_d\Delta t^d\lambda^d)u_0$. Thus, $R(\Delta tL)^nu_0 \to \infty$ for $n \to \infty$ if Δt is big enough.
2. Zero is the only eigenvalue of L: Then, there is a vector u_0 such that $Lv \neq 0$ but $L^2v = 0$ (consider the Jordan canonical form). Thus, $R(\Delta tL)^nu_0 = (1 + n\alpha_1\Delta tL)u_0 \to \infty$ for $n \to \infty$ if Δt is big enough. □

In practice, due to rounding errors, it is reasonable to expect a blowup for almost all initial data.

5.1 Nonlinear Problems

It is natural to ask if a result like Theorem 5.1 holds when L is allowed to be nonlinear. It seems quite natural to expect that the answer is yes. However, we have the following result:

Theorem 5.2. *There exists an explicit Runge–Kutta method and non-trivial function $f(u)$ such that the numerical solution of $u'(t) = f(u)$ remains bounded as $n \to \infty$ for every step size Δt and for every initial value.*

Proof. We prove this result by constructing an example – the only example of which we are presently aware. We take the explicit midpoint Runge–Kutta method

$$y_2 = u^n + \frac{\Delta t}{2}f(u^n), \qquad u^{n+1} = u^n + \Delta tf(y_2), \quad (48)$$

and the ODEs

$$\begin{pmatrix} u_1'(t) \\ u_2'(t) \end{pmatrix} = \frac{1}{u_1^2 + u_2^2} \begin{pmatrix} -u_2 \\ u_1 \end{pmatrix}. \tag{49}$$

Direct calculation of the change in energy over a step (using (7)) reveals that it is constant:

$$\|u_+\|^2 - \|u_0\|^2 = \langle f(u_0),\, f(y_2) \rangle - \langle f(y_2),\, f(y_2) \rangle = 0. \tag{50}$$

□

In fact, we can write explicitly the solution obtained for the example in the proof. For the general initial value $u_0 = r_0(\cos(\theta_0), \sin(\theta_0))^T$, the numerical solution of (49) obtained with the explicit midpoint RK method is

$$u^n = r_0 \begin{pmatrix} \cos(\theta_0 + n\theta_h) \\ \sin(\theta_0 + n\theta_h) \end{pmatrix}, \quad \text{where} \quad \theta_h = \arccos \frac{r^2 - \frac{\Delta t^2}{4r^2}}{r^2 + \frac{\Delta t^2}{4r^2}}. \tag{51}$$

This example is quite remarkable, and naturally leads one to wonder if others like it exist. If we assume the numerical energy is constant, then the problem must also be energy conservative since the Runge–Kutta scheme converges to the analytical solution (if the right hand side is locally Lipschitz continuous). In $\mathbb{R}^2$, every energy conservative problem is of the form

$$\begin{pmatrix} u_1'(t) \\ u_2'(t) \end{pmatrix} = g(u) \begin{pmatrix} -u_2 \\ u_1 \end{pmatrix}, \tag{52}$$

where g is a scalar valued function. We have the following uniqueness result.

Theorem 5.3. *Let a consistent two-stage explicit Runge–Kutta method and a function $g\colon \mathbb{R}^2 \to \mathbb{R}$ be given. Consider the ODE* (52) *and suppose that the numerical solution satisfies $\|u^n\| = \|u_0\|$ for all step sizes Δt and all n.*

a) If g is not identically zero, then the Runge–Kutta method must be the explicit midpoint method (48).

b) If $g\colon \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}$ is analytic, then g must be a scalar multiple of $u \mapsto \|u\|^{-2}$.

Proof. a) For brevity, let $r = g(u_0)$ and $s = g(y_2) = g(u_0 + a_{21}\Delta t f(u_0))$.

$$\begin{aligned} \langle f_1,\, f_1 \rangle &= \|u_0\|^2\, r^2, \\ \langle f_2,\, f_2 \rangle &= \|u_0\|^2\, s^2(1 + a_{21}^2 \Delta t^2 r^2), \\ \langle f_1,\, f_2 \rangle &= \|u_0\|^2\, rs. \end{aligned} \tag{53}$$

Thus energy is conserved if and only if (after dividing through by $\|u_0\|^2$)

$$2b_2 a_{21} rs = b_1^2 r^2 + b_2^2 s^2 (1 + a_{21}^2 \Delta t^2 r^2) + 2b_1 b_2 rs = 0. \tag{54}$$

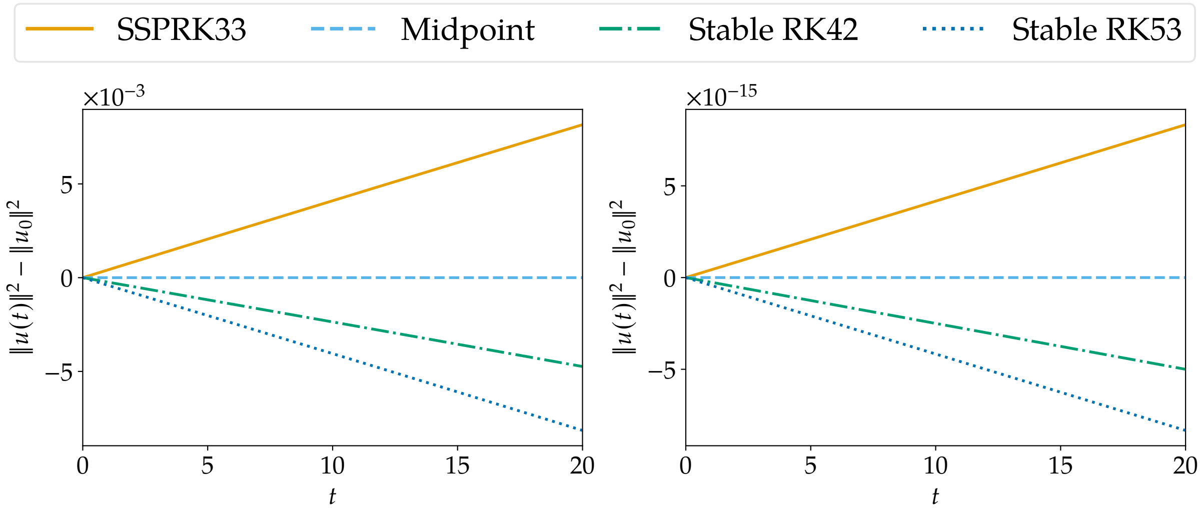
This is a quadratic equation in s, which has real roots only if

$$b_2^2 r^2 (b_1 - a_{21})^2 \geq b_1^2 b_2^2 r^2 (1 + a_{21}^2 \Delta t^2 r^2). \tag{55}$$

This cannot hold for all Δt unless the term involving Δt vanishes. So we must have $b_1 b_2 a_{21} r = 0$. The case $r = 0$ is ruled out by assumption, while $b_2 = 0$ implies $r = 0$ by (54). Taking $a_{21} = 0$ implies (by (54)) that the ratio r/s is equal to a constant independent of Δt or u, which is not possible. Thus we must have $b_1 = 0$; consistency then requires $b_2 = 1$. This implies that $s = 0$ (trivial) or that

$$g(u + a_{21}\Delta t f(u))(1 + a_{21}^2 \Delta t^2 (g(u))^2) = 2a_{21} g(u). \tag{56}$$

Considering $\Delta t \to 0$, we find that necessarily $a_{21} = 1/2$; the resulting method is the midpoint Runge–Kutta method (48).



(a) $\Delta t = 10^{-1}$. (b) $\Delta t = 10^{-5}$.

Figure 2: Evolution of the energy of numerical solutions for the nonlinear problem (49).

b) It suffices to consider $g \neq 0$. Expanding (56) with $a_{21} = 1/2$ as required by a) yields

$$\sum_{k\geq 0} \frac{1}{k!} 2^{-k} \Delta t^k g^{(k)}(f, \dots, f)\big(1 + 2^{-2}\Delta t^2 g^2\big) = g. \tag{57}$$

This is equivalent to

$$\frac{1}{2}\Delta t g' f + \sum_{k\geq 2} 2^{-k}\Delta t^k \left(\frac{1}{k!} g^{(k)}(f, \dots, f) + \frac{1}{(k-2)!} g^2 g^{(k-2)}(f, \dots, f)\right) = 0. \tag{58}$$

Since this has to hold for all Δt,

$$g' f = 0, \quad \forall k \geq 2\colon g^{(k)}(f, \dots, f) = -\frac{k!}{(k-2)!} g^2 g^{(k-2)}(f, \dots, f). \tag{59}$$

This infinite set of conditions determines g uniquely up to a scalar multiple. Indeed, using polar coordinates (r, θ) for u, the condition $g' f = 0$ implies that g does not depend on the angle θ, i.e. that g is radially symmetric. Hence, g can be considered as a function depending only on the radius r in (56). Expanding this analytic function $(0, \infty) \to \mathbb{R}$ in (56), all derivatives at an arbitrary point are fixed. Hence, g is determined up to a multiplicative factor. □

Remark 5.4. Theorem 5.3 holds also for $\mathbb{R}^3$ instead of $\mathbb{R}^2$. Indeed, the action of an arbitrary skew-symmetric matrix is equivalent to the cross product with an associated vector in $\mathbb{R}^3$. The span of this vector is irrelevant for the considered problem. ◁

5.2 Numerical Experiments

Here, some numerical experiments using the ODE (49) are performed. The explicit midpoint method (48), the third order strong stability preserving method SSPRK33 of [31], and the energy stable second and third order methods given in Examples 4.10 and 4.14 are applied with constant time steps $\Delta t \in \{10^{-1}, 10^{-5}\}$. The methods are implemented using quadruple precision numbers `Float128` in the package `DifferentialEquations.jl` [23] in Julia [1]. The source code for these numerical experiments is available at [26].

The results displayed in Figure 2 confirm the analytical results: The energy grows monotonically for SSPRK33, stays constant for the midpoint rule and decays for the energy stable methods.

Remark 5.5. The first terms of the expansion (26) for the test problem (49) and the second order method of Example 4.10 are

$$-\frac{1}{4}\Delta t^4 \|f' f\|^2 + O(\Delta t^5) = -\frac{1}{4}\Delta t^4 \|u\|^{-6} + O(\Delta t^5). \tag{60}$$

Hence, it can be verified easily that the method is conditionally energy stable. Similarly, the first terms for the third order method of Example 4.14 are

$$-\frac{5}{12}\Delta t^4 \|f' f\|^2 + O(\Delta t^5) = -\frac{5}{12}\Delta t^4 \|u\|^{-6} + O(\Delta t^5). \tag{61}$$

Thus, this method is energy stable, too. Additionally, this explains why the third order method is more (nearly twice as) dissipative than the second order one in Figure 2. ◁

6 Summary and Conclusions

As we have seen, explicit Runge–Kutta methods that are conditionally energy stable for *all* nonlinear autonomous semibounded problems are rare, but they do exist. The existence of explicit energy stable methods for non-autonomous problems (even in the linear setting) is still an open question. Any explicit energy stable method for non-autonomous problems must be confluent, since otherwise it would need to be unconditionally stable (see Theorem 3.15). Nevertheless, there is at least a second order accurate scheme that is energy stable for a restricted class of relevant problems (see Theorem 3.12).

For nonlinear autonomous problems, our analysis is based on the series expansion (26) of the change in energy. Besides deriving necessary conditions, Proposition 4.12 and Remark 4.13 list sufficient conditions and approaches that can be used to create energy stable second and third order methods (see Theorem 4.11 and Corollary 4.15). This approach could also be used to construct methods that are energy stable for a particular problem, if one knows which elementary differentials of f vanish.

Some of our results seem at first glance surprising or counter-intuitive. A common intuition is that linear problems are *easier* than nonlinear problems. But if explicit time dependence is allowed, the linear setting becomes much more challenging: notice that the derived necessary conditions for energy stability in this case are more restrictive than for autonomous nonlinear problems. In a similar vein, a *particular* nonlinear ODE may be easier to deal with than even any linear autonomous ODE, as demonstrated by Theorem 5.2 (showing unconditional stability for a specific nonlinear ODE and explicit RK method) and Theorem 5.1 (showing that no explicit RK method is unconditionally stable for any non-trivial linear problem).

In the literature, the assumption of non-confluence is often merely technical and not necessary. In contrast, in our study of energy stability for non-autonomous problems we have found that confluence is an important property, and that certain confluent Runge–Kutta methods (more specifically, methods with no unique quadrature node) can have stability properties that are impossible for non-confluent methods.

While developing many answers, this article has also revealed several open questions and directions of further research. First of all, an obvious question concerns the possibility of fourth or higher order explicit Runge–Kutta that are conditionally energy stable (at least for autonomous problems). From a practical point of view, it would be interesting to perform a computational optimization of energy stable Runge–Kutta methods and compare them to state of the art schemes that are not energy stable.

More theoretically interesting questions concern the independence of the elementary differentials for conservative problems and the existence or uniqueness of unconditionally stable explicit Runge–Kutta methods and associated nonlinear right hand sides. For instance, if it were possible to choose the values of the elementary differentials independently while also choosing f to conserve energy, then it would be possible (in principle) to construct, for each Runge–Kutta method (including all explicit methods), a problem for which that method is unconditionally energy conservative.

Acknowledgements

Research reported in this publication was supported by the King Abdullah University of Science and Technology (KAUST). The first author was partially supported by the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft) under Grant SO 363/14-1.

References

[1] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah. "Julia: A Fresh Approach to Numerical Computing." In: *SIAM Review* 59.1 (2017), pp. 65–98. DOI: 10.1137/141000671. arXiv: 1411.1607 [cs.MS].

[2] P. D. Boom and D. W. Zingg. "High-order implicit time-marching methods based on generalized summation-by-parts operators." In: *SIAM Journal on Scientific Computing* 37.6 (2015), A2682–A2709. DOI: 10.1137/15M1014917.

[3] K. Burrage and J. C. Butcher. "Stability criteria for implicit Runge-Kutta methods." In: *SIAM Journal on Numerical Analysis* 16.1 (1979), pp. 46–57. DOI: 10.1137/0716004.

[4] J. C. Butcher. *Numerical Methods for Ordinary Differential Equations*. Chichester: John Wiley & Sons Ltd, 2016.

[5] M. Calvo, M. Laburta, J. Montijano, and L. Rández. "Projection methods preserving Lyapunov functions." In: *BIT Numerical Mathematics* 50.2 (2010), pp. 223–241. DOI: 10.1007/s10543-010-0259-3.

[6] M. Calvo, D. Hernández-Abreu, J. I. Montijano, and L. Rández. "On the Preservation of Invariants by Explicit Runge–Kutta Methods." In: *SIAM Journal on Scientific Computing* 28.3 (2006), pp. 868–885. DOI: 10.1137/04061979X.

[7] M. Crouzeix. "Sur la B-stabilité des méthodes de Runge-Kutta." In: *Numerische Mathematik* 32.1 (1979), pp. 75–82. DOI: 10.1007/BF01397651.

[8] G. Dahlquist and R. Jeltsch. *Generalized disks of contractivity for explicit and implicit Runge-Kutta methods*. Technical Report TRITA-NA-7906. Departement of Numerical Analysis and Computing Sciences, The Royal Institute of Technology, S-100 44 Stockholm 70, Sweden: Institut för Numerisk Analysis, KTH Stockholm, Jan. 1979.

[9] D. C. D. R. Fernández, J. E. Hicken, and D. W. Zingg. "Review of summation-by-parts operators with simultaneous approximation terms for the numerical solution of partial differential equations." In: *Computers & Fluids* 95 (2014), pp. 171–196. DOI: 10.1016/j.compfluid.2014.02.016.

[10] L. Friedrich, G. Schnücke, A. R. Winters, D. C. D. R. Fernández, G. J. Gassner, and M. H. Carpenter. "Entropy Stable Space–Time Discontinuous Galerkin Schemes with Summation-by-Parts Property for Hyperbolic Conservation Laws." In: *Journal of Scientific Computing* 80.1 (2019), pp. 175–222. DOI: 10.1007/s10915-019-00933-2. arXiv: 1808.08218 [math.NA].

[11] G. J. Gassner. "A Skew-Symmetric Discontinuous Galerkin Spectral Element Discretization and Its Relation to SBP-SAT Finite Difference Methods." In: *SIAM Journal on Scientific Computing* 35.3 (2013), A1233–A1253. DOI: 10.1137/120890144.

[12] B. Gustafsson. *High Order Difference Methods for Time Dependent PDE*. Vol. 38. Springer Series in Computational Mathematics. Berlin Heidelberg: Springer Science & Business Media, 2007. DOI: 10.1007/978-3-540-74993-6.

[13] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*. Vol. 31. Springer Series in Computational Mathematics. Berlin Heidelberg: Springer-Verlag, 2006. DOI: 10.1007/3-540-30666-8.

[14] E. Hairer, S. P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I: Nonstiff Problems*. Vol. 8. Springer Series in Computational Mathematics. Berlin Heidelberg: Springer-Verlag, 2008. DOI: 10.1007/978-3-540-78862-1.

[15] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems*. Vol. 14. Springer Series in Computational Mathematics. Berlin Heidelberg: Springer-Verlag, 2010. DOI: 10.1007/978-3-642-05221-7.

[16] I. Higueras. "Monotonicity for Runge-Kutta Methods: Inner Product Norms." In: *Journal of Scientific Computing* 24.1 (2005), pp. 97–117. DOI: 10.1007/s10915-004-4789-1.

[17] D. I. Ketcheson. "Highly Efficient Strong Stability-Preserving Runge-Kutta Methods with Low-Storage Implementations." In: *SIAM Journal on Scientific Computing* 30.4 (2008), pp. 2113–2136. DOI: 10.1137/07070485X.

[18] D. I. Ketcheson. *Relaxation Runge–Kutta Methods: Conservation and Stability for Inner-Product Norms*. May 2019. arXiv: 1905.09847 [math.NA].

[19] H.-O. Kreiss and J. Lorenz. *Initial-Boundary Value Problems and the Navier-Stokes Equations*. Vol. 47. Classics in Applied Mathematics. Philadelphia: SIAM, 2004.

[20] H.-O. Kreiss and G. Scherer. "Finite Element and Finite Difference Methods for Hyperbolic Partial Differential Equations." In: *Mathematical Aspects of Finite Elements in Partial Differential Equations*. Ed. by C. de Boor. New York: Academic Press, 1974, pp. 195–212.

[21] J. Nordström and M. Björck. "Finite volume approximations and strict stability for hyperbolic problems." In: *Applied Numerical Mathematics* 38.3 (2001), pp. 237–255. DOI: 10.1016/S0168-9274(01)00027-7.

[22] J. Nordström and T. Lundquist. "Summation-by-parts in time." In: *Journal of Computational Physics* 251 (2013), pp. 487–499. DOI: 10.1016/j.jcp.2013.05.042.

[23] C. Rackauckas and Q. Nie. "DifferentialEquations.jl – A Performant and Feature-Rich Ecosystem for Solving Differential Equations in Julia." In: *Journal of Open Research Software* 5.1 (2017), p. 15. DOI: 10.5334/jors.151.

[24] H. Ranocha. *On Strong Stability of Explicit Runge-Kutta Methods for Nonlinear Semibounded Operators*. Nov. 2018. arXiv: 1811.11601 [math.NA].

[25] H. Ranocha. "Some Notes on Summation by Parts Time Integration Methods." In: *Results in Applied Mathematics* 1 (June 2019), p. 100004. DOI: 10.1016/j.rinam.2019.100004. arXiv: 1901.08377 [math.NA].

[26] H. Ranocha and D. I. Ketcheson. *EnergyStabilityExplicitRungeKuttaNonlinearNonautonomous. Energy Stability of Explicit Runge–Kutta Methods for Non-autonomous or Nonlinear Problems*. https://github.com/ranocha/EnergyStabilityExplicitRungeKuttaNonlinearNonautonomous. Sept. 2019. DOI: 10.5281/zenodo.3464243.

[27] H. Ranocha and P. Öffner. "L_2 Stability of Explicit Runge-Kutta Schemes." In: *Journal of Scientific Computing* 75.2 (May 2018), pp. 1040–1056. DOI: 10.1007/s10915-017-0595-4.

[28] H. Ranocha, P. Öffner, and T. Sonar. "Summation-by-parts operators for correction procedure via reconstruction." In: *Journal of Computational Physics* 311 (Apr. 2016), pp. 299–328. DOI: 10.1016/j.jcp.2016.02.009. arXiv: 1511.02052 [math.NA].

[29] H. Ranocha, M. Sayyari, L. Dalcin, M. Parsani, and D. I. Ketcheson. *Relaxation Runge–Kutta Methods: Fully-Discrete Explicit Entropy-Stable Schemes for the Euler and Navier–Stokes Equations*. May 2019. arXiv: 1905.09129 [math.NA].

[30] J. T. Schwartz. *Nonlinear functional analysis*. Vol. 4. Notes on Mathematics and its Applications. New York: Gordon and Breach Science Publishers Inc., 1969.

[31] C.-W. Shu and S. Osher. "Efficient implementation of essentially non-oscillatory shock-capturing schemes." In: *Journal of Computational Physics* 77.2 (1988), pp. 439–471. DOI: 10.1016/0021-9991(88)90177-5.

[32] Z. Sun and C.-W. Shu. "Stability of the fourth order Runge-Kutta method for time-dependent partial differential equations." In: *Annals of Mathematical Sciences and Applications* 2.2 (2017), pp. 255–284. DOI: 10.4310/AMSA.2017.v2.n2.a3.

[33] Z. Sun and C.-W. Shu. *Strong Stability of Explicit Runge-Kutta Time Discretizations*. Submitted to SIAM Journal on Numerical Analysis. Nov. 2018. arXiv: 1811.10680 [math.NA].

[34] M. Svärd and J. Nordström. “Review of summation-by-parts schemes for initial-boundary-value problems.” In: *Journal of Computational Physics* 268 (2014), pp. 17–38. DOI: `10.1016/j.jcp.2014.02.031`.

[35] E. Tadmor. “From Semidiscrete to Fully Discrete: Stability of Runge-Kutta Schemes by the Energy Method II.” In: *Collected Lectures on the Preservation of Stability under Discretization*. Ed. by D. J. Estep and S. Tavener. Vol. 109. Proceedings in Applied Mathematics. Philadelphia: Society for Industrial and Applied Mathematics, 2002, pp. 25–49.