Annals of Mathematical Sciences and Applications
Volume 2, Number 2, 255–284, 2017

Stability of the fourth order Runge-Kutta method for time-dependent partial differential equations*

Zheng Sun and Chi-Wang Shu

In this paper, we analyze the stability of the fourth order Runge-Kutta method for integrating semi-discrete approximations of time-dependent partial differential equations. Our study focuses on linear problems and covers general semi-bounded spatial discretizations. A counter example is given to show that the classical four-stage fourth order Runge-Kutta method can not preserve the one-step strong stability, even though the ordinary differential equation system is energy-decaying. But with an energy argument, we show that the strong stability property holds in two steps under an appropriate time step constraint. Based on this fact, the stability extends to general well-posed linear systems. As an application, we utilize the results to examine the stability of the fourth order Runge-Kutta approximations of several specific method of lines schemes for hyperbolic problems, including the spectral Galerkin method and the discontinuous Galerkin method.

AMS 2000 subject classifications: 65L06, 65M12, 65M20.
Keywords and phrases: Stability analysis, Runge-Kutta method, method of lines, spectral Galerkin method, discontinuous Galerkin method.

1. Introduction

In common practice, the method of lines for solving time-dependent partial differential equations (PDEs) starts with a spatial discretization to reduce the problem to an ordinary differential equation (ODE) system solely dependent on time, then a suitable ODE solver is used for the time integration. The issue of stability arises from such a procedure, whether a stable semi-discrete scheme will still be stable after coupling with the time discretization. In this paper, we focus on the stability of the fourth order Runge-Kutta (RK) method in this context.

The model problem for stability analysis is usually chosen as a well-posed linear system $\partial_t u = L(x, t, \partial_x)u$. But firstly, let us consider a simpler

*Research is supported by ARO grant W911NF-15-1-0226 and NSF grant DMS-1418750.

problem $\partial_t u = L(x, \partial_x)u$ with $L + L^\top \le 0$. (Here the superscript $\top$ stands for the adjoint, to be distinguished from T, which stands for the transpose. But for matrices, they are the same under the usual dot product in $\mathbb{R}^m$.) Its semi-discrete scheme corresponds to an autonomous ODE system

$$\frac{d}{dt}u_N(t) = L_N u_N(t).$$

Here L_N is a constant matrix and the parameter N relates with the degree of freedom in the spatial discretization. Suppose L_N inherits the semi-negativity $L_N + L_N^\top \le 0$. Then $\frac{d}{dt}\|u_N(t)\|^2 = (L_N u_N, u_N) + (u_N, L_N u_N) \le 0$, where we temporarily assume $(\cdot,\cdot)$ to be the usual dot product and $\|\cdot\|$ to be the induced norm. Therefore the solution to the semi-discrete scheme is strongly stable. We are interested in whether this strong stability will be preserved after the four-stage fourth order RK time discretization. In other words, with τ being the time step and

$$u_N^{n+1} = P_4(\tau L_N)u_N^n = \left(I + \tau L_N + \frac{1}{2}(\tau L_N)^2 + \frac{1}{6}(\tau L_N)^3 + \frac{1}{24}(\tau L_N)^4\right)u_N^n,$$

we wonder whether $\|u_N^{n+1}\| \le \|u_N^n\|$, or equivalently, whether the operator norm of $P_4(\tau L_N)$ is bounded by 1. If this is true, the stability extends to general semi-bounded systems $\partial_t u = L(x, t, \partial_x)u$ by using a standard argument.

For L_N being a scalar and τ being sufficiently small, the proposition above can be justified by analyzing the region of absolute stability, see Chapter IV.2 in [15], for example. A natural attempt is to apply the eigenvalue analysis to extend the result to systems. However, this technique can only be used for normal matrices L_N, namely $L_N L_N^\top = L_N^\top L_N$. If re-norming is allowed, the method also covers diagonalizable L_N. (But it would still be useless if the diagonalizing matrix is ill-conditioned, see [10].) While for general systems, the naive eigenvalue analysis may not be sufficient. We refer to Chapter 17.1 in [8] for a specific example.

The analysis for the general systems initiates from the coercive problems, namely $L_N + L_N^\top \le -\eta L_N^\top L_N$ for some positive constant η. In [10], Levy and Tadmor used the energy method to prove the strong stability of the third order and the fourth order RK schemes under the coercivity condition and the time step constraint $\tau \le c\eta$, where $c = \frac{3}{50}$ for the third order scheme and $c = \frac{1}{62}$ for the fourth order scheme. Later in [6] and [14], a simpler proof was provided, with the time step constraint being significantly relaxed to $c = 1$, which extends to general s-stage s-th order linear Runge-Kutta schemes. The authors in [6] pointed out that the coercivity condition implies the strong stability of the forward Euler scheme when $\tau \leq \eta$. The strong stability of the high order RK schemes follows from the fact that they can be rewritten as convex combinations of the forward Euler steps. These results coincides with the earlier work on contractivity analysis of the numerical solutions to ODE systems, see [12] and [9]. In their theory for contractivity, or strong stability in our context, a circle condition is assumed, which is essentially equivalent to the strong stability assumption for the forward Euler scheme in [6].

In general, the coercivity condition may not hold for method of lines schemes arising from purely hyperbolic problems. Hence it is still imperative to analyze schemes which only satisfy a semi-negativity condition $L_N + L_N^\top \leq 0$. In [14], for the third order RK scheme, Tadmor successfully removed the coercivity assumption and proved that the third order RK scheme is strongly stable if $L_N + L_N^\top \leq 0$ and $\tau\|L_N\| \leq 1$. But the strong stability for the fourth order RK method remains open, which is the main issue we are concerned with in this paper.

We find that, although the strong stability of the fourth order RK method passes the examination of the scalar equation, and can be extended to normal systems, it does NOT hold in general. More specifically, we have the following counter example in Proposition 1.1, whose proof is given in Appendix A.

Proposition 1.1. *Let* $L_N = -\begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$. *Then* $L_N + L_N^T \leq 0$ *and* $\|P_4(\tau L_N)\| > 1$ *for* τ *sufficiently small.*

Interestingly, despite the negative result on the one-step performance, the strong stability property holds in two time steps. In other words, with the constraint $\tau\|L_N\| \leq c_0$ for some positive constant c_0, $\|P_4(\tau L_N)^2\| \leq 1$ and hence $\|u_N^{n+2}\| \leq \|u_N^n\|$. This is still sufficient to justify the stability after long time integration. It can also be used to obtain the stability $\|u_N^n\| \leq K(t^n)\|u_N^0\|$ for semi-bounded and time-dependent L_N. We will apply the results to study the stability of different spatial discretizations coupled with the fourth order RK approximation.

The paper is organized as follows. In Section 2 we exhibit our main results, and prove the stability of the fourth order RK method with semi-negative and semi-bounded L_N. The case for L_N dependent on time is also discussed. In Section 3, we apply our results to several different spatial discretizations. We specifically focus on Galerkin methods as a complementary of [14], including the spectral Galerkin method and the discontinuous Galerkin method. Finally, we give our concluding remarks in Section 4.

2. Stability of the fourth order RK method

In this section, we analyze the stability of the fourth order RK schemes. For simplicity, we assume everything to be real, but the approach extends to the complex spaces. To facilitate our later discussion on applications to spatial discretizations based on Galerkin methods, we discuss over a general Hilbert space. We avoid clearly characterizing the method of lines scheme in this general setting. But it will not lead to ambiguities, since it only serves as a motivation to derive the fourth order RK iteration, and will not be used in the stability analysis. To reduce to ODE systems, one simply sets $\mathcal{V} = \mathbb{R}^m$ and $(\cdot,\cdot)$ to be the usual dot product. We also remark that, by using the inner product $(u,v)_H = u^T H v$ with H being a symmetric positive definite matrix, $L_N + L_N^\top \leq 0$ is equivalent to $L_N^T H + H L_N \leq 0$, which is consistent with the assumption in equation (6) of [14].

2.1. Notations and the main results

We denote by $\mathcal{V}$ a real Hilbert space equipped with the inner product $(\cdot,\cdot)$. The induced norm is defined as $\|\cdot\| = \sqrt{(\cdot,\cdot)}$. Consider a method of lines scheme defined on $\mathcal{V}$.

$$\partial_t u_N = L_N u_N, \tag{1}$$

where $u_N(\cdot,t) \in \mathcal{V}$ and L_N is a bounded linear operator on $\mathcal{V}$. Suppose L_N is independent of t, then the four-stage fourth order RK approximation can be written as

$$u_N^{n+1} = P_4(\tau L_N) u_N^n, \quad P_4(\tau L_N) = I + \tau L_N + \frac{1}{2}(\tau L_N)^2 + \frac{1}{6}(\tau L_N)^3 + \frac{1}{24}(\tau L_N)^4. \tag{2}$$

When $L_N = L_N(t)$ depends on time, we denote by

$$u_N^{n+1} = R_\tau(t) u_N^n \tag{3}$$

and the specific form of $R_\tau(t)$ will be introduced latter.

For simplicity, we drop all the subscripts N in the remaining parts of the section.

Given an operator L, the operator norm is defined as $\|L\| = \sup_{\|v\|=1} \|Lv\|$. We denote by $\mathcal{B}(\mathcal{V}) = \{L : \|L\| < +\infty\}$ the collection of bounded linear operators on $\mathcal{V}$. For any $L \in \mathcal{B}(\mathcal{V})$, there exists a unique $L^\top \in \mathcal{B}(\mathcal{V})$ such that $(Lw, v) = (w, L^\top v)$. $L^\top$ is referred as the adjoint of L.

If $(v, (L + L^\top)v) \leq 0$, $\forall v \in \mathcal{V}$, denoted by $L + L^\top \leq 0$, then we say L is semi-negative. A special case is when $L + L^\top = 0$, L is referred to be skew-symmetric. If $L + L^\top \leq 2\mu I$ for some positive number μ, namely $L - \mu I$ is semi-negative, then we say L is semi-bounded.

We denote by $[w, v] = -(w, (L + L^\top)v)$ and $[\![w]\!] = \sqrt{[w, w]}$. $[w, v]$ is a bilinear form on $\mathcal{V}$. For $L + L^\top \leq 0$, $[w, v]$ is a semi-inner-product, and one has the Cauchy-Schwarz inequality, $[w, v] \leq [\![w]\!][\![v]\!]$.

To simplify our notation, we define $\mathcal{L} = \tau L$. This notation will be used in the intermediate lemmas and the proofs. For clearness, we restore the notation τL in our main results.

Here we list our stability results on the fourth order RK approximation. The details will be given in Theorem 2.1, Theorem 2.2, Theorem 2.3, Corollary 2.1 and Corollary 2.2 respectively. In the following statements, c_0 and K refer to some positive real numbers.

Our main theorem states that

(i) Suppose $L + L^\top \leq 0$. (2) is strongly stable in two steps. $\|u^{n+2}\| \leq \|u^n\|$, if $\tau\|L\| \leq c_0$.

As a corollary of (1), the stability for semi-bounded and time-dependent operator L have also been proved by using a perturbation analysis and a frozen-coefficient argument.

(ii) Suppose $L + L^\top \leq 2\mu I$. (2) is stable, $\|u^n\| \leq K(t^n)\|u^0\|$, if $\tau \leq \frac{c_0}{\|L\|+\mu}$.

(iii) Suppose $L = L(t)$ satisfies certain regularity assumptions, and $L + L^\top \leq 2\mu I$. Then (3) is stable, $\|u^n\| \leq K(t^n)\|u^0\|$, under an appropriate time step constraint.

As we know, according to the eigenvalue analysis, for normal and semi-negative L, the one-step strong stability will hold under a proper time-step restriction. We also use the energy argument to prove the one-step results, as corollaries of the two-step stability analysis.

(iv) Suppose $L + L^\top = 0$. (2) is strongly stable, $\|u^{n+1}\| \leq \|u^n\|$, if $\tau\|L\| \leq 2\sqrt{2}$.

(v) Suppose $L + L^\top \leq 0$, $LL^\top = L^\top L$. (2) is strongly stable, $\|u^{n+1}\| \leq \|u^n\|$, if $\tau\|L\| \leq c_0$.

For simplicity, we assume the time steps to be uniform in (i), (ii) and (iii). However, these results are essentially built on the two-step performance of the RK time integrator. One only needs the time steps to be uniform for every two steps, namely, $\tau_{2k+1} = \tau_{2k+2}$. As for (iv) and (v), they fit for general time step sizes.

2.2. An energy equality

The first thing we would do is to derive an energy equality, which would be useful in understanding the subtle change of the norm of the solution after one time step. To this end, we introduce several identities in Lemma 2.1, and then use them to derive the equality in Lemma 2.2.

Lemma 2.1.

$$(\mathcal{L}v, v) = -\frac{\tau}{2}[\![v]\!]^2, \tag{4}$$

$$(\mathcal{L}^2 v, v) = -\|\mathcal{L}v\|^2 - \tau[\mathcal{L}v, v], \tag{5}$$

$$(\mathcal{L}^3 v, v) = \frac{\tau}{2}[\![\mathcal{L}v]\!]^2 - \tau[\mathcal{L}^2 v, v], \tag{6}$$

$$(\mathcal{L}^4 v, v) = \|\mathcal{L}^2 v\|^2 + \tau[\mathcal{L}^2 v, \mathcal{L}v] - \tau[\mathcal{L}^3 v, v]. \tag{7}$$

Proof. (i) $(\mathcal{L}v, v) = \frac{1}{2}(\mathcal{L}v, v) + \frac{1}{2}(v, \mathcal{L}v) = \frac{1}{2}(v, (\mathcal{L} + \mathcal{L}^\top)v) = -\frac{\tau}{2}[\![v]\!]^2$.
(ii) $(\mathcal{L}^2 v, v) = (\mathcal{L}v, \mathcal{L}^\top v) = -(\mathcal{L}v, \mathcal{L}v) + (\mathcal{L}v, (\mathcal{L}+\mathcal{L}^\top)v) = -\|\mathcal{L}v\|^2 - \tau[\mathcal{L}v, v]$.
(iii) $(\mathcal{L}^3 v, v) = (\mathcal{L}^2 v, \mathcal{L}^\top v) = -(\mathcal{L}(\mathcal{L}v), \mathcal{L}v) + (\mathcal{L}^2 v, (\mathcal{L} + \mathcal{L}^\top)v) = \frac{\tau}{2}[\![\mathcal{L}v]\!]^2 - \tau[\mathcal{L}^2 v, v]$, where we have used (4) in the last equality.
(iv) $(\mathcal{L}^4 v, v) = (\mathcal{L}^3 v, \mathcal{L}^\top v) = -(\mathcal{L}^2(\mathcal{L}v), \mathcal{L}v) + (\mathcal{L}^3 v, (\mathcal{L} + \mathcal{L}^\top)v) = \|\mathcal{L}^2 v\|^2 + \tau[\mathcal{L}^2 v, \mathcal{L}v] - \tau[\mathcal{L}^3 v, v]$, where (5) is used in the last equality. □

Lemma 2.2 (Energy equality)**.**

$$\|u^{n+1}\|^2 - \|u^n\|^2 = Q_1(u^n),$$

where

$$Q_1(u^n) = \frac{1}{576}\|\mathcal{L}^4 u^n\|^2 - \frac{1}{72}\|\mathcal{L}^3 u^n\|^2 + \tau \sum_{i,j=0}^{3} \alpha_{ij}[\mathcal{L}^i u^n, \mathcal{L}^j u^n], \tag{8}$$

and

$$A = (\alpha_{ij})_{4\times 4} = -\begin{pmatrix} 1 & 1/2 & 1/6 & 1/24 \\ 1/2 & 1/3 & 1/8 & 1/24 \\ 1/6 & 1/8 & 1/24 & 1/48 \\ 1/24 & 1/24 & 1/48 & 1/144 \end{pmatrix}.$$

Proof. Taking an inner product of (2) with u^n, one has

$$\frac{1}{2}\|u^{n+1}\|^2 - \frac{1}{2}\|u^n\|^2 - \frac{1}{2}\|u^{n+1} - u^n\|^2 = ((\mathcal{L} + \frac{1}{2}\mathcal{L}^2 + \frac{1}{6}\mathcal{L}^3 + \frac{1}{24}\mathcal{L}^4)u^n, u^n).$$

Applying (4)–(7) with $v = u^n$, we obtain

$$\begin{aligned}(9)\qquad &\|u^{n+1}\|^2 - \|u^n\|^2 \\ &= \|u^{n+1} - u^n\|^2 - \|\mathcal{L}u^n\|^2 + \frac{1}{12}\|\mathcal{L}^2u^n\|^2 - \tau[\![u^n]\!]^2 - \tau[\mathcal{L}u^n, u^n] \\ &+ \frac{\tau}{6}[\![\mathcal{L}u^n]\!]^2 - \frac{\tau}{3}[\mathcal{L}^2u^n, u^n] + \frac{\tau}{12}[\mathcal{L}^2u^n, \mathcal{L}u^n] - \frac{\tau}{12}[\mathcal{L}^3u^n, u^n].\end{aligned}$$

Using (4)–(6) with $v = \mathcal{L}u^n$, the first two terms on the right can be expanded.

$$\begin{aligned}(10)\qquad &\|u^{n+1} - u^n\|^2 - \|\mathcal{L}u^n\|^2 \\ &= (u^{n+1} - u^n - \mathcal{L}u^n, u^{n+1} - u^n + \mathcal{L}u^n) \\ &= \|u^{n+1} - u^n - \mathcal{L}u^n\|^2 + ((\frac{1}{2}\mathcal{L}^2 + \frac{1}{6}\mathcal{L}^3 + \frac{1}{24}\mathcal{L}^4)u^n, 2\mathcal{L}u^n) \\ &= \|u^{n+1} - u^n - \mathcal{L}u^n\|^2 - \frac{1}{3}\|\mathcal{L}^2u^n\|^2 - \frac{\tau}{2}[\![\mathcal{L}u^n]\!]^2 \\ &- \frac{\tau}{3}[\mathcal{L}^2u^n, \mathcal{L}u^n] + \frac{\tau}{24}[\![\mathcal{L}^2u^n]\!]^2 - \frac{\tau}{12}[\mathcal{L}^3u^n, \mathcal{L}u^n].\end{aligned}$$

Substitute (10) into (9), one has

$$\begin{aligned}(11)\qquad &\|u^{n+1}\|^2 - \|u^n\|^2 = \|u^{n+1} - u^n - \mathcal{L}u^n\|^2 - \frac{1}{4}\|\mathcal{L}^2u^n\|^2 - \tau[\![u^n]\!]^2 \\ &- \tau[\mathcal{L}u^n, u^n] - \frac{\tau}{3}[\![\mathcal{L}u^n]\!]^2 - \frac{\tau}{3}[\mathcal{L}^2u^n, u^n] - \frac{\tau}{4}[\mathcal{L}^2u^n, \mathcal{L}u^n] \\ &- \frac{\tau}{12}[\mathcal{L}^3u^n, u^n] + \frac{\tau}{24}[\![\mathcal{L}^2u^n]\!]^2 - \frac{\tau}{12}[\mathcal{L}^3u^n, \mathcal{L}u^n].\end{aligned}$$

Similar as before, one can use (4) and (5) with $v = \mathcal{L}^2u^n$ to calculate the difference of square terms on the right.

$$\begin{aligned}(12)\qquad &\|u^{n+1} - u^n - \mathcal{L}u^n\|^2 - \frac{1}{4}\|\mathcal{L}^2u^n\|^2 \\ &= (u^{n+1} - u^n - \mathcal{L}u^n - \frac{1}{2}\mathcal{L}^2u^n, u^{n+1} - u^n - \mathcal{L}u^n + \frac{1}{2}\mathcal{L}^2u^n) \\ &= \|u^{n+1} - u^n - \mathcal{L}u^n - \frac{1}{2}\mathcal{L}^2u^n\|^2 + ((\frac{1}{6}\mathcal{L}^3 + \frac{1}{24}\mathcal{L}^4)u^n, \mathcal{L}^2u^n)\end{aligned}$$

$$= \|u^{n+1} - u^n - \mathcal{L}u^n - \frac{1}{2}\mathcal{L}^2u^n\|^2 - \frac{1}{24}\|\mathcal{L}^3u^n\|^2 - \frac{\tau}{12}[\![\mathcal{L}^2u^n]\!]^2 - \frac{\tau}{24}[\mathcal{L}^3u^n, \mathcal{L}^2u^n].$$

Once again, we plug (12) into (11) to get

$$\begin{aligned}
&\|u^{n+1}\|^2 - \|u^n\|^2 \\
&\quad= \|u^{n+1} - u^n - \mathcal{L}u^n - \frac{1}{2}\mathcal{L}^2u^n\|^2 - \frac{1}{24}\|\mathcal{L}^3u^n\|^2 \\
&\quad- \tau[\![u^n]\!]^2 - \tau[\mathcal{L}u^n, u^n] - \frac{\tau}{3}[\![\mathcal{L}u^n]\!]^2 - \frac{\tau}{3}[\mathcal{L}^2u^n, u^n] - \frac{\tau}{4}[\mathcal{L}^2u^n, \mathcal{L}u^n] \\
&\quad- \frac{\tau}{12}[\mathcal{L}^3u^n, u^n] - \frac{\tau}{24}[\![\mathcal{L}^2u^n]\!]^2 - \frac{\tau}{12}[\mathcal{L}^3u^n, \mathcal{L}u^n] - \frac{\tau}{24}[\mathcal{L}^3u^n, \mathcal{L}^2u^n].
\end{aligned}$$

Finally, note that

$$\begin{aligned}
&\|u^{n+1} - u^n - \mathcal{L}u^n - \frac{1}{2}\mathcal{L}^2u^n\|^2 - \frac{1}{36}\|\mathcal{L}^3u^n\|^2 \\
&\quad= \|u^{n+1} - u^n - \mathcal{L}u^n - \frac{1}{2}\mathcal{L}^2u^n - \frac{1}{6}\mathcal{L}^3u^n\|^2 + (\frac{1}{24}\mathcal{L}^4u^n, \frac{1}{3}\mathcal{L}^3u^n) \\
&\quad= \frac{1}{576}\|\mathcal{L}^4u^n\|^2 - \frac{\tau}{144}[\![\mathcal{L}^3u^n]\!]^2.
\end{aligned}$$

Therefore

$$\begin{aligned}
Q_1(u^n) = &\frac{1}{576}\|\mathcal{L}^4u^n\|^2 - \frac{1}{72}\|\mathcal{L}^3u^n\|^2 - \tau[\![u^n]\!]^2 - \tau[\mathcal{L}u^n, u^n] - \frac{\tau}{3}[\![\mathcal{L}u^n]\!]^2 \\
&- \frac{\tau}{3}[\mathcal{L}^2u^n, u^n] - \frac{\tau}{4}[\mathcal{L}^2u^n, \mathcal{L}u^n] - \frac{\tau}{12}[\mathcal{L}^3u^n, u^n] - \frac{\tau}{24}[\![\mathcal{L}^2u^n]\!]^2 \\
&- \frac{\tau}{12}[\mathcal{L}^3u^n, \mathcal{L}u^n] - \frac{\tau}{24}[\mathcal{L}^3u^n, \mathcal{L}^2u^n] - \frac{\tau}{144}[\![\mathcal{L}^3u^n]\!]^2,
\end{aligned}$$

which can be rewritten as (8). ☐

According to Lemma 2.2, the energy change $Q_1(u^n)$ consists of two parts, the numerical dissipation $\frac{1}{576}\|\mathcal{L}^4u^n\|^2 - \frac{1}{72}\|\mathcal{L}u^n\|^2$ and the generalized quadratic form $\tau\sum_{i,j=0}^4 \alpha_{ij}[\mathcal{L}u^i, \mathcal{L}u^j]$. When L is skew-symmetric, the quadratic form is simply 0. Hence one can obtain the one-step strong stability.

Corollary 2.1. *Suppose $L + L^\top = 0$, then the fourth order RK approximation (2) to the method of lines scheme (1) is strongly stable, $\|u^{n+1}\| \leq \|u^n\|$, if $\tau\|L\| \leq 2\sqrt{2}$.*

Proof. $L + L^\top = 0$ implies $[w, v] = 0, \forall v, w$. Hence when $\tau\|L\| \le 2\sqrt{2}$,

$$\|u^{n+1}\|^2 - \|u^n\|^2 = Q_1(u^n) = \frac{1}{576}\|\mathcal{L}^4 u^n\|^2 - \frac{1}{72}\|\mathcal{L}^3 u^n\|^2$$
$$\le (\frac{1}{576}\|\mathcal{L}\|^2 - \frac{1}{72})\|\mathcal{L}^3 u^n\|^2 \le 0. \qquad \square$$

2.3. Stability for semi-negative L

According to Lemma 2.2, for non-skew-symmetric L, we would need to absorb the quadratic form with the help of the numerical dissipation. One can see that the high order terms $\mathcal{L}^i u^n$ with $i \ge 3$ are easy to control. But there are no other terms to bound u^n, $\mathcal{L}u^n$ and $\mathcal{L}^2 u^n$. Our only hope is that $\tau \sum_{ij=0}^{2} \alpha_{ij}[\mathcal{L}u^i, \mathcal{L}u^j]$ itself is negative-definite. Lemma 2.3 indicates that, to check the negativity of the generalized quadratic form, one only needs to examine the coefficient matrix, as that for the polynomials. In Lemma 2.4, we prove a sufficient condition for strong stability, once $\sum_{ij=0}^{2} \alpha_{ij}[\mathcal{L}u^i, \mathcal{L}u^j]$ is negative-definite, the energy change will be non-positive when $\|\mathcal{L}\|$ is sufficiently small.

Lemma 2.3. *Suppose L is semi-negative. Let $M = (m_{ij})$ be a real symmetric semi-negative definite matrix, then $\sum_{i,j} m_{ij}[u_i, u_j] \le 0$.*

Proof. Suppose $M = S^T \Lambda S$, where $S = (s_{ij})$ is an orthogonal matrix and Λ is a diagonal matrix with the diagonal elements $\lambda_i \le 0$. Note that $m_{ij} = \sum_k \lambda_k s_{ki} s_{kj}$. Therefore

$$\sum_{i,j} m_{ij}[u_i, u_j] = \sum_{i,j} (\sum_k \lambda_k s_{ki} s_{kj})[u_i, u_j] = \sum_k \lambda_k (\sum_{i,j} s_{ki} s_{kj}[u_i, u_j])$$
$$= \sum_k \lambda_k [\sum_i s_{ki} u_i, \sum_j s_{kj} u_j] = \sum_k \lambda_k [\![\sum_i s_{ki} u_i]\!]^2 \le 0. \quad \square$$

Lemma 2.4. *Suppose L is semi-negative. Let*

$$Q(u) = \alpha\|\mathcal{L}^3 u\|^2 + \tau \sum_{i,j=0}^{a} \alpha_{ij}[\mathcal{L}^i u, \mathcal{L}^j u],$$

where $\alpha_{ij} = \alpha_{ji}$ and $a \ge 2$. If $\alpha < 0$ and $A_2 = \begin{pmatrix} \alpha_{00} & \alpha_{01} & \alpha_{02} \\ \alpha_{10} & \alpha_{11} & \alpha_{12} \\ \alpha_{20} & \alpha_{21} & \alpha_{22} \end{pmatrix}$ is negative-definite, then there exists $c_0 > 0$, such that $Q(u) \le 0$ as long as $\|\mathcal{L}\| \le c_0$.

Proof. Let $-\varepsilon$ be the largest eigenvalue of the matrix A_2. Then $A_2 + \varepsilon I$ is semi-negative definite. According to Lemma 2.3,

(13)
$$\sum_{i,j=0}^{2} \alpha_{ij}[\mathcal{L}^i u, \mathcal{L}^j u] = \sum_{i,j=0}^{2} (\alpha_{ij} + \varepsilon\delta_{ij})[\mathcal{L}^i u, \mathcal{L}^j u] - \varepsilon([\![u]\!]^2 + [\![\mathcal{L}u]\!]^2 + [\![\mathcal{L}^2 u]\!]^2)$$
$$\leq -\varepsilon([\![u]\!]^2 + [\![\mathcal{L}u]\!]^2 + [\![\mathcal{L}^2 u]\!]^2).$$

By Cauchy-Schwartz and the arithmetic-geometric mean inequality,

$$\sum_{\max(i,j)>2}^{a} \alpha_{ij}[\mathcal{L}^i u, \mathcal{L}^j u] \leq \varepsilon([\![u]\!]^2 + [\![\mathcal{L}u]\!]^2 + [\![\mathcal{L}^2 u]\!]^2) + \sum_{i=3}^{a} \tilde{\alpha}_i [\![\mathcal{L}^i u]\!]^2$$
$$\leq \varepsilon([\![u]\!]^2 + [\![\mathcal{L}u]\!]^2 + [\![\mathcal{L}^2 u]\!]^2)$$
$$+ \sum_{i=3}^{a} 2\tilde{\alpha}_i \|L\| \|\mathcal{L}\|^{2(i-3)} \|\mathcal{L}^3 u\|^2, \tag{14}$$

where we have used the fact $[\![v]\!]^2 \leq 2\|L\|\|v\|^2$ in the last inequality, and $\tilde{\alpha}_i$ are some non-negative constants depending on ε and α_{ij}. Using (13) and (14), if $\|\mathcal{L}\| = \tau\|L\| \leq c_0$, one has

$$Q(u) \leq (\alpha + \sum_{i=3}^{a} 2\tilde{\alpha}_i c_0^{2(i-3)+1}) \|\mathcal{L}^3 u\|^2.$$

Since α is negative, $Q(u)$ is non-positive as long as c_0 is sufficiently small. $\square$

Unfortunately, $Q_1(u^n)$ does not satisfy the assumptions in Lemma 2.4. Actually, the three eigenvalues of $A_2 = -\begin{pmatrix} 1 & 1/2 & 1/6 \\ 1/2 & 1/3 & 1/8 \\ 1/6 & 1/8 & 1/24 \end{pmatrix}$ are 0.00560618, −0.0793266 and −1.30128. This motivates us to disprove the strong stability of the fourth order RK method and we end up with the counter example in Proposition 1.1. So we turn to seek the power-boundedness of $P_4(\mathcal{L})$. Surprisingly, though $Q_1(u^n)$ itself fails to pass Lemma 2.4, $Q_1(u^{n+1}) + Q_1(u^n)$ succeeds. Hence we obtain the two-step strong stability in Theorem 2.1.

Theorem 2.1 (Two-step strong stability for semi-negative L)**.** *Suppose $L + L^\top \leq 0$, then the fourth order RK approximation* (2) *to the method of lines scheme* (1) *is strongly stable in two steps, $\|u^{n+2}\| \leq \|u^n\|$, if $\tau\|L\| \leq c_0$.*

Proof. Let $u^{n+2} = P_4(\mathcal{L})u^{n+1}$. Then

$$\|u^{n+2}\|^2 - \|u^{n+1}\|^2 = Q_1(u^{n+1}).$$

Substitute (2) into $Q_1(u^{n+1})$ and rewrite the quadratic form in terms of u^n. By direct calculation, one can obtain

$$Q_1(u^{n+1}) = \frac{1}{576}\|\mathcal{L}^4 u^{n+1}\|^2 - \frac{1}{72}\|\mathcal{L}^3 u^{n+1}\|^2 + \tau \sum_{i,j=0}^{7} \tilde{\alpha}_{ij}[\mathcal{L}^i u^n, \mathcal{L}^j u^n],$$

where

$$\tilde{A}_2 = (\tilde{\alpha}_{ij})_{3\times 3} = -\begin{pmatrix} 1 & 3/2 & 7/6 \\ 3/2 & 7/3 & 15/8 \\ 7/6 & 15/8 & 37/24 \end{pmatrix}.$$

The complete coefficient matrix $\tilde{A}$ is given in Appendix B. While according to Lemma 2.2,

$$Q_1(u^n) = \frac{1}{576}\|\mathcal{L}^4 u^n\|^2 - \frac{1}{72}\|\mathcal{L}^3 u^n\|^2 + \tau \sum_{i,j=0}^{3} \hat{\alpha}_{ij}[\mathcal{L}^i u^n, \mathcal{L}^j u^n],$$

where

$$\hat{A}_2 = (\hat{\alpha}_{ij})_{3\times 3} = -\begin{pmatrix} 1 & 1/2 & 1/6 \\ 1/2 & 1/3 & 1/8 \\ 1/6 & 1/8 & 1/24 \end{pmatrix}.$$

When $\|\mathcal{L}\| \leq 2$,

$$\frac{1}{576}\|\mathcal{L}^4 u^{n+1}\|^2 - \frac{1}{72}\|\mathcal{L}^3 u^{n+1}\|^2 \leq 0,$$
$$\frac{1}{576}\|\mathcal{L}^4 u^n\|^2 - \frac{1}{72}\|\mathcal{L}^3 u^n\|^2 \leq -\frac{1}{144}\|\mathcal{L}^3 u^n\|^2.$$

Hence

$$\begin{aligned}\|u^{n+2}\|^2 - \|u^n\|^2 &= Q_1(u^{n+1}) + Q_1(u^n) \\ &\leq -\frac{1}{144}\|\mathcal{L}^3 u^n\|^2 + \tau \sum_{i,j=0}^{7} \alpha_{ij}[\mathcal{L}^i u^n, \mathcal{L}^j u^n],\end{aligned}$$

where

$$A_2 = (\alpha_{ij})_{3\times 3} = \tilde{A}_2 + \hat{A}_2 = -\begin{pmatrix} 2 & 2 & 4/3 \\ 2 & 8/3 & 2 \\ 4/3 & 2 & 19/12 \end{pmatrix}.$$

The three eigenvalues of A_2 are -5.73797, -0.499093 and -0.0129329. Hence A_2 is negative definite, and one could complete the proof by applying Lemma 2.4. $\square$

Remark 2.1. *There is another interpretation of the two-step strong stability. We treat u^{n+1} as a intermediate stage and consider u^{n+2} as the solution of an eight-stage fourth order RK scheme composed by two four-stage fourth order RK time stepping. Then Theorem 2.1 states this new RK scheme is strongly stable. This coincides with Remark 3 of Section 4 in [14], where the author conjectured the fourth order RK scheme can preserve the strong stability if allowing additional stages.*

Remark 2.2. *Here we only focus on the existence of the constant c_0 instead of giving a specific estimate. The argument above is far from sharp. To obtain a reasonable bound for c_0, one should expand $\frac{1}{576}\|\mathcal{L}^4 u^{n+1}\|^2 - \frac{1}{72}\|\mathcal{L}^3 u^{n+1}\|^2$ to separate the quadratic form involving $[\cdot,\cdot]$, and carefully go through all the constants. To this end, a generalized version of Lemma 2.1 is also needed. We leave these to interested readers.*

Realizing the fact that, for a normal operator G, $\|G^2\| \leq 1$ implies $\|G\| \leq 1$, we prove the one-step strong stability for normal and semi-negative L.

Corollary 2.2 (Strong stability for normal and semi-negative L)**.** *Suppose $LL^\top = L^\top L$ and $L + L^\top \leq 0$, then the fourth order RK approximation (2) to the method of lines scheme (1) is strongly stable, $\|u^{n+1}\| \leq \|u^n\|$, if $\tau\|L\| \leq c_0$ for some constant c_0.*

Proof. According to Theorem 2.1, we have $\|P_4(\mathcal{L})^2\| \leq 1$ if $\|\mathcal{L}\| \leq c_0$. When $\mathcal{L}$ is normal, $P_4(\mathcal{L})$ is also normal, $P_4(\mathcal{L})P_4(\mathcal{L})^\top = P_4(\mathcal{L})^\top P_4(\mathcal{L})$. Hence for any $u \in \mathcal{V}$,

$$\begin{aligned}\|P_4(\mathcal{L})u\|^2 &= (P_4(\mathcal{L})u, P_4(\mathcal{L})u) = (u, P_4(\mathcal{L})^\top P_4(\mathcal{L})u)\\ &= (u, P_4(\mathcal{L})P_4(\mathcal{L})^\top u) = \|P_4(\mathcal{L})^\top u\|^2.\end{aligned}$$

Therefore, under the same constraints, $\|\mathcal{L}\| \leq c_0$, one has

$$\begin{aligned}\|P_4(\mathcal{L})\|^2 &= \sup_{\|u\|=1} \|P_4(\mathcal{L})u\|^2 = \sup_{\|u\|=1} (u, P_4(\mathcal{L})^\top P_4(\mathcal{L})u)\\ &\leq \sup_{\|u\|=1} \|P_4(\mathcal{L})^\top P_4(\mathcal{L})u\| = \sup_{\|u\|=1} \|P_4(\mathcal{L})^2 u\| = \|P_4(\mathcal{L})^2\| \leq 1.\end{aligned}$$

In other words, $\|u^{n+1}\| \leq \|u^n\|$ if $\|\mathcal{L}\| \leq c_0$. $\square$

Remark 2.3. *By using scalar eigenvalue analysis, the sharp bound of c_0 in Corollary 2.2 is $2\sqrt{2}$, as that in Corollary 2.1.*

2.4. Stability for semi-bounded L

We then apply the perturbation analysis to obtain the stability for semi-bounded L. In the following proof, one actually needs to assume the time steps satisfy $\tau_{2k+1} = \tau_{2k+2}$ in order to apply Theorem 2.1. For simplicity, we use a uniform time stepping, namely $\tau_k = \tau$. The same assumption will be used in Section 2.5.

Theorem 2.2. *Suppose $L + L^\top \leq 2\mu I$, then the fourth order RK approximation (2) to the method of lines scheme (1) is stable, $\|u^n\| \leq K(t^n)\|u^0\|$, if $\tau \leq \frac{c_0}{\|L\|+\mu}$ for some constant c_0. Here $K(t^n)$ is a constant depending on μ and t^n.*

Proof. Note that

$$P_4(\tau L)^2 = P_4(\tau(L - \mu I) + \tau\mu I)^2 = P_4(\tau(L - \mu I))^2 + \tau\mu P(\tau\mu, \tau(L - \mu I)),$$

where $P(s, G)$ is of the form $P(s, G) = \sum_{i,j\geq 0,\ i+j\leq 7} \alpha_{ij} s^i G^j$. Since $L+L^\top \leq 2\mu I$, $(L-\mu I)+(L-\mu I)^\top \leq 0$. According to Theorem 2.1, $\|P_4(\tau(L-\mu I))^2\| \leq 1$ if $\tau\|L - \mu I\| \leq c_0$. Under a stricter restriction $\tau \leq \frac{c_0}{\|L\|+\mu}$, one also has

$$\|P(\tau\mu, \tau(L-\mu I))\| \leq \sum_{i,j\geq 0,\ i+j\leq 7} |\alpha_{ij}||\tau\mu|^i\|\tau(L-\mu I)\|^j \leq \sum_{i,j\geq 0,\ i+j\leq 7} |\alpha_{ij}|c_0^{i+j}.$$

Therefore for $\tau \leq \frac{c_0}{\|L\|+\mu}$, $\|P_4(\tau L)^2\| \leq 1+c\mu\tau$, where $c = \sum_{i,j\geq 0,\ i+j\leq 7} |\alpha_{ij}|\times c_0^{i+j}$. With this one-step estimate, one has

$$\|u^{2k}\| \leq \|P_4(\tau L)^2\|^k\|u^0\| \leq (1 + c\mu\tau)^{\frac{t^{2k}}{2\tau}}\|u^0\| \leq K(t^{2k})\|u^0\|,$$

and

$$\|u^{2k+1}\| \leq \|P_4(\tau L)\|\|u^{2k}\| \leq P_4(c_0)K(t^{2k})\|u^0\|,$$

which prove the stability. □

2.5. Stability for semi-bounded and time-dependent L

We conclude this section by extending the results to L dependent on time. One should note in this case, the fourth order RK scheme can no longer be written as the truncated exponential in (2). Also different fourth order RK time integrators are no longer equivalent. One can use the classical four-stage fourth order RK scheme in (15) as an example for our stability analysis. However, our proof does not rely on this specific form and can be used for general cases. The classical four-stage fourth order RK scheme is

$$\begin{aligned} k_1 &= L(t^n)u^n, \\ k_2 &= L(t^{n+\frac{1}{2}})(u^n + \frac{\tau}{2}k_1), \\ k_3 &= L(t^{n+\frac{1}{2}})(u^n + \frac{\tau}{2}k_2), \\ k_4 &= L(t^{n+1})(u^n + \tau k_3), \\ u^{n+1} &= u^n + \frac{\tau}{6}(k_1 + 2k_2 + 2k_3 + k_4), \end{aligned} \tag{15}$$

where $t^{n+\frac{1}{2}} = t^n + \frac{\tau}{2}$. As a short hand notation, let us denote by

$$u^{n+1} = R_\tau(t^n)u^n,$$

where

$$\begin{aligned} R_\tau(t^n) = I &+ \frac{\tau}{6}\left(L(t^n) + 4L(t^{n+\frac{1}{2}}) + L(t^{n+1})\right) \\ &+ \frac{\tau^2}{6}\left(L(t^n)L(t^{n+\frac{1}{2}}) + L(t^{n+\frac{1}{2}})^2 + L(t^{n+\frac{1}{2}})L(t^{n+1})\right) \\ &+ \frac{\tau^3}{12}\left(L(t^n)L(t^{n+\frac{1}{2}})^2 + L(t^{n+\frac{1}{2}})^2L(t^{n+1})\right) \\ &+ \frac{\tau^4}{24}L(t^n)L(t^{n+\frac{1}{2}})^2L(t^{n+1}). \end{aligned} \tag{16}$$

The two-step approximation can be written as $u^{n+2} = R_\tau(t^{n+1})R_\tau(t^n)u^n$.

We also assume that L satisfies the Lipschitz continuity condition

$$\|L(t_2) - L(t_1)\| \le \eta|t_2 - t_1|, \tag{17}$$

with $\sup_t \|L(t)\| \le \eta < +\infty$ for some constant η.

Lemma 2.5. *Suppose L satisfies (17). Then*

$$\|\prod_{i=1}^m L(t + \alpha_i\tau) - (L(t))^m\| \le \sum_{i=1}^m |\alpha_i|\eta^m\tau.$$

Proof. We prove by induction. According to assumption (17), the lemma holds for $m = 1$. Suppose it holds for m, then

$$
\begin{aligned}
&\|\prod_{i=1}^{m+1} L(t+\alpha_i\tau) - (L(t))^{m+1}\| \\
&\quad \le \|\prod_{i=1}^{m+1} L(t+\alpha_i\tau) - L(t)\prod_{i=1}^{m} L(t+\alpha_i\tau)\| \\
&\quad + \|L(t)\prod_{i=1}^{m} L(t+\alpha_i\tau) - (L(t))^{m+1}\| \\
&\quad \le \|L(t+\alpha_{m+1}\tau) - L(t)\| \prod_{i=1}^{m} \|L(t+\alpha_i\tau)\| \\
&\quad + \|L(t)\|\|\prod_{i=1}^{m} L(t+\alpha_i\tau) - (L(t))^{m}\| \\
&\quad \le |\alpha_{m+1}|\eta\tau\cdot\eta^m + \eta\cdot(\sum_{i=1}^{m}|\alpha_i|)\eta^m\tau \\
&\quad = \sum_{i=1}^{m+1}|\alpha_i|\eta^{m+1}\tau,
\end{aligned}
$$

where we have used the inductive assumption and $\eta > \sup_t \|L(t)\|$ in the second last line. Hence the lemma holds for any positive integer m. ☐

Theorem 2.3. *Suppose $L = L(t)$ and $L + L^\top \le 2\mu I$ for some positive number μ. L satisfies the Lipschitz continuity condition (17). Then the four-stage fourth order RK approximation (15) to the method of lines scheme (1) is stable, $\|u^n\| \le K(t^n)\|u^0\|$, if $\tau \le \frac{c_0}{\eta+\mu}$. Here $K(t^n)$ is some constant depending on μ and t^n.*

Proof.

$$
\begin{aligned}
P_4(\tau L)^2 = I &+ 2\tau L + 2(\tau L)^2 + \frac{4}{3}(\tau L)^3 + \frac{2}{3}(\tau L)^4 \\
&+ \frac{1}{4}(\tau L)^5 + \frac{5}{72}(\tau L)^6 + \frac{1}{72}(\tau L)^7 + \frac{1}{576}(\tau L)^8.
\end{aligned} \tag{18}
$$

And

$$
R_\tau(t^{n+1})R_\tau(t^n) = \sum_{i=0}^{8}\sum_j \alpha_{ij}\prod_{k=1}^{i}(\tau L(t^n + \tilde{\alpha}_{ijk}\tau)), \tag{19}
$$

with $\{\sum_j \alpha_{ij}\}_{i=0}^{8}$ take values of 1, 2, 2, $\frac{4}{3}$, $\frac{2}{3}$, $\frac{1}{4}$, $\frac{5}{72}$, $\frac{1}{72}$ and $\frac{1}{576}$ as those in $P_4(\tau L)^2$. (This must be true since when L is independent of time, (19) reduces to (18).) Applying Lemma 2.5, one has

$$\begin{aligned}\|R_\tau(t^{n+1})R_\tau(t^n) - P_4(\tau L)^2\| &= \|\sum_{i=0}^{8}\sum_j \alpha_{ij}\tau^i(\prod_{k=1}^{i} L(t+\tilde{\alpha}_{ijk}\tau) - (L(t))^i)\| \\ &\le \sum_{i=0}^{8}\sum_j |\alpha_{ij}|\tau^i \|\prod_{k=1}^{i} L(t+\tilde{\alpha}_{ijk}\tau) - (L(t))^i\| \\ &\le \tau \sum_{i=0}^{8}\sum_j |\alpha_{ij}|(\sum_{k}^{i} |\tilde{\alpha}_{ijk}|)(\tau\eta)^i = \tilde{c}\tau.\end{aligned}$$

On the other hand, $\|P_4(\tau L)^2\| \le 1 + c\mu\tau$ if $\tau(\eta+\mu) \le c_0$, hence

$$\|R_\tau(t^{n+1})R_\tau(t^n)\| \le \|P_4(\tau L)^2\| + \|R_\tau(t^{n+1})R_\tau(t^n) - P_4(\tau L)^2\| \le 1 + (c\mu + \tilde{c})\tau.$$

As we have done in Theorem 2.2, one can prove $\|u^n\| \le K(t^n)\|u^0\|$, if $\tau \le \frac{c_0}{\eta+\mu}$. □

Remark 2.4. *For different four-stage fourth order RK schemes, the values of α_{ij} and $\tilde{\alpha}_{ijk}$ are different. But $\sum_j \alpha_{ij}$ are always the same. Hence the proof above would still work.*

3. Applications to hyperbolic problems

In this section, we apply the previous results to several semi-discrete approximations of hyperbolic problems, and justify their stability after coupling with the fourth order RK time integrator. The time step restriction is also referred as the CFL condition in this context. In [14], Tadmor has provided detailed examples for spatial discretization based on the nodal-value formulation, including the finite difference method and spectral collocation method. Complementary to [14], we will mainly focus on Galerkin methods, including the global approach, the spectral Galerkin method, and the local approach, the finite element discontinuous Galerkin method.

The Galerkin methods are spatial discretization techniques based on the weak formulation. Consider the initial value problem,

$$\begin{cases} \partial_t u(x,t) = L(x,t,\partial_x)u(x,t), & (x,t) \in \Omega \times (0,T), \\ u(x,0) = u_0(x), & x \in \Omega. \end{cases}$$

The Galerkin methods seeks a solution of the form $u_N(x,t) = \sum_{k=0}^{N} c_k(t) \times \phi_k(x)$. Here $\{\phi_k(x)\}_{k=0}^{N}$ forms the basis of a finite dimensional space $\mathcal{V}$ on Ω, equipped with the inner product $(\cdot,\cdot)$. In addition, one requires that

$$\begin{cases} (\partial_t u_N, v_N) = B(u_N, v_N), & \forall v_N \in \mathcal{V}, \\ \quad u_N(x,0) = Pu_0. \end{cases}$$

Here B is a bilinear form on $\mathcal{V}$ with $B(u_N, v_N)$ approximating (Lu_N, v_N), and P is the projection to $\mathcal{V}$.

For spectral Galerkin method, $\mathcal{V}$ is chosen as the span of the trigonometric functions or polynomials over the whole domain. Since the functions in $\mathcal{V}$ are sufficiently smooth, one can directly set $B(u_N, v_N) = (Lu_N, v_N)$. Hence the method can be written as

$$\partial_t u_N = L_N u_N = PLu_N = PLPu_N.$$

It suffices to check whether $L_N = PLP$ satisfies the conditions in Section 2 and examine $\|L_N\|$ to obtain the time step constraint.

For discontinuous Galerkin method, $\mathcal{V}$ is a piecewise polynomial space based on an appropriate partition of the domain Ω. The associated bilinear form is involved with the so-called numerical flux. We will explain it in the latter part of this section.

3.1. Spectral Galerkin method

The following examples are based on Example 3.8 and Example 8.3, 8.4 in [7]. The estimate of $\|L_N\|$ relies on the inverse inequalities on the finite dimensional trigonometric or polynomial spaces, which we refer to [11] for details. For interested readers, we also refer to the same books [7] and [11] for a systematic setup of the spectral methods.

3.1.1. Fourier method. Consider the linear problem

$$\partial_t u = a(x,t)\partial_x u + b(x,t)u, \qquad t \in (0,T),\ x \in (0,2\pi), \tag{20}$$

with periodic boundary conditions. We assume that the coefficients $a(x,t)$ and $b(x,t)$ are smooth in both x and t, and are periodic with respect to x.

Under these circumstances, $L = a(x,t)\partial_x + b(x,t)I$. Then for any smooth and periodic functions $v(x)$ and $w(x)$,

$$\begin{aligned}(L^\top v, w) = (v, Lw) &= (v, (a\partial_x + b)w) = (-\partial_x(av) + bv, w) \\ &= ((-\partial_x a + 2b)\, v, w) - (Lv, w).\end{aligned}$$

Hence $L^\top = (-\partial_x a + 2b)I - L$. Here $(\cdot,\cdot)$ is specified as the L^2 inner product on $[0, 2\pi]$.

In the Fourier Galerkin method, we set $\mathcal{V} = \text{span}\{\sin(kx), \cos(kx)\}_{k=0}^N$, and choose $(\cdot,\cdot)$ to be the same L^2 inner product. Noting that P is self-adjoint, one has

$$L_N^\top = PL^\top P = P((-(\partial_x a) + 2b)I - L)P = P(-(\partial_x a) + 2b)P - L_N.$$

Hence

$$L_N + L_N^\top = P(-(\partial_x a) + 2b)P \leq (-(\partial_x a) + 2b)I \leq (\sup_{x,t} |\partial_x a| + 2\sup_{x,t} |b|)I,$$

and L_N is semi-bounded.

On the other hand, by using the inverse inequality for trigonometric functions $\|\phi'\| \leq N\|\phi\|$, $\forall \phi \in \mathcal{V}$ and the fact $\|P\| \leq 1$, one has

$$\|L_N\| \leq \|a(\cdot,t)\partial_x + b(\cdot,t)I\| \leq \sup_{x,t} |a|N + \sup_{x,t} |b|,$$

and

$$\begin{aligned}\|L_N(t_2) - L_N(t_1)\| &\leq \|(a(\cdot,t_2) - a(\cdot,t_1))\partial_x + (b(\cdot,t_2) - b(\cdot,t_1))I\| \\ &\leq (\sup_{x,t} |\partial_t a|N + \sup_{x,t} |\partial_t b|)|t_2 - t_1|.\end{aligned}$$

Hence, the Lipschitz condition holds for $\eta = \max\{\sup_{x,t} |a|N + \sup_{x,t} |b|, \sup_{x,t} |\partial_t a|N + \sup_{x,t} |\partial_t b|\}$. We obtain the following corollary of Theorem 2.3.

Corollary 3.1. *Consider the Fourier Galerkin approximation of the linear problem (20) with the fourth-order RK time discretization,*

$$u_N^{n+1} = R_\tau(t^n)u_N^n.$$

Here $R_\tau(t^n)$ is defined in (16), and $L_N = P(a(x,t)\partial_x + b(x,t)I)P$. This fully-discrete scheme is stable,

$$\|u_N^n\| \leq K(t^n)\|u_N^0\|,$$

under the CFL condition $\tau(\max\{\sup_{x,t} |a|N + \sup_{x,t} |b|, \sup_{x,t} |\partial_t a|N + \sup_{x,t} |\partial_t b|\} + \frac{1}{2}\sup_{x,t} |\partial_x a| + \sup_{x,t} |b|) \leq c_0$, for some constant c_0.

Remark 3.1. *When a and b are constant, $L_N^\top = 2bI - L_N$. Then $L_N^\top$ commutes with L_N and L_N is normal. Furthermore, if $b \leq 0$, $L_N^\top + L_N = 2bI \leq 0$, L_N is also semi-negative. One can use Corollary 2.2 to prove the* *one-step strong stability under a sufficiently small time step. In particular, when a is a constant and $b = 0$, L_N is skew-symmetric, we can get a specific estimate of the time step size from Corollary 2.1.*

3.1.2. Polynomial method. Consider the linear advection equation

$$\partial_t u = \partial_x u, \qquad t \in (0,T),\ x \in (-1,1), \tag{21}$$

with the inflow boundary condition $u(1,t) = 0$.

In the polynomial Galerkin method, we set $\mathcal{V} = \{\phi(x) \in \mathrm{span}\{x^k\}_{k=0}^N | \phi(1) = 0\}$ and $(\cdot,\cdot)_w$ is the weighted L^2 inner product, namely $(f,g)_w = \int_{-1}^1 w(x)f(x)g(x)dx$. The corresponding norm is denoted as $\|\cdot\|_w$. We denote by P_w the projection to $\mathcal{V}$ under $(\cdot,\cdot)_w$.

Different weight functions $w(x)$ correspond to different polynomial methods in the literature. Let us consider the Jacobi method as an example. Here $w(x) = (1+x)^\alpha(1-x)^\beta$ with $\alpha \ge 0$ and $-1 < \beta \le 0$. When $\alpha = \beta = 0$, $w(x) = 1$, and the method is also referred as the Legendre method.

It can be shown that L_N is semi-negative in this setting. For any $u_N \in \mathcal{V}$, one has

$$\begin{aligned}(u_N,(L_N + L_N^\top)u_N)_w &= (P_w u_N, LP_w u_N)_w + (LP_w u_N, P_w u_N)_w \\ &= 2(u_N, Lu_N)_w = 2\int_{-1}^1 w u_N \partial_x u_N dx = \int_{-1}^1 w\partial_x u_N^2 dx \\ &= -\int_{-1}^1 w' u_N^2 dx - w(-1)u_N(-1,t)^2 \le -\int_{-1}^1 w' u_N^2 dx.\end{aligned}$$

For $\alpha > 0$ and $\beta < 0$,

$$\begin{aligned}w'(x) &= \alpha(1+x)^{\alpha-1}(1-x)^\beta - \beta(1+x)^\alpha(1-x)^{\beta-1} \\ &= (1+x)^{\alpha-1}(1-x)^{\beta-1}(\alpha(1-x) - \beta(1+x)) \ge 0.\end{aligned}$$

Similarly, one can prove $w'(x) \ge 0$ for other cases. Hence $L_N + L_N^\top \le 0$.

Furthermore, we apply the inverse inequality for Jacobi polynomials to obtain $\|L_N\|_w \le CN^2$ for some constant C. (Refer to Theorem 3.34 in [11].) Therefore, by applying Theorem 2.1, one has the following corollary.

Corollary 3.2. *Consider the Jacobi Galerkin approximation of the advection equation (21) with the fourth-order RK time discretization,*

$$u_N^{n+1} = P_4(\tau L_N)u_N^n, \qquad L_N = P_w(\partial_x)P_w.$$

This fully-discrete scheme is strongly stable in two steps,

$$\|u_N^{n+2}\|_w \leq \|u_N^n\|_w,$$

under the CFL condition $\tau \leq CN^{-2}$, *where* C *depends on the constant in the inverse inequality.*

3.2. Discontinuous Galerkin method

A detailed discussion of the example in Section 3.2.1 can be found in [13]. For a systematic setup of the discontinuous Galerkin method for solving conservation laws, one can refer to a series of paper [4], [3], [2], [1] and [5] by Cockburn et al. To be consistent with the past literatures, we switch the subscript N to h in this section, where h corresponds to the mesh size. Also, we will use the bold fonts for vectors and matrices in this section.

3.2.1. Multi-dimensional scalar equation. Consider the linear scalar conservation law

$$\partial_t u(\boldsymbol{x}, t) = \nabla \cdot (\boldsymbol{\beta}(\boldsymbol{x}) u(\boldsymbol{x}, t)), \qquad \boldsymbol{x} \in \Omega \subset \mathbb{R}^d, t \in (0, T), \tag{22}$$

with the periodic boundary conditions. Here $\boldsymbol{\beta}$ is a smooth function satisfying the divergence-free condition, $\nabla \cdot \boldsymbol{\beta}(\boldsymbol{x}) = 0$.

Suppose $\mathcal{K} = \{K\}$ is a quasi-uniform partition of the domain Ω. We denote by h the largest diameter of the elements. The collection of cell interfaces is denoted by $\mathcal{E}$. Let us define $\mathcal{V} = \{v \in L^2(\Omega) : v|_K \in \mathcal{P}^p(K), \forall K \in \mathcal{K}\}$, where $\mathcal{P}^p(K)$ is the space of polynomials of degree no more than p on K. As for $(\cdot, \cdot)$, we use the L^2 inner product on Ω.

In discontinuous Galerkin method, one seeks a solution satisfying

$$\int_\Omega \partial_t u_h v_h dx = \sum_K (-\int_K u_h \nabla \cdot (\boldsymbol{\beta} v_h) dx + \int_{\partial K} \widehat{u}_h v_h \boldsymbol{\beta} \cdot \boldsymbol{n} dl), \qquad \forall v_h \in \mathcal{V},$$

where $\widehat{u}_h$ is the numerical flux, approximating the trace of u along the edges. A popular choice is to obey the unwinding principle. In our case, that is $\widehat{u}_h = u_h^+$, where $u_h^+ = \lim_{\varepsilon \to 0^+} u_h(\boldsymbol{x} + \varepsilon \boldsymbol{\beta})$. We also use $u_h^- = \lim_{\varepsilon \to 0^+} u_h(\boldsymbol{x} - \varepsilon \boldsymbol{\beta})$ to represent the trace of u_h from the downwind side. When $\boldsymbol{\beta}(\boldsymbol{x})$ is parallel to the cell interfaces, $u_h^\pm$ are not well-defined. But in this case, $\boldsymbol{\beta}(\boldsymbol{x}) \cdot \boldsymbol{n} = 0$, hence the value of $u_h^\pm$ will not make any difference.

Let us introduce the short hand notation,

$$\mathcal{H}_{\boldsymbol{\beta}}^{+}(w,v) = \sum_{K\in\mathcal{K}} \Big(\int_K w\nabla\cdot(\boldsymbol{\beta}v)dx - \int_{\partial K} w^{+}v\boldsymbol{\beta}\cdot\boldsymbol{n}dl \Big). \tag{23}$$

Then the scheme becomes, find $u_h \in \mathcal{V}$, such that

$$(\partial_t u_h, v_h) = -\mathcal{H}_{\boldsymbol{\beta}}^{+}(u_h, v_h), \qquad \forall v_h \in \mathcal{V}.$$

$\mathcal{H}_{\boldsymbol{\beta}}^{+}$ has the following properties. We refer to [13] for details.

Lemma 3.1. *For any $w_h, v_h \in \mathcal{V}$, we have*
(1) $\mathcal{H}_{\boldsymbol{\beta}}^{+}(v_h,v_h) = \frac{1}{2}[\![v_h]\!]_{\boldsymbol{\beta}}^2$, where $[\![v_h]\!]_{\boldsymbol{\beta}} = \sqrt{\sum_{e\in\mathcal{E}}\int_e (v_h^{+}-v_h^{-})^2|\boldsymbol{\beta}\cdot\boldsymbol{n}|dl}$.
(2) $|\mathcal{H}_{\boldsymbol{\beta}}^{+}(w_h,v_h)| \le Ch^{-1}\|w_h\|\|v_h\|$, for some constant C depending on $\boldsymbol{\beta}$ and the constant in the inverse estimate.

By defining

$$(L_h u_h, v_h) = -\mathcal{H}_{\boldsymbol{\beta}}^{+}(u_h, v_h), \tag{24}$$

we can also rewrite the scheme as

$$\partial_t u_h = L_h u_h.$$

Using Lemma 3.1, one has

$$(v_h, (L_h + L_h^{\top})v_h) = 2(L_h v_h, v_h) = -2\mathcal{H}_{\boldsymbol{\beta}}^{+}(v_h,v_h) = -[\![v_h]\!]_{\boldsymbol{\beta}}^2 \le 0, \tag{25}$$

$$\|L_h v_h\|^2 = -\mathcal{H}_{\boldsymbol{\beta}}^{+}(v_h, L_h v_h) \le Ch^{-1}\|v_h\|\|L_h v_h\|. \tag{26}$$

Hence L_h is semi-negative and $\|L_h\| \le Ch^{-1}$. The stability of the fully discretized scheme now follows as a corollary of Theorem 2.1.

Corollary 3.3. *Consider the discontinuous Galerkin approximation of (22) with the fourth-order RK time discretization,*

$$u_h^{n+1} = P_4(\tau L_h)u_h^n, \qquad L_h \text{ defined in (23) and (24)}.$$

This fully-discrete scheme is strongly stable in two steps,

$$\|u_h^{n+2}\| \le \|u_h^n\|,$$

under the CFL condition $\tau \le Ch$, where C depends on $\boldsymbol{\beta}$ and the constant in the inverse estimate.

3.2.2. Multi-dimensional system. We now consider the symmetric hyperbolic system

$$\partial_t \boldsymbol{u}(\boldsymbol{x},t) = \sum_{i=1}^{d} \boldsymbol{A}_i \partial_{x_i} \boldsymbol{u}(\boldsymbol{x},t),$$
$$\boldsymbol{x} = (x_1, ..., x_d) \in \Omega \subset \mathbb{R}^d, \qquad \boldsymbol{u} = (u^1, ..., u^m)^T \in \mathbb{R}^m, \tag{27}$$

where $\boldsymbol{A}_i$ are $m \times m$ constant real symmetric matrices. For simplicity, we assume Ω to be a hypercube in $\mathbb{R}^d$ and apply the periodic boundary conditions.

Again, we denote by $\mathcal{K} = \{K\}$ a quasi-uniform partition of the domain Ω with the mesh size h. And $\mathcal{E}$ is the collection of the cell interfaces. The space $\mathcal{V}$ is chosen as

$$\mathcal{V} = \{\boldsymbol{v} \in [L^2(\Omega)]^m | \boldsymbol{v} = (v^1, ..., v^m)^T, v^k|_K \in \mathcal{P}^p(K), \forall K \in \mathcal{K}\},$$

with the inner product $(\boldsymbol{w}, \boldsymbol{v}) = \sum_{K\in\mathcal{K}} (\boldsymbol{w}, \boldsymbol{v})_K = \sum_{K\in\mathcal{K}} \int_K \boldsymbol{w} \cdot \boldsymbol{v} dx$ and the induced norm $\|\cdot\| = \sqrt{(\cdot,\cdot)}$. We also use $\langle \boldsymbol{w}, \boldsymbol{v}\rangle_e = \int_e \boldsymbol{w} \cdot \boldsymbol{v} dl$ for the integration along the cell interfaces.

To set up the discontinuous Galerkin approximation, we firstly write down the weak formulation of (27)

$$(\partial_t \boldsymbol{u}, \boldsymbol{v})_K = -\sum_{i=1}^{d} (\boldsymbol{A}_i \boldsymbol{u}, \partial_{x_i} \boldsymbol{v})_K + \sum_{e\in\partial K} \langle (\sum_{i=1}^{d} n_{e,K}^i \boldsymbol{A}_i) \boldsymbol{u}, \boldsymbol{v}\rangle_e, \tag{28}$$

where $\boldsymbol{n}_{e,K} = (n_{e,K}^1, ..., n_{e,K}^d)$ is the outward unit normal vector to the edge e in K. Since $\boldsymbol{A}_i$ are symmetric, for each e there is an orthogonal matrix $\boldsymbol{S}_e$ such that $\boldsymbol{\Lambda}_e = \mathrm{diag}(\lambda_e^1, ..., \lambda_e^m) = \boldsymbol{S}_e (\sum_{i=1}^d n_{e,K}^i \boldsymbol{A}_i) \boldsymbol{S}_e^T$ is diagonal. The numerical scheme can be obtained by applying the upwind flux for each eigen-component, namely

$$(\partial_t \boldsymbol{u}_h, \boldsymbol{v}_h)_K = -\sum_{i=1}^{d} (\boldsymbol{A}_i \boldsymbol{u}_h, \partial_{x_i} \boldsymbol{v}_h)_K + \sum_{e\in\partial K} \langle \boldsymbol{\Lambda}_e \widehat{\boldsymbol{S}_e \boldsymbol{u}_h}, \boldsymbol{S}_e \boldsymbol{v}_h\rangle_e, \tag{29}$$

where $\widehat{\boldsymbol{S}_e \boldsymbol{u}_h} = ((\widehat{\boldsymbol{S}_e \boldsymbol{u}_h})^1, ..., (\widehat{\boldsymbol{S}_e \boldsymbol{u}_h})^m)^T$ and

$$(\widehat{\boldsymbol{S}_e \boldsymbol{u}_h})^k = \begin{cases} (\boldsymbol{S}_e \boldsymbol{u}_h^{int})^k, & \lambda_e^k < 0, \\ (\boldsymbol{S}_e \boldsymbol{u}_h^{ext})^k, & \lambda_e^k > 0. \end{cases} \tag{30}$$

Here, $\boldsymbol{u}_h^{int}(\boldsymbol{x},t) = \lim_{\boldsymbol{y}\to\boldsymbol{x},\boldsymbol{y}\in K} \boldsymbol{u}_h(\boldsymbol{x},t)$ and $\boldsymbol{u}_h^{ext}(\boldsymbol{x},t) = \lim_{\boldsymbol{y}\to\boldsymbol{x},\boldsymbol{y}\in K^c} \boldsymbol{u}_h(\boldsymbol{x},t)$. When $\lambda_e^k = 0$, $(\widehat{\boldsymbol{S}_e\boldsymbol{u}_h})^k$ will not contribute to the integration, hence one can avoid defining it in this case. Also note that $\sum_{i=1}^d n_{e,K}^i \boldsymbol{A}_i = -\sum_{i=1}^d n_{e,K'}^i \boldsymbol{A}_i$ for $K' \cap K = e$, and the same $\boldsymbol{S}_e$ should be used on both sides. Then $\widehat{\boldsymbol{S}_e\boldsymbol{u}_h}$ defined in K and K' are the same, and the numerical flux is single-valued on the cell interfaces. Furthermore, for the scalar case, (30) coincides with the previous definition in Section 3.2.1.

As before, we define

$$\begin{aligned} \mathcal{H}(\boldsymbol{w}_h,\boldsymbol{v}_h) &= \sum_{K\in\mathcal{K}} \mathcal{H}_K(\boldsymbol{w}_h,\boldsymbol{v}_h) \\ &= \sum_{K\in\mathcal{K}} \left(\sum_{i=1}^d (\boldsymbol{A}_i\boldsymbol{w}_h, \partial_{x_i}\boldsymbol{v}_h)_K - \sum_{e\in\partial K} \langle \boldsymbol{\Lambda}_e \widehat{\boldsymbol{S}_e\boldsymbol{w}_h}, \boldsymbol{S}_e\boldsymbol{v}_h\rangle_e \right). \end{aligned} \tag{31}$$

$\mathcal{H}$ has the following properties, and its proof can be found in Appendix C.

Lemma 3.2. *$\forall \boldsymbol{v}_h, \boldsymbol{w}_h \in \mathcal{V}$, $\mathcal{H}(\boldsymbol{v}_h,\boldsymbol{v}_h) \geq 0$ and $|\mathcal{H}(\boldsymbol{w}_h,\boldsymbol{v}_h)| \leq Ch^{-1}\|\boldsymbol{w}_h\|\|\boldsymbol{v}_h\|$, where C depends on $\boldsymbol{A}_i$ and the constant in the inverse estimate.*

With the definition of $\mathcal{H}$, the scheme (29) can be written as

$$\partial_t \boldsymbol{u}_h = L_h\boldsymbol{u}_h, \qquad (L_h\boldsymbol{u}_h, \boldsymbol{v}_h) = -\mathcal{H}(\boldsymbol{u}_h,\boldsymbol{v}_h), \qquad \forall \boldsymbol{v}_h \in \mathcal{V}. \tag{32}$$

Similar to those in (25) and (26), Lemma 3.2 implies that L_h is semi-negative and $\|L_h\| \leq Ch^{-1}$. Hence, one can use Theorem 2.1 to obtain the following corollary.

Corollary 3.4. *Consider the discontinuous Galerkin approximation of (27) with the fourth-order RK time discretization,*

$$\boldsymbol{u}_h^{n+1} = P_4(\tau L_h)\boldsymbol{u}_h^n, \qquad L_h \text{ defined in (30), (31) and (32).}$$

The fully-discrete scheme is strongly stable in two steps,

$$\|\boldsymbol{u}_h^{n+2}\| \leq \|\boldsymbol{u}_h^n\|,$$

under the CFL condition $\tau \leq Ch$. Here C depends on A_i and the constant in the inverse estimate.

Remark 3.2. *The stability extends to symmetrizable hyperbolic systems with constants coefficients. More specifically, if there is a symmetric positive-definite matrix $\boldsymbol{H}$ such that $\boldsymbol{H}\boldsymbol{A}_i$ are symmetric, then under an appropriate*

CFL condition $\tau \le Ch^{-1}$, one has $\|\boldsymbol{u}_h^{n+2}\|_{\boldsymbol{H}} \le \|\boldsymbol{u}_h^n\|_{\boldsymbol{H}}$, where $\|\boldsymbol{v}\|_{\boldsymbol{H}} = \sum_{K\in\mathcal{K}} \int_K \boldsymbol{v}^T \boldsymbol{H}\boldsymbol{v} dx$.

Remark 3.3. *For symmetric hyperbolic systems with variable coefficients, one can follow similar lines to show L_h is semi-bounded, and prove the stability of the fully discretized scheme. But the extension to symmetrizable hyperbolic systems with variable coefficients is non-trivial.*

4. Concluding remarks

We analyze the stability of the fourth order RK method for integrating method of lines schemes solving the well-posed linear PDE system $\partial_t u = L(x,t,\partial_x)u$. The issue of strong stability is of special interests. We consider the ODE system $\frac{d}{dt}u_N = L_N u_N$, with $L_N + L_N^\top \le 0$ and L_N independent of time. When L_N is normal, the strong stability of the fourth order RK approximation has already been justified by the scalar eigenvalue analysis. But for non-normal L_N, we provide a counter example to show that the strong stability can not be preserved whatever small time step we choose. However, the strong stability can actually be obtained in two steps. We prove $\|u_N^{n+2}\| \le \|u_N^n\|$ under the time step constraint $\tau\|L_N\| \le c_0$ for some constant c_0. This can also be interpreted as the strong stability of the eight-stage fourth order RK method composed by two steps of the four-stage method. Then, based on this fact, we extend the stability results to general semi-bounded linear systems after using a perturbation analysis and a frozen-coefficient argument. Finally, we apply the results to justify the stability of the fully discretized schemes combining the fourth order RK method and different spatial discretizations, including the spectral Galerkin method and the discontinuous Galerkin method. The corresponding CFL time step restrictions are obtained.

Appendix A. Proof of Proposition 1.1

Proof. (1) To prove $L_N + L_N^T \le 0$.

$$L_N + L_N^T = -\begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix} = -2\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 \end{pmatrix}.$$

Hence for any $u = (u_1, u_2, u_3)^T \in \mathbb{R}^3$,

$$u^T(L_N + L_N^T)u = -2(u_1 + u_2 + u_3)^2 \le 0.$$

(2) To prove $\|P_4(\tau L_N)\| > 1$.

By definition,

$$P_4(\tau L_N)$$
$$= \begin{pmatrix} 1-\tau+\frac{\tau^2}{2}-\frac{\tau^3}{6}+\frac{\tau^4}{24} & -2\tau+2\tau^2-\tau^3+\frac{\tau^4}{3} & -2\tau+4\tau^2-3\tau^3+\frac{4\tau^4}{3} \\ 0 & 1-\tau+\frac{\tau^2}{2}-\frac{\tau^3}{6}+\frac{\tau^4}{24} & -2\tau+2\tau^2-\tau^3+\frac{\tau^4}{3} \\ 0 & 0 & 1-\tau+\frac{\tau^2}{2}-\frac{\tau^3}{6}+\frac{\tau^4}{24} \end{pmatrix}.$$

Since $\|P_4(\tau L_N)\| = \sqrt{\lambda_{\max}(P_4(\tau L_N)^T P_4(\tau L_N))}$ is the square root of the largest eigenvalue of $P_4(\tau L_N)^T P_4(\tau L_N)$, we define $f(\lambda,\tau) = \det(\lambda I - P_4(\tau L_N)^T P_4(\tau L_N))$. It suffices to show, for any sufficiently small τ, $f(\cdot,\tau)$ has a root larger than 1.

After direct calculation, one can obtain

$$f(\lambda,\tau) = \lambda^3 + c_2\lambda^2 + c_1\lambda + c_0,$$

where

$$c_0 = -\left(1-\tau+\frac{1}{2}\tau^2-\frac{1}{6}\tau^3+\frac{1}{24}\tau^4\right)^6,$$
$$\begin{aligned} c_1 = 3 - 12\tau + 36\tau^2 - 72\tau^3 + 100\tau^4 - \frac{211\tau^5}{2} + \frac{1069\tau^6}{12} \\ - \frac{257\tau^7}{4} + \frac{4081\tau^8}{96} - \frac{3827\tau^9}{144} + \frac{2225\tau^{10}}{144} - \frac{1133\tau^{11}}{144} \\ + \frac{1895\tau^{12}}{576} - \frac{3709\tau^{13}}{3456} + \frac{197\tau^{14}}{768} - \frac{851\tau^{15}}{20736} + \frac{385\tau^{16}}{110592}, \end{aligned}$$
$$c_2 = -3 + 6\tau - 18\tau^2 + 36\tau^3 - 46\tau^4 + \frac{163\tau^5}{4} - \frac{589\tau^6}{24} + \frac{75\tau^7}{8} - \frac{385\tau^8}{192}.$$

One can see

$$\begin{aligned} f(1,\tau) = -\frac{8\tau^9}{27} + \frac{\tau^{10}}{3} - \frac{\tau^{11}}{9} - \frac{59\tau^{12}}{216} + \frac{341\tau^{13}}{864} - \frac{247\tau^{14}}{864} + \frac{1435\tau^{15}}{10368} \\ - \frac{19\tau^{16}}{384} + \frac{2299\tau^{17}}{165888} - \frac{4757\tau^{18}}{1492992} + \frac{79\tau^{19}}{124416} - \frac{107\tau^{20}}{995328} + \frac{179\tau^{21}}{11943936} \\ - \frac{13\tau^{22}}{7962624} + \frac{\tau^{23}}{7962624} - \frac{\tau^{24}}{191102976}, \end{aligned}$$
$$\begin{aligned} f(2,\tau) = 1 + 6\tau - 18\tau^2 + 36\tau^3 - 38\tau^4 + \frac{67\tau^5}{4} + \frac{371\tau^6}{24} - \frac{289\tau^7}{8} \\ + \frac{7007\tau^8}{192} - \frac{11609\tau^9}{432} + \frac{2273\tau^{10}}{144} - \frac{383\tau^{11}}{48} + \frac{5213\tau^{12}}{1728} - \frac{2345\tau^{13}}{3456} \end{aligned}$$

$$-\frac{203\tau^{14}}{6912}+\frac{673\tau^{15}}{6912}-\frac{5087\tau^{16}}{110592}+\frac{2299\tau^{17}}{165888}-\frac{4757\tau^{18}}{1492992}+\frac{79\tau^{19}}{124416}$$
$$-\frac{107\tau^{20}}{995328}+\frac{179\tau^{21}}{11943936}-\frac{13\tau^{22}}{7962624}+\frac{\tau^{23}}{7962624}-\frac{\tau^{24}}{191102976}.$$

Hence, for any sufficiently small τ,

$$f(1,\tau)=-\frac{8\tau^8}{27}+\mathcal{O}(\tau^9)<0, \qquad f(2,\tau)=1+6\tau+\mathcal{O}(\tau^2)>1.$$

Note $f(\lambda,\tau)$ is a polynomial with respect to λ and τ, hence it is continuous. By the intermediate value theorem, there exists $\lambda(\tau)\in(1,2)$ such that $f(\lambda(\tau),\tau)=0$. Therefore

$$\|P_4(\tau L_N)\|\geq\sqrt{\lambda(\tau)}>1. \qquad \square$$

Remark A.1. *We also provide the numerical examination of Proposition 1.1. The values of $\|P_4(\tau L_N)\|-1$ and $\|P_4(\tau L_N)^2\|-1$ with a decreasing sequence of τ are listed in Table 1. As we can see, for this specific case, the one-step strong stability does fail, but the two-step strong stability holds.*

Table 1: The numerical examination of Proposition 1.1 and Theorem 2.1

τ	$\|P_4(\tau L_N)\|-1$	$\|P_4(\tau L_N)^2\|-1$
0.5	1.2794e-3	-1.0428e-3
0.2	8.3857e-6	-1.8788e-5
0.1	2.2173e-7	-6.6719e-7
0.05	6.3437e-9	-2.2029e-8
0.02	6.1504e-11	-2.3254e-10
0.01	1.8868e-12	-7.3375e-12

Appendix B. Coefficient matrix $\tilde{A}$

$$\tilde{A}=-\begin{pmatrix} 1 & 3/2 & 7/6 & 5/8 & 1/4 & 5/72 & 1/72 & 1/576 \\ 3/2 & 7/3 & 15/8 & 25/24 & 31/72 & 1/8 & 5/192 & 1/288 \\ 7/6 & 15/8 & 37/24 & 127/144 & 3/8 & 65/576 & 7/288 & 1/288 \\ 5/8 & 25/24 & 127/144 & 25/48 & 131/576 & 61/864 & 1/64 & 1/432 \\ 1/4 & 31/72 & 3/8 & 131/576 & 175/1728 & 37/1152 & 25/3456 & 5/4608 \\ 5/72 & 1/8 & 65/576 & 61/864 & 37/1152 & 1/96 & 11/4608 & 5/13824 \\ 1/72 & 5/192 & 7/288 & 1/64 & 25/3456 & 11/4608 & 23/41472 & 7/82944 \\ 1/576 & 1/288 & 1/288 & 1/432 & 5/4608 & 5/13824 & 7/82944 & 1/82944 \end{pmatrix}.$$

Appendix C. Proof of Lemma 3.2

Proof. (1) To prove $\mathcal{H}(\boldsymbol{v}_h, \boldsymbol{v}_h) \geq 0$.

We denote by $\boldsymbol{v}_{e,h} = \boldsymbol{S}_e \boldsymbol{v}_h$ and $\widehat{\boldsymbol{v}}_{e,h} = \widehat{\boldsymbol{S}_e \boldsymbol{v}_h}$. Since $\boldsymbol{A}_i$ are symmetric. Integrating by parts, one has

(33)
$$\sum_{i=1}^{d} (\boldsymbol{A}_i \boldsymbol{v}_h, \partial_{x_i} \boldsymbol{v}_h)_K = \sum_{e \in \partial K} \langle \frac{1}{2} \sum_{i=1}^{d} n_{e,K}^i \boldsymbol{A}_i \boldsymbol{v}_h, \boldsymbol{v}_h \rangle_e = \sum_{e \in \partial K} \langle \frac{1}{2} \boldsymbol{\Lambda}_e \boldsymbol{v}_{e,h}, \boldsymbol{v}_{e,h} \rangle_e.$$

Substitute (33) into (31), then we obtain

$$\mathcal{H}(\boldsymbol{v}_h, \boldsymbol{v}_h) \;=\; \sum_{K \in \mathcal{K}} \sum_{e \in \partial K} \langle \boldsymbol{\Lambda}_e (\frac{1}{2} \boldsymbol{v}_{e,h} - \widehat{\boldsymbol{v}}_{e,h}), \boldsymbol{v}_{e,h} \rangle_e.$$

Each cell interface e is shared by two elements. We sit in either side and call $\boldsymbol{v}_{e,h}$ defined in this element to be $\boldsymbol{v}_{e,h}^{int}$. $\boldsymbol{v}_{e,h}$ defined from the other side is referred as $\boldsymbol{v}_{e,h}^{ext}$. Then the summation can be rewritten as

$$\begin{aligned} &\mathcal{H}(\boldsymbol{v}_h, \boldsymbol{v}_h) \\ &= \sum_{e \in \mathcal{E}} \left(\langle \boldsymbol{\Lambda}_e (\frac{1}{2} \boldsymbol{v}_{e,h}^{int} - \widehat{\boldsymbol{v}}_{e,h}), \boldsymbol{v}_{e,h}^{int} \rangle_e + \langle -\boldsymbol{\Lambda}_e (\frac{1}{2} \boldsymbol{v}_{e,h}^{ext} - \widehat{\boldsymbol{v}}_{e,h}), \boldsymbol{v}_{e,h}^{ext} \rangle_e \right) \\ &= \sum_{e \in \mathcal{E}} (\frac{1}{2} \langle \boldsymbol{\Lambda}_e \boldsymbol{v}_{e,h}^{int}, \boldsymbol{v}_{e,h}^{int} \rangle_e - \frac{1}{2} \langle \boldsymbol{\Lambda}_e \boldsymbol{v}_{e,h}^{ext}, \boldsymbol{v}_{e,h}^{ext} \rangle_e + \langle \boldsymbol{\Lambda}_e \widehat{\boldsymbol{v}}_{e,h}, \boldsymbol{v}_{e,h}^{ext} - \boldsymbol{v}_{e,h}^{int} \rangle_e). \end{aligned}$$

By checking the definition of $\widehat{\boldsymbol{v}}_{e,h}$, one can prove that

$$\mathcal{H}(\boldsymbol{v}_h, \boldsymbol{v}_h) = \frac{1}{2} \sum_{e \in \mathcal{E}} \langle \text{abs}(\boldsymbol{\Lambda}_e)(\boldsymbol{v}_{e,h}^{ext} - \boldsymbol{v}_{e,h}^{int}), (\boldsymbol{v}_{e,h}^{ext} - \boldsymbol{v}_{e,h}^{int}) \rangle_e,$$

where $\text{abs}(\boldsymbol{\Lambda}_e) = \text{diag}(|\lambda_e^1|, ..., |\lambda_e^m|)$. Hence $\mathcal{H}(\boldsymbol{v}_h, \boldsymbol{v}_h)$ is non-negative.

(2) To prove $|\mathcal{H}(\boldsymbol{w}_h, \boldsymbol{v}_h)| \leq C h^{-1} \|\boldsymbol{w}_h\| \|\boldsymbol{v}_h\|$.

We will use the notation $|\cdot|$ for different meanings. For scalars, $|a|$ is the absolute value of a; for vectors, $|\boldsymbol{a}|$ stands for the Euclidean norm of $\boldsymbol{a}$; and for matrices, $|\boldsymbol{A}|$ is the operator norm of $\boldsymbol{A}$. We also use the notation $\|\cdot\|_K = \sqrt{(\cdot,\cdot)_K}$.

By using the Cauchy-Schwarz inequality and the inverse estimate, one has

$$
\begin{aligned}
|\sum_{K\in\mathcal{K}}\sum_{i=1}^{d}(\boldsymbol{A}_i\boldsymbol{w}_h,\partial_{x_i}\boldsymbol{v}_h)_K| &\leq \sum_{K\in\mathcal{K}}\sum_{i=1}^{d}|\boldsymbol{A}_i|\|\boldsymbol{w}_h\|_K\|\partial_{x_i}\boldsymbol{v}_h\|_K \\
&\leq \left(Ch^{-1}\sum_{i=1}^{d}|\boldsymbol{A}_i|\right)\sum_{K\in\mathcal{K}}\|\boldsymbol{w}_h\|_K\|\boldsymbol{v}_h\|_K \\
&\leq \left(Ch^{-1}\sum_{i=1}^{d}|\boldsymbol{A}_i|\right)\|\boldsymbol{w}_h\|\|\boldsymbol{v}_h\|.
\end{aligned}
\tag{34}
$$

And

$$
\begin{aligned}
&|\sum_{K\in\mathcal{K}}\sum_{e\in\partial K}\langle\boldsymbol{\Lambda}_e\widehat{\boldsymbol{S}_e\boldsymbol{w}_h},\boldsymbol{S}_e\boldsymbol{v}_h\rangle_e| \\
&\qquad\leq \sup_{k,e}|\lambda_e^k|\sum_{K\in\mathcal{K}}\sum_{e\in\partial K}\sqrt{\int_e|\widehat{\boldsymbol{S}_e\boldsymbol{w}_h}|^2dl}\sqrt{\int_e|\boldsymbol{S}_e\boldsymbol{v}_h|^2dl}.
\end{aligned}
$$

Note that $\boldsymbol{S}_e$ are orthogonal, $|\boldsymbol{S}_e\boldsymbol{v}_h| = |\boldsymbol{v}_h|$. By using the inverse inequality, we obtain

$$
\sqrt{\int_e|\boldsymbol{S}_e\boldsymbol{v}_h|^2dl} = \sqrt{\int_e|\boldsymbol{v}_h|^2dl} \leq Ch^{-\frac{1}{2}}\|\boldsymbol{v}_h\|_K,
$$

and

$$
\sqrt{\int_e|\widehat{\boldsymbol{S}_e\boldsymbol{w}_h}|^2dl} \leq \sqrt{\int_e|\boldsymbol{w}_h^{int}|^2+|\boldsymbol{w}_h^{ext}|^2dl} \leq Ch^{-\frac{1}{2}}\|\boldsymbol{w}_h\|_{\mathcal{N}(K)},
$$

where $\mathcal{N}(K)$ is the union of K and its neighboring elements. Hence

$$
\begin{aligned}
|\sum_{K\in\mathcal{K}}\sum_{e\in\partial K}\langle\boldsymbol{\Lambda}_e\widehat{\boldsymbol{S}_e\boldsymbol{w}_h},\boldsymbol{S}_e\boldsymbol{v}_h\rangle_e| \leq& Ch^{-1}\sup_{k,e}|\lambda_e^k|\sum_{K\in\mathcal{K}}\|\boldsymbol{w}_h\|_{\mathcal{N}(K)}\|\boldsymbol{v}_h\|_K \\
\leq& Ch^{-1}\sup_{k,e}|\lambda_e^k|\sqrt{\sum_{K\in\mathcal{K}}\|\boldsymbol{w}_h\|^2_{\mathcal{N}(K)}}\sqrt{\sum_{K\in\mathcal{K}}\|\boldsymbol{v}_h\|^2_K} \\
\leq& Ch^{-1}\sup_{k,e}|\lambda_e^k|\|\boldsymbol{w}_h\|\|\boldsymbol{v}_h\|.
\end{aligned}
\tag{35}
$$

Combining (34) and (35), we get

$$
|\mathcal{H}(\boldsymbol{w}_h,\boldsymbol{v}_h)| \leq Ch^{-1}\|\boldsymbol{w}_h\|\|\boldsymbol{v}_h\|. \qquad \square
$$

References

[1] B. Cockburn, S. Hou, and C.-W. Shu. The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws IV: the multidimensional case. *Mathematics of Computation*, 54:545–581, 1990. MR1010597

[2] B. Cockburn, S.-Y. Lin, and C.-W. Shu. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws III: one-dimensional systems. *Journal of Computational Physics*, 84:90–113, 1989. MR1015355

[3] B. Cockburn and C.-W. Shu. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws II: general framework. *Mathematics of Computation*, 52:411–435, 1989. MR0983311

[4] B. Cockburn and C.-W. Shu. The Runge-Kutta local projection P^1-discontinuous-Galerkin finite element method for scalar conservation laws. *RAIRO-Modélisation Mathématique et Analyse Numérique*, 25:337–361, 1991. MR1103092

[5] B. Cockburn and C.-W. Shu. The Runge-Kutta discontinuous Galerkin method for conservation laws V: multidimensional systems. *Journal of Computational Physics*, 141:199–224, 1998. MR1619652

[6] S. Gottlieb, C.-W. Shu, and E. Tadmor. Strong stability-preserving high-order time discretization methods. *SIAM Review*, 43:89–112, 2001. MR1854647

[7] J. S. Hesthaven, S. Gottlieb, and D. Gottlieb. *Spectral Methods for Time-dependent Problems.* Cambridge University Press, Cambridge, 2007. MR2333926

[8] A. Iserles. *A First Course in the Numerical Analysis of Differential Equations.* Cambridge University Press, Cambridge, 2009. MR2478556

[9] J. F. B. M. Kraaijevanger. Contractivity of Runge-Kutta methods. *BIT Numerical Mathematics*, 31:482–528, 1991. MR1127488

[10] D. Levy and E. Tadmor. From semidiscrete to fully discrete: stability of Runge-Kutta schemes by the energy method. *SIAM Review*, 40:40–73, 1998. MR1612565

[11] J. Shen, T. Tang, and L.-L. Wang. *Spectral Methods: Algorithms, Analysis and Applications.* Springer, Heidelberg, 2011. MR2867779

[12] M. N. Spijker. Contractivity in the numerical solution of initial value problems. *Numerische Mathematik*, 42:271–290, 1983. MR0723625

[13] Z. Sun and C.-W. Shu. Stability analysis and error estimates of Lax-Wendroff discontinuous Galerkin methods for linear conservation laws. *ESAIM: Mathematical Modelling and Numerical Analysis*, to appear. DOI: http://dx.doi.org/10.1051/m2an/2016049.

[14] E. Tadmor. From semidiscrete to fully discrete: stability of Runge-Kutta schemes by the energy method II. *Collected Lectures on the Preservation of Stability under Discretization*, Lecture Notes from Colorado State University Conference, Fort Collins, CO, 2001 (D. Estep and S. Tavener, eds.), Proceedings in Applied Mathematics, SIAM, 109:25–49, 2002. MR2026662

[15] G. Wanner and E. Hairer. *Solving Ordinary Differential Equations II.* Springer-Verlag, Berlin, 1991. MR1111480

Zheng Sun
Division of Applied Mathematics
Brown University
Providence, RI 02912
USA
E-mail address: zheng_sun@brown.edu

Chi-Wang Shu
Division of Applied Mathematics
Brown University
Providence, RI 02912
USA
E-mail address: shu@dam.brown.edu

Received August 12, 2016