ISSN 0915-633X

NATIONAL INSTITUTE FOR FUSION SCIENCE

Rooted Tree Analysis of Runge-Kutta Methods with
Exact Treatment of Linear Terms

S. Koikari

(Received - Jan. 13, 2004)

NIFS-790 Jan. 2004

TOKI, JAPAN

Rooted tree analysis of Runge–Kutta methods with exact treatment of linear terms

S. Koikari

Theory and Computer Simulation Center, National Institute for Fusion Science, Oroshi-cho 322-6, Toki-shi 509-5292, Japan

Abstract

We investigate a class of time discretization schemes called "ETD Runge–Kutta methods," where the linear terms of an ordinary differential equation are treated rigorously, while the other terms are numerically integrated by a one-step method. These schemes, proposed by previous authors, can be regarded as modified Runge–Kutta methods whose coefficients are matrices instead of scholars. From this viewpoint, we reexamine the notion of consistency, convergence and order to provide a mathematical foundation for new methods. Applying the rooted tree analysis, expansion theorems of both the strict and numerical solutions are proved, and two types of order conditions are defined. Several classes of formulas with up to four stages that satisfy the conditions are derived, and it is shown that the power series of matrices, employed as their coefficients, are well characterized by the requirement of the stage order.

Key words: ETD Runge–Kutta methods; Exact treatment of linear part; Rooted tree analysis; Integrating factor methods; Spectral methods

1 Introduction

When we numerically solve an advection-diffusion equation such as the Navier–Stokes equations by spectral methods [4,9], the partial differential equation, which describes the spatiotemporal structure of vector fields, is decomposed by a set of basis functions which are orthogonal and complete on the spatial domain considered. The resultant equations are simultaneous ordinary differential equations for the mode amplitudes and can be written in a general form as

$$y'(x) = \Lambda\, y(x) + f(y(x))\ , \tag{1}$$

where x is time, $y(x)$ is an N dimensional vector, the elements of which are the mode amplitudes, $\Lambda\, y(x)$ is a linear term and $f(y(x))$ is a nonlinear term.

The matrix Λ is typically the representation of a spatial operator such as the Laplacian ∇^2 and is a source of stiffness in most cases of spectral simulations. To circumvent this difficulty and enable efficient and stable calculations, the semi-implicit scheme has widely been used, where the linear term is integrated by the Crank–Nicolson formula (implicit) and the other terms are treated by the Adams–Bashforth formula (explicit). Although the semi-implicit scheme has been applied successfully to a certain range of problems, it is also known that some limitations exist. For example, schemes of this type with third and higher order accuracy have usually been avoided, since they are essentially linear multistep methods.

Recently, new methods have been proposed as viable alternatives to this conventional one by G. Beylkin et al. [3], B. Alpert et al. [1] and S. M. Cox and P. C. Matthews [10,11], where linear terms are treated rigorously without a conversion of the dependent variable. To accomplish this task, they transform the differential equation (1) to an integral equation by multiplying $e^{-\Lambda x}$ from the left and integrating it from x_n to x_{n+1}. The transformed equation is

$$y(x_{n+1}) = e^{\Lambda h}\, y(x_n) + e^{\Lambda h} \int_0^h e^{-\Lambda \tau} f(y(x_n + \tau))\, d\tau \quad (x_{n+1} := x_n + h)\,. \quad (2)$$

The concept of the new methods is that the exponential function and the integral in eq. (2) are treated rigorously as far as possible, while the function $f(y(x_n+\tau))$ is approximated by a polynomial of τ. There are two possible ways of approximating $f(y(x_n+\tau))$. The first is to use linear multistep methods and the second is to use one-step methods, i.e., Runge–Kutta methods. The former is called "exact treatment of linear part" ("ELP") [1,3] or "exponential time differencing" ("ETD") [10,11] by previous authors; therefore, we denote it as "ELP/ETD" in this paper. The latter is named "ETD Runge–Kutta" by the authors [10,11] and we adopt the same name in the present paper. One of the reasons why these new methods are promising is, in addition to the possibility of higher order schemes, that they have some affinity with the semi-implicit method. There exists no essential difference between the semi-implicit method and the two-step ELP/ETD method except the order of approximation of the exponential function. This point is clarified in Section 2.

The basic theory of ELP/ETD methods has already been established [3,10]; however, the counterpart of ETD Runge–Kutta methods does not exist in the literature yet. The main purpose of this paper is to provide a mathematical foundation of ETD Runge–Kutta methods through a theoretical investigation. As the preparation for mathematical treatment, we represent ETD Runge–Kutta methods by Butcher tableaux [7] whose elements themselves are matrices instead of scalars (Section 3). This representation enables us to clearly understand and rigorously define what the new methods are. By applying the rooted tree analysis [5,7] to the methods, we establish the expansion theorems for both the strict and numerical solutions (Section 4). Using these expansion theorems, we define two types of order condition, the strong one and the weak one. Finally, in Section 5, we derive several classes of formulas with up to four stages that fulfil the order conditions. It is shown that power series of matrices used as their weights are well characterized by additional conditions for the stage order.

2 ELP/ETD methods and the semi-implicit scheme

In this section, we show that the semi-implicit scheme can be derived from an ELP/ETD method when it is combined with Padé approximants to the exponential function.

If the function $f(y(x_n+\tau))$ is approximated by a first-order polynomial

$$f(y(x_n+\tau)) \sim f_n + \frac{f_n - f_{n-1}}{h}\tau \quad (f_n := f(y(x_n))) \ , \tag{3}$$

eq. (2) can be integrated to give

$$y(x_{n+1}) \sim Q_0(\Lambda h) y(x_n) + hQ_1(\Lambda h) f_n + hQ_2(\Lambda h)(f_n - f_{n-1}) \ , \tag{4}$$

where Q_0 , Q_1 and Q_2 are

$$Q_0(\Lambda h) := \mathrm{e}^{\Lambda h}; \quad Q_1(\Lambda h) := \frac{\mathrm{e}^{\Lambda h} - I}{\Lambda h} \quad \text{and} \quad Q_2(\Lambda h) := \frac{\mathrm{e}^{\Lambda h} - I - \Lambda h}{\Lambda^2 h^2} \ , \tag{5}$$

respectively. The approximation (4) defines a two-step ELP/ETD formula

$$y_{n+1} = Q_0(\Lambda h) y_n + hQ_1(\Lambda h) f_n + hQ_2(\Lambda h)(f_n - f_{n-1}) \ . \tag{6}$$

If the exponential function in eq. (6) is approximated by the Padé approximant with second-order accuracy [2], i.e.,

$$\mathrm{e}^{\Lambda h} = Q_0(\Lambda h) \sim \left(I - \frac{\Lambda h}{2}\right)^{-1} \left(I + \frac{\Lambda h}{2}\right) \ , \tag{7}$$

$Q_1(\Lambda h)$ and $Q_2(\Lambda h)$ are consistently approximated as

$$Q_1(\Lambda h) \sim \left(I - \frac{\Lambda h}{2}\right)^{-1} \quad \text{and} \quad Q_2(\Lambda h) \sim \frac{1}{2}\left(I - \frac{\Lambda h}{2}\right)^{-1} . \tag{8}$$

Substituting these expressions into eq. (6), we obtain the semi-implicit scheme

$$y_{n+1} = \left(I - \frac{\Lambda h}{2}\right)^{-1} \left\{ \left(I + \frac{\Lambda h}{2}\right) y_n + h\left(\frac{3}{2} f_n - \frac{1}{2} f_{n-1}\right) \right\} . \tag{9}$$

From the above derivation, we can see that the semi-implicit scheme (9) and the two-step ELP/ETD method (6) have no essential difference except the order of approximation of the exponential function $e^{\Lambda h}$.

Instead of the rigorous calculation as in ELP/ETD methods, it may be possible to approximate $e^{\Lambda h}$ by A-acceptable $2p$th order Padé approximants ($p > 1$). Such approximants can be derived from Runge–Kutta methods as their stability functions; therefore, in such a scheme, the Dahlquist second barrier does not exist as far as the linear term is concerned. It can be interpreted as a combined scheme, where the linear term is integrated by an implicit Runge–Kutta method and the nonlinear term is integrated by an explicit linear multistep method. If an explicit Runge–Kutta method is employed for the nonlinear term, a semi-implicit Runge–Kutta method is obtained. From this viewpoint, ETD Runge–Kutta methods are regarded as semi-implicit Runge–Kutta methods whose implicit part has the infinite-order accuracy.

3 Power series of matrices and Butcher tableaux

3.1 Power series of matrices

ETD Runge–Kutta formulas with up to four stages were proposed and evaluated by previous authors [10,11]. A two-stage formula with second-order accuracy is denoted by "ETD2RK" and is derived in their paper as follows. First, by substituting the 0th order approximation $f(y(x_n + \tau)) \sim f_n$ in eq. (2) and integrating it, we obtain a stage value as

$$\xi_2 = e^{\Lambda h} y_n + \frac{e^{\Lambda h} - I}{\Lambda} f_n \,. \tag{10}$$

After evaluating $f(\xi_2)$, we can construct a first-order approximation

$$f \sim f_n + \frac{x - x_n}{h}(f(\xi_2) - f_n) \,. \tag{11}$$

Then, the integration of eq. (2) with the improved approximation (11) yields

$$y_{n+1} = \xi_2 + \frac{e^{\Lambda h} - I - \Lambda h}{\Lambda^2 h}\{f(\xi_2) - f_n\}. \tag{12}$$

We use notations that are slightly different from those in their paper so as to explicitly indicate that Λ and $e^{\Lambda h}$ are matrices. A three-stage formula with third-order accuracy "ETD3RK" and a four-stage formula with fourth-order accuracy "ETD4RK" are also given.

A common feature of these formulas, i.e., ETD2RK, ETD3RK and ETD4RK, is that their weights used in calculations of the stage values $\{\xi_i\}$ and the final result y_{n+1} are linear combinations of power series of matrices,

$$Q_n(Z) := \sum_{k=0}^{\infty} \frac{Z^k}{(k+n)!} = \frac{I}{Z^n}\sum_{k=n}^{\infty}\frac{Z^k}{k!} = \frac{I}{Z^n}\left(e^Z - \sum_{k=0}^{n-1}\frac{Z^k}{k!}\right), \tag{13}$$

where Z is an arbitrary $N \times N$ matrix. This class of power series was first introduced by G. Beylkin et al. [3] in their theoretical investigation of ELP/ETD methods. $Q_0(Z)$ is e^Z and $Q_k(Z)$ for an arbitrary $k \in \mathbb{N}$ is also an absolutely convergent power series without negative powers. Since $\lim_{(Z)\to O} Q_k(Z) = 1/(k!)$, $Q_k(Z)$ can be considered to be an $O(1)$ quantity in the context of the asymptotic analysis $h \to 0$.

Using the power series defined above, ETD2RK is written as

$$\xi_1 = y_n\,, \tag{14a}$$
$$\xi_2 = Q_0(\Lambda h)\,y_n + hQ_1(\Lambda h)\,f(\xi_1)\,, \tag{14b}$$
$$y_{n+1} = Q_0(\Lambda h)\,y_n + h\{Q_1(\Lambda h) - Q_2(\Lambda h)\}f(\xi_1) + hQ_2(\Lambda h)f(\xi_2)\,. \tag{14c}$$

ETD3RK and ETD4RK can be represented in a similar way. This fact leads us to introduce Butcher tableaux into ETD Runge–Kutta methods.

3.2 *Butcher tableaux*

The s-stage non-ETD Runge–Kutta method to solve an ordinary differential equation

$$u'(x) = g(x, u(x)) \tag{15}$$

is specified by column vectors $\boldsymbol{b}$, $\boldsymbol{c}$ and a matrix $A = (a_{ij})$. The set $(A, \boldsymbol{b}, \boldsymbol{c})$ of an explicit scheme defines a procedure,

$$\xi_i := u_n + h\sum_{j=1}^{i-1} a_{ij}\, g(x + c_j h\,,\ \xi_j) \quad (i = 1, 2, \cdots, s)\,, \tag{16a}$$

$$u_{n+1} := u_n + h\sum_{i=1}^{s} b_i\, g(x + c_i h\,,\ \xi_i)\,, \tag{16b}$$

and is usually written in the tabular form $\begin{array}{c|c} \boldsymbol{c} & A \\ \hline & \boldsymbol{b}^T \end{array}$, which is called a "Butcher tableau."

To represent ETD Runge–Kutta methods in the same tabular form, we redefine A and $\boldsymbol{b}$ as a matrix and a vector, the elements of which are themselves

$$
\begin{array}{c|cc}
0 & & \\
1 & Q_1(\Lambda h) & \\
\hline
 & Q_1 - Q_2 & Q_2
\end{array}
$$

$$
\begin{array}{c|ccc}
0 & & & \\
\frac{1}{2} & \frac{1}{2}Q_1(\Lambda h/2) & & \\
1 & -Q_1(\Lambda h) & 2Q_1(\Lambda h) & \\
\hline
 & Q_1 - 3Q_2 + 4Q_3 & 4Q_2 - 8Q_3 & -Q_2 + 4Q_3
\end{array}
$$

$$
\begin{array}{c|cccc}
0 & & & & \\
\frac{1}{2} & \frac{1}{2}Q_1(\Lambda h/2) & & & \\
\frac{1}{2} & O & \frac{1}{2}Q_1(\Lambda h/2) & & \\
1 & Q_1(\Lambda h) - Q_1(\Lambda h/2) & O & Q_1(\Lambda h/2) & \\
\hline
 & Q_1 - 3Q_2 + 4Q_3 & 2Q_2 - 4Q_3 & 2Q_2 - 4Q_3 & -Q_2 + 4Q_3
\end{array}
$$

Fig. 1. ETD2RK, ETD3RK and ETD4RK represented by Butcher tableaux. Q_k without an argument means $Q_k(\Lambda h)$ and O is a zero matrix.

matrices, α_{ij} and β_i, given by power series as

$$\alpha_{ij} = \sum_{k=0}^{\infty} \alpha_{ij}^{(k)} (\Lambda c_i h)^k \qquad (\alpha_{ij}^{(k)} \in \mathbb{R}) , \tag{17a}$$

$$\beta_i = \sum_{k=0}^{\infty} \beta_i^{(k)} (\Lambda h)^k \qquad (\beta_i^{(k)} \in \mathbb{R}) , \tag{17b}$$

where the superscripts (k) on α_{ij} and β_i are merely indices and do not signify the power or derivative. We denote the redefined matrix (α_{ij}) and the vector (β_i) by $\mathcal{A}$ and $\boldsymbol{\beta}$, respectively. The set $(\mathcal{A}, \boldsymbol{\beta}, \boldsymbol{c})$ of an explicit scheme defines a procedure,

$$\xi_i := \mathrm{e}^{\Lambda c_i h} y_n + h \sum_{j=1}^{i-1} \alpha_{ij}\, f(\xi_j) \quad (i = 1, 2, \cdots, s) , \tag{18a}$$

$$y_{n+1} := \mathrm{e}^{\Lambda h} y_n + h \sum_{i=1}^{s} \beta_i f(\xi_i) , \tag{18b}$$

and can be written in the same tabular form as in non-ETD Runge–Kutta methods. Using modified Butcher tableaux, ETD2RK, ETD3RK and ETD4RK can be written as shown in Fig. 1.

For a succinct description, we introduce several symbols here to represent the coefficients of lower order terms by

$$
\begin{aligned}
a_{ij} &:= \alpha_{ij}^{(0)} = \lim_{(\Lambda c_i h)\to O} \alpha_{ij} \,, && \text{(19a)}\\
\gamma_{ij} &:= \alpha_{ij}^{(1)} = \lim_{(\Lambda c_i h)\to O} \frac{\alpha_{ij} - a_{ij}}{\Lambda c_i h} \,, && \text{(19b)}\\
\delta_{ij} &:= \alpha_{ij}^{(2)} = \lim_{(\Lambda c_i h)\to O} \frac{\alpha_{ij} - a_{ij} - \gamma_{ij}\Lambda c_i h}{(\Lambda c_i h)^2} \,, && \text{(19c)}\\
b_i &:= \beta_i^{(0)} = \lim_{(\Lambda h)\to O} \beta_i \,. && \text{(19d)}
\end{aligned}
$$

In the following sections, we assume that α_{ij} and β_i are arbitrary power series defined by eqs. (17a) and (17b). Additional assumptions such as that they are linear combinations of Q_k are not made a priori as far as possible.

4 Rooted tree analysis and order conditions

To derive the order conditions of Runge–Kutta methods, we must expand both the strict and numerical solutions in terms of step width h and compare their lower order terms. The rooted tree analysis is a common device for conducting such operations systematically in non-ETD Runge–Kutta methods [7]. We apply the tree analysis to ETD Runge–Kutta methods in this section.

4.1 Consistency

Before examining the order conditions, we require the consistency, i.e., the stage values $\{\xi_i\}$ and the final result y_{n+1} are rigorously calculated when $f(y(x))$ in eq. (1) is a constant vector f_c. In such a case, the integral equation (2) becomes

$$
y(x_n + t) = \mathrm{e}^{\Lambda t}\, y(x_n) + \mathrm{e}^{\Lambda t} \int_0^t \mathrm{e}^{-\Lambda\tau}\, \mathrm{d}\tau\, f_c \,. \tag{20}
$$

This can be integrated to give

$$
\begin{aligned}
y(x_n + hc_i) &= Q_0^i y(x_n) + hc_i Q_1^i f_c \,, && \text{(21a)}\\
y(x_n + h\;\;) &= Q_0 y(x_n) + hQ_1 f_c \,, && \text{(21b)}
\end{aligned}
$$

where $Q_k^i := Q_k(\Lambda c_i h)$. On the other hand, an s-stage ETD Runge–Kutta method, when applied to eq. (20), gives

$$
\xi_i = Q_0^i y_n + h \sum_{j=1}^{i-1} \alpha_{ij} f(\xi_j) = Q_0^i y_n + h \left(\sum_{j=1}^{i-1} \alpha_{ij} \right) f_c \,, \tag{22a}
$$

$$y_{n+1} = Q_0 y_n + h \sum_{i=1}^{s} \beta_i f(\xi_i) \ = Q_0 y_n + h \left(\sum_{i=1}^{s} \beta_i \right) f_c \, . \tag{22b}$$

By comparing these expressions, we see that

$$\sum_{j=1}^{i-1} \alpha_{ij} = c_i Q_1^i \quad \text{and} \quad \sum_{i=1}^{s} \beta_i = Q_1 \tag{23}$$

are sufficient conditions for consistency. Note that a consistent method can solve eq. (20) exactly even if $\Lambda = O$.

We assume throughout this paper that the consistency is always satisfied.

4.2 Elementary differentials

In this section, we represent the kth derivative of $f(y(x))$ with respect to x by a weighted sum of the elementary differentials over a set of rooted trees.

If the right-hand side of the differential equation (1) is interpreted as a function $g(y(x)) := \Lambda y(x) + f(y(x))$, the equation takes a general form $y' = g(y(x))$. The formula for the kth derivative of this $g(y(x))$ is already known [7] and is written as

$$\frac{\mathrm{d}^k}{\mathrm{d}x^k} g(y(x)) = \sum_{t \in T_{k+1}} \alpha(t) G(t)(y(x)) \ \ (0 \le \forall k \in \mathbb{Z}) \, , \tag{24}$$

where T_{k+1} is a set of all rooted trees with $k+1$ vertices, $\alpha(t)$ is a function over the set of trees defined by

$$\alpha(t) = \frac{r(t)!}{\gamma(t)\sigma(t)} \quad \Big(r(t) : \text{order}, \ \gamma(t) : \text{density}, \ \sigma(t) : \text{symmetry} \Big) \, , \tag{25}$$

and $G(t)$ is a map from the set of trees to the elementary differentials defined in terms of the Fréchet derivatives $g^{(m)}$ as

$$G(\tau)(y) := g(y) \qquad (\tau \text{ being a tree with one vertex}) \, , \tag{26a}$$

$$G([t_1, \cdots, t_m])(y) := g^{(m)}(y)(G(t_1)(y), G(t_2)(y), \cdots, G(t_m)(y)) \, . \tag{26b}$$

This definition is used recursively. The lower order Fréchet derivatives, $g^{(1)}$, $g^{(2)}$, $g^{(3)}$ and $g^{(4)}$, are also denoted by g', g'', g''' and g'''', respectively.

Let us introduce a subset of T_{k+1} by

$$\widehat{T}_{k+1} := \{ [t] \ ; \ t \in T_k \} \, , \tag{27}$$

where $[\,]$ is an operation to add a new root to the original tree t, i.e., to add a new vertex as a root connected to the original root by an additional edge. We denote the inverse of $[\,]$ by $\{\,\}$ which is applicable only to the elements of $\widehat{T}_{k+1}$. The equality $\alpha(t) = \alpha(\{t\})$ holds for all trees in $\widehat{T}_{k+1}$, because

$$\alpha(t) = \frac{(r(\{t\})+1)!}{(r(\{t\})+1)\gamma(\{t\})1!\sigma(\{t\})^1} = \frac{r(\{t\})!}{\gamma(\{t\})\sigma(\{t\})} = \alpha(\{t\})\,. \qquad (28)$$

Theorem 1 *Define a map F_Λ from the set of rooted trees to the elementary differentials of $f(y)$ as*

$$F_\Lambda(t) := \begin{cases} f & (t \in T_1 \Leftrightarrow t = \tau)\,, \\ G(t) & (t \in T_{k+1} - \widehat{T}_{k+1}\,,\ k \in \mathbb{N})\,, \\ G(t) - \Lambda G(\{t\}) & (t \in \widehat{T}_{k+1}\,,\ k \in \mathbb{N})\,, \end{cases} \qquad (29)$$

then the kth derivative of $f(y(x))$ with respect to x is written as

$$\frac{\mathrm{d}^k}{\mathrm{d}x^k} f(y(x)) = \sum_{t \in T_{k+1}} \alpha(t) F_\Lambda(t)(y(x))\,. \qquad (30)$$

Proof. When $k = 0$, the theorem is obvious from the definitions; therefore, we consider the case where $k \geq 1$. Differentiating k times the both sides of the equation $g(y(x)) = \Lambda y(x) + f(y(x))$, we obtain

$$\begin{aligned} \frac{\mathrm{d}^k}{\mathrm{d}x^k} f(y(x)) &= \frac{\mathrm{d}^k}{\mathrm{d}x^k} g(y(x)) - \Lambda \frac{\mathrm{d}^k}{\mathrm{d}x^k} y(x) \\ &= \frac{\mathrm{d}^k}{\mathrm{d}x^k} g(y(x)) - \Lambda \frac{\mathrm{d}^{k-1}}{\mathrm{d}x^{k-1}} g(y(x))\,. \end{aligned} \qquad (31)$$

Substitution of eq. (24) into this equation leads to

$$\frac{\mathrm{d}^k}{\mathrm{d}x^k} f(y(x)) = \sum_{t \in T_{k+1}} \alpha(t) G(t)(y(x)) - \Lambda \sum_{t \in T_k} \alpha(t) G(t)(y(x))\,. \qquad (32)$$

The summation in the first term on the right-hand side can be divided into two parts. The first is over the subset $T_{k+1} - \widehat{T}_{k+1}$ and the second is over the subset $\widehat{T}_{k+1}$. After this division and a few calculations using $\alpha(t) = \alpha(\{t\})$, the right-hand side may be rewritten as

$$\sum_{t \in T_{k+1} - \widehat{T}_{k+1}} \alpha(t) G(t)(y(x)) + \sum_{t \in \widehat{T}_{k+1}} \alpha(t) \left(G(t) - \Lambda G(\{t\})\right)(y(x))\,. \qquad (33)$$

This is the definition of F_Λ. □

The map F_Λ can be described as an algorithm:

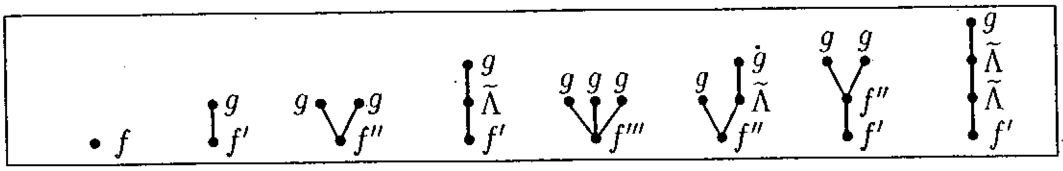

Fig. 2. Trees with up to four vertices in the map F_Λ

Algorithm 1 (calculation of F_Λ)

- *Attach $f^{(m)}$ to the root if it has m children.*
- *Attach $g = \Lambda y + f$ to leaves.*
- *Attach $\widetilde{\Lambda} := \Lambda + f' = g'$ to vertices with one child.*
- *Attach $f^{(m)}$ to vertices with m (two or more) children.*

These items must be executed sequentially from the top of this list, and once something is attached to a vertex, nothing more should be reattached to it in the subsequent steps. If we utilize the map G, a more simple description is possible.

Algorithm 2 (calculation of F_Λ using G)

- *Attach the Fréchet derivatives of g to vertices using the map G.*
- *Replace g attached to the τ with f.*
- *Replace $g' = \Lambda + f'$ attached to the root with f'.*

The last item of this list corresponds to the calculation $G(t) - \Lambda G(\{t\})$ in the theorem. In other words, $G(t) - \Lambda G(\{t\})$ removes Λ that is attached to the root. In Fig. 2, we show the trees with up to four vertices to which the elementary differentials are attached by the map F_Λ. By regarding a child as an argument of its parent, we can construct the compositions of multilinear mappings from these trees. The compositions are f, $f'g$, $f''(g,g)$, $f'\widetilde{\Lambda}g$, $f'''(g,g,g)$, $f''(g,\widetilde{\Lambda}g)$, $f'f''(g,g)$ and $f'\widetilde{\Lambda}\widetilde{\Lambda}g$ from the left.

From the Theorem 1 proved and described here, we see that the Taylor expansion of the function $f(y(x_n+\tau))$ around $\tau = 0$ is

$$f(y(x_n+\tau)) = \sum_{k=0}^{m-1} \frac{\tau^k}{k!} \sum_{t \in T_{k+1}} \alpha(t) F_\Lambda(t)(y(x_n)) + O(\tau^m) \,. \tag{34}$$

Substituting this expansion into eq. (2) and integrating it term by term, we obtain the expansion of the strict solution as

$$y(x_{n+1}) - \mathrm{e}^{\Lambda h} y(x_n) = \sum_{k=0}^{m-1} h^{k+1} Q_{k+1}(\Lambda h) \sum_{t \in T_{k+1}} \alpha(t) F_\Lambda(t)(y(x_n)) + O(h^{m+1}) \,. \tag{35}$$

4.3 Expansion of numerical solution

In non-ETD Runge–Kutta methods, the scalar weights $\{a_{ij}\}$ commute with the Fréchet derivatives $\{g^{(m)}\}$. This fact enables separate definitions of elementary differentials and elementary weights. On the other hand, the weight matrices $\{\alpha_{ij}\}$ do not commute with the Fréchet derivatives $\{f^{(m)}\}$ in ETD Runge–Kutta methods; therefore, it is impossible to define elementary differentials and weights separately and a combined definition becomes necessary. We provide such a definition here.

Definition 2 (weighted elementary differentials)
"Weighted elementary differentials" $W_i^D(t)$, $W_i^S(t)$ and $W^E(t)$ for $\forall t \in T$ and $1 \le \forall i \le s$ are defined by

$$W_i^S(\tau) := c_i Q_1^i g \,, \tag{36a}$$
$$W_i^D(\tau) := f \,, \tag{36b}$$
$$W_i^S([t_1 t_2 \cdots t_m]) := \sum_{j=1}^{i-1} \alpha_{ij} W_j^D([t_1 t_2 \cdots t_m]) \,, \tag{36c}$$
$$W_i^D([t_1 t_2 \cdots t_m]) := f^{(m)}(W_i^S(t_1), W_i^S(t_2), \cdots , W_i^S(t_m)) \,, \tag{36d}$$
$$W^E(t) := \sum_{i=1}^{s} \beta_i W_i^D(t) \,. \tag{36e}$$

These expressions are used recursively.

The expansion theorem of the numerical solution can be formulated by these maps, W_i^D, W_i^S and W^E.

Theorem 3 *An s-stage explicit ETD Runge–Kutta formula that satisfies the consistency conditions is expanded as*

$$\xi_i = y_n + \sum_{t \in T_{\le m}} \frac{h^{r(t)}}{\sigma(t)} W_i^S(t)(y_n) + O(h^{m+1}) \,, \tag{37a}$$

$$hf(\xi_i) = \sum_{t \in T_{\le m}} \frac{h^{r(t)}}{\sigma(t)} W_i^D(t)(y_n) + O(h^{m+1}) \,, \tag{37b}$$

$$y_{n+1} = \mathrm{e}^{\Lambda h} y_n + \sum_{t \in T_{\le m}} \frac{h^{r(t)}}{\sigma(t)} W^E(t)(y_n) + O(h^{m+1}) \tag{37c}$$

$$(1 \le \forall i \le s \,,\ T_{\le m} := \{t \in T \,;\ r(t) \le m\}) \,.$$

Proof. Equation (37c) becomes trivial once the expansion (37b) is proved; therefore, we prove eqs. (37a) and (37b) for all i ($1 \le \forall i \le s$) by mathematical induction.

First, we consider the case $i = 1$. $W_1^S(t)$ is O for all $t \in T$ because $c_1 = 0$ and $\alpha_{1j} = O$. $W_1^D(t)$ is f when $t = \tau$ and is O for the other $t \in T$, since $W_1^S(t)$ in the definition of W_1^D is O. Thus, eqs. (37a) and (37b) give the expressions $\xi_1 = y_n + O(h^{m+1})$ and $hf(\xi_1) = hf + O(h^{m+1})$, which are consistent with the definitions $\xi_1 = y_n$ and $hf(\xi_1) = hf$.

Suppose that eqs. (37a) and (37b) are valid for $1, 2, \cdots, i \in \mathbb{N}$.

The stage value ξ_{i+1} then becomes

$$
\begin{aligned}
\xi_{i+1} &= \mathrm{e}^{\Lambda c_{i+1} h} y_n + \sum_{j=1}^{i} \alpha_{i+1,j} hf(\xi_j) \\
&= \mathrm{e}^{\Lambda c_{i+1} h} y_n + \sum_{t \in T_{\leq m}} \frac{h^{r(t)}}{\sigma(t)} \sum_{j=1}^{i} \alpha_{i+1,j} W_j^D(t)(y_n) + O(h^{m+1}) \\
&= I y_n + h c_{i+1} Q_1^{i+1} \Lambda y_n + h c_{i+1} Q_1^{i+1} f + \\
&\quad \sum_{t \in (T_{\geq 2} \cap T_{\leq m})} \frac{h^{r(t)}}{\sigma(t)} W_{i+1}^S(t)(y_n) + O(h^{m+1}) \quad (\because \text{consistency}) \\
&= y_n + \sum_{t \in T_{\leq m}} \frac{h^{r(t)}}{\sigma(t)} W_{i+1}^S(t)(y_n) + O(h^{m+1}) \\
&\quad \left(\because\ g = \Lambda y_n + f\ ,\ W_{i+1}^S(\tau)(y_n) = c_{i+1} Q_1^{i+1} g\right) .
\end{aligned}
\tag{38}
$$

Therefore, the expansion (37a) of the stage value ξ_{i+1} is proved.
As for the the expansion (37b) of the stage difference $hf(\xi_{i+1})$, we must expand the expression

$$
hf\left(y_n + \sum_{t \in T_{\leq m-1}} \frac{h^{r(t)}}{\sigma(t)} W_{i+1}^S(t)(y_n) + O(h^m)\right) . \tag{$*$}
$$

Let M be the cardinal number of the set $T_{\leq m-1}$, then $T_{\leq m-1}$ can be written as $\{t_1, t_2, \cdots, t_M\}$ by numbering all trees from 1 to M. Since the Fréchet derivatives are multilinear mappings and their images are independent of the order of their arguments, the above equation can be written as

$$
\begin{aligned}
(*) \;=\; &\sum_{k=0}^{m-1} \sum_{(k_1,\cdots,k_M) \in \mathcal{M}_k} \frac{h^{1+\sum_{j=1}^{M} k_j r(t_j)}}{\prod_{j=1}^{M} k_j! \sigma(t_j)^{k_j}} \times \\
&f^{(k)}\overbrace{(\underbrace{W_{i+1}^S(t_1), \cdots, W_{i+1}^S(t_1)}_{k_1}, \cdots, \underbrace{W_{i+1}^S(t_M), \cdots, W_{i+1}^S(t_M)}_{k_M})}^{k} + O(h^{m+1}) ,
\end{aligned}
\tag{39}
$$

where $\mathcal{M}_k$ is a subset of $\mathbb{Z}^M := \overbrace{\mathbb{Z} \times \mathbb{Z} \times \cdots \times \mathbb{Z}}^{M}$ defined by

$$\mathcal{M}_k := \left\{ (k_1, \cdots, k_M) \; ; \; 1 \leq \forall j \leq M \, , \, 0 \leq k_j \in \mathbb{Z} \, , \, \sum_{j=1}^{M} k_j = k \right\} . \tag{40}$$

Let t be a tree constructed from $t_1, t_2, \cdots, t_M$ as

$$t := \left[t_1^{k_1} t_2^{k_2} \cdots t_M^{k_M} \right] = [\underbrace{t_1 \cdots t_1}_{k_1} \underbrace{t_2 \cdots t_2}_{k_2} \cdots \underbrace{t_M \cdots t_M}_{k_M}] . \tag{41}$$

Then, it can be seen that the Fréchet derivative $f^{(k)}$ in the summation of the expression (39) is $W_{i+1}^D(t)$. From the definitions of the order $r(t)$ and the symmetry $\sigma(t)$, it is obvious that

$$1 + \sum_{j=1}^{M} k_j r(t_j) = r(t) \quad \text{and} \quad \prod_{j=1}^{M} \left(k_j! \sigma(t_j)^{k_j} \right) = \sigma(t) . \tag{42}$$

Since the set of rooted trees defined by

$$\begin{aligned} [T_{\leq m-1}] := \Big\{ \left[t_1^{k_1} t_2^{k_2} \cdots t_M^{k_M} \right] \; ; (k_1, k_2, \cdots, k_M) \in \mathcal{M}_k \, , \\ 0 \leq k \leq m-1 \, , t_1, t_2, \cdots, t_M \in T_{\leq m-1} \Big\} \end{aligned} \tag{43}$$

contains all the elements of $T_{\leq m}$ and the terms that correspond to the trees in $[T_{\leq m-1}] - T_{\leq m}$ are $O(h^{m+1})$, the summation over the set $\bigcup_{k=0}^{m-1} \mathcal{M}_k$ in eq. (39) can be replaced by the summation over the set $T_{\leq m}$ when the $O(h^{m+1})$ terms can be neglected. Then the expression (39) becomes

$$\sum_{t \in T_{\leq m}} \frac{h^{r(t)}}{\sigma(t)} W_{i+1}^D(t)(y_n) + O(h^{m+1}) , \tag{44}$$

which validates eq. (37b). □

The weighted elementary differentials defined above can be calculated by the following algorithm.

Algorithm 3 (weighted elementary differentials)

- *Attach indices $i, j, k, l, \cdots$ to each vertex of a rooted tree t. The index i should be attached to the root.*
- *Attach α_{ij} to the edge $[i, j] = \overset{j}{\underset{i}{|}}$, α_{jk} to the edge $[j, k] = \overset{k}{\underset{j}{|}}$, $\cdots$.*
- *Attach the Fréchet derivative $f^{(m)}$ to vertices that have m children $(1 \leq m)$.*
- *Attach g to leaves of a tree $t \neq \tau$ and attach f to the root of τ.*
- *Construct a composition of (multi)linear mappings, so that the edges be the arguments of the parent and that the child be the argument of the edge.*

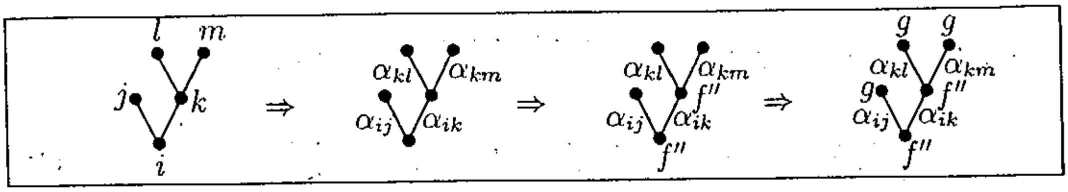

Fig. 3. Calculation of a weighted elementary differential

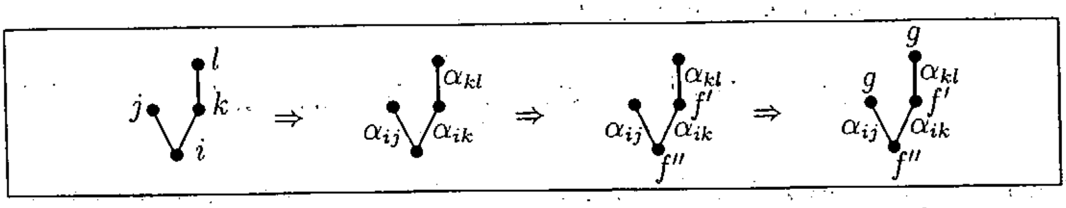

Fig. 4. Another example of a weighted elementary differential

- *Sum this composition over each index except i, i.e., $\sum_{j=1}^{s}\sum_{k=1}^{s}\sum_{l=1}^{s}\cdots$. The resulting sum is $W_i^D(t)$.*
- *Calculate $\sum_{i=1}^{s}\beta_i W_i^D(t) = W^E(t)$.*

This algorithm is executed on a tree as shown in Fig. 3. The composition constructed from this tree corresponds to $f''(\alpha_{ij}g\,,\,\alpha_{ik}f''(\alpha_{kl}g\,,\,\alpha_{km}g))$. The summation over each index leads to

$$W_i^D(t) = \sum_{j=1}^{s}\sum_{k=1}^{s}\sum_{l=1}^{s}\sum_{m=1}^{s} f''(\alpha_{ij}g\,,\,\alpha_{ik}f''(\alpha_{kl}g\,,\,\alpha_{km}g))$$

$$= \sum_{k=1}^{s} f''\left(c_iQ_1^ig\,,\,\alpha_{ik}f''(c_kQ_1^kg\,,\,c_kQ_1^kg)\right) \quad (\because \text{consistency})\,; \tag{45a}$$

$$W^E(t) = \sum_{i=1}^{s}\sum_{k=1}^{s}\beta_i f''\left(c_iQ_1^ig\,,\,\alpha_{ik}f''(c_kQ_1^kg\,,\,c_kQ_1^kg)\right)\,. \tag{45b}$$

By expanding the power series β_i, Q_1^i, Q_1^k and α_{ik} in the last equation and neglecting the first and higher order terms, we obtain

$$\left(\sum_{i=1}^{s}\sum_{k=1}^{s} b_ic_ia_{ik}c_k^2\right) f''(g, f''(g,g))\,. \tag{46}$$

The coefficient part is an elementary weight $\Phi(t)$ and the composition part corresponds to an elementary differential $F(t)$ in the theory of non-ETD Runge–Kutta methods [7].

Reexpansion of the power series in the expression of $W_i^D(t)$ gives its complete expansion. For example, applying the algorithm to the tree $[\tau[\tau]]$ as depicted in Fig. 4, we can calculate $W_i^D(t)$ as

$$W_i^D(t) = \sum_{j=1}^{s}\sum_{k=1}^{s}\sum_{l=1}^{s} f''(\alpha_{ij}g\,,\,\alpha_{ik}f'\alpha_{kl}g) = \sum_{k=1}^{s} f''(c_iQ_1^ig\,,\,\alpha_{ik}f'c_kQ_1^kg)\,. \tag{47}$$

If we choose the m_1th order term in Q_1^i, the m_2th order term in α_{ik} and the m_3th order term in Q_1^k, i.e.,

$$\frac{(\Lambda c_i h)^{m_1}}{(m_1+1)!}, \quad \alpha_{ik}^{(m_2)}(\Lambda c_i h)^{m_2} \quad \text{and} \quad \frac{(\Lambda c_k h)^{m_3}}{(m_3+1)!}, \tag{48}$$

respectively, we can see that the term

$$h^{m_1+m_2+m_3} \sum_{k=1}^{s} \underbrace{\frac{c_i^{m_1+1}}{(m_1+1)!}}_{c_i Q_1^i} \overbrace{c_i^{m_2}\alpha_{ik}^{(m_2)}}^{\alpha_{ik}} \underbrace{\frac{c_k^{m_3+1}}{(m_3+1)!}}_{c_k Q_1^k} \overbrace{f''(\Lambda^{m_1} g,\ \Lambda^{m_2} f' \Lambda^{m_3} g)}^{\text{composed linear mappings}} \tag{49}$$

is included in the complete expansion of $W_i^D(t)$. By adding terms over all possible and necessary sets (m_1, m_2, m_3), we obtain the truly truncated expansion of $W_i^D(t)$.

4.4 Order conditions

Now that the expansions of the strict and numerical solutions are established, we can define order conditions by comparing their lower order terms. In the following, the terms up to pth order of a power series $S(h)$ are denoted by $[S(h)]_{\le p}$ and the pth order terms are denoted by $[S(h)]_{=p}$. The equality $y(x_n) = y_n$ is assumed.

Definition 4 (strong pth order conditions)
If the terms up to pth order of the difference $y(x_{n+1}) - e^{\Lambda h} y(x_n)$,

$$\sum_{k=0}^{p-1} h^{k+1} Q_{k+1}(\Lambda h) \sum_{t \in T_{k+1}} \alpha(t) F_\Lambda(t)(y(x_n)) , \tag{50}$$

and those of the difference $y_{n+1} - e^{\Lambda h} y_n$ calculated by an ETD Runge–Kutta formula ,

$$\sum_{t \in T_{\le p}} \frac{h^{r(t)}}{\sigma(t)} \sum_{i=1}^{s} \beta_i \left[W_i^D(t)(y_n) \right]_{\le p - r(t)} , \tag{51}$$

coincide with each other term-by-term, the formula is said to have strong pth order accuracy.

In this definition, Q_k in eq. (50) and β_i in eq. (51) are not expanded and are treated as $O(1)$ quantities, while the weighted elementary differentials $W_i^D(t)(y_n)$, i.e., $hf(\xi_i)$, are completely expanded.

Definition 5 (weak pth order conditions)
Suppose that an ETD Runge–Kutta formula satisfies the strong $(p-1)$th order

conditions. If the reexpanded and truncated pth order terms of the numerical solution ,

$$\sum_{t\in T_p}\frac{h^{r(t)}}{\sigma(t)}\sum_{i=1}^{s}[\beta_i]_{=0}\left[W_i^D(t)(y_n)\right]_{=(p-r(t))}\ , \tag{52}$$

and those of the strict solution ,

$$h^p\left[Q_p(\Lambda h)\right]_{=0}\sum_{t\in T_p}\alpha(t)F_\Lambda(t)(y(x_n))\ , \tag{53}$$

coincide with each other term-by-term, the formula is said to have weak pth order accuracy.

Note that Q_p and β_i in the pth order terms are expanded and their first and higher order terms are neglected in the definition of the weak order condition.

In terms of these conditions, ETD2RK has the strong second-order accuracy, ETD3RK has the weak third-order accuracy and ETD4RK has the weak fourth-order accuracy (Section 5).

When an ETD Runge–Kutta formula satisfies the weak pth order conditions, it converges to the corresponding non-ETD Runge–Kutta formula with the same order accuracy in the limit $(\Lambda h) \to O$. This means that in the case where the matrix Λ is O or negligible, the numerical solution is calculated by a standard non-ETD Runge–Kutta formula or its neighbors.

4.5 Convergence

Convergence of a time discretization scheme is usually defined as follows:

Definition 6 (convergence)
Suppose that a numerical integration is performed over a time interval $[a,b]$ by a scheme using a step width $h := (b-a)/N$. Let e_n be an error vector at the nth step ($x = a + nh$) and assume that the initial value has no error, i.e., $e_0 = 0$. If the error vectors $\{e_n\}$ converge to 0 in the limit of infinitesimal step width, i.e.,

$$0 \le \forall n \le N\ ,\ \lim_{N\to\infty}\|e_n\| = 0\ , \tag{54}$$

then the scheme is said to be convergent.

We check the convergence of ETD Runge–Kutta methods in this section. The proof for non-ETD Runge–Kutta methods [12] is slightly modified and applied to ETD methods. We assume that the functions $f(y)$, $y(x)$ and $\alpha_{ij}(h)$ are continuously differentiable at least p-times with respect to their arguments, and that $\{\alpha_{ij}\}$ and $\{\beta_i\}$ are bounded operators.

When an ETD Runge–Kutta formula satisfies the strong pth order conditions, the error $E_{n+1}(h)$ defined, in terms of the strict solution $\widehat{y}_n := y(x_n)$ and $\widehat{y}_{n+1} := y(x_{n+1})$, by

$$\widehat{\xi}_i := Q_0^i \widehat{y}_n + h \sum_{j=1}^{i-1} \alpha_{ij} f(\widehat{\xi}_j) \quad (i = 1, 2, \cdots, s) , \tag{55a}$$

$$\widehat{y}_{n+1} = Q_0 \widehat{y}_n + h \sum_{i=1}^{s} \beta_i f(\widehat{\xi}_i) + E_{n+1}(h) , \tag{55b}$$

is an $O(h^{p+1})$ quantity, i.e.,

$$0 < \exists h_0,\ \exists G \in \mathbb{R},\ 0 < \forall h < h_0,\ \|E_{n+1}\| < G h^{p+1} . \tag{56}$$

Let e_n be the difference $y_n - \widehat{y}_n$ and assume that the initial value is correct, i.e., $e_0 := 0$. By subtracting the numerical solution from eq. (55b), we can see that

$$\|e_{n+1}\| \le \|Q_0\| \|e_n\| + hL \sum_{i=1}^{s} \|\beta_i\| \|\xi_i - \widehat{\xi}_i\| + G h^{p+1} \quad (0 < \forall h < h_0) , \tag{57}$$

where L is the Lipschitz constant of the Lipschitz continuous function f.

Following is the proof for two-stage formulas with the strong second-order accuracy. No essential difference exists in the proof for three or more stage formulas with higher order accuracy. Substituting the evaluations,

$$\|\xi_1 - \widehat{\xi}_1\| = \|e_n\| , \tag{58a}$$

$$\|\xi_2 - \widehat{\xi}_2\| \le (\|Q_0^2\| + hL\|\alpha_{21}\|)\|e_n\| , \tag{58b}$$

into the inequality (57), we obtain

$$\begin{aligned} \|e_{n+1}\| &\le \{\|Q_0\| + hL(\|\beta_1\| + \|\beta_2\|\|Q_0^2\|) + h^2L^2\|\beta_2\|\|\alpha_{21}\|\}\|e_n\| + Gh^3 \\ &= \|Q_0\| \left(1 + hL\frac{\delta_1}{\|Q_0\|} + h^2L^2\frac{\delta_2}{\|Q_0\|}\right) \|e_n\| + Gh^3 , \end{aligned} \tag{59}$$

where δ_1 and δ_2 are positive constants independent of $h \in (0, h_0)$. By a suitable choice of $h_1 (< h_0)$, this can be rewritten as

$$\|e_{n+1}\| \le \|Q_0\|(1 + hL')\|e_n\| + Gh^3 \quad (0 < \forall h < h_1) , \tag{60}$$

where L' is a constant multiple of the Lipschitz constant L. Repeatedly applying this inequality to itself, we obtain

$$\begin{aligned} \|e_n\| &\le \sum_{m=0}^{n-1} \|Q_0\|^m (1 + hL')^m G h^3 \ \le\ e^{Nh\|\Lambda\|} \sum_{m=0}^{n-1} (1 + hL')^m G h^3 \\ &\le e^{Nh\|\Lambda\|} (e^{nhL'} - 1) \frac{G}{L'} h^2 \ \le\ e^{(\|\Lambda\| + L')(b-a)} \frac{G}{L'} h^2 \quad (0 \le \forall n \le N) . \end{aligned} \tag{61}$$

Convergence

$$0 \le \forall n \le N \,, \ \lim_{N\to\infty} \|e_n\| = 0 \tag{62}$$

follows from the inequality (61) since the coefficient of h^2 is independent of $h \in (0, h_1)$.

5 Derivation of formulas

5.1 Two-stage formulas

Two-stage ETD Runge–Kutta methods can be represented by a tableau as

$$\begin{array}{c|cc} 0 & & \\ c_2 & \alpha_{21} & \\ \hline & \beta_1 & \beta_2 \end{array}$$

To satisfy the consistency condition, α_{21} must be $c_2 Q_1^2$. The expansion of the difference $y_{n+1} - Q_0 y_n$ becomes

$$\begin{aligned} & h(\beta_1+\beta_2)f + h^2\beta_2 c_2 f'g \\ & + h^3\frac{1}{2}\beta_2 c_2^2 f'\widetilde{\Lambda} g + h^3\frac{1}{2}\beta_2 c_2^2 f''(g,g) - h^3\frac{1}{2}\beta_2 c_2^2 f'f'g + O(h^4)\,, \end{aligned} \tag{63}$$

where $\widetilde{\Lambda} := \Lambda + f'$. Comparing these expressions with the expansion of the strict solution,

$$\begin{aligned} y(x_{n+1}) = \mathrm{e}^{\Lambda h} y(x_n) + hQ_1(\Lambda h)f + h^2 Q_2(\Lambda h) f'g + \\ h^3 Q_3(\Lambda h) f'\widetilde{\Lambda} g + h^3 Q_3(\Lambda h) f''(g,g) + O(h^4)\,, \end{aligned} \tag{64}$$

we obtain the strong second-order conditions as

$$\beta_1 + \beta_2 = Q_1 \quad \text{and} \quad \beta_2 c_2 = Q_2\,. \tag{65}$$

These equations can be solved to give

$$\beta_1 = Q_1 - \frac{1}{c_2} Q_2 \quad \text{and} \quad \beta_2 = \frac{1}{c_2} Q_2\,. \tag{66}$$

Therefore, the two-stage formula with the strong second-order accuracy is represented by a Butcher tableau as

$$
\begin{array}{c|cc}
0 & & \\
c_2 & c_2 Q_1^2 & \\
\hline
 & Q_1 - \frac{1}{c_2} Q_2 & \frac{1}{c_2} Q_2
\end{array}
$$

A choice of the parameter $c_2 := 1$ yields ETD2RK.

From the above derivation, we can see that the power series α_{21}, β_1 and β_2 are well characterized by the consistency and order conditions.

The difference in the third-order terms between the numerical and strict solutions is

$$
h^3 \left\{ \left(\frac{c_2}{2} Q_2 - Q_3 \right) \left(f' \tilde{\Lambda} g + f''(g, g) \right) - \frac{c_2}{2} Q_2 f' f' g \right\} . \tag{67}
$$

By neglecting the first and higher order terms in the power series Q_2 and Q_3, we obtain

$$
h^3 \left\{ \left(\frac{c_2}{4} - \frac{1}{6} \right) \left(f' \tilde{\Lambda} g + f''(g, g) \right) - \frac{c_2}{4} f' f' g \right\} . \tag{68}
$$

When $c_2 = 2/3$, the first term in the brace brackets vanishes and the rest becomes $-(h^3/6)\, f' f' g$.

5.2 *Three-stage formulas*

Three-stage ETD Runge–Kutta methods can be represented by a tableau as

$$
\begin{array}{c|ccc}
0 & & & \\
c_2 & \alpha_{21} & & \\
c_3 & \alpha_{31} & \alpha_{32} & \\
\hline
 & \beta_1 & \beta_2 & \beta_3
\end{array}
$$

If we require the consistency conditions,

$$
\alpha_{21} = c_2 Q_1^2 \quad \text{and} \quad \alpha_{31} + \alpha_{32} = c_3 Q_1^3 , \tag{69}
$$

the difference $y_{n+1} - Q_0\, y_n$ can be expanded as

$$
\begin{aligned}
& h(\beta_1 + \beta_2 + \beta_3) f + h^2 (\beta_2 c_2 + \beta_3 c_3) f' g \\
& + h^3 \frac{1}{2} \left(\beta_2 c_2^2 + \beta_3 c_3^2 \right) (f' \Lambda g + f''(g, g)) + h^3 \beta_3 a_{32} c_2 f' f' g + O(h^4) .
\end{aligned} \tag{70}
$$

The strong third-order conditions for the numerical result y_{n+1} are derived from the comparison between the expansion of the numerical solution (70) and that of the strict solution (64), and can be written as

$$\beta_1 + \beta_2 + \beta_3 = Q_1 \,, \tag{71a}$$

$$\beta_2 c_2 + \beta_3 c_3 = Q_2 \,, \tag{71b}$$

$$\beta_2 c_2^2 + \beta_3 c_3^2 = 2Q_3 \,, \tag{71c}$$

$$\beta_3 a_{32} c_2 = Q_3 \,. \tag{71d}$$

Three-stage formulas that satisfy all these equations simultaneously do not exist; therefore, we replace the last condition with its constant term $b_3 a_{32} c_2 = 1/6$. This means that the relaxed condition is imposed on the term $h^3 f' f' g$ which is proportional to Λ, while the more stringent ones are imposed as before on the terms $h^3 f' \Lambda g$ and $h^3 f''(g,g)$ which are proportional to Λ^2. Suppose that $c_2 \neq c_3$, then the upper three equations, (71a), (71b) and (71c), can be solved to give

$$\beta_1 = Q_1 - \frac{c_2 + c_3}{c_2 c_3} Q_2 + \frac{2}{c_2 c_3} Q_3 \,, \tag{72a}$$

$$\beta_2 = \frac{c_3}{c_2(c_3 - c_2)} Q_2 - \frac{2}{c_2(c_3 - c_2)} Q_3 \,, \tag{72b}$$

$$\beta_3 = -\frac{c_2}{c_3(c_3 - c_2)} Q_2 + \frac{2}{c_3(c_3 - c_2)} Q_3 \,, \tag{72c}$$

and the constant b_3 becomes

$$b_3 = \lim_{(\Lambda h) \to O} \beta_3 = \frac{2 - 3c_2}{6 c_3 (c_3 - c_2)} \,. \tag{73}$$

Substitution of b_3 into the equation $b_3 a_{32} c_2 = 1/6$ leads to

$$a_{32} = \frac{c_3(c_3 - c_2)}{c_2(2 - 3c_2)} \,. \tag{74}$$

5.2.1 *A class including ETD3RK*

In the preceding section, the power series α_{21} is completely determined by the consistency condition, while α_{31} and α_{32} are not. The expression of a_{32}, derived from one of the weak third-order conditions, specifies only the constant term of the power series α_{32}.

If α_{32} is required to be a constant multiple of Q_1^3 from some other considerations, then α_{31} and α_{32} are completely determined as

$$\alpha_{32} = a_{32} Q_1^3 \quad \text{and} \quad \alpha_{31} = (c_3 - a_{32}) Q_1^3 \,. \tag{75}$$

This defines a class of schemes parameterized by c_2 and c_3. A choice of these parameters $(c_2, c_3) := (1/2, 1)$ gives ETD3RK.

5.2.2 *A class derived from the stage order conditions*

In this section, we consider the case where the stage value ξ_3 is required to have the strong second-order accuracy. This requirement can be written as

$$\alpha_{32} c_2 = c_3^2 Q_2^3 \,. \tag{76}$$

This equation with the consistency condition can be solved to give

$$\alpha_{32} = \frac{c_3^2}{c_2} Q_2^3 \quad \text{and} \quad \alpha_{31} = c_3 Q_1^3 - \frac{c_3^2}{c_2} Q_2^3 \,. \tag{77}$$

Although the parameters c_2 and c_3 are still arbitrary, forms of the power series α_{31} and α_{32} are almost characterized. To be consistent with a_{32} already derived from the weak third order conditions, the relation

$$\frac{c_3^2}{2c_2} = a_{32} = \frac{c_3(c_3 - c_2)}{c_2(2 - 3c_2)} \tag{78}$$

must be satisfied, namely, $c_3 = 2/3$. Thus, we obtain a class of schemes parameterized by c_2.

5.3 *Four-stage formulas*

Four-stage ETD Runge–Kutta methods define the numerical procedure,

$$\xi_i = \mathrm{e}^{\Lambda c_i h} y_n + h \sum_{j=1}^{i-1} \alpha_{ij} f(\xi_j) \quad (i = 1, 2, 3, 4)\,, \tag{79a}$$

$$y_{n+1} = \mathrm{e}^{\Lambda h} y_n + h \sum_{i=1}^{4} \beta_i f(\xi_i)\,. \tag{79b}$$

Since it is impossible to satisfy the strong fourth-order conditions by the four-stage methods, we consider the weak fourth-order conditions in this section.

The expansion of the difference $y_{n+1} - Q_0 y_n$ of a consistent method with four stages is shown in Fig. 5 and the weak fourth-order conditions derived from this expansion are shown in Fig. 6. The equation including γ_{32}, γ_{42} and γ_{43} has no counterpart in the fourth-order conditions of non-ETD Runge–Kutta methods. Therefore, the weak fourth-order conditions cannot be satisfied by a simple replacement of a_{ij} in the non-ETD formula with $\alpha_{ij} := a_{ij} Q_1^i$.

5.3.1 *A class including ETD4RK*

The class of formulas that satisfy the weak fourth-order conditions includes ETD4RK. In addition to the order conditions, we require that $\alpha_{31} = \alpha_{42} =$

$$
\begin{aligned}
\text{first-order} \quad &: \quad h(\beta_1+\beta_2+\beta_3+\beta_4)f \\
\text{second-order} \quad &: \quad h^2(\beta_2c_2+\beta_3c_3+\beta_4c_4)f'g \\
\text{third-order} \quad &: \quad h^3\tfrac{1}{2}(\beta_2c_2^2+\beta_3c_3^2+\beta_4c_4^2)f''(g,g)+ \\
&\quad\ \ h^3\tfrac{1}{2}(\beta_2c_2^2+\beta_3c_3^2+\beta_4c_4^2)f'\Lambda g+ \\
&\quad\ \ h^3(\beta_3a_{32}c_2+\beta_4a_{42}c_2+\beta_4a_{43}c_3)f'f'g \\
\text{fourth-order} \quad &: \quad h^4\tfrac{1}{6}(\beta_2c_2^3+\beta_3c_3^3+\beta_4c_4^3)f'\Lambda^2g+ \\
&\quad\ \ h^4\tfrac{1}{2}(\beta_2c_2^3+\beta_3c_3^3+\beta_4c_4^3)f''(g,\Lambda g)+ \\
&\quad\ \ h^4\tfrac{1}{6}(\beta_2c_2^3+\beta_3c_3^3+\beta_4c_4^3)f'''(g,g,g)+ \\
&\quad\ \ h^4\tfrac{1}{2}(\beta_3a_{32}c_2^2+\beta_4a_{42}c_2^2+\beta_4a_{43}c_3^2)f'f'\Lambda g+ \\
&\quad\ \ h^4\tfrac{1}{2}(\beta_3a_{32}c_2^2+\beta_4a_{42}c_2^2+\beta_4a_{43}c_3^2)f'f''(g,g)+ \\
&\quad\ \ h^4(\beta_3c_3\gamma_{32}c_2+\beta_4c_4\gamma_{42}c_2+\beta_4c_4\gamma_{43}c_3)f'\Lambda f'g+ \\
&\quad\ \ h^4(\beta_3c_3a_{32}c_2+\beta_4c_4a_{42}c_2+\beta_4c_4a_{43}c_3)f''(g,f'g)+ \\
&\quad\ \ h^4\beta_4a_{43}a_{32}c_2f'f'f'g
\end{aligned}
$$

Fig. 5. Expansion of the difference of a four-stage formula

$$
\begin{aligned}
&\text{(consistency)}\left\{
\begin{aligned}
\alpha_{21} &= c_2Q_1^2 \\
\alpha_{31}+\alpha_{32} &= c_3Q_1^3 \\
\alpha_{41}+\alpha_{42}+\alpha_{43} &= c_4Q_1^4
\end{aligned}\right. \\
&\text{(strong third)}\left\{
\begin{aligned}
\beta_1+\beta_2+\beta_3+\beta_4 &= Q_1 \\
\beta_2c_2+\beta_3c_3+\beta_4c_4 &= Q_2 \\
\beta_2c_2^2+\beta_3c_3^2+\beta_4c_4^2 &= 2Q_3 \\
\beta_3a_{32}c_2+\beta_4a_{42}c_2+\beta_4a_{43}c_3 &= Q_3
\end{aligned}\right. \\
&\text{(weak fourth)}\left\{
\begin{aligned}
b_2c_2^3+b_3c_3^3+b_4c_4^3 &= \tfrac{1}{4} \\
b_3a_{32}c_2^2+b_4a_{42}c_2^2+b_4a_{43}c_3^2 &= \tfrac{1}{12} \\
b_3c_3\gamma_{32}c_2+b_4c_4\gamma_{42}c_2+b_4c_4\gamma_{43}c_3 &= \tfrac{1}{24} \\
b_3c_3a_{32}c_2+b_4c_4a_{42}c_2+b_4c_4a_{43}c_3 &= \tfrac{1}{8} \\
b_4a_{43}a_{32}c_2 &= \tfrac{1}{24}
\end{aligned}\right.
\end{aligned}
$$

Fig. 6. Weak fourth-order conditions

O and that the scheme converges to the classical four-stage Runge–Kutta formula with fourth-order accuracy in the limit $(\Lambda h) \to O$. These requirements determine the matrices α_{21}, α_{32} and $\{\beta_i \; ; \; 1 \le i \le 4\}$ as

$$\alpha_{21} = \alpha_{32} = \frac{1}{2} Q_1 \left(\Lambda \frac{h}{2} \right) , \tag{80a}$$

$$\beta_1 = Q_1 - 3Q_2 + 4Q_3, \quad \beta_2 = \beta_3 = 2Q_2 - 4Q_3, \quad \beta_4 = -Q_2 + 4Q_3 \, . \tag{80b}$$

The weights $\{\alpha_{4j} \; ; \; j = 1, 3\}$ for the fourth stage are not determined yet. The conditions for the fourth stage are given by

$$\alpha_{41} = Q_1(\Lambda h) - \alpha_{43} \; , \quad a_{43} = 1 \quad \text{and} \quad \gamma_{43} = \frac{1}{4} \, . \tag{81}$$

Only the constant and first-order terms of α_{43} are specified, while the other terms remain arbitrary. Some other conditions are necessary for a better characterization. If we define $\alpha_{43} := Q_1(\Lambda h/2)$, then we can obtain ETD4RK.

5.3.2 A class derived from the stage order conditions

In addition to the consistency and the weak fourth-order conditions, we require that ξ_3 and ξ_4 have the strong second-order accuracy and that the strong third-order conditions are partially satisfied by ξ_4. The second item means that among the third order terms, $h^3 f''(g,g)$, $h^3 f' \Lambda g$ and $h^3 f' f' g$, the strong order conditions for the first and second terms are satisfied. These additional conditions can be written as

$$\alpha_{32} c_2 = c_3^2 Q_2^3 \, , \tag{82a}$$

$$\alpha_{42} c_2 + \alpha_{43} c_3 = c_4^2 Q_2^4 \, , \tag{82b}$$

$$\alpha_{42} c_2^2 + \alpha_{43} c_3^2 = 2 c_4^3 Q_3^4 \, . \tag{82c}$$

All these conditions, i.e., the consistency, the weak fourth order and the additional ones mentioned in this section, can be solved to give a unique formula when $c_2 \neq c_3$ and $c_3 \neq c_4$. The following is the formula:

$$\begin{array}{c|cccc}
0 & & & & \\
\frac{1}{3} & \frac{1}{3}Q_1^2 & & & \\
\frac{1}{2} & \frac{1}{2}Q_1^3 - \frac{3}{4}Q_2^3 & \frac{3}{4}Q_2^3 & & \\
1 & Q_1^4 - 5Q_2^4 + 12Q_3^4 & 9Q_2^4 - 36Q_3^4 & -4Q_2^4 + 24Q_3^4 & \\
\hline
 & Q_1 - 3Q_2 + 4Q_3 & O & 4Q_2 - 8Q_3 & -Q_2 + 4Q_3
\end{array}$$

The limit scheme $\lim_{(\Lambda h) \to O} (\mathcal{A}, \boldsymbol{\beta}, \boldsymbol{c}) =: (A, \boldsymbol{b}, \boldsymbol{c})$ satisfies the fourth-order conditions of non-ETD Runge–Kutta methods.

6 Concluding remarks

We have investigated ETD Runge–Kutta methods proposed by previous authors. The main purpose of the present study is to provide a mathematical foundation of the new methods.

First, we have shown that the new schemes can be represented by Butcher tableaux whose elements are themselves matrices instead of scalars. Each element of a tableau is defined by a power series of the matrix Λ that appears in the linear term of a differential equation $y' = \Lambda y + f(y(x))$. This representation enables us to view the new methods from a perspective of the theory of Runge–Kutta methods.

To establish the order conditions, we have applied the rooted tree analysis to the new methods. The essential difference from the conventional Runge–Kutta theory resides in the fact that the coefficient matrices of a scheme, i.e., $\{\alpha_{ij}\}$ and $\{\beta_i\}$, do not commute with the Frechét derivatives $f^{(m)}$. From this fact, separate definitions of elementary differentials and weights, as in the conventional theory, becomes impossible. Therefore, we have provided combined definitions, i.e., weighted elementary differentials. In terms of these differentials, the expansion theorems and the order conditions have been established. Convergence of a scheme that satisfies the order conditions was also confirmed. Furthermore, we saw that an ETD Runge–Kutta formula with the weak pth order accuracy is a conventional Runge–Kutta formula with the same order accuracy in the limit $\Lambda \to O$.

Finally, with regard to the derivation of the new ETD Runge–Kutta formulas, we saw that one of the difficulties is the characterization of the power series α_{ij} and β_i without arbitrary assumptions. It was shown that imposing the stage order conditions is one of the ways to overcome this difficulty.

The concept of ETD Runge–Kutta methods may be combined with other time discretization schemes, e.g., general linear methods [6,7], pseudo Runge–Kutta methods [8,13], two-step Runge–Kutta methods [14] and so on. We will continue to study such composite schemes in the future.

Acknowledgements

The author would like to express his gratitude to Prof. S. Kida and Dr. S. Goto for their suggestions and discussions. He is also grateful to S. Iwasaki for the various support to this research.

References

[1] B. Alpert, G. Beylkin, D. Gines and L. Vozovoi, Adaptive Solution of Partial Differential Equations in Multiwavelet Bases, J. Comput. Phys. 182 (2002) 149-190.

[2] G. A. Baker, Jr., Essentials of Padé Approximants, Academic Press, New York, 1975

[3] G. Beylkin, J. M. Keiser and L. Vozovoi, A New Class of Time Discretization Schemes for the Solution of Nonlinear PDEs, J. Comput. Phys. 147 (1998) 362-387.

[4] J. P. Boyd, Chebyshev and Fourier Spectral Methods, Dover Publications, New York, 2001.

[5] J. C. Butcher, Coefficients for the study of Runge–Kutta integration processes, J. Austral. Math. Soc. 3 (1963) 185-201.

[6] J. C. Butcher, On the convergence of numerical solutions to ordinary differential equations, Math. Comp. 20 (1966) 1-10

[7] J. C. Butcher, Numerical Methods for Ordinary Differential Equations, John Wiley & Sons, Chichester, 2003.

[8] G. D. Byrne and R. J. Lambert, Pseudo–Runge–Kutta methods involving two points, J. Assoc. Comput. Mach. 13 (1966) 114-123

[9] C. Canuto, M. Y. Hussaini, A. Quarteroni and T. A. Zang, Spectral Methods in Fluid Dynamics, Springer-Verlag, New York, 1988.

[10] S. M. Cox and P. C. Matthews, Exponential Time Differencing for Stiff Systems, J. Comput. Phys. 176 (2002) 430-455.

[11] S. M. Cox and P. C. Matthews, Instability and localisation of patterns due to a conserved quantity, Physica D 175 (2003) 196-219.

[12] T. Mitsui, Numerical methods for differential equations I (in Japanese), Iwanami Shoten, Tokyo, 1993.

[13] M. Nakashima, Pseudo Runge–Kutta Processes, Math. Sci. Res. Ins. publ. 23 (1987) 583-611.

[14] N. H. Cong and T. Mitsui, Collocation-based two-step Runge–Kutta methods, Japan J. Industr. Appl. Math. 13 (1996) 171-183.

Recent Issues of NIFS Series

NIFS-767 R. Kanno, N. Nakajima, M. Okamoto and T. Hayashi
Computational Study of Three Dimensional MHD Equilibrium with m/n=1/1 Island
Dec. 2002

NIFS-768 M. Yagi, S.-I. Itoh, M. Kawasaki, K. Itoh and A. Fukuyama
Multiple-Scale Turbulence and Bifurcaion
Jan. 2003

NIFS-769 S.-I. Itoh, K. Itoh and S. Toda
Statistical Theory of L-H Transition and its Implication to Threshold Database
Jan. 2003

NIFS-770 K. Itoh
Summary: Theory of Magnetic Confinement
Jan. 2003

NIFS-771 S.-I. Itoh, K. Itoh and S. Toda
Statistical Theory of L-H Transition in Tokamaks
Jan. 2003

NIFS-772 M. Stepic, L. Hadzievski and M.M. Skoric
Modulation Instability in Two-dimensional Nonlinear Schrodinger Lattice Models with Dispersion and Long-range Interactions
Jan. 2003

NIFS-773 M.Yu. Isaev, K.Y. Watanabe, M. Yokoyama and K. Yamazaki
The Effect of Hexapole and Vertical Fields on α-particle Confinement in Heliotron Configurations
Mar. 2003

NIFS-774 K. Itoh, S.-I. Itoh, F. Spineanu, M.O. Vlad and M. Kawasaki
On Transition in Plasma Turbulence with Multiple Scale Lengths
May 2003

NIFS-775 M. Vlad, F. Spineanu, K. Itoh, S.-I. Itoh
Intermittent and Global Transitions in Plasma Turbulence
July 2003

NIFS-776 Y. Kondoh, M. Kondo, K. Shimoda, T. Takahashi and K. Osuga
Innovative Direct Energy Conversion Systems from Fusion Output Thermal Power to the Electrical One with the Use of Electronic Adiabatic Processes of Electron Fluid in Solid Conductors.
July 2003

NIFS-777 S.-I. Itoh, K. Itoh and M. Yagi
A Novel Turbulence Trigger for Neoclassical Tearing Modes in Tokamaks
July 2003

NIFS-778 T. Utsumi, J. Koga, T. Yabe, Y. Ogata, E. Matsunaga, T. Aoki and M. Sekine
Basis Set Approach in the Constrained Interpolation Profile Method
July 2003

NIFS-779 Oleg I. Tolstikhin and C. Namba
CTBC A Program to Solve the Collinear Three-Body Coulomb Problem: Bound States and Scattering Below the Three-Body Disintegration Threshold
Aug. 2003

NIFS-780 Contributions to 30th European Physical Society Conference on Controlled Fusion and Plasma Physics
(St.Petersburg, Russia, 7-11 July 2003) from NIFS
Aug. 2003

NIFS-781 Ya. I. Kolesnichenko, K. Yamazaki, S. Yamamoto, V.V. Lutsenko, N. Nakajima, Y. Narushima, K. Toi, Yu. V. Yakovenko
Interplay of Energetic Ions and Alfvén Modes in Helical Plasmas
Aug. 2003

NIFS-782 S.-I. Itoh, K. Itoh and M. Yagi
Turbulence Trigger for Neoclassical Tearing Modes in Tokamaks
Sep. 2003

NIFS-783 F. Spineanu, M. Vlad, K. Itoh, H. Sanuki and S.-I. Itoh
Pole Dynamics for the Flierl-Petviashvili Equation and Zonal Flow
Sep. 2003

NIFS-784 R. Smirnov, Y. Tomita, T. Takizuka, A. Takayama, Yu. Chutov
Particle Simulation Study of Dust Particle Dynamics in Sheaths
Oct. 2003

NIFS-785 T.-H. Watanabe and H. Sugama
Kinetic Simulation of Steady States of Ion Temperature Gradient Driven Turbulence with Weak Collisionality
Nov. 2003

NIFS-786 K. Itoh, K. Hallatschek, S. Toda, H. Sanuki and S.-I. Itoh
Coherent Structure of Zonal Flow and Nonlinear Saturation
Dec. 2003

NIFS-787 S.I. Itoh, K. Itoh, M. Yagi and S. Toda
Statistical Theory for Transition and Long-time Sustainment of Improved Confinement State
Dec. 2003

NIFS-788 A. Yoshizawa, S.-I. Itoh, K. Itoh and N. Yokoi
Dynamos and MHD Theory of Turbulence Suppression
Dec. 2003

NIFS-789 V.D. Pustovitov
Pressure-induced Shift of the Plasma in a Helical System with Ideally Conducting Wall
Jan. 2004

NIFS-790 S. Koikari
Rooted Tree Analysis of Runge-Kutta Methods with Exact Treatment of Linear Terms
Jan. 2004