Runge-Kutta Methods for Linear Ordinary Differential Equations

David W. Zingg and Todd T. Chisholm
University of Toronto Institute for Aerospace Studies

RIACS Technical Report 97.07
July 1997
Runge-Kutta Methods
for Linear Ordinary Differential Equations

David W. Zingg and Todd T. Chisholm
University of Toronto Institute for Aerospace Studies

The Research Institute for Advanced Computer Science is operated by Universities Space Research Association, The American City Building, Suite 212, Columbia, MD 21044, (410)730-2656

Work reported herein was supported by the National Aeronautics and Space Administration under Contract NAS 2-13721 to the Universities Space Research Association (USRA).
RUNGE-KUTTA METHODS
FOR LINEAR ORDINARY DIFFERENTIAL EQUATIONS

D.W. ZINGG AND T.T. CHISHOLM

Abstract

Three new Runge-Kutta methods are presented for numerical integration of systems of linear inhomogeneous ordinary differential equations (ODEs) with constant coefficients. Such ODEs arise in the numerical solution of the partial differential equations governing linear wave phenomena. The restriction to linear ODEs with constant coefficients reduces the number of conditions which the coefficients of the Runge-Kutta method must satisfy. This freedom is used to develop methods which are more efficient than conventional Runge-Kutta methods. A fourth-order method is presented which uses only two memory locations per dependent variable, while the classical fourth-order Runge-Kutta method uses three. This method is an excellent choice for simulations of linear wave phenomena if memory is a primary concern. In addition, fifth- and sixth-order methods are presented which require five and six stages, respectively, one fewer than their conventional counterparts, and are therefore more efficient. These methods are an excellent option for use with high-order spatial discretizations.

Introduction

We consider the numerical integration of large linear inhomogenous systems of ordinary differential equations in the form

\[ \frac{du}{dt} = Au - g(t) \]  

where \( A \) is an \( M \) by \( M \) matrix whose elements depend on neither \( u \) nor \( t \), and \( u \) and \( g(t) \) are vectors of length \( M \). Such essentially autonomous systems arise in the numerical solution of partial differential equations (PDEs) governing linear wave phenomena after application of a spatial discretization such as a finite-difference, finite-volume, or finite-element method. Examples of such PDEs are the linearized Euler equations governing acoustic waves and the Maxwell equations governing electromagnetic waves. The elements of \( A \) depend on the PDE and the spatial discretization. The inhomogeneous term \( g(t) \) is associated with either a source term or the boundary conditions. In the context of wave propagation, the system of ODEs is
often mildly stiff with the eigenvalues of $A$ typically lying near the imaginary axis.

The system of ODEs arising from the application of a spatial discretization to a system of PDEs can be very large, especially in three-dimensional simulations. Consequently, the constraints on the methods used for integrating these systems are somewhat different from those which have driven much of the development of numerical methods for initial value problems. Due to their high accuracy and modest memory requirements, explicit Runge-Kutta methods have become popular for simulations of wave phenomena [5,6,7,15,17]. Third- and fourth-order methods requiring only two memory locations per dependent variable are particularly useful [3,13,14]. This property is easily achieved by a third-order Runge-Kutta method [14], but an additional stage is required for a fourth-order method [3]. Since the primary cost of the integration is in the evaluation of the derivative function, and each stage requires a function evaluation, the additional stage represents a significant increase in expense. For the same reason, error checking is generally not performed when solving very large systems of ODEs arising from the discretization of PDEs.

There have been several attempts to develop efficient methods for integrating linear systems of ODEs [4,9,10,11]. The basic premise of these methods is that the major cost in evaluating the derivative function is in forming the matrix $A$ and the vector $g(t)$. In the application considered here, the simulation of linear wave phenomena, the matrix $A$ is never explicitly formed or stored. Hence the methods previously proposed for linear systems are not appropriate for this application.

It is well known that a Runge-Kutta method with $p$ stages has an order of accuracy not exceeding $p$ [1,2]. For $p\leq 4$, methods of order $p$ can be derived with $p$ stages. However, fifth- and sixth-order methods require at least six and seven stages, respectively. Nine stages are required for seventh-order accuracy and eleven for eighth-order accuracy [1]. Since the cost for our application is roughly proportional to the number of stages, this represents a significant limitation of higher-order Runge-Kutta methods.

Several authors have considered various approximations to reduce the number of stages and the storage requirements of high-order Runge-Kutta methods. Shanks [12] was able to develop schemes with a reduced number of stages by requiring only that the accuracy conditions be approximately satisfied. Zingg et al. [16,17] propose methods with low storage requirements which are of high order for linear homogeneous ODEs but second-order otherwise. A similar idea was proposed previously by Lorenz [8].
In this paper, we develop Runge-Kutta methods specifically for linear ODEs with constant coefficients. By removing the constraints imposed by nonlinearity in the derivative function, high-order Runge-Kutta methods can be derived which are more efficient in some respect than the classical methods. In the next section, we present a fourth-order method which requires less memory than the classical fourth-order Runge-Kutta method. We then present fifth- and sixth-order methods requiring fewer derivative function evaluations per time step than fifth- and sixth-order Runge-Kutta methods applicable to nonlinear problems.

**General Form of an Explicit Runge-Kutta Method**

Without loss of generality, we consider the following scalar ODE:

\[
\frac{du}{dt} = f(t, u)
\]  

(2)

A general \(p\)-stage explicit Runge-Kutta method can be written as

\[
k_1 = f(t_n, u_n)
\]

\[
k_i = f(t_n + c_i h, u_n + h \sum_{j=1}^{i-1} a_{ij} k_j) \quad i = 2, \ldots, p
\]

(3)

\[
u_{n+1} = u_n + h \sum_{i=1}^{p} b_i k_i
\]

where \(h = \Delta t\) is the time step, \(t_n = nh\), and \(u_n\) is an approximation to \(u(t_n)\).

**Low-Storage Fourth-Order Method**

We consider first the case \(p = 4\). With the constraints

\[
c_2 = a_{21}
\]

\[
c_3 = a_{31} + a_{32}
\]

\[
c_4 = a_{41} + a_{42} + a_{43}
\]

(4)

there remain ten parameters. For fourth-order accuracy, there are eight conditions which must be satisfied. Four of these arise even for linear homogeneous constant-coefficient ODEs. A further three conditions must be met if the ODEs are inhomogeneous. The final condition is
associated with non-constant coefficients or nonlinearity. Therefore, fourth-order Runge-Kutta methods are a two-parameter family of which the classical method is a particular choice.

If we restrict our attention to linear constant-coefficient ODEs, the number of conditions is reduced to seven. These are

\[
\begin{align*}
\sum_{i=1}^{4} b_i &= 1 \\
\sum_{i=2}^{4} c_i b_i &= 1/2 \\
c_2 a_{32} b_3 + b_4 (c_2 a_{42} + c_3 a_{43}) &= 1/6 \\
c_2 a_{32} a_{43} b_4 &= 1/24 \\
\sum_{i=2}^{4} b_i c_i^2 &= 1/3 \\
\sum_{i=2}^{4} b_i c_i^3 &= 1/4 \\
b_3 c_2^2 a_{32} + b_4 (c_2^2 a_{42} + c_3^2 a_{43}) &= 1/12
\end{align*}
\]

The reduction in the number of conditions to be satisfied does not permit us to reduce the number of stages. However, we can obtain reduced storage requirements.

Following the approach of Wray [14], the requirement that only two memory locations be used imposes the following three constraints:

\[
\begin{align*}
b_1 &= a_{41} = a_{31} \\
b_2 &= a_{42}
\end{align*}
\]

With these constraints, only two memory locations are required for both the dependent variable and the value of the time derivative. Hence the method requires minimal storage even when compact or spectral methods are used for the spatial discretization. With the memory locations denoted A and B, the method proceeds as follows.

1. Initially, \( u_n \) is stored in A, and B is empty.
2. The term \( k_1 = f(t_n, u_n) \) is evaluated and stored in B.
3. The quantity $u_n + ha_{31}k_1$, is calculated and stored in A.

4. The quantity $u_n + ha_{21}k_1$ is calculated and stored in B.

5. The term $k_2 = f(t_n + c_2 h, u_n + ha_{21}k_1)$ is evaluated and stored in B.

6. The contents of the two memory locations are linearly combined to form $u_n + h (a_{31}k_1 + a_{32}k_2)$, which is stored in B.

7. With $a_{41} = a_{31}$, another linear combination gives $u_n + h (a_{41}k_1 + a_{42}k_2)$, which is stored in A.

8. The term $k_3 = f(t_n + c_3 h, u_n + h(a_{31}k_1 + a_{32}k_2))$ is evaluated and stored in B.

9. The contents of the two memory locations are linearly combined to form $u_n + h (a_{41}k_1 + a_{42}k_2 + a_{43}k_3)$, which is stored in B.

10. With $b_1 = a_{41}$ and $b_2 = a_{42}$, another linear combination gives $u_n + h (b_1 k_1 + b_2 k_2 + b_3 k_3)$, which is stored in A.

11. The term $k_4 = f(t_n + c_4 h, u_n + h(a_{41}k_1 + a_{42}k_2 + a_{43}k_3))$ is evaluated and stored in B.

12. The contents of the two memory locations are linearly combined to form $u_{n+1}$.

With the additional constraints imposed by the low-storage requirement, we are left with seven parameters to satisfy the seven conditions given in eq. (5). Although this system may possess more than one solution, the only solution we have found is

\[ a_{21} = c_2 = 0.69631521002413, \quad c_3 = 0.29441651741, \]
\[ c_4 = 0.82502163765, \quad b_1 = a_{41} = a_{31} = 0.07801567728325, \]
\[ a_{32} = 0.21640084013679, \quad b_2 = a_{42} = 0.04708870117112, \]
\[ a_{43} = 0.69991725920066, \quad b_3 = 0.47982272993855, \]
\[ b_4 = 0.39507289160708 \]

**Five-Stage Fifth-Order Method**

For the case $p=5$, we have, in addition to the constraints given in eq. (4), the following condition:
\[ c_5 = a_{51} + a_{52} + a_{53} + a_{54} \]  \hspace{1cm} (7)

Consequently, adding the fifth stage has produced five additional parameters for a total of fifteen. The coefficients must satisfy the following eleven conditions in order to produce fifth-order accuracy for linear constant-coefficient ODEs:

\[ \sum_{i=1}^{5} b_i = 1 \]

\[ \sum_{i=2}^{5} c_i b_i = 1/2 \]

\[ c_2 a_{32} b_3 + b_4 (c_2 a_{42} + c_3 a_{43}) + b_5 (c_2 a_{52} + c_3 a_{53} + c_4 a_{54}) = 1/6 \]

\[ c_2 a_{32} a_{43} b_4 + b_5 [c_2 (a_{42} a_{54} + a_{32} a_{53}) + c_3 a_{43} a_{54}] = 1/24 \]

\[ c_2 a_{32} a_{43} a_{54} b_5 = 1/120 \]

\[ \sum_{i=2}^{5} b_i c_i^2 = 1/3 \]

\[ \sum_{i=2}^{5} b_i c_i^3 = 1/4 \]

\[ b_3 c_2^2 a_{32} + b_4 (c_2^2 a_{42} + c_3^2 a_{43}) + b_5 (c_2^2 a_{52} + c_3^2 a_{53} + c_4^2 a_{54}) = 1/12 \]

\[ \sum_{i=2}^{5} b_i c_i^4 = 1/5 \]

\[ (b_4 a_{42} + b_3 a_{32} + b_5 a_{52}) c_2^3 + (b_5 a_{53} + b_4 a_{43}) c_3^3 + b_5 a_{54} c_4^3 = 1/20 \]

\[ b_5 [a_{54}(a_{42} c_2^2 + a_{43} c_3^2) + a_{53} a_{32} c_2^2] + b_4 a_{43} a_{32} c_2^2 = 1/60 \]

Thus a four-parameter family of solutions is obtained. Several different criteria can be applied in order to choose a method from this family. The following values have been found by minimizing the \( L_2 \) norm of a vector containing the coefficients of the method:
Six-Stage Sixth-Order Method

With $p=6$, the following condition must be satisfied in addition to the constraints given in eqs. (4) and (7):

\[ c_6 = a_{61} + a_{62} + a_{63} + a_{64} + a_{65} \]  

(9)

Therefore, there remain twenty-one free coefficients. The requirement of sixth-order accuracy for linear constant-coefficient ODEs produces the following sixteen conditions:

\[ \sum_{i=1}^{6} b_i = 1 \]

\[ \sum_{i=2}^{6} c_i b_i = 1/2 \]

\[ c_2 a_{32} b_3 + b_4 (c_2 a_{42} + c_3 a_{43}) + b_5 (c_2 a_{52} + c_3 a_{53} + c_4 a_{54}) \]

\[ + b_6 (c_2 a_{62} + c_3 a_{63} + c_4 a_{64} + c_5 a_{65}) = 1/6 \]

\[ c_2 a_{32} a_{43} b_4 + b_5 [c_2 (a_{42} a_{54} + a_{32} a_{53}) + c_3 a_{43} a_{54}] \]

\[ + b_6 [a_{65} (a_{54} c_4 + a_{53} c_3 + a_{52} c_2) + a_{64} (a_{43} c_3 + a_{42} c_2) + a_{63} a_{32} c_2] = 1/24 \]
Using the same criterion as for the fifth-order method, the following coefficients have been chosen from the five-parameter family of solutions to the above conditions (again possibly nonunique):
\[ a_{21} = c_2 = 0.15, \quad c_3 = 0.36, \quad c_4 = 0.57, \]
\[ c_5 = 0.75, \quad c_6 = 0.90, \quad a_{32} = 0.45818181818182, \]
\[ a_{42} = 0.09769454545455, \quad a_{52} = 0.10861879806510, \]
\[ a_{62} = 0.20874226393025, \quad b_2 = 0.24971305394585, \]
\[ a_{43} = 0.48766666666667, \quad a_{53} = 0.04655817933320, \]
\[ a_{63} = 0.12686271445897, \quad b_3 = 0.11278150363005, \]
\[ a_{54} = 0.44703799502007, \quad a_{64} = 0.02734417934727, \]
\[ b_4 = 0.35718962665957, \quad a_{65} = 0.37591957583530, \]
\[ b_5 = -0.00478351095633, \quad b_6 = 0.24659027402511 \]
\[ b_1 = 0.03850905269576 \]

with

\[ a_{31} = c_3 - a_{32}, \quad a_{41} = c_4 - a_{42} - a_{43}, \]
\[ a_{51} = c_5 - a_{52} - a_{53} - a_{54}, \]
\[ a_{61} = c_6 - a_{62} - a_{63} - a_{64} - a_{65} \]

**Stability Contours**

The stability contours of the three new methods are shown in Fig. 1. Satisfaction of the first four conditions in eq. (5) ensures that the new fourth-order method has the same stability contour as the classical fourth-order Runge-Kutta method. Similarly, the stability contours of the five-stage fifth-order method and the six-stage sixth-order method are uniquely defined and do not depend on which members of the respective families are selected.

Although the stable regions of the fifth- and sixth-order methods are somewhat larger than that of the fourth-order method, the increase is not sufficient to compensate for the cost of the additional stages. Therefore, the fourth-order method is a better choice if the time step is limited by stability considerations. The stable regions of the fifth- and sixth-order methods do not include the imaginary axis. Systems with pure imaginary eigenvalues are obtained when central differencing is applied to the spatial derivatives in partial differential equations governing wave
propagation phenomena with no physical dissipation, in the absence of boundary conditions. However, Zingg et al. [17] have demonstrated that by adding a small amount of numerical dissipation to the spatial discretization, stable schemes can be obtained using such methods. The amount of dissipation required is sufficiently low that the overall accuracy of the scheme is not compromised. The stability contour of the method successfully used in [7] for simulations of the propagation and scattering of electromagnetic waves is identical to that of the present sixth-order method.

Fourier Error Analysis

Using Fourier analysis we can determine the errors produced by an integration method when applied to a linear homogeneous ODE. Since our interest is in wave propagation, we consider a scalar ODE of the form

\[ \frac{du}{dt} = i\omega u \]  

(11)

where \( \omega \) is a real constant. The Runge-Kutta methods developed here produce a solution in the form

\[ u_n = \sigma^n u_0 \]  

(12)

where

\[ \sigma = \sum_{k=0}^{p} \frac{1}{k!} (i\omega)^k \]  

(13)

and \( p \) is the number of stages. The local amplitude and phase errors are determined from \( \sigma \) as follows

\[ er_a = |\sigma| - 1 \]  

(14)

\[ er_p = \frac{\tan^{-1}(\sigma_r/\sigma_i)}{\omega h} + 1 \]  

(15)

where \( \sigma_r \) and \( \sigma_i \) denote the real and imaginary parts of \( \sigma \).

Figs. 2 and 3 show the local amplitude and phase errors produced by the three new methods. In order to account for the number of stages, the errors are plotted versus \( \omega h / p \). Hence the errors shown are for approximately equal computational effort. Since the time step is thus proportional to \( p \), the amplitude error shown is \( |\sigma|^{1/p} - 1 \). The figures show that each increase in the order of the method produces an increase in accuracy even though the extra work has been accounted for.
Hence the fifth- and sixth-order methods can be more efficient than the fourth-order method if a sufficiently accurate spatial discretization is used.

\section*{An Example}

In order to demonstrate the validity and correctness of the above derivations, we apply the new methods and the classical fourth-order Runge-Kutta method to a sample inhomogeneous linear scalar ODE given by

\[
\frac{du}{dt} = \lambda u + ae^{\mu t}
\]  

(16)

The exact solution is

\[
u(t) = u(0)e^{\lambda t} + \frac{a(e^{\mu t} - e^{\lambda t})}{\mu - \lambda}
\]

(17)

when \(\mu \neq \lambda\). For the example, we use \(\lambda = -1 + 5i, \mu = i, a = 10,\) and \(u(0) = 1\). With these parameters, the exact solution at \(t = 2\) is \(-2.60430984499756 - 0.20741391939986i\). In the table below, the magnitude of the errors obtained at \(t = 2\) are compared for time step sizes ranging from 0.4 to 0.05. The classical fourth-order method is designated RK4, the new low-storage fourth-order method, RK4L, the new fifth-order method, RK5, and the new sixth-order method, RK6. In addition to showing the error magnitude, the table also shows the order of the method estimated from the error at the time step indicated and the next larger time step. The main results of the table are that the new methods approach the expected order of accuracy as the time step size is reduced and that the low-storage fourth-order method is as accurate as the classical method for inhomogeneous linear ODEs.

<table>
<thead>
<tr>
<th>h</th>
<th>RK4 error</th>
<th>order</th>
<th>RK4L error</th>
<th>order</th>
<th>RK5 error</th>
<th>order</th>
<th>RK6 error</th>
<th>order</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>0.3437</td>
<td>-</td>
<td>0.3323</td>
<td>-</td>
<td>0.2761</td>
<td>-</td>
<td>0.0664</td>
<td>-</td>
</tr>
<tr>
<td>0.2</td>
<td>0.0343</td>
<td>3.33</td>
<td>0.0340</td>
<td>3.29</td>
<td>0.0059</td>
<td>5.54</td>
<td>8.63e-4</td>
<td>6.27</td>
</tr>
<tr>
<td>0.1</td>
<td>0.0020</td>
<td>4.11</td>
<td>0.0020</td>
<td>4.10</td>
<td>1.70e-4</td>
<td>5.12</td>
<td>1.24e-5</td>
<td>6.12</td>
</tr>
<tr>
<td>0.05</td>
<td>1.19e-4</td>
<td>4.06</td>
<td>1.19e-4</td>
<td>4.05</td>
<td>5.11e-6</td>
<td>5.06</td>
<td>1.87e-7</td>
<td>6.06</td>
</tr>
</tbody>
</table>

Table 1. Sample computations using the new methods.
Conclusions

Three new Runge-Kutta methods have been presented for the integration of linear systems of ODEs with constant coefficients. If the time step size is limited by stability, then the new fourth-order method is the most suitable of the new methods. This method requires less memory than the classical fourth-order Runge-Kutta method and less computational effort than the low-storage method proposed in [3]. If the time step is limited by accuracy, and memory is a secondary concern, then the new fifth- and sixth-order methods present an efficient new alternative. Since the expense of the methods is roughly proportional to the number of stages for the problems of interest here, the new fifth- and sixth-order methods are significantly more efficient than their counterparts for nonlinear ODEs. The sixth-order method is a particularly good choice for use with high-order spatial discretizations.

References


Figure 1: Stability contours for the fourth-order (---), fifth-order (- - -), and sixth-order (· · ·) methods.

Figure 2: Amplitude error produced by the fourth-order (---), fifth-order (- - -), and sixth-order (· · ·) methods.
Figure 3: Phase error produced by the fourth-order (---), fifth-order (-- -), and sixth-order (···) methods.