111-02
9993
P-17

NASA Technical Memorandum 109111

Third-Order 2N-Storage Runge-Kutta Schemes with Error Control

Mark H. Carpenter
Langley Research Center, Hampton, Virginia

Christopher A. Kennedy
University of California, San Diego, La Jolla, California

(NASA-TM-109111) THIRD-ORDER 2N-STORAGE RUNGE-KUTTA SCHEMES WITH ERROR CONTROL (NASA. Langley Research Center) 17 p

N94-34204

Unclas

G3/02 0009993

June 1994

National Aeronautics and
Space Administration
Langley Research Center
Hampton, Virginia 23681-0001

THIRD-ORDER 2N-STORAGE RUNGE-KUTTA SCHEMES WITH ERROR CONTROL

Mark H. Carpenter *, Christopher A. Kennedy †

Abstract

A family of four-stage third-order explicit Runge-Kutta schemes is derived that requires only two storage locations and has desirable stability characteristics. Error control is achieved by embedding a second-order scheme within the four-stage procedure. Certain schemes are identified that are as efficient and accurate as conventional embedded schemes of comparable order and require fewer storage locations.

Section 1: Introduction

Runge-Kutta (RK) embedding techniques are an effective means of solving non-stiff ordinary differential equations (ODE's). (See references [1], [2], [3], [4] for examples of high-order RK schemes that utilize embedding.) Embedding utilizes two formulas of orders p and q ($p \neq q$) to calculate the evolution of the solution in time. By comparing the two solutions at each time step, an estimate of the temporal error can be determined and can be used to adjust the time step. For example, if $q = p + 1$, then the difference between the pth- and qth-order solutions provides a measure of the error committed in using formula p. The two solutions cannot be advanced without computational overhead; however, this overhead can be minimized by requiring both formulas to have similar coefficients and storage requirements. In a best-case scenario, no additional work or storage is necessary for the implementation of the embedded scheme. To date, most embedded schemes have been optimized in terms of accuracy and efficiency, with little regard to storage requirements.

For the ODE's that result from the semidiscretization of partial differential equations (PDE's) (fluid mechanics, for example), the overriding consideration is often the availability of fast memory.

*Aerospace Engineer, Theoretical Flow Physics Branch, NASA Langley Research Center, Hampton, VA 23681.
†Department of Applied Mechanics and Engineering Sciences, University of California, San Diego, La Jolla, CA 92093.

A numerical integration technique that minimizes memory storage is essential and can be formulated with a RK methodology. Williamson [5] showed that all second-order and some third-order explicit RK schemes can be cast in a 2N-storage format, where N is the dimension of the system of ODE's. He also showed that the four-stage fourth-order RK schemes cannot be put into 2N-storage format. Carpenter et al. [6] showed that a fourth-order 2N-storage scheme could be achieved with five stages (this scheme is abbreviated RK[5,4]-2N). In addition, they showed that tuned four-stage third-order schemes (RK[4,3]-2N) are significantly more efficient than conventional three-stage third-order schemes (RK[3,3]-2N) for certain problems.

The principal motivation of this work is to derive a family of schemes that requires 2N storage and that has the capability to monitor temporal integration error. Specifically, a four-stage third-order RK scheme that satisfies the 2N-storage constraint is sought (RK[4,3]-2N). To accomplish error control, a three-stage second-order 2N-storage RK scheme (RK[3,2]-2N) is embedded within the first three stages. (The resulting scheme is abbreviated RK[4,3(2)]-2N.) The family is then tuned so that a desirable stability envelope is achieved for the four-stage scheme. In section 2, we describe the conventional and the 2N-storage RK nomenclature. In section 3, we derive a new family of four-stage third-order 2N-storage schemes with an embedded formula. We then optimize the family for maximal stability characteristics. In section 4, the time-step control procedure is described. In section 5, the efficacy of the new schemes is demonstrated on several test problems, and then conclusions are drawn in regard to the utility of the new schemes.

Section 2: Runge-Kutta Nomenclature

We are concerned with the numerical solution of the initial value problem

$$\frac{dU}{dt} = F[t, U(t)]; \quad U(t_0) = U_0$$

Assume that the discrete approximation is made with an M-stage explicit RK scheme which includes an embedded scheme within the M-stage procedure. The implementation over a time step h is accomplished by

$$\begin{aligned}
k_1 &= F(t_0, U^0) \\
k_i &= F\left(t_0 + c_i h,\ U^0 + h\sum_{j=1}^{i-1} a_{i,j} k_j\right) \qquad\qquad i = 2, ..., M \\
\hat{U}^1 &= U^0 + h\sum_{j=1}^{M} \hat{b}_j k_j \\
U^1 &= U^0 + h\sum_{j=1}^{M} b_j k_j
\end{aligned}$$

where $U^0 = U(t_0)$ and $\hat{U}^1$ and U^1 are the solutions at time level $n+1$ of order p and q, respectively. The fixed scalars $a_{i,j}, b_j, \hat{b}_j, c_i$ are the coefficients of the RK formula and, for a four-stage third-order scheme, must satisfy the equations [7]

$$\sum_{i=1}^{4} b_i = 1 \quad ; \quad \sum_{i=1}^{4} b_i c_i = \frac{1}{2} \quad ; \quad \sum_{i=1}^{4} b_i c_i^2 = \frac{1}{3} \quad ; \quad \sum_{i,j=1}^{4} b_i a_{i,j} c_j = \frac{1}{6} \quad ; \quad \sum_{i=1}^{4} \hat{b}_i = 1 \quad ; \quad \sum_{i=1}^{4} \hat{b}_i c_i = \frac{1}{2} \tag{1}$$

and

$$c_i = \sum_{j=1}^{4} a_{i,j} \qquad i = 1,4 \tag{2}$$

The last two constraints in equation (1) ensure second-order accuracy of the three-stage embedded scheme.

To devise low-storage RK schemes, Williamson [5] exploits the technique of leaving useful information in the storage register. Each successive stage is written onto the same register without zeroing the previous value. Thus, the M-stage algorithm becomes

$$\begin{aligned} dU_j &= A_j dU_{j-1} + hF(U_j) \\ U_j &= U_{j-1} + B_j dU_j \qquad j = 1, M \end{aligned}$$

So that the algorithm is self-starting, $A_1 = 0$. Only the dU and U vectors must be stored, which results in a 2N-storage algorithm.

The following relations, first presented by Williamson, [5] describe the dependence between the 2N-storage variables A_j and B_j and the conventional RK variables $a_{i,j}, b_j$, and c_i:

$$\begin{aligned} B_j &= a_{j+1,j} & (j \neq M) \\ B_M &= b_M & \\ A_j &= (b_{j-1} - B_{j-1})/b_j & (j \neq 1, b_j \neq 0) \\ A_j &= (a_{j+1,j-1} - c_j)/B_j & (j \neq 1, b_j = 0) \end{aligned} \tag{3}$$

The precise values of A_j and B_j that are required to yield a higher order scheme remain to be determined.

In terms of the Butcher array (see reference [7] for details), the relationship between the conventional RK scheme and the 2N-storage RK scheme can be expressed as

$$\begin{array}{c|cccc} 0 & & & & \\ c_2 & a_{2,1} & & & \\ c_3 & a_{3,1} & a_{3,2} & & \\ c_4 & \hat{b}_1 & \hat{b}_2 & \hat{b}_3 & \\ \hline & b_1 & b_2 & b_3 & b_4 \end{array} \quad =$$

$$
\begin{array}{c|cccc}
0 & & & & \\
B_1 & B_1 & & & \\
B_1 + B_2(A_2+1) & A_2B_2 + B_1 & B_2 & & \\
B_1 + B_2(A_2+1) + B_3[A_3(A_2+1)+1] & A_2(A_3B_3 + B_2) + B_1 & A_3B_3 + B_2 & B_3 & \\
\hline
 & A_2[A_3(A_4B_4 + B_3) + B_2] + B_1 & A_3(A_4B_4 + B_3) + B_2 & A_4B_4 + B_3 & B_4
\end{array}
$$

Note that this form assumes four explicit stages, with the coefficients $a_{4,j}$, $j = 1,4$ replaced by $\hat{b}_j$, $j = 1,4$.

Section 3: Four-Stage Third-Order RK Schemes With Embedding

The solution to the four-stage third-order RK scheme is formed by solving eight equations in fourteen variables and is, in general, a six-parameter family. The variables $a_{i,j}, b_j, c_j$ are changed to A_j, B_j using the relationships defined in equations (3). Specifically, the values $a_{2,1}$, $a_{3,2}$, $\hat{b}_3$, b_4, b_3, b_2, and b_1 are expressed in terms of the values B_1, B_2, B_3, B_4, A_4, A_3, and A_2, respectively. By definition, the conditions $c_i = \sum_{j=1}^{4} a_{i,j}$ will automatically be satisfied. The only conditions that are not immediately satisfied involve $a_{3,1}$, $\hat{b}_1$, and $\hat{b}_2$. These three conditions provide three additional constraints on the system, and eliminate three of the six degrees of freedom. If a three-stage second-order scheme is embedded, then two additional constraints are provided and the number of degrees of freedom is reduced to one. A solution that involves one free parameter can be written in Butcher array form as

$$
\begin{array}{c|cccc}
0 & 0 & 0 & 0 & 0 \\
\frac{3c_3-2}{6c_3-3} & \frac{3c_3-2}{6c_3-3} & 0 & 0 & 0 \\
c_3 & -\frac{6c_3^2-8c_3+3}{6c_3-4} & \frac{3(2c_3-1)^2}{6c_3-4} & 0 & 0 \\
1 & -\frac{(2c_3-1)\left(12c_3^2-18c_3+7\right)}{(6c_3-4)(X)} & \frac{\left(18c_3^2-24c_3+9\right)(2c_3-1)^2}{(6c_3-4)(X)} & -\frac{c_3-1}{X} & 0 \\
\hline
 & \frac{(3c_3-1)\left(12c_3^2-18c_3+7\right)}{(18c_3-12)(Y)} & \frac{(3c_3-3)(2c_3-1)^2}{(6c_3-4)(Y)} & -\frac{3c_3-2}{(6c_3-6)(Y)} & \frac{c_3\left(12c_3^2-18c_3+7\right)}{(6c_3-6)(Y)}
\end{array}
$$

where $X = (12\,c_3{}^3 - 24\,c_3{}^2 + 16\,c_3 - 3)$ and $Y = (6\,c_3{}^2 - 6\,c_3 + 1)$. Written in 2N-storage form, the arrays A_j and B_j become

$$
\begin{array}{c|c}
A_1 & B_1 \\
A_2 & B_2 \\
A_3 & B_3 \\
A_4 & B_4
\end{array}
=
\begin{array}{c|c}
0 & \frac{3c_3-2}{6c_3-3} \\
-\frac{36c_3^3-48c_3^2+18c_3-1}{9(2c_3-1)^3} & \frac{3(2c_3-1)^2}{6c_3-4} \\
\frac{(9c_3-9)(2c_3-1)^3}{3c_3-2} & -\frac{c_3-1}{X} \\
-\frac{1}{X} & \frac{c_3\left(12c_3^2-18c_3+7\right)}{(6c_3-6)(Y)}
\end{array}
$$

with $c_3 \neq \frac{1}{2}, \frac{2}{3}, 1, \frac{\sqrt{(3)}\pm 3}{6}$, or $\frac{2-\sqrt[3]{\frac{5}{4}}}{3}$.

We now use efficiency and accuracy considerations to isolate specific values of c_3 that exhibit desirable characteristics. The linear stability of the four-stage third-order schemes is governed by

$\mathbf{G} = 1+\mathbf{Z}+\frac{1}{2}\mathbf{Z}^2+\frac{1}{6}\mathbf{Z}^3+\sum_{i,j,k=1}^{4} b_i a_{ij} a_{jk} c_k \mathbf{Z}^4$. (See Butcher [7] for details.) Because $\sum_{i,j,k=1}^{4} b_i a_{ij} a_{jk} c_k = -\frac{(2\,c_3-1)c_3\left(12\,c_3{}^2-18\,c_3+7\right)}{12XY}$ for all values of c_3, the linear stability envelope is found by solving $|\mathbf{G}| \leq 1$, where

$$\mathbf{G} = 1 + \mathbf{Z} + \frac{1}{2}\mathbf{Z}^2 + \frac{1}{6}\mathbf{Z}^3 - \frac{(2\,c_3-1)\,c_3\,\left(12\,c_3{}^2 - 18\,c_3 + 7\right)}{12XY}\mathbf{Z}^4$$

Carpenter et al. [6] reported that the four-stage third-order RK schemes (RK[4,3]) have desirable stability characteristics (for use with hyperbolic PDE's) for $\sum_{i,j,k=1}^{4} b_i a_{ij} a_{jk} c_k \leq \frac{1}{24}$, with optimal values near $\frac{1}{24}$. For values greater than $\frac{1}{24}$, the RK[4,3] schemes abruptly become unstable. For values less than $\frac{1}{24}$, a gradual decrease in stability occurs. Figure 1 illustrates this behavior by showing the stability of the scalar hyperbolic equation $u_t + u_x = 0$. The spatial derivative u_x is approximated with a periodic sixth-order compact spatial operator. The RK[4,3] schemes are used as the temporal discretization. In this test problem, the spatial eigenvalues are purely imaginary and are representative of the behavior of many spatial operators used for hyperbolic PDE's.

Note that a linear fourth-order scheme is recovered if $\sum_{i,j,k=1}^{4} b_i a_{ij} a_{jk} c_k = \frac{1}{24}$. Thus, the linear stability is identical to that encountered with four-stage fourth-order RK schemes. The value of c_3 for which $\sum_{i,j,k=1}^{4} b_i a_{ij} a_{jk} c_k = \frac{1}{24}$ is $c_3 = \frac{1+\sqrt[3]{\frac{5}{4}}}{3}$.

If rational numbers are desirable, then a scheme for which the stability is nearly identical to the linear fourth-order case is $c_3 = \frac{86}{125}$, or in Butcher form

$$\begin{array}{c|cccc}
0 & 0 & 0 & 0 & 0 \\
\frac{8}{141} & \frac{8}{141} & 0 & 0 & 0 \\
\frac{86}{125} & -\frac{5251}{2000} & \frac{6627}{2000} & 0 & 0 \\
1 & -\frac{271837625}{17364752} & \frac{347983777}{17364752} & \frac{609375}{1085297} & 0 \\
\hline
 & -\frac{615391}{215952} & \frac{258453}{71984} & -\frac{62500}{526383} & \frac{198961}{526383}
\end{array}$$

for which the 2N-storage array becomes

$$\begin{array}{c|c}
A_1 & B_1 \\
A_2 & B_2 \\
A_3 & B_3 \\
A_4 & B_4
\end{array}
=
\begin{array}{c|c}
0 & \frac{8}{141} \\
-\frac{756391}{934407} & \frac{6627}{2000} \\
-\frac{36441873}{15625000} & \frac{609375}{1085297} \\
-\frac{1953125}{1085297} & \frac{198961}{526383}
\end{array}$$

For this scheme, $\sum_{i,j,k=1}^{4} b_i a_{ij} a_{jk} c_k = \frac{1168895875}{29296507218} \approx \frac{1}{25.06}$. Note that this scheme is very close to the embedded four-stage third-order RK pair RK[4,3(2)]-3N proposed by Sharp et al. [8] (which is not a

2N-storage scheme):

$$
\begin{array}{c|cccc}
0 & 0 & 0 & 0 & 0 \\
\frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 \\
\frac{3}{5} & -\frac{387}{625} & \frac{762}{625} & 0 & 0 \\
1 & \frac{1}{5} & \frac{1}{5} & \frac{3}{5} & 0 \\
\hline
\hat{b} & \frac{1}{10} & \frac{2}{5} & \frac{1}{4} & \frac{1}{4} \\
b & -\frac{1}{90} & \frac{8}{15} & \frac{5}{18} & \frac{1}{5}
\end{array}
$$

for which $\sum_{i,j,k=1}^{4} b_i a_{ij} a_{jk} c_k = \frac{1143}{31250} \approx \frac{1}{27.34}$.

For linear problems, the optimal RK[4,3(2)]-2N scheme is the fourth-order case $c_3 = \frac{1+\sqrt[3]{\frac{5}{4}}}{3}$. The optimization of the accuracy for nonlinear problems is more difficult because the optimization is problem dependent. We can, however, use heuristic arguments to identify certain schemes as less desirable. Verner [9] and Prince et al. [4] cite several theoretical considerations that should be used in determining desirable high-order RK schemes and schemes with embedding. Those theoretical considerations that are relevant to this work are

I Each intermediate time level ($c_i, i = 1,4$) should be in the interval [0,1] to control the effect of rapidly changing derivatives.

II To minimize roundoff error, the $c_i, i = 1,4$ should be reasonably distinct to avoid large b_i and $a_{i,j}$ values.

III The weights of the $b_j, j = 1,4$ should be positive.

IV Coefficients should incorporate rational numbers that require a small number of digits.

V The leading-order truncation terms $||\tau^{q+1}||$ should be small.

VI The leading-order truncation terms of the embedded scheme $||\hat{\tau}^{p+1}||$ should dominate its error.

VII None of the leading-order truncation terms of the embedded scheme $||\hat{\tau}^{p+1}||$ should vanish, which ensures that each will contribute to the local error estimate δ_{n+1}.

The first four conditions are relevant to the accuracy of any RK scheme. The last three pertain to embedded schemes.

In the RK[4,3(2)]-2N family of schemes, no values of $0 \leq c_3 \leq 1$ exist for which $0 \leq b_j \leq 1$ for all values of $j = 1,4$. Constraint I and IV, therefore, can not be satisfied simultaneously. For reasons of stability and accuracy, all practical RK[4,3(2)]-2N schemes will have values of $c_3 \approx \frac{1+\sqrt[3]{\frac{5}{4}}}{3}$. The coefficient matrices $a_{i,j}$, b_j, and c_j are well behaved for these values of c_3, and the truncations terms

τ^{q+1} and $\hat{\tau}^{p+1}$ have satisfactory properties. The coefficient matrices become ill-conditioned near the singular values $c_3 \approx \frac{1}{2}, \frac{2}{3}, 1, \frac{\sqrt{(3)}\pm 3}{6}$, and $\frac{2-\sqrt[3]{\frac{5}{4}}}{3}$. Restrictive stability domains preclude all these values for c_3, except the point $c_3 \rightarrow \frac{2}{3}$. Values of c_3 near $\frac{2}{3}$ should not be used, in spite of the acceptable stability domain.

Section 4: Time-Step Control

For many systems of ODE's that result from the semidiscretization of PDE's, the temporal error does not need to be monitored at each time step because the maximum stable time step yields reasonable levels of temporal accuracy. Under these circumstances, the prominent concern in choosing a time-advancement scheme is the efficiency of the scheme. (Efficiency is defined as the absolute size of the stability envelope relative to the number of stages of the scheme.) Carpenter et al. [6] demonstrated that the efficiency of the RK[4,3]-2N schemes was as good or better than the RK[3,3]-2N schemes. The required work per time step is increased by one third, but the increased stability domain more than offsets the increased computational cost per step. This result is consistent with those obtained by Sharp et al. [8] for other high-order RK schemes. The embedded RK[4,3(2)]-2N schemes have the same stability envelope as the RK[4,3]-2N scheme and will, therefore, exhibit the same stability advantage over the RK[3,3]-2N schemes.

Many nonstiff equations exist for which the time step must be determined by the solution error and not by stability considerations. Embedded RK schemes provide a means of adjusting the time step during the calculation to achieve a desired temporal error level. The quantity $\delta^{n+1} = U^1 - \hat{U}^1$ is the local error at time $n + 1$ in the pth- order formula. Given δ^{n+1}, the widely used formula quoted by Hull et al. [10] $h_{n+1} = \kappa\, h_n \left\{\frac{\epsilon}{||\delta_{n+1}||_\infty}\right\}^{\frac{1}{p+1}}$, can be used to adjust the time step to control the error per step, where ϵ is the solution error tolerance. (For the schemes proposed here, $p = 2$ and $\kappa = 0.95$ is a satisfactory constant.)

An advantage of the RK[4,3(2)]-2N schemes is that $\delta^{n+1} = U^1 - \hat{U}^1$ is available without the additional expense of computing it and it does not require any additional storage. Specifically, the estimate $\delta^{n+1} = U^1 - \hat{U}^1 = B_4 dU_4$. This results from the fact that the scheme is second- and third-order accurate after the third and fourth stages, respectively. The difference between the two solutions must be the leading-order error term in the second-order formula (assuming the scheme satisfies condition VI listed above).

A disadvantage is that the time step can be adjusted only after the error has been committed because of the 2N-storage constraint. For conventional embedded RK schemes, if the prescribed error tolerance is exceeded, then the entire step is repeated with a smaller time step until the specified tolerance is met. This option is not available for the 2N-storage scheme because the original solution vector at time level n is overwritten by the intermediate solution vectors. The subsequent time step can be adjusted, but the error incurred during the "failed" step is accumulated into the solution vector for all time. In practice, adjustment of the error tolerance to a lower level eliminates this problem at the expense of more computational cost.

Section 5: Accuracy and Efficiency of RK[4,3(2)] Embedded RK Schemes

Three test problems are used to compare the new embedded RK[4,3(2)]-2N schemes with error control. The first is a nonlinear ODE used by Dormand et al. [3] to test the accuracy of various RK schemes. The ODE is defined by $y' = y\cos(x)$, $y(0) = 1$ on the interval $0 \leq x \leq 20$, with the exact solution $y(x) = \exp^{\sin(x)}$. Figure 2 shows a convergence study for the following schemes: the second-order RK[3,2]-2N (the embedded scheme in RK[4,3(2)]-2N), Williamson's [5] RK[3,3]-2N, the embedded RK[4,3(2)]-2N(a) $\left(c_3 = \frac{1+\sqrt[3]{\frac{5}{4}}}{3}\right)$, the embedded RK[4,3(2)]-2N(b) ($c_3 = \frac{86}{125}$), and the classical fourth-order RK[4,4]-3N scheme. All schemes approach the exact solution at their theoretical rate. Note that the Williamson RK[3,3]-2N scheme is nearly indistinguishable from the two embedded RK[4,3(2)]-2N(a) and RK[4,3(2)]-2N(b) schemes. (Williamson claims the RK[3,3]-2N scheme is optimal in terms of error.) This study verifies the nonlinear accuracy of the newly developed 2N-storage schemes for ODE's. Although the third-order scheme exhibits comparable accuracy, the Williamson RK[3,3]-2N scheme requires one-third fewer function evaluations than the RK[4,3(2)]-2N schemes but has a considerably smaller stability envelope.

The second problem is from the class D orbit equations used by Hull et al. [10] to test the accuracy and efficiency of ODE solvers. The problem is defined by

$$\begin{aligned}
\frac{dy_1}{dt} &= y_3, & y_1(0) &= 1-\epsilon, \\
\frac{dy_2}{dt} &= y_4, & y_2(0) &= 0 \\
\frac{dy_3}{dt} &= -\frac{y_1}{r^3}, & y_3(0) &= 0 \\
\frac{dy_4}{dt} &= -\frac{y_2}{r^3}, & y_4(0) &= \sqrt{\frac{1+\epsilon}{1-\epsilon}}
\end{aligned}$$

with $r^2 = y_1^2 + y_2^2$. The exact solution is

$$\begin{aligned}
y_1 &= \cos(u) - \epsilon, & y_2 &= \sqrt{1-\epsilon^2}\sin(u) \\
y_3 &= \frac{-\sin(u)}{1-\epsilon\cos(u)}, & y_4 &= \frac{\sqrt{1-\epsilon^2}\cos(u)}{1-\epsilon\cos(u)}
\end{aligned}$$

with $u - \epsilon\sin(u) - t = 0$. The eccentricity of the orbit is governed by ϵ and becomes singular for values of $\epsilon \to 1$. A value of $\epsilon = 0.9$ provides a severe test of an embedded scheme's error control capabilities.

We begin by testing the error control features in the 2N-storage scheme. The conventional RK schemes reintegrates a rejected time step by using the solution information stored at time level n; the

2N-storage schemes do not have this capability. The effects of this error accumulation on the long-time accuracy of the solution was tested by comparing the RK[4,3(2)]-2N(a) scheme in low-storage and conventional RK format. In the conventional format, the rejected time steps were reintegrated, to provide a direct comparison between the two approaches. The results on the orbit problem indicate that the accumulation of error could be controlled by decreasing the constant κ in the error control expression. For values of $\kappa \leq 0.9$, the percentage of rejected steps becomes insignificant, and both methods give the same accuracy and efficiency.

As a final test on the orbit problem, the new embedded RK[4,3(2)]-2N(a) scheme was compared with the embedded RK[4,3]-3N scheme reported by Sharp et al. [8] Figure 3 shows a logarithmic comparison of the error versus the number of time steps for each method. Both schemes have comparable absolute stability envelopes, and the efficiency for a given accuracy is slightly better for the new 2N-storage scheme.

The last problem is the solution of the linear hyperbolic equation defined by

$$\begin{aligned} u_t + u_x &= 0, \quad 0 \leq x \leq 1, t \geq 0 && (4) \\ u(0,t) &= \sin 2\pi(-t), \quad t \geq 0 && (5) \\ u(x,0) &= \sin 2\pi(x), \quad 0 \leq x \leq 1 && (6) \end{aligned}$$

The exact solution is

$$u(x,t) = \sin 2\pi(x-t), \quad 0 \leq x \leq 1, t \geq 0$$

and is a model for the class of problems that the embedded RK[4,3(2)]-2N schemes were developed to integrate.

Equations (4) through (6) are solved with four different embedded RK[4,3(2)]-2N schemes; specifically, the cases $c_3 = \frac{1+\sqrt[3]{\frac{5}{4}}}{3}$, $c_3 = \frac{432}{625}$, $c_3 = \frac{86}{125}$, and $c_3 = \frac{62}{100}$. In all cases, the spatial operator used is the sixth-order compact scheme developed by Carpenter et al. [11] and shown to be formally sixth-order accurate. The physical boundary condition is imposed by solving the differentiated boundary condition on the boundary with the RK procedure. This technique was shown by Carpenter et al. [12] to yield a fourth-order temporally accurate procedure. Specifically, the boundary condition is $d^3u(0,t)/dt^3 = g'''(t)$, where g is the physical boundary condition at the inflow plane. The CFL's that govern the stability of the hyperbolic problem range from $\sqrt{2}$ for $c_3 = \frac{1+\sqrt[3]{\frac{5}{4}}}{3}$ to 1.16 for $c_3 = \frac{62}{100}$.

After grid refinement with a vanishingly small CFL, all schemes recover the theoretical spatial sixth-order accuracy. The leading-order error terms for values of the CFL near the theoretical maximum are dominated by the temporal error components. On a specific grid, temporal refinement showed third-order temporal accuracy. Fourth-order temporal accuracy was obtained for the specific value $c_3 = \frac{1+\sqrt[3]{\frac{5}{4}}}{3}$. Table 1 shows the results from a grid-refinement study performed with a CFL of 1. Only the cases $c_3 = \frac{1+\sqrt[3]{\frac{5}{4}}}{3}$ and $c_3 = \frac{62}{100}$ are shown because they bracket the behavior of the other two schemes.

Table 1. Grid Refinement for Embedded RK[4,3(2)]-2N Schemes with CFL = 1

	$c_3 = \frac{1+\sqrt[3]{\frac{5}{4}}}{3}$				$c_3 = \frac{62}{100}$			
Grid	$\log L_2$	Rate	$\log L_\infty$	Rate	$\log L_2$	Rate	$\log L_\infty$	Rate
41	-3.695		-3.022		-2.815		-2.884	
81	-4.888	3.96	-3.931	3.01	-3.745	3.09	-3.786	3.00
161	-6.092	4.00	-4.836	3.00	-4.660	3.04	-4.689	3.00
321	-7.296	4.00	-5.739	3.00	-5.567	3.01	-5.592	3.00
641	-8.501	4.00	-6.642	3.00	-6.472	3.01	-6.495	3.00

The $\log L_2$ column represents the logarithm base 10 of the L_2 solution error, and the L_∞ column represents the maximal error in the solution as calculated by the embedded scheme. For the value $c_3 = \frac{1+\sqrt[3]{\frac{5}{4}}}{3}$, the embedded scheme overpredicts the solution error on this linear problem. For the values $c_3 \neq \frac{1+\sqrt[3]{\frac{5}{4}}}{3}$, a direct correlation between the solution error and the predicted error from the embedded scheme.

Table 2 shows the results from a temporal refinement study on a grid of 161 points. Two values of c_3 are used to show the trends of the study. Fourth-order temporal accuracy was obtained for the specific value $c_3 = \frac{1+\sqrt[3]{\frac{5}{4}}}{3}$. Third-order temporal accuracy was observed for $c_3 \neq \frac{1+\sqrt[3]{\frac{5}{4}}}{3}$.

Table 2. Temporal Refinement of Embedded RK[4,3(2)]-2N Schemes

	$c_3 = \frac{1+\sqrt[3]{\frac{5}{4}}}{3}$				$c_3 = \frac{62}{100}$			
CFL	$\log L_2$	Rate	$\log L_\infty$	Rate	$\log L_2$	Rate	$\log L_\infty$	Rate
1	-6.169		-4.836		-4.751		-4.689	
$\frac{1}{2}$	-7.353	3.93	-5.739	3.00	-5.653	3.00	-5.593	3.00
$\frac{1}{4}$	-7.733	1.26	-6.642	3.01	-6.543	2.96	-6.500	3.00
$\frac{1}{8}$	-7.731	-	-7.546	2.99	-7.331	2.62	-7.399	3.00
$\frac{1}{16}$	-7.730	-	-8.448	3.00	-7.677	1.15	-8.301	3.00
$\frac{1}{32}$	-7.730	-	-9.352	3.00	-7.724	0.16	-9.205	3.00

Note that the solution error asymptotes to a value dictated by the spatial error terms; the embedded error converges at a rate of at least three (independent of the CFL used). This study shows that for the cases in which $c_3 \neq \frac{1+\sqrt[3]{\frac{5}{4}}}{3}$ temporal error can be controlled by monitoring the embedded error. When the embedded scheme with a value $c_3 = \frac{1+\sqrt[3]{\frac{5}{4}}}{3}$ is used on linear problems, the temporal error will be overpredicted.

These three test problems demonstrate the efficacy of the embedded RK[4,3(2)]-2N schemes for linear and nonlinear ODE's for which time control is important. The behavior of the new schemes is similar to other embedded RK schemes of comparable order that exist in the literature. If the overriding concern in the choice of integrator is the reduction of storage and temporal error control is necessary, then the newly developed embedded RK[4,3(2)]-2N schemes are clearly advantageous over existing 2N-storage schemes.

Conclusions

A class of new four-stage third-order RK[4,3(2)]-2N Runge-Kutta (RK) schemes are derived that require only two storage locations. The class has a three-stage second-order 2N-storage RK scheme embedded within the first three stages. A comparison of the second- and third-order solutions can give an estimate of the temporal error at each time step. The subsequent time step can then be adjusted to achieve a desired error control. A particular scheme is identified that has the desirable efficiency characteristics for hyperbolic and parabolic initial (boundary) value problems. In the inviscid and viscous limits, this new RK[4,3(2)]-2N storage scheme has comparable accuracy for a given step size and has a larger allowable stability domain than the RK[3,3]-2N scheme advocated by Williamson. The absolute stability of the new schemes is comparable to that achieved with the classical four-stage fourth-order RK scheme. Numerical tests are presented that verify these results on nonlinear ordinary differential equations (ODE's) and linear hyperbolic equations.

Acknowledgments

The second author would like to acknowledge financial support provided while in residence at NASA Langley Research Center, Theoretical Flow Physics Branch, under contract NAG-1-1193.

References

[1] Fehlberg, E. "Classical fifth, sixth, seventh and eight order Runge-Kutta formulas with stepsize control", NASA TR R-287, (1968).

[2] Shampine, L.F., and Watts, H.A. : "Global error estimation for ordinary differential equations", ACM TOMS, 2, No. 2, (1976), pp. 172-186.

[3] Dormand, J.R., and Prince, P.J., "A family of embedded Runge-Kutta formulae," *J. Comp. & Appl. Math.*, **6**, 1, (1980), pp. 19-26.

[4] Prince, P.J. and Dormand, J.R., "High order embedded Runge-Kutta formulae," *J. Comp. & Appl. Math.*, **7**, 1, 67 (1981), pp. 67-75.

[5] Williamson, J.H., "Low-storage Runge-Kutta schemes," *J. Comp. Phys.*, **35**, 48 (1980), pp. 48-56.

[6] Carpenter, M.H., and Kennedy, C.A., "Fourth-Order 2N-Storage Runge-Kutta Schemes," submitted to *SIAM J. Sci. Comp.*, Feb, (1994).

[7] Butcher, J.C., *The Numerical Analysis of Ordinary Differential Equations: Runge-Kutta and General Linear Methods*, John Wiley & Sons, Chichester (1987).

[8] Sharp, P.W. and Smart, E., "Explicit Runge-Kutta Pairs with One More Derivative Evaluation than the Minimum," *SIAM J. Sci. Comp.*, **14**, 2 (1993), pp. 338-348.

[9] Verner, J.H., "Explicit Runge-Kutta Methods with Estimates of the Local Truncation Error," *SIAM J. Numer. Anal.*, 15 (1978), pp. 772-790.

[10] Hull, T.E., Enright, W.H., Fellen, B.M., and Sedgwick, A.E. "Comparing numerical methods for Ordinary Differential Equations", *SIAM J. Numer. Anal.*, **9**, No. 4, 603-637, 1972.

[11] Carpenter, M.H., Gottlieb, D., and Abarbanel, S., "The Stability of Numerical Boundary Treatments for Compact High-Order Finite-Difference Schemes," *J. Comp. Phy*, **108**, 2, (1993), pp. 272-295.

[12] Carpenter, M.H., Gottlieb, D., Abarbanel, S. and Don, W.-S. " The Theoretical Accuracy of Runge-Kutta Time Discretizations for the Initial Boundary Value Problem: A Careful Study of the Boundary Error," NASA CR-191561, ICASE Report No. 93-83, (1993). Submitted to *SIAM J. Numer. Anal.*.

FIGURE 1. CFL dependence on parameter $\frac{1}{\alpha}$ for RK[4,3] schemes where $\alpha = \sum_{i,j,k=1}^{4} b_i a_{ij} a_{jk} c_k$.

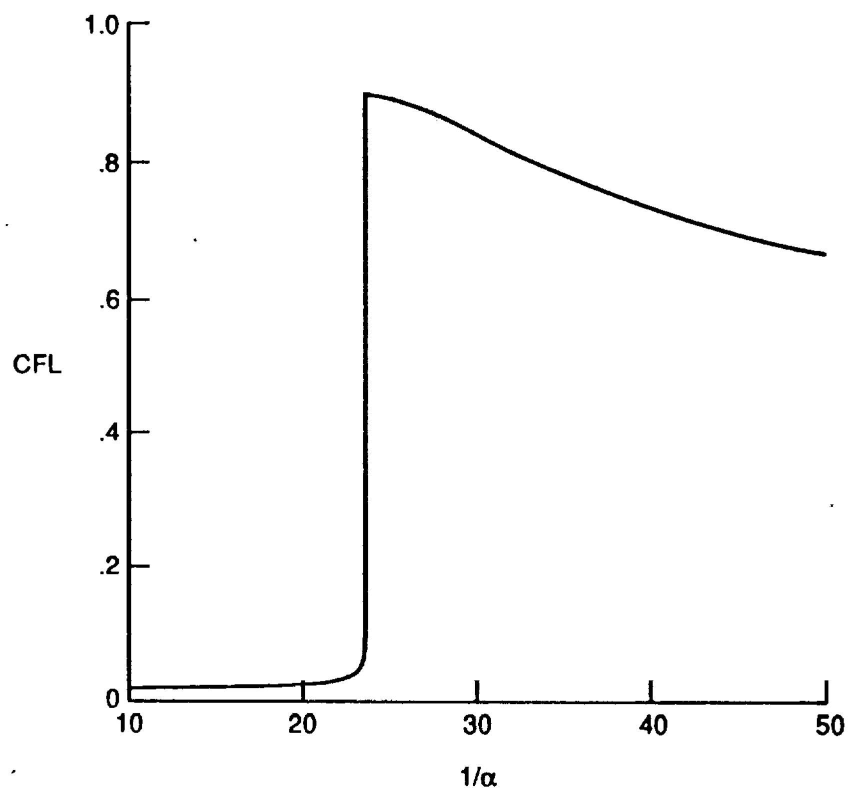


FIGURE 2. Comparison of convergence between RK[4,3(2)]-2N schemes and conventional RK[3,3]-2N and RK[4,4]-3N schemes. Problem is ODE defined by $y' = y\ \cos(x)$.

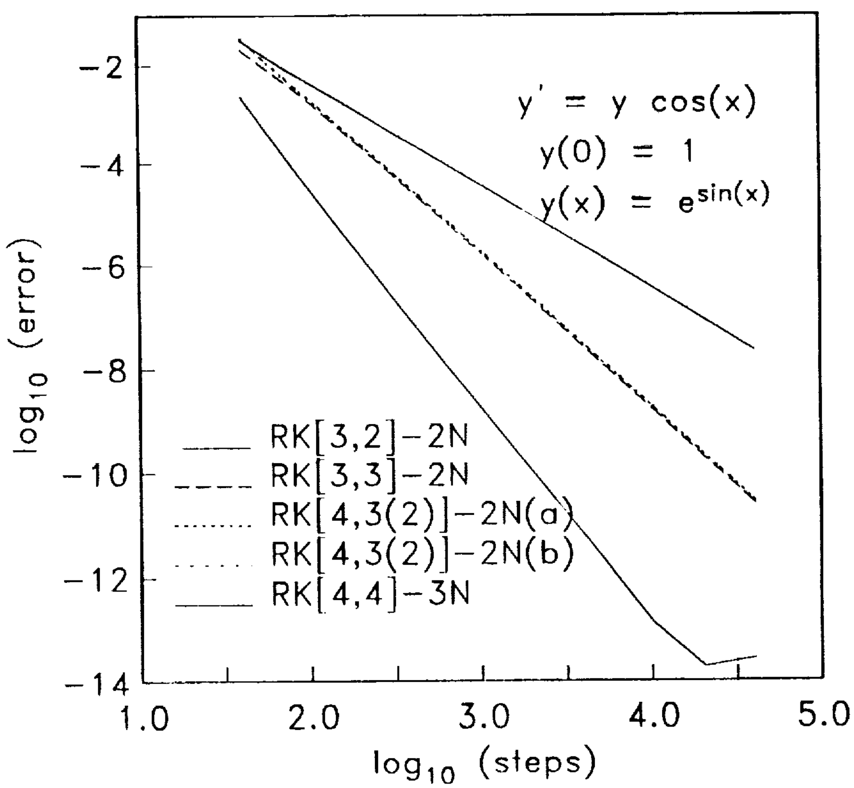
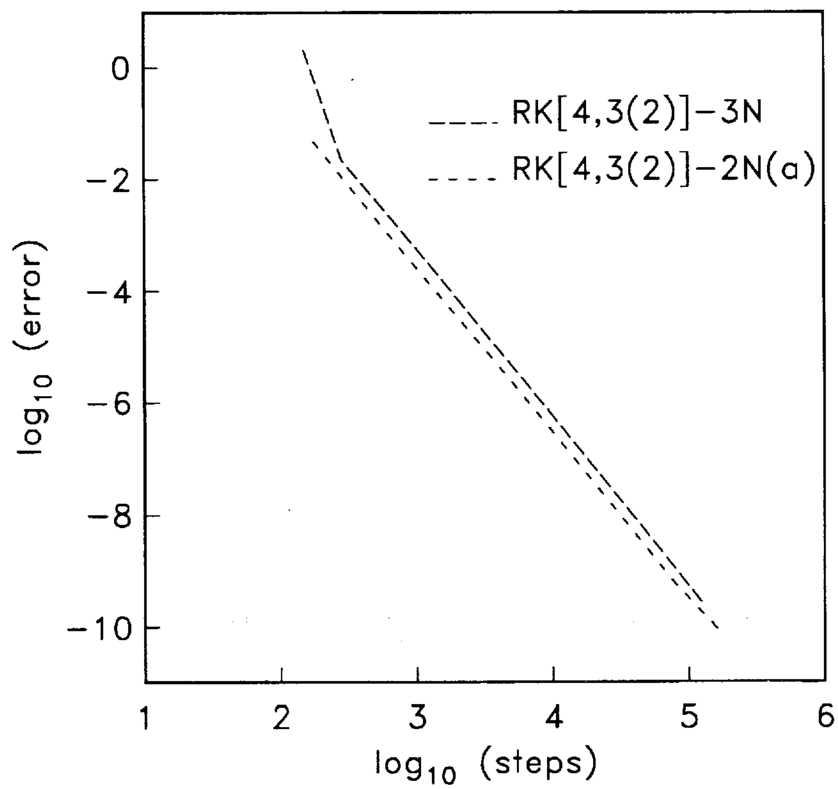


FIGURE 3. Comparison of convergence between RK[4,3(2)]-2N scheme and existing RK[4,3(2)]-3N scheme on orbit problem.



REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	June 1994	Technical Memorandum

4. TITLE AND SUBTITLE
Third-Order 2N-Storage Runge-Kutta Schemes with Error Control

5. FUNDING NUMBERS
505-70-62-13

6. AUTHOR(S)
Mark H. Carpenter
Christopher A. Kennedy

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
NASA Langley Research Center
Hampton, VA 23681-0001

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration
Langley Research Center
Hampton, VA 23681

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
NASA TM-109111

11. SUPPLEMENTARY NOTES
Carpenter: Langley Research Center, Hampton, VA.
Kennedy: University of California at San Diego, LaJolla, CA.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Unclassified - Unlimited

Subject Category: 02

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
A family of four-stage third-order explicit Runge-Kutta schemes is derived that requires only two storage locations and has desirable stability characteristics. Error control is achieved by embedding a second-order scheme within the four-stage procedure. Certain schemes are identified that are as efficient and accurate as conventional embedded schemes of comparable order and require fewer storage locations.

14. SUBJECT TERMS
Runge-Kutta, low-storage, error control, embedding

15. NUMBER OF PAGES
16

16. PRICE CODE
A03

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited