University of Amsterdam

## UvA-DARE (Digital Academic Repository)

## Approximating Runge-Kutta matrices by triangular matrices

Hoffmann, W.; de Swart, J.J.B.
DOI
10.1007/BF02510217

Publication date
1997
Document Version
Final published version
Published in
Bit : numerical mathematics

Link to publication

Citation for published version (APA):
Hoffmann, W., \& de Swart, J. J. B. (1997). Approximating Runge-Kutta matrices by triangular matrices. Bit : numerical mathematics , 37(2), 346-354. https://doi.org/10.1007/BF02510217

## General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

## Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

# APPROXIMATING RUNGE-KUTTA MATRICES BY TRIANGULAR MATRICES * 

W. HOFFMANN ${ }^{1}$ and J. J. B. DE SWART ${ }^{2} \dagger$<br>${ }^{1}$ Department of Mathematics and Computer Science, University of Amsterdam<br>Kruislaan 403, 1098 SJ Amsterdam, The Netherlands. email: walter@fwi.uva.nl<br>${ }^{2}$ Department of Numerical Mathematics, CWI, P.O. Box 94079<br>1090 GB Amsterdam, The Netherlands. email: jacques@cwi.nl


#### Abstract

. The implementation of implicit Runge-Kutta methods requires the solution of large systems of non-linear equations. Normally these equations are solved by a modified Newton process, which can be very expensive for problems of high dimension. The recently proposed triangularly implicit iteration methods for ODE-IVP solvers [5] substitute the Runge-Kutta matrix $A$ in the Newton process for a triangular matrix $T$ that approximates $A$, hereby making the method suitable for parallel implementation. The matrix $T$ is constructed according to a simple procedure, such that the stiff error components in the numerical solution are strongly damped. In this paper we prove for a large class of Runge-Kutta methods that this procedure can be carried out and that the diagonal entries of $T$ are positive. This means that the linear systems that are to be solved have a non-singular matrix.


AMS subject classification: Primary: 65L06, Secondary: 15A23.
Key words: Numerical analysis, Runge--Kutta methods, Matrix analysis.

## 1 Introduction and motivation.

For solving the stiff initial value problem

$$
y^{\prime}(t)=f(t, y(t)), \quad y\left(t_{0}\right)=y_{0}, \quad y, f \in \mathbf{R}^{d}, \quad t_{0} \leq t \leq t_{e}
$$

one of the most powerful methods is an implicit Runge-Kutta (RK) method. In such a method we have to solve every time step a system of non-linear equations of the form

$$
\begin{equation*}
R\left(Y_{n}\right)=0 ; \quad R\left(Y_{n}\right):=Y_{n}-(e \otimes I) y_{n-1}-h_{n}(A \otimes I) F\left(Y_{n}\right) \tag{1.1}
\end{equation*}
$$

where $A$ denotes the $s \times s$ matrix containing the parameters of the $s$-stage RK method, $y_{n-1}$ the approximation to $y\left(t_{n-1}\right), e$ is the $s$-dimensional vector with unit entries, $I$ is the $d \times d$ identity matrix, $h_{n}$ is the step size $t_{n}-t_{n-1}$ and $\otimes$

[^0]denotes the Kronecker product. The $s$ components $Y_{n, i}$ of the $s d$-dimensional solution vector $Y_{n}$ represent $s$ numerical approximations to the $s$ exact solution vectors $y\left(t_{n-1}+c_{i} h_{n}\right)$; here, $c$ denotes the abscissa vector and $i$ ranges from 1 to $s$. Furthermore, for any vector $X=\left(X_{i}\right), F(X)$ contains the derivative values $\left(f\left(X_{i}\right)\right)$. It is assumed that the components of $c$ are distinct and positive.

Once we have solved (1.1), we obtain the step point value $y_{n} \approx y\left(t_{n}\right)$ by the formula

$$
y_{n}=y_{n-1}+h_{n}\left(b^{T} \otimes I\right) F\left(Y_{n}\right)
$$

where $b$ is a vector of dimension $s$ containing method parameters.
To solve (1.1), in general one uses a Newton-type iteration scheme of the form

$$
\begin{equation*}
\left(I-B \otimes h_{n} J_{n}\right) \Delta Y_{n}^{(j+1)}=-R\left(Y_{n}^{(j)}\right) ; Y_{n}^{(j+1)}=Y_{n}^{(j)}+\Delta Y_{n}^{(j+1)} \tag{1.2}
\end{equation*}
$$

where $J_{n}$ is an approximation to the Jacobian of the right hand side function $f$ at $t_{n-1}, Y_{n}^{(0)}$ is the initial iterate to be provided by some predictor formula and $B$ is an $s \times s$ matrix that defines the type of Newton iteration. To get insight in the convergence behaviour of (1.2), we apply the scheme to the scalar test equation $y^{\prime}=\lambda y$. Defining the iteration error $\epsilon_{n}^{(j)}$ by $Y_{n}^{(j)}-Y_{n}$, we see from (1.1) and (1.2) that these errors are amplified by the matrix $Z$ defined by

$$
Z(z)=z(I-z B)^{-1}(A-B) ; \quad z:=\lambda h_{n} .
$$

We introduce the stiff and non-stiff amplification matrices of scheme (1.2), notation $Z_{\infty}(B)$ and $Z_{0}(B)$, respectively, by

$$
Z_{\infty}(B):=\lim _{|z| \rightarrow \infty} Z(z)=I-B^{-1} A
$$

and

$$
Z_{0}(B):=\lim _{|z| \rightarrow 0} \frac{Z(z)}{|z|}=A-B
$$

Choosing $B=A$ would lead to the modified Newton process, for which $Z(z)=$ 0 for all $z$. However, the computation of $Y_{n}^{(j)}$ now requires the solution of a linear system of dimension $s d$. For high-dimensional problems this requires a lot of computational effort. Several attempts have been made to reduce these costs by selecting matrices $B$ different from $A$.

In [1], Cooper and Butcher propose the choice $B=P$, where $P$ is a matrix that has a one-point spectrum. By performing a similarity transformation to (1.2) they arrive at the scheme

$$
\begin{align*}
P Q & =Q L  \tag{1.3}\\
\left(I-L \otimes h_{n} J_{n}\right) \Delta X_{n}^{(j+1)} & =-\left(Q^{-1} \otimes I\right) R\left(Y_{n}^{(j)}\right) \\
Y_{n}^{(j+1)} & =Y_{n}^{(j)}+(Q \otimes I) \Delta X_{n}^{(j+1)}
\end{align*}
$$

where $L$ and $Q$ are lower triangular and orthogonal matrices, respectively, that define the Schur decomposition of $P$. Since the diagonal entries of $L$ are equal, implementing (1.3) requires only one $L U$-decomposition of dimension $d$.

In [4], the authors select $B=D$, where $D$ is a diagonal matrix. Iteration scheme (1.2) is now suitable for implementation on an $s$ processor machine, since the $s$ components of $Y_{n}^{(j)}$ can be computed independently. The matrix $D$ is constructed such that $\rho\left(Z_{\infty}(D)\right)=0$, where $\rho(\cdot)$ denotes the spectral radius function. This method was called PDIRK, Parallel Diagonal-implicit Iterated Runge-Kutta.

Recently, in [5], a mixture of the two strategies described above was presented and given the name PTIRK, Parallel Triangularly-implicit Iterated RungeKutta. Here, the matrix $B$ was identified with a lower triangular matrix $T$ such that $A=T U$ is the Crout decomposition of $A$, i.e., $U$ is unit upper triangular. One easily verifies that for this $T$ the stiff amplification matrix $Z_{\infty}(T)$ is strictly upper triangular. Throughout this paper, $T$ will always denote this special lower triangular matrix. This choice of $B$ yields, just like in PDIRK, a stiff amplification matrix that has a zero spectral radius. However, the new strategy leads to an amplification matrix $Z(z)$ that has a much smaller departure from normality than the amplification matrix in PDIRK. Consequently, the amplification after several iterations, i.e., the norm of the powers of $Z(z)$, is now considerably smaller (see [5, Table 3.1]). Suppose that all diagonal entries of $T$ are distinct and that the eigenvalue decomposition of $T$ is given by $T Q=Q D$, where $D$ is diagonal and $Q$ non-singular. Applying a similarity transformation in an analogous way as in [1], we arrive at the scheme

$$
\begin{align*}
T Q & =Q D, \\
\left(I-D \otimes h_{n} J_{n}\right) \Delta X_{n}^{(j+1)} & =-\left(Q^{-1} \otimes I\right) R\left(Y_{n}^{(j)}\right),  \tag{1.4}\\
Y_{n}^{(j+1)} & =Y_{n}^{(j)}+(Q \otimes I) \Delta X_{n}^{(j+1)} .
\end{align*}
$$

It is clear that the $s$ components of $Y_{n}^{(j)}$ can be computed in parallel. The only additional costs of (1.4) with respect to PDIRK are the appliance of the transformations $(Q \otimes I)$ and $\left(Q^{-1} \otimes I\right)$.

In order to ensure the non-singularity of the matrix $\left(I-D \otimes h_{n} J_{n}\right)$ in (1.4), the positiveness of the diagonal entries of $D$ is required. In [5] the positiveness of $D$ was proved for $s \leq 5$ and conjectured for $s>5$. The main scope of this paper is to prove this conjecture. This will be done in Section 3, using operator theory.

The outline of the rest of the paper is as follows. Section 2 gives some preliminaries to the conjecture. In Section 4 we prove for $s=2$, that the choice $B=T$ made in PTIRK is in some sense optimal.

## 2 Preliminaries.

The $s \times s$ matrix $A$ belonging to the RK collocation method with abscissa vector $c$ has the form [3, p.82],

$$
A=C V R V^{-1}
$$

where $C=\operatorname{diag}\left\{c_{1}, c_{2}, \ldots, c_{s}\right\}, R=\operatorname{diag}\{1,1 / 2, \ldots, 1 / s\}$ and $V$ is the Vandermonde matrix generated by $c$, i.e.,

$$
V=\left(\begin{array}{cccc}
1 & c_{1} & \ldots & c_{1}^{s-1} \\
\vdots & \vdots & & \vdots \\
1 & c_{s} & \ldots & c_{s}^{s-1}
\end{array}\right)
$$

Here, the abscissae $c_{i}$ have to be distinct. In the sequel the abscissae are also supposed to be positive. Without loss of generality, we assume that the RK method is written such that $c_{1}<c_{2}<\ldots<c_{s}$. Let $A=T U$ denote the Crout decomposition of $A$. The diagonal entries $t_{k k}$ of $T$ satisfy the formula [5]

$$
\begin{equation*}
t_{k k}=\frac{\left|A_{k}\right|}{\left|A_{k-1}\right|} \tag{2.1}
\end{equation*}
$$

where $\left|A_{j}\right|$ denotes the determinant of the $j$ th principal sub-matrix of $A$ and $\left|A_{0}\right|:=1$. From (2.1) we see that the existence of the Crout decomposition immediately follows from the positiveness of $t_{k k}$.

In [5] the authors proved the positiveness of $t_{k k}, k \in\{1,2, \ldots, s\}$, for $s \leq 5$ in the following way: first they showed that $\left|A_{1}\right|$ and $\left|A_{s}\right|$ are positive (for general $s$ ); then the positiveness of the remaining $\left|A_{2}\right|, \ldots,\left|A_{s-1}\right|$ was demonstrated by computing them explicitly; this approach does not lead to a proof for general $s$.

Another idea is to investigate whether the matrix $V R V^{-1}$ is positive definite. By using the result that every positive definite matrix has an $L U$-decomposition with positive diagonal entries [2, p. 140], the proof of the conjecture would then easily follow, realizing that $T=C L$, where $L$ is the lower triangular matrix in the Crout decomposition of $V R V^{-1}$. However, the following example shows that $V R V^{-1}$ is not always positive definite: If $s=3, c=(1 / 3,1 / 2,2 / 3)^{T}$ and $x=(1,-3,-7)^{T}$, then $x^{T} V R V^{-1} x=-11$.

In the following section the proof of the conjecture will be given by considering $V R V^{-1}$ as the matrix of an operator on the space of polynomials of degree less than $s$ with respect to a basis of Lagrange polynomials.

## 3 Proof of the conjecture.

Theorem 3.1. Let $V$ be the $s \times s$ Vandermonde matrix generated by $c_{1}, c_{2}, \ldots$, $c_{s}$, where $0<c_{1}<c_{2}<\cdots<c_{s}$, let $R$ be the diagonal matrix diag $(1,1 / 2, \ldots$, $1 / s$ ). There exist a lower triangular matrix $L$, and unit upper triangular matrix $U$, such that $L U=V R V^{-1}$. The diagonal entries of $L$ are positive.

Notice that from this theorem it immediately follows that for any $s \times s$ RK collocation matrix $A$ with positive distinct abscissae, there exists a lower triangular matrix $T$ with positive diagonal entries such that $Z_{\infty}(T)$ is strictly upper triangular, by setting $T=C L$.

Proof. Let $\mathbf{P}_{s}$ be the $s$-dimensional linear space of polynomials of degree less than $s$ with real coefficients, and $\mathcal{C}$ the canonical basis for $\mathbf{P}_{s}$, i.e.,

$$
\mathcal{C}=\left\{1, x, \ldots, x^{s-1}\right\}
$$

Define the operator $H: \mathbf{P}_{s} \rightarrow \mathbf{P}_{s}$ by $H(p)=q$ where $q$ is defined by

$$
q(x)=\frac{1}{x} \int_{0}^{x} p(t) d t
$$

We use the notation $\operatorname{mat}(H)_{\mathcal{C}}$ for the matrix of the operator $H$ with respect to the basis $\mathcal{C}$. It can be easily verified that

$$
\operatorname{mat}(H)_{\mathcal{C}}=R
$$

We denote the $k$ th Lagrange polynomial with respect to $c_{1}, c_{2}, \ldots, c_{s}$ by $l_{k}$ :

$$
l_{k}(x)=\prod_{i \neq k} \frac{x-c_{i}}{c_{k}-c_{i}} ; \quad k \in\{1,2, \ldots, s\}
$$

Notice that $l_{k}$ is of degree $s-1$ and thus an element of $\mathbf{P}_{s}$. The Lagrange polynomials define also a basis for $\mathbf{P}_{s}$, which will be denoted by $\mathcal{L}$ :

$$
\mathcal{L}=\left\{l_{1}, l_{2}, \ldots, l_{s}\right\}
$$

We write $\mathcal{C}_{\mathcal{L}}$ for the matrix that expresses the canonical basis in the Lagrange basis. Since for every $m \in\{0,1, \ldots, s-1\}$ the equality

$$
x^{m}=c_{1}^{m} l_{1}+c_{2}^{m} l_{2}+\ldots+c_{s}^{m} l_{s}
$$

should hold, it can be seen that $\mathcal{C}_{\mathcal{L}}=V$. Consequently, the matrix of the operator $H$ with respect to the basis $\mathcal{L}$ is given by

$$
\operatorname{mat}(H)_{\mathcal{L}}=\mathcal{C}_{\mathcal{L}} \cdot \operatorname{mat}(H)_{\mathcal{C}} \cdot \mathcal{C}_{\mathcal{L}}^{-1}=V R V^{-1}=: B
$$

If $\left(H\left(l_{k}\right)\right)_{\mathcal{L}}$ denotes the image under $H$ of $l_{k}$ with respect to the basis $\mathcal{L}$, then

$$
\left(H\left(l_{k}\right)\right)_{\mathcal{L}}=B e_{k}=\left(\begin{array}{c}
\beta_{1 k} \\
\vdots \\
\beta_{n k}
\end{array}\right)
$$

where $e_{k}$ is the $k$ th canonical basis vector of $\mathbf{R}^{s}$ and $\left(\beta_{i j}\right)=B$. We claim that $\beta_{11}>0$. To see this, notice that $H\left(l_{1}\right)$ is a polynomial with coefficient $\beta_{11}$ in the direction of $l_{1}$. Since $l_{k}\left(c_{1}\right)=0$ for $k>1$, it is clear that

$$
\left(H\left(l_{1}\right)\right)\left(c_{1}\right)=\beta_{11} .
$$

With respect to the value of $l_{1}$ in zero, we observe that $l_{1}\left(c_{1}\right)=1$, and that all its roots are to the right of $c_{1}$; therefore $l_{1}$ is positive on $\left[0, c_{1}\right]$, which implies

$$
\left(H\left(l_{1}\right)\right)\left(c_{1}\right)=\frac{1}{c_{1}} \int_{0}^{c_{1}} l_{1}(t) d t>0
$$

Consequently, $\beta_{11}>0$.

It is now possible to define $v_{1 k}:=-\beta_{1 k} / \beta_{11} ; k \in\{2, \ldots, s\}$. From this definition it follows that, for $k>1$,

$$
\left(H\left(l_{k}+v_{1 k} l_{1}\right)\right)_{\mathcal{L}}=B\left(e_{k}+v_{1 k} e_{1}\right)=B\left(\begin{array}{c}
v_{1 k} \\
0 \\
\vdots \\
0 \\
1 \\
0 \\
\vdots \\
0
\end{array}\right)=\left(\begin{array}{c}
0 \\
\beta_{2 k}^{(1)} \\
\vdots \\
\beta_{n k}^{(1)}
\end{array}\right) .
$$

Assuming $\beta_{22}^{(1)} \neq 0$, we are able to define $v_{2 k}:=-\beta_{2 k}^{(1)} / \beta_{22}^{(1)} ; k \in\{3, \ldots, s\}$, such that

$$
\left(H\left(l_{k}+v_{2 k} l_{2}+v_{1 k} l_{1}\right)\right)_{\mathcal{L}}=B\left(\begin{array}{c}
v_{1 k} \\
v_{2 k} \\
0 \\
\vdots \\
0 \\
1 \\
0 \\
\vdots \\
0
\end{array}\right)=\left(\begin{array}{c}
0 \\
0 \\
\beta_{3 k}^{(2)} \\
\vdots \\
\beta_{n k}^{(2)}
\end{array}\right) .
$$

Continuing this procedure, we finally arrive at

$$
\left(H\left(\sum_{i=1}^{k} v_{i k} l_{i}\right)\right)_{\mathcal{L}}=B u_{k}=r_{k}
$$

where

$$
v_{i k}=\left\{\begin{array}{cc}
-\frac{\beta_{i k}^{(i-1)}}{\beta_{i i}^{(i-1)}} & \text { for } \quad i<k \\
1 & \text { for } \quad i=k
\end{array}\right.
$$

(defining $\beta_{i j}^{(0)}=\beta_{i j}$ ) and $u_{k}$ and $r_{k}$ are vectors defined by

$$
u_{k}=\left(\begin{array}{c}
v_{1 k} \\
\vdots \\
v_{k-1, k} \\
1 \\
0 \\
\vdots \\
0
\end{array}\right) \quad \text { and } \quad r_{k}=\left(\begin{array}{c}
0 \\
\vdots \\
0 \\
\beta_{k k}^{(k-1)} \\
\vdots \\
\beta_{n k}^{(k-1)}
\end{array}\right) .
$$

If we can show that $\beta_{k k}^{(k-1)}>0$ for $k \in\{2,3, \ldots, s\}$, we have demonstrated that the procedure outlined above can be carried out. By observing that $u_{k}$ and $r_{k}$ are columns of matrices $\widetilde{U}$ and $L$, respectively, for which the relation $B \widetilde{U}=L$ holds, we then have proved Theorem 3.1 using $U$ for $\widetilde{U}^{-1}$.


Figure 3.1: Sketch of $U_{k}(x)$.
The vectors $u_{k}$ and $r_{k}$ can be considered as polynomials in $\mathbf{P}_{s}$ with respect to the basis $\mathcal{L}$. Moreover, $r_{k}$ is the image of $u_{k}$ under the operator $H$ :

$$
H\left(u_{k}\right)=r_{k}
$$

Since $r_{k}\left(c_{k}\right)=\beta_{k k}^{(k-1)}$, we have to prove that $r_{k}\left(c_{k}\right)>0$. We define the polynomial $U_{k}$ of degree $s+1$ by

$$
U_{k}(x)=\int_{0}^{x} u_{k}(t) d t
$$

Notice that $U_{k}(0)=0$ and, for $x>0$, the sign of $r_{k}$ equals the sign of $U_{k}$ (the latter holds since $\left.U_{k}=x r_{k}\right)$. Since $l_{k}\left(c_{i}\right)=0$ for $i<k$ and $r_{k}$ has only components in the direction of $l_{j}$ with $j \geq k$, we see that $r_{k}\left(c_{i}\right)=0$ for $i<k$ and consequently

$$
U_{k}\left(c_{i}\right)=0 \quad \text { for } \quad i<k
$$

This means that $u_{k}$ (being the derivative of $U_{k}$ ) has $k-1$ zeros in the interval ( $0, c_{k-1}$ ). All components of $u_{k}$ in the direction of the last $s-k$ Lagrange polynomials are zero. Consequently, $u_{k}\left(c_{i}\right)=0$ for $i>k$, so that $u_{k}$ has $s-k$ zeros in the interval $\left[c_{k+1}, c_{k}\right]$.

We now consider 2 cases (see also Figure 3.1):

$$
\begin{align*}
& u_{k}\left(c_{k-1}\right)>0  \tag{3.1}\\
& u_{k}\left(c_{k-1}\right)<0 \tag{3.2}
\end{align*}
$$

Remark that, since all $c_{i}$ are distinct, $U_{k}$ has a single zero in $c_{k-1}$, so that the situation $u_{k}\left(c_{k-1}\right)=0$ does not arise. Suppose that (3.2) holds. Since $u_{k}\left(c_{k}\right)=1$, the polynomial $u_{k}$ should have a zero in the interval ( $c_{k-1}, c_{k}$ ). In that case, $u_{k}$ has $(k-1)+(s-k)+1=s$ zeros. However, the degree of $u_{k}$ is only $s-1$, proving that only situation (3.1) can occur, and $u_{k}>0$ on ( $c_{k-1}, c_{k}$ ).

From $U_{k}\left(c_{k-1}\right)=0$, it now follows that $U_{k}\left(c_{k}\right)>0$. Since $r_{k}$ has the same sign as $U_{k}$, we have proved the theorem.

## 4 Is PTIRK optimal?

In this section we investigate the optimality of the matrix $T$ in PTIRK. Since the number of parameters becomes too large to handle conveniently for $s>2$, we restrict ourselves here to methods with 2 implicit stages, i.e., $s=2$.

In the class of lower triangular matrices, $T$ is optimal in the sense that it leads to the smallest stiff amplification matrix measured in the infinity norm:

THEOREM 4.1. If $L$ is a $2 \times 2$ lower triangular matrix, then

$$
\left\|Z_{\infty}(L)\right\|_{\infty} \geq\left\|Z_{\infty}(T)\right\|_{\infty}
$$

Proof. Write $L^{-1}=\left(l_{i j}\right)$ with $l_{12}=0$. Then

$$
Z_{\infty}(L)=\left(\begin{array}{cc}
1+\frac{l_{1,1} c_{1}\left(-2 c_{2}+c_{1}\right)}{2\left(c_{2}-c_{1}\right)} & \frac{l_{1,1} c_{1}{ }^{2}}{2\left(c_{2}-c_{1}\right)} \\
* & *
\end{array}\right)
$$

Define for $x>0$ :

$$
g(x)=\left|1+\frac{c_{1}\left(-2 c_{2}+c_{1}\right)}{2\left(c_{2}-c_{1}\right)} x\right|+\frac{c_{1}^{2}}{2\left(c_{2}-c_{1}\right)} x
$$

Then $g(x) \geq g\left(x_{\text {min }}\right)=c_{1} /\left(2 c_{2}-c_{1}\right)$, where

$$
x_{\min }=2\left(c_{2}-c_{1}\right) /\left(c_{1}\left(2 c_{2}-c_{1}\right)\right)
$$

Since $\left\|Z_{\infty}(T)\right\|_{\infty}=g\left(x_{\min }\right)$, it follows that $\left\|Z_{\infty}(L)\right\|_{\infty} \geq\left\|Z_{\infty}(T)\right\|_{\infty}$.
For two well-known stiffly accurate RK methods with 2 implicit stages, it is possible to show that in the class of lower triangular matrices that lead to a 'small' stiff amplification matrix, $T$ is optimal in the sense that it has the smallest non-stiff amplification matrix, again measured in the infinity norm:

Theorem 4.2. If $L$ is a $2 \times 2$ lower triangular matrix with the property that $\rho\left(Z_{\infty}(L)\right)=0$, then, for the 2-stage Radau IIA, and the 3-stage Lobatto IIIA method,

$$
\left\|Z_{0}(L)\right\|_{\infty} \geq\left\|Z_{0}(T)\right\|_{\infty}
$$

Proof. Write $A=\left(a_{i j}\right)$ and $L=\left(l_{i j}\right)$ with $l_{12}=0$. Then $\left\|Z_{0}(L)\right\|_{\infty}=$ $\max \left(m_{1}, m_{2}\right)$, where $m_{1}$ and $m_{2}$ are given by

$$
m_{1}=\left|a_{11}-l_{11}\right|+\left|a_{12}\right| \quad \text { and } \quad m_{2}=\left|a_{21}-l_{21}\right|+\left|a_{12}-l_{22}\right|
$$

Let $J$ be the interval such that if $l_{11} \notin J$, then $m_{1}>\left\|Z_{0}(T)\right\|_{\infty}$. Notice that $J$ only depends on $c$. From $\sigma\left(Z_{\infty}(L)\right)=0$ it follows that $\operatorname{trace}\left(Z_{\infty}(L)\right)=$ $\operatorname{det}\left(Z_{\infty}(L)\right)=0$. Using these two equations, it is possible to express $l_{21}$ and $l_{22}$, and thus $m_{2}$, in $l_{11}$. We have to prove that for $l_{11} \in J, m_{2} \geq\left\|Z_{0}(T)\right\|_{\infty}$. We treat the two methods separately.

Radau IIA.

$$
\begin{aligned}
& c=(1 / 3,1)^{T},\left\|Z_{0}(T)\right\|_{\infty}=3 / 20, J=[7 / 20,29 / 60], \text { and } \\
& \qquad m_{2}\left(l_{11}\right)=\left|\frac{3}{4}+\frac{-24 l_{1,1}+5+18 l_{1,1}^{2}}{6 l_{1,1}}\right|+\left|\frac{1}{4}-\frac{1}{6 l_{1,1}}\right| .
\end{aligned}
$$

It can be verified that $\min _{l_{11} \in J}\left(m_{2}\left(l_{11}\right)\right)=m_{2}\left(t_{11}\right)=3 / 20$.
Lobatto IIIA.

$$
\begin{aligned}
& c=(0,1 / 2,1)^{T},\left\|Z_{0}(T)\right\|_{\infty}=1 / 12, J=[7 / 24,3 / 8], \text { and } \\
& \qquad m_{2}\left(l_{11}\right)=\left|\frac{2}{3}+\frac{-12 l_{1,1}+2+12 l_{1,1}^{2}}{3 l_{1,1}}\right|+\left|\frac{1}{6}-\frac{1}{12 l_{1,1}}\right| .
\end{aligned}
$$

The reader is invited to check that $\min _{l_{11} \in J}\left(m_{2}\left(l_{11}\right)\right)=m_{2}\left(t_{11}\right)=1 / 12$.

## Acknowledgements.

With great pleasure we acknowledge the insight of Hendrik W. Lenstra who shared with the first author the idea to use the materialization of the operator $H$ on different bases for the purpose of the proof of Theorem 3.1.

## REFERENCES

1. G. J. Cooper and J. C. Butcher, An iteration scheme for implicit Runge-Kutta methods, IMA J. Numer. Anal., 3 (1983), pp. 127-140.
2. G. H. Golub and C. F. van Loan, Matrix Computations, Johns Hopkins University Press, Baltimore and London, 2nd ed., 1989.
3. E. Hairer and G. Wanner, Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems, Springer-Verlag, 1991.
4. P. J. van der Houwen and B. P. Sommeijer, Iterated Runge-Kutta methods on parallel computers, SIAM J. Sci. Stat. Comput., 12 (1991), pp. 1000-1028.
5. P. J. van der Houwen and J. J. B. de Swart, Triangularly implicit iteration methods for ODE-IVP solvers, SIAM J. Sci. Comput., 18 (1997), pp. 41-55.

[^0]:    *Received September 1995.
    ${ }^{\dagger}$ The research reported in this paper was supported by STW (Dutch Foundation for Technical Sciences).

