Centrum voor Wiskunde en Informatica Centre for Mathematics and Computer Science
W.H. Hundsdorfer

A note on monotonicity of a Rosenbrock method

# A Note on Monotonicity of a Rosenbrock Method 

W.H. Hundsdorfer<br>Centre for Mathematics and Computer Science<br>P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

For a dissipative differential equation with stationary solution $u^{*}$, the difference between any solution $U(t)$ and $u^{*}$ is nonincreasing with $t$. In this note we present necessary and sufficient conditions in order for a similar monotonicity property to hold for numerical approximations computed from a Rosenbrock method. Our results also provide global convergence results for some modifications of Newton's method.

1980 Mathematics Subject Classification: 65L20, 65H10
Key Words \& Phrases: stiff differential equations, monotonicity, Rosenbrock methods, nonlinear algebraic equations, modified Newton methods.

## 1. Introduction

Consider an initial value problem in $\mathbb{R}^{m}$

$$
\begin{align*}
& U^{\prime}(t)=f(U(t)) \quad(t \geqslant 0)  \tag{1.1a}\\
& U(0)=u_{0} \tag{1.1b}
\end{align*}
$$

whose solution $U(t)$ tends to a stationary solution $u^{*} \in \mathbb{R}^{m}$. For the numerical solution of (1.1) we consider the Rosenbrock method

$$
\begin{equation*}
u_{n+1}=u_{n}+\left(I-h \theta f^{\prime}\left(u_{n}\right)\right)^{-1} h f\left(u_{n}\right) \tag{1.2}
\end{equation*}
$$

where $\theta$ is a positive parameter, $h>0$ is the stepsize and the vectors $u_{n} \in \mathbb{R}^{m}$ approximate $U\left(t_{n}\right)$, $t_{n}=n h(n=0,1,2, \ldots)$.

Assume the function $f$ is dissipative with respect to an inner product $<\cdot,>$ on $\mathbb{R}^{m}$ $\left(<f(\tilde{u})-f(u), \tilde{u}-u>\leqslant 0\right.$ for all $\left.\tilde{u}, u \in \mathbb{R}^{m}\right)$ and let $\|x\|=<x, x>^{1 / 2}$ (for $x \in \mathbb{R}^{m}$ ). This assumption implies that the difference $\|\tilde{U}(t)-U(t)\|$ of any two solutions of the differential equation (1.1a) is nonincreasing with $t$. The corresponding property for the numerical approximations, $\left\|\tilde{u}_{n+1}-u_{n+1}\right\| \leqslant\left\|\tilde{u}_{n}-u_{n}\right\|$, only holds under additional, rather restrictive conditions on $f$ (see e.g. [3]). In this note we look at the less exacting monotonicity property

$$
\begin{equation*}
\left\|u_{n+1}-u^{*}\right\| \leqslant\left\|u_{n}-u^{*}\right\|, \tag{1.3}
\end{equation*}
$$

and we shall present conditions on $f$ which are necessary and sufficient for (1.3) to hold with arbitrary stepsize $h$. Under somewhat stronger conditions on $f$ the convergence of $u_{n}$ to $u^{*}$ can be guaranteed. These results are relevant to stiff ordinary differential equations and to partial differential equations (via the method of lines) since neither the Lipschitz constant of $f$ nor the dimension $m$ are involved.

The monotonicity property (1.3) is of particular interest if scheme (1.2) is regarded as a timemarching procedure for finding stationary solutions. The scheme has been used for this purpose in [4] with $\theta=1$ (and with an approximation to the Jacobian matrix $f^{\prime}\left(u_{n}\right)$; cf.(2.1)). We note that in such a situation (1.2) can be considered as a modified Newton procedure for solving $f(u)=0$. By introducing $\omega=1 / \theta$ and $\lambda=1 / h \theta$ we can rewrite (1.2) as

$$
u_{n+1}=u_{n}-\omega\left(f^{\prime}\left(u_{n}\right)-\lambda I\right)^{-1} f\left(u_{n}\right)
$$

in which $\omega>0$ can be viewed as a relaxation parameter and $\lambda>0$ ensures that $f^{\prime}\left(u_{n}\right)-\lambda I$ is

## Report NM-R8606

Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands
nonsingular whenever $f$ is dissipative (see [5; sect.5.4,7.1]).

## 2. Monotonicity for Numerical Approximations

Besides the Rosenbrock method (1.2) we also consider the more general linearly implicit scheme

$$
\begin{equation*}
u_{n+1}=u_{n}+\left(I-h \theta J\left(u_{n}\right)\right)^{-1} h f\left(u_{n}\right) \tag{2.1}
\end{equation*}
$$

where $J\left(u_{n}\right)$ is an $m \times m$ matrix. Further we shall use the following notation. By $L\left(\mathbb{R}^{m}\right)$ we denote the space of linear operators on $\mathbb{R}^{m}$. If $\|\cdot\|$ is a norm on $\mathbb{R}^{m}$, the corresponding operator norm on $L\left(\mathbb{R}^{m}\right)$ is also denoted by $\|\cdot\|$, and $\mu \cdot \cdot]$ will stand for the logarithmaic norm (cf.[2]).

Consider the following set of assumptions (2.2)-(2.6), with $\delta>0$ and $\epsilon \geqslant 0$, which will be denoted by $\left(A_{1}\right)$.

$$
\begin{equation*}
m \in \mathbb{N} \text { and }\|\cdot\| \text { is a norm on } \mathbb{R}^{m} \text { generated by an inner product }<\cdot,> \tag{2.2}
\end{equation*}
$$

$$
\begin{align*}
& f: \mathbb{R}^{m} \rightarrow \mathbb{R}^{m}, u^{*} \in \mathbb{R}^{m} \text { is a zero of } f, \text { and } J: \mathbb{R}^{m} \rightarrow L\left(\mathbb{R}^{m}\right) ;  \tag{2.3}\\
& \left\{\begin{array}{l}
D=\left\{u: u \in \mathbb{R}^{m},\left\|u-u^{*}\right\|<\delta\right\}, f \text { is continuously differentiable on } D \\
\text { and } J \text { is continuous on } D ;
\end{array}\right. \tag{2.4}
\end{align*}
$$

$\left\{\begin{array}{l}\text { for any } u \in D \text { we have } \mu\left[f^{\prime}(u)\right] \leqslant 0 \\ \text { and } J(u) \text { does not have positive real eigenvalues; }\end{array}\right.$
$\left\{\begin{array}{l}\text { for all } u, v \in D \text { there is a } E(u, v) \in L\left(\mathbb{R}^{m}\right) \text { such that } \\ f^{\prime}(v)=J(u)(I+E(u, v)),\|E(u, v)\| \leqslant \epsilon .\end{array}\right.$
Further $\left(A_{2}\right)$ will stand for these assumptions (2.2)-(2.6) together with

$$
\begin{equation*}
J(u)=f^{\prime}(u)(\text { for all } u \in D) \tag{2.7}
\end{equation*}
$$

In (2.4) continuously differentiable means that that the matrix of partial derivatives $f^{\prime}(u)=\left(\partial f_{i}(u) / \partial u_{j}\right)$ exists and depends continuously on $u$. The condition $\mu\left[f^{\prime}(u)\right] \leqslant 0$ on $D$ is equivalent to requiring that $f$ is dissipative on $D$ (see e.g.[5, sect.5.4]). The condition in (2.6) states that the relative difference between $f^{\prime}(v)$ and $J(u)$ is bounded by $\epsilon$; in case $J(u)$ is regular it reads $\left\|J(u)^{-1}\left(f^{\prime}(v)-J(u)\right)\right\| \leqslant \epsilon$. Thus, in a relative sense, the variation of $f^{\prime}$ on $D$ may not be too large and $J(u)$ has to approximate $f^{\prime}(u)$ accurately enough.

In order to formulate our main results we define the real functions $\psi_{k}(k=1,2)$ on the interval $[1 / 2, \infty)$ by

$$
\begin{align*}
& \psi_{1}(\theta)=\min \{2 \theta-1,1\}  \tag{2.8}\\
& \psi_{2}(\theta)=\min \{2 \theta-1, \sqrt{(2 \theta-1) / \theta}\} \tag{2.9}
\end{align*}
$$

Theorem 2.1. Let $h$ and $\delta$ be positive, and $k$ equal to 1 or 2 . We have $\left\|u_{n+1}-u^{*}\right\| \leqslant\left\|u_{n}-u^{*}\right\|$ (whenever $u_{n} \in D$ and $\left(A_{k}\right)$ is valid) iff $\theta \geqslant 1 / 2$ and $\epsilon \leqslant \psi_{k}(\theta)$.

This theorem is an extension of a result by M.N. SPIJKER and the present author, see [7; sect. 4]. The proof will be given in the next section. The restriction $\theta \geqslant 1 / 2$ in this theorem is not surprising since the methods with $\theta<1 / 2$ are not $A$-stable. For $\theta=1 / 2$ we see that the monotonicity property only holds for linear problems $(\epsilon=0)$.
Under slightly stronger conditions on $f$ it can be shown that $\left\|u_{n+1}-u^{*}\right\|<\left\|u_{n}-u^{*}\right\|$ (for $\left.u_{n} \in D, u_{n} \neq u^{*}\right)$. This leads to the following result which will also be proved in section 3 .
Theorem 2.2. Let $h$ and $\delta$ be positive, and $k$ equal to 1 or 2 . Assume $u_{0} \in D, \theta \geqslant 1 / 2, \epsilon \leqslant \psi_{k}(\theta)$ and
$\left(A_{k}\right)$. Assume in addition that either $\epsilon<\psi_{k}(\theta)$ and $J(u)$ is regular (for all $u \in D$ ) or $\mu\left[f^{\prime}(u)\right]<0$ (for all $u \in D$ ). Then $u^{\star}$ is the unique zero of f in $D$ and $\lim _{n \rightarrow \infty} u_{n}=u^{*}$.

## 3. Proof of the Monotonicity Results

### 3.1. Preliminaries

In order to prove the theorems of section 2 we first derive some technical results. Consider arbitrary $A, B \in L\left(\mathbb{R}^{m}\right)$ with $m \in \mathbb{N}$, and suppose $\|\cdot\|$ is a norm on $\mathbb{R}^{m}$ generated by an inner product $<, \cdot>$. For any $C \in L\left(\mathbb{R}^{m}\right)$ we denote by $C^{*}$ its adjoint with respect to this inner product ( $<C x, y>=<x, C^{\star} y>$ for all $x, y \in \mathbb{R}^{m}$ ).

The relation $\|B x\| \leqslant \gamma\|A x\|$ (for all $x \in \mathbb{R}^{m}$ ), with $\gamma>0$ given, implies the existence of a $C \in L\left(\mathbb{R}^{m}\right)$ such that $B=C A,\|C\| \leqslant \gamma$; if $A$ is regular we can take $C=B A^{-1}$ and for singular $A$ the inverse $A^{-1}$ can be replaced by the generalized inverse of $A$ (see e.g. [1; ch.8]). Since $\left\|C^{*}\right\|=\|C\|$ for any $C \in L\left(\mathbb{R}^{m}\right)$ one easily arrives at the following result.

Lemma 3.1. Let $\gamma>0$. We have $\left\|B^{*} x\right\| \leqslant \gamma\left\|A^{*} x\right\|\left(\right.$ for all $\left.x \in \mathbb{R}^{m}\right)$ iff $B=A C$ for some $C \in L\left(\mathbb{R}^{m}\right)$ with $\|C\| \leqslant \gamma$.

Consider the following statements, with $\theta \geqslant \frac{1}{2}$ and $\epsilon \geqslant 0$,

$$
\begin{align*}
& B=A\left(I+E_{1}\right) \text { for some } E_{1} \in L\left(\mathbb{R}^{m}\right) \text { with }\left\|E_{1}\right\| \leqslant \epsilon  \tag{3.1a}\\
& A=B\left(I+E_{2}\right) \text { for some } E_{2} \in L\left(\mathbb{R}^{m}\right) \text { with }\left\|E_{2}\right\| \leqslant \epsilon \tag{3.1b}
\end{align*}
$$

and

$$
\begin{equation*}
B=\theta A(I+F) \text { for some } F \in L\left(\mathbb{R}^{m}\right) \text { with }\|F\| \leqslant 1 \tag{3.2}
\end{equation*}
$$

Lemma 3.2. (3.1a) implies (3.2) iff $\epsilon \leqslant \psi_{1}(\theta)$.
Proof. Assuming (3.1a) and $\epsilon \leqslant \psi_{1}(\theta)$ we set $F=\theta^{-1}\left[E_{1}+(1-\theta) I\right]$, in which case $B=\theta A(I+F)$ and

$$
\|F\| \leqslant \theta^{-1}(\epsilon+|1-\theta|) \leqslant \theta^{-1}\left(\psi_{1}(\theta)+|1-\theta|\right)=1 .
$$

To construct a counterexample in case $\epsilon>\psi_{1}(\theta)$ we first consider the simple scalar (complex) example $A=a, B=b$ with $a, b \in \mathbb{C}$. The condition in (3.la) corresponds to

$$
\begin{equation*}
|b-a| \leqslant \epsilon|a| \tag{3.3}
\end{equation*}
$$

and (3.2) corresponds to

$$
\begin{equation*}
|b-\theta a| \leqslant \theta|a| \tag{3.4}
\end{equation*}
$$

By simple geometrical arguments it follows that for $\epsilon>\psi_{1}(\theta)$ there exist $a, b \in \mathbb{C}$ satisfying (3.3) but violating (3.4).

These considerations on $\mathbb{C}$ lead to a counterexample with $A=A_{1}$ and $B=B_{1} \in L\left(\mathbb{R}^{2}\right)$,

$$
A_{1}=\left[\begin{array}{ll}
\operatorname{Re} a & -\operatorname{Im} a \\
\operatorname{Im} a & \operatorname{Re} a
\end{array}\right], B_{1}=\left[\begin{array}{ll}
\operatorname{Re} b & -\operatorname{Im} b \\
\operatorname{Im} b & \operatorname{Re} b
\end{array}\right]
$$

and with $\|\cdot\|$ the Euclidian norm on $\mathbb{R}^{2}$ (and the corresponding spectral norm on $L\left(\mathbb{R}^{2}\right)$ ).
Lemma 3.3. (3.1a) \& (3.1b) implies (3.2) iff $\epsilon \leqslant \psi_{2}(\theta)$.

Proof. Assume (3.1a), (3.1b) and $\epsilon \leqslant \psi_{2}(\theta)$. To show that (3.2) holds it is, in view of lemma 3.1, sufficient to consider the remaining case $1<\epsilon^{2} \leqslant(2 \theta-1) / \theta$. From lemma 3.1 it follows that for any $x \in \mathbb{R}^{m}$

$$
\begin{aligned}
& \left\|B^{*} x\right\|^{2}-2<A^{\star} x, B^{*} x>+\left\|A^{\star} x\right\|^{2} \leqslant \epsilon^{2}\left\|A^{*} x\right\|^{2}, \\
& \left\|B^{\star} x\right\|^{2}-2<A^{*} x, B^{\star} x>+\left\|A^{*} x\right\|^{2} \leqslant \epsilon^{2}\left\|B^{*} x\right\|^{2} .
\end{aligned}
$$

Combining these inequalities we obtain

$$
<A^{*} x, B^{*} x>\geqslant\left(1-\frac{1}{2} \epsilon^{2}\right)\left\|B^{*} x\right\|^{2}
$$

From our assumption on $\epsilon$ it follows that

$$
\left\|B^{*} x\right\|^{2} \leqslant 2 \theta<\dot{A}^{\star} x, B^{*} x>
$$

and hence

$$
\left\|\left(B^{\star}-\theta A^{\star}\right) x\right\| \leqslant \theta\left\|A^{\star} x\right\| .
$$

Statement (3.2) now follows by again applying lemma 3.1.
Now assume $\epsilon>\psi_{2}(\theta)$ and $1 / 2 \leqslant \theta \leqslant 1$. Then we obtain a (scalar real) counterexample by taking $m=1, A=-1, B=-1-\epsilon$.

Finally assume $\epsilon>\psi_{2}(\theta)$ and $\theta>1$. Let $\xi \in((2 \theta-1) / \theta, 2)$ such that $\xi<\epsilon^{2}$, and take $a, b \in \mathbb{C}$ such that $b / a$ equals $(1-\xi / 2)+i \sqrt{\xi(1-\xi / 4)}$. Then $|b-a| \leqslant \epsilon|a|$ and $|b-a| \leqslant \epsilon|b|$ but $|b-\theta a|>\theta|a|$. As in the proof of lemma 3.2 such $a, b \in \mathbb{C}$ lead to $A_{2}, B_{2} \in L\left(\mathbb{R}^{2}\right)$ such that for $A=A_{2}, B=B_{2}$ the statements (3.1a), (3.1b) hold whereas (3.2) is violated.

We note that in the above counterexamples which prove the necessity of $\epsilon \leqslant \psi_{k}(\theta)$ we can choose the $a, b \in \mathbb{C}$ such that $\operatorname{Re} a \leqslant 0, \quad \operatorname{Re} b \leqslant 0$. This leads to $A_{k}, B_{k} \in L\left(\mathbb{R}^{2}\right)$ satisfying $\mu\left[A_{k}\right] \leqslant 0, \mu\left[B_{k}\right] \leqslant 0$ (for $k=1,2$ ).
The following lemma is a slight generalization of results in [6] and [7; lemma 4.3].
Lemma 3.5. Assume $I-\lambda \theta A$ is regular for all $\lambda>0$. We have $\left\|I+(I-\lambda \theta A)^{-1} \lambda B\right\| \leqslant 1$ (for all $\lambda>0$ ) iff $\mu[B] \leqslant 0$ and (3.2) holds.

Proof. Let $C=B-\theta A$. Then $I+(I-\lambda \theta A)^{-1} \lambda B=(I-\lambda \theta A)^{-1}(I+\lambda C)$, and it follows that $\left\|I+(I-\lambda \theta A)^{-1} \lambda B\right\| \leqslant 1$ iff

$$
\left\|\left(I+\lambda C^{\star}\right) x\right\| \leqslant\left\|\left(I-\lambda \theta A^{*}\right) x\right\|\left(\text { for all } x \in \mathbb{R}^{m}\right)
$$

The latter inequality can be written as

$$
2 \lambda<B^{*} x, x>+\lambda^{2}\left\|C^{*} x\right\|^{2} \leqslant \lambda^{2}\left\|\theta A^{*}\right\|^{2} \text { (for all } x \in \mathbb{R}^{m} \text { ). }
$$

Clearly this holds for all $\lambda>0$ iff

$$
<B x, x>\leqslant 0 \text { and }\left\|C^{\star} x\right\| \leqslant\left\|\theta A^{*} x\right\| \text { (for all } x \in \mathbb{R}^{m} \text { ). }
$$

Application of lemma 3.1 completes the proof.

### 3.2. The proof of theorem 2.1

For $u \in D$ we define

$$
\begin{equation*}
\sigma(u)=\int_{0}^{1}\left\|I+(I-h \theta J(u))^{-1} h f^{\prime}\left(u^{*}+\tau\left(u-u^{*}\right)\right)\right\| d \tau \tag{3.5}
\end{equation*}
$$

Since for any $u_{n} \in D$

$$
\left\|u_{n+1}-u^{*}\right\|=\left\|u_{n}-u^{*}+\left(I-h \theta J\left(u_{n}\right)\right)^{-1} h\left(f\left(u_{n}\right)-f\left(u^{*}\right)\right)\right\|
$$

it follows by the mean-value theorem that

$$
\begin{equation*}
\left\|u_{n+1}-u^{*}\right\| \leqslant \sigma\left(u_{n}\right)\left\|u_{n}-u^{*}\right\| . \tag{3.6}
\end{equation*}
$$

Application of the lemmas 3.2, 3.3 and 3.5 with $A=h J\left(u_{n}\right)$ and $B=h f^{\prime}\left(u^{*}+\tau\left(u_{n}-u^{*}\right)\right)$ shows the sufficiency of $\epsilon \leqslant \psi_{k}(\theta)$ for having $\left\|u_{n+1}-u^{*}\right\| \leqslant\left\|u_{n}-u^{*}\right\|$ in case $\left(A_{k}\right)$ holds, $k=1,2$. The necessity will be proved by some counterexamples.

A counterexample in case $\left(A_{1}\right)$ holds, $\epsilon>\psi_{1}(\theta)$ is given by $h J(u)=\lambda A_{1}, h f(u)=\lambda B_{1} u$ (for $u \in \mathbb{R}^{2}$ ) with $A_{1}, B_{1}$ as in the proof of lemma $3.2, \lambda>0$ and $\|\cdot\|$ the Euclidean norm on $\mathbb{R}^{2}$. With $u^{*}=0, u_{n} \in \mathbb{R}^{2}$, we obtain

$$
\begin{aligned}
& \left\|u_{n+1}-u^{*}\right\|=\left\|\left(I+\left(I-\lambda \theta A_{1}\right)^{-1} \lambda B_{1}\right)\left(u_{n}-u^{*}\right)\right\|= \\
& =\left\|I+\left(I-\lambda \theta A_{1}\right)^{-1} \lambda B_{1}\right\|\left\|u_{n}-u^{*}\right\|>\left\|u_{n}-u^{*}\right\|
\end{aligned}
$$

provided $\lambda>0$ is suitably chosen (see lemma 3.5 ).
Next we give a scalar (real) example for $\epsilon>\psi_{2}(\theta), 1 / 2 \leqslant \theta \leqslant 1$ in case $\left(A_{2}\right)$ is valid. This counterexample is similar to one given by Sandberg and Shichman [6].

Take, for convenience, $h=1, \delta>1$ and $u^{\star}=0, u_{0}=1$. Let $\eta \in(2 \theta-1, \epsilon)$ and $f(u)=\lambda g(u)$ (for $u \in \mathbb{R}$ ) with $\lambda>0$ and $g: \mathbb{R} \rightarrow \mathbb{R}$ a continuously differentiable function such that

$$
\begin{aligned}
& g(0)=0, g^{\prime}(u) \in[-1-\epsilon,-1](\text { for all } u \in \mathbb{R}) \\
& g(u)=-u+\eta(\text { for } u \leqslant 1), g(u)=-u-\eta(\text { for } u \geqslant 1) .
\end{aligned}
$$

Such an $f$ meets the conditions imposed in $\left(A_{2}\right)$. Further we have

$$
u_{1}=(1+\lambda \theta)^{-1}(1+\lambda(\theta-1-\eta))
$$

and thus $\left|u_{1}-u^{*}\right|$ tends to $\theta^{-1}(\eta+1-\theta)>1=\left|u_{0}-u^{*}\right|$ for $\lambda \rightarrow \infty$.
Finally we assume $\epsilon>\psi_{2}(\theta), \theta>1$. For this we construct a complex, scalar counterexample, which can as before be converted to a real one by identifying $\mathbb{C}$ with $\mathbb{R}^{2}$ in the usual way. Suppose $(2 \theta-1) / \theta<\xi<\min \left\{2, \epsilon^{2}\right\}$ and let $a, b \in \mathbb{C}$ be such that $\operatorname{Re} a<0, \operatorname{Re} b<0$ and $b / a=(1-\xi / 2)+i \sqrt{\xi(1-\xi / 4)}$ (as in the proof of lemma 3.3). Then $|b-a|<\epsilon|a|,|b-a|<\epsilon|b|$ but $|b-\theta a|>\theta|a|$, and thus for $\lambda>0$ suitably chosen $\left|1+(1-\lambda \theta a)^{-1} \lambda b\right|>1$ (see lemma 3.5). We put $\alpha=\lambda a$ and $\beta=\lambda b$.

Let $D$ be the unit disk in $\mathbb{C}, h=1, u^{*}=0$, and define

$$
\begin{aligned}
& f(u)=\Phi(0)(\alpha-\beta)+\alpha u+\Phi(u)(\beta-\alpha) \\
& \Phi(u)=-\frac{2 k}{k+1}\left(\frac{1}{2}-\frac{1}{2} u\right)^{1+1 / k}
\end{aligned}
$$

for $u \in \mathbb{C}$, where $k \in \mathbb{N}$ is to be specified later. Then $f(0)=0$ and

$$
\begin{aligned}
& f^{\prime}(u)=\alpha+\phi(u)(\beta-\alpha) \\
& \phi(u)=\left(\frac{1}{2}-\frac{1}{2} u\right)^{1 / k}
\end{aligned}
$$

The image of $D$ under $\phi$ tends to the interval $(0,1)$ on the real axis if $k \rightarrow \infty$. By using this property it can be shown that, for $k$ sufficiently large, the conditions on $f$ in $\left(A_{2}\right)$ are satisfied. Moreover, since $f^{\prime}(1)=\alpha$ and $f(1)$ tends to $\beta$ for $k \rightarrow \infty$,

$$
\left|1+\left(1-\theta f^{\prime}(1)\right)^{-1} f(1)\right|>1
$$

provided $k$ is sufficiently large. It follows that, for such $k$ and $u_{0}$ close to 1 ,

$$
\left|u_{1}-u^{\star}\right|>\left|u_{0}-u^{\star}\right|
$$

### 3.3. The proof of theorem 2.2

First we show that under the assumptions of theorem 2.2 the function $\sigma$, defined by (3.5), satisfies $\sigma(u)<1$ (for all $u \in D$ ). Examination of the proof of lemma 3.5 shows that for $A, B \in L\left(\mathbb{R}^{m}\right)$ satisfying (3.2) and $\mu[B] \leqslant 0$ we have

$$
\left\|I+(I-\lambda \theta A)^{-1} \lambda B\right\|<1(\text { for all } \lambda>0)
$$

provided we assume in addition either

$$
\mu[B]<0
$$

or

$$
A \text { is regular and } B=\theta A(I+F),\|F\|<1
$$

Further it is easily seen, by regarding the proofs of lemma 3.2 and lemma 3.3, that if $A$ is regular and we have (3.1a) with $\epsilon<\psi_{1}(\theta)$ or (3.1a), (3.1b) with $\epsilon<\psi_{2}(\theta)$ then there is an $F \in L\left(\mathbb{R}^{m}\right)$ such that $B=\theta A(I+F),\|F\|<1$. By setting $A=h J(u), B=h f^{\prime}\left(u^{\star}+\tau\left(u-u^{*}\right)\right)$ it follows that the assumptions of theorem 2.2 imply $\sigma(u)<1$ on $D$.

The function $\sigma$ is continuous on $D$. Therefore we obtain for arbitrary $u_{0} \in D$

$$
\left\|u_{n}-u^{\star}\right\| \leqslant s_{0}^{n}\left\|u_{0}-u^{*}\right\|
$$

with $s_{0}=\max \left\{\sigma(u): u \in D,\left\|u-u^{*}\right\| \leqslant\left\|u_{0}-u^{*}\right\|\right\}<1$. From this it is clear that $u^{*}$ is the unique zero of $f$ in $D$ and that the $u_{n}$ converge to $u^{*}$ for $n \rightarrow \infty$.

## Acknowledgement

The author is grateful to professor M.N. Spijker for the valuable and stimulating discussions.

## References

[1] A. Ben-Israel, T.N.E. Greville Generalized inverses: theory and applications, Wiley-Interscience, New York, 1974.
[2] K. Dekker, J.G. Verwer, Stability of Runge-Kutta methods for stiff nonlinear differential equations, North-Holland, Amsterdam, 1984.
[3] W.H. HuNDSDORFER, The numerical solution of nonlinear stiff initial value problems, CWI Tract 12, Centre for Math. and Comp. Sc., Amsterdam, 1985.
[4] B. van Leer, W.A. Mulder, Relaxation methods for hyperbolic equations, Report 84-20, Dept. of Math.and Informatics, Delft Univ. of Techn., Delft, 1984.
[5] J.M. Ortega, W.C. Rheinboldt, Iterative solution of nonlinear equations in several variables, Academic Press, New York, 1970.
[6] I.W. Sandberg, H. Shichman, Numerical Integration of systems of stiff nonlinear initial value problems, B.S.T.J. 47, 511-527 (1968).
[7] M.N. SpiJker, A note on monotonicity and contractivity in the numerical solution of stiff differential equations, Report 1985-25, Inst. of Appl. Math. and Comp. Sc., Univ. of Leiden, 1985.

