stichting
mathematisch
centrum

MC

AFDELING NUMERIEKE WISKUNDE
(DEPARTMENT OF NUMERICAL MATHEMATICS)

NW 32/76 AUGUSTUS

C. DEN HEIJER

ITERATIVE SOLUTION OF NONLINEAR EQUATIONS BY IMBEDDING METHODS

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

AMS(MOS) subject classification scheme (1970): 65J05

Iterative solution of nonlinear equations by imbedding methods

by

C. den Heijer

ABSTRACT

A class of stationary iterative methods for solving nonlinear equations is constructed. This is done by an imbeddingstechnique. The local convergence behaviour of these methods is investigated. Furthermore the concept of the radius of convergence of an iterative method is introduced. This is a measure of how far from the true solution a startingpoint is allowed to be, the generated sequence still being convergent.

The radius of convergence of Newton's method is given. Furthermore it is proved that all the members of a subclass of the iterative methods constructed here have a greater radius of convergence than Newton's method.

KEY WORDS & PHRASES: *nonlinear equations, imbedding methods, stationary iterative methods, local convergence, radius of convergence.*

1. INTRODUCTION

1.1. *The problem*

Let X be a Hilbertspace, and F: X → X a nonlinear operator. In this report we shall be concerned with iterative methods for solving the equation

$$(1.1.1)\qquad F(x) = 0.$$

Suppose that x^* ∈ X is the solution of (1.1.1). A well-known method for solving (1.1.1) is *Newton's method* defined by

$$(1.1.2)\qquad \begin{aligned} &\text{given } x_0 \in X, \\ &x_{k+1} = x_k - F'(x_k)^{-1}F(x_k), \qquad k = 0,1,\ldots; \end{aligned}$$

where $F'(x)$ denotes the Fréchet-derivative of F at x. However, if the starting point x_0 is not close to x^*, then the sequence $\{x_k\}$ defined in (1.1.2) need not converge to x^*. In that case *imbedding methods* have been shown to be more effective than Newton's method. In these methods, (1.1.1) is transformed into an *initial value problem*. This is done as follows:
Given an operator K: X × X → X such that

$$(1.1.3)\qquad K(x,x) = 0, \qquad \text{for all } x \in X.$$

K may be dependent on F.
Let $x_0 \in X$ be a (bad) initial guess at x^*, and define

$$(1.1.4)\qquad H(t,x) = (1-t)K(x,x_0) + tF(x), \qquad t \in [0,1], \qquad x \in X.$$

Thus we have

$$(1.1.5)\qquad \begin{aligned} &H(0,x_0) = 0, \\ &H(1,x) \equiv F(x). \end{aligned}$$

Suppose that $H(t,x) = 0$ has, for any $t \in [0,1]$, a unique solution $x(t)$, i.e.

$$H(t,x(t)) = 0, \quad x(t) \text{ unique}, \ t \in [0,1]. \tag{1.1.6}$$

Note that

$$\begin{aligned} x(0) &= x_0, \\ x(1) &= x^*. \end{aligned} \tag{1.1.7}$$

Differentiation with respect to t yields

$$H_1(t,x(t)) + H_2(t,x(t))\dot{x}(t) = 0, \tag{1.1.8}$$

where H_1 and H_2 are the partial Fréchet-derivatives of H with respect to t and x respectively and $\dot{x}(t)$ denotes $\frac{d}{dx}\, x(t)$. If

$$g: [0,1] \to \mathbb{R} \tag{1.1.9}$$

is a (given) real function, then (1.1.6) and (1.1.8) yield

$$H_1(t,x(t)) + H_2(t,x(t))\dot{x}(t) + g(t)H(t,x(t)) = 0, \quad t \in [0,1]. \tag{1.1.10}$$

If we assume that $H_2(t,x(t))$ is invertible for $t \in [0,1]$ then according to (1.1.7) and (1.1.10), the curve $x(t)$ satisfies the following initial value problem

$$\begin{aligned} \dot{x}(t) &= -\,H_2(t,x(t))^{-1}[H_1(t,x(t))+g(t)H(t,x(t))], \quad t \in [0,1], \\ x(0) &= x_0. \end{aligned} \tag{1.1.11}$$

With rather weak assumptions about F,K,g and x_0, solving (1.1.11) is equivalent to solving (1.1.1). (cf. [5]). Now, solving (1.1.11) with a (given) Runge-Kutta method, the calculated approximation to $x(1) = x^*$ is x_1, in short $x_1 = G(x_0)$.

Repeat this procedure, i.e. solve (1.1.11), taking x_1 instead of x_0: $x_2 = G(x_1)$, etc.

The iterating function G is determined by K,g, the Runge-Kutta method (and of course F).

In short

$$G(x) \equiv G(x;K,g,\text{"Runge-Kutta method"}). \tag{1.1.12}$$

The problem we are concerned with is, how the convergence behaviour of iterating functions G of type (1.1.12) is. We first give an example.

<u>EXAMPLE</u>. Take

$$K(x,y) = F(x) - F(y),$$
$$g(t) \equiv 0.$$

Let $x_0 \in X$, then

$$H(t,x) = (1-t)[F(x)-F(x_0)] + tF(x), \tag{1.1.13}$$

and (1.1.1) is transformed into the initial value problem

$$\begin{aligned} \dot{x}(t) &= - F'(x(t))^{-1}F(x_0), \\ x(0) &= x_0. \end{aligned} \tag{1.1.14}$$

Solving (1.1.14) with Euler's method taking a stepsize $h = \frac{1}{N}$, where N is a natural number, then

$$x_1 = y_N, \tag{1.1.15a}$$

where

$$\begin{aligned} y_0 &= x_0 \\ y_i &= y_{i-1} - \frac{1}{N} F'(y_{i-1})^{-1}F(x_0), \qquad i = 1,\dots,N. \end{aligned} \tag{1.1.15b}$$

y_i is an approximation to $x(\frac{i}{N})$, $i = 0,1,\ldots,N$.

The iterating function G is now defined by

(1.1.16a) $G(x) = y_N(x)$,

where

(1.1.16b)
$$y_o(x) = x$$
$$y_i(x) = y_{i-1}(x) - \frac{1}{N} F'(y_{i-1}(x))^{-1}F(x), \qquad i = 1,\ldots,N.$$

In chapter 2 we introduce some conventions. The radius of convergence of an iterative method is also introduced. This is a measure to indicate how far a starting point x_0 of an iterative process is allowed to be from x^*, while the generated sequence $x_0,x_1,x_2,\ldots$ still converges to x^*. We end this chapter with some elementary results which will be used subsequently.

In chapter 3 we give an explicit expression of the iterating function to be considered, in terms of F,K,g and the Runge-Kutta method.

In chapter 4 we investigate the restrictions to be imposed on K in order to prevent the construction of iterative methods with a zero radius of convergence.

In chapter 5, the radius of convergence of Newton's method is given.

Finally, in chapter 6 we construct a class of iterative methods whose members all have a greater radius of convergence than Newton's method.

Test results for the methods considered here will be given in a following report.

2. NOTATIONS, CONVENTIONS AND SOME ELEMENTARY RESULTS

2.1. *Conventions and Notations*

From now on the following conventions hold:
X is a real Hilbertspace, with innerproduct $(\cdot,\cdot)$, and norm $\|\cdot\| = (\cdot,\cdot)^{\frac{1}{2}}$.
If $A: D \to X$, $D \subset X$, then $A'(x)$ denotes the *Fréchet-derivative* of A at x, for $x \in$ interior (D).
Let $X_1,\ldots,X_{n+1}$ be Hilbertspaces and $X_0 = X_1 \times X_2 \times\ldots\times X_n$ the

productspace. If $G: D \to X_{n+1}$, $D \subset X_0$, then for $x = (x_1,\ldots,x_n) \in$ interior (D), $G_i(x)$ denotes the *partial Fréchet-derivative* of G with respect to x_i at x, $i = 1,\ldots,n$.
Let $x: [0,1] \to X$, then $\dot{x}(t)$ denotes $\frac{d}{dt}\, x(t)$, $t \in [0,1]$.
For a formal definition of these concepts, see [1].

For $x \in X$ and $\rho > 0$, $B(x,\rho) = \{y \mid y \in X, \|y-x\| < \rho\}$. Furthermore, if $V \subset X$ is a subset of X, then $\bar{V}$ denotes the *closure* of V.

2.2. *Iterative methods*

Let

(2.2.1) $\quad \mathcal{G} = \{G \mid G: D \to X,\ D \subset X\}$

and

(2.2.2) $\quad \mathcal{F}^* = \{F \mid F: D \to X,\ D \subset X$ and the equation (1.1.1) has a unique solution$\}$.

For given $F \in \mathcal{F}^*$, x^* will always denote the unique solution of (1.1.1).

Let $\{G_k\} = G_0,G_1,\ldots$, where $G_k \in \mathcal{G}$ has domain $D_k \subset X$, $k = 0,1,\ldots$.
Then

(2.2.3) $\quad D(\{G_k\}) = \{x_0 \mid$ there exists a sequence $\{x_k\}$ such that $x_k \in D_k$ and $x_{k+1} = G_k(x_k)$, $k = 0,1,\ldots\}$.

For a (given) subset $\mathcal{F}_0 \subset \mathcal{F}^*$ let

(2.2.4) $\quad \mathcal{M}_0 = \{M \mid M: \mathcal{F}_0 \to \mathcal{G}\}$.

Any sequence $\{M_k\} = M_0,M_1,\ldots$ with $M_k \in \mathcal{M}_0$, $k = 0,1,\ldots$, is called an *iterative method* (applicable to $\mathcal{F}_0$).

To any iterative method $\{M_k\}$ and $F \in \mathcal{F}_0$ the related *iterative process* $(\{M_k\},F)$ is defined by

(2.2.5a) $\quad x_{k+1} = G_k(x_k), \qquad k = 0,1,\ldots;$

where

(2.2.5b) $\quad G_k = M_k(F), \qquad k = 0,1,\ldots .$

The *starting point* x_0 of (2.2.5a) should be an element of $D(\{G_k\})$ in order to prevent the iterative process breaking off prematurely.
Given an iterative process $(\{M_k\},F)$ and $x_0 \in D(\{G_k\})$, then the sequence $\{x_k\}$ *generated by* x_0 and the iterative process $(\{M_k\},F)$ is, of course defined by (2.2.5).

Let $F \in \mathcal{F}_0$, $\{M_k\}$ be an iterative method applicable to $\mathcal{F}_0$, $G_k = M_k(F)$, $k = 0,1,\ldots$.
Then the *region of convergence* $S = S(\{M_k\},F)$ of the iterative process $(\{M_k\},F)$ is defined by

(2.2.6) $\quad S = \{x_0 \mid x_0 \in D(\{G_k\})$ and the sequence $\{x_k\}$ generated by x_0 and $(\{M_k\},F)$ converges to $x^*\}$.

If $x^* \in$ interior (S) then the iterative process $(\{M_k\},F)$ is said to be *locally convergent*.

Let $(\{M_k\},F)$ be a locally convergent iterative process.
If a neighbourhood V of x^* and a $\delta > 0$ exists such that

1. $V \subset S$,
2. for all $x_0 \in V$ the sequence $\{x_k\}$ generated by x_0 and $(\{M_k\},F)$ satisfies

$$\|x_{k+1} - x^*\| \leq \delta \|x_k - x^*\|^2, \qquad k = 0,1,\ldots;$$

then the iterative process $(\{M_k\},F)$ is said to be *locally, quadratically convergent*.
If a neighbourhood V of x^* exists such that

1. $V \subset S$,
2. for all $x_0 \in V$ the sequence $\{x_k\}$ generated by x_0 and $(\{M_k\},F)$ satisfies

$$\|x_{k+1} - x^*\| \leq \|x_k - x^*\|, \qquad k = 0,1,\ldots;$$

then the iterative process $(\{M_k\},F)$ is said to be *locally, monotonically convergent*.

2.3. *The radius of convergence*

Let $F \in \mathcal{F}^*$.
As pointed out in the previous chapter, we are interested in iterative methods $\{M_k\}$, such that the related interative processes $(\{M_k\},F)$ generate sequences $\{x_k\}$ that converge to x^*, even if x_0 is not close to x^*.

In order to be able to compare iterative methods by this criterion, we introduce the following definitions.

Let $\mathcal{F}_0 \subset \mathcal{F}^*$.

DEFINITION 2.3.1. For $F \in \mathcal{F}_0$ and iterative method $\{M_k\}$ (applicable to $\mathcal{F}_0$),

$$r(\{M_k\},F) = \sup\{\rho \mid B(x^*,\rho) \subset S(\{M_k\},F)\}$$

is called the *radius of convergence of the iterative process* $(\{M_k\},F)$.

DEFINITION 2.3.2. For an iterative method $\{M_k\}$ (applicable to $\mathcal{F}_0$),

$$r(\{M_k\}) = \inf_{F\in\mathcal{F}_0} r(\{M_k\},F)$$

is called the *radius of convergence of the iterative method* $\{M_k\}$ with respect to $\mathcal{F}_0$.

It is clear that, the larger $r(\{M_k\})$ is for an iterative method $\{M_k\}$, the better the convergence behaviour will be for the iterative processes generated by it.

2.4. *Stationary iterative methods*

Let $\mathcal{F}_0 \subset \mathcal{F}^*$.
In this report we restrict our attention to *stationary iterative methods* $\{M_k\}$. This means that $M_k = M$, $k = 0,1,\ldots$.

Let $F \in \mathcal{F}_0$, $\{M\}$ be a (stationary) iterative method (applicable to $\mathcal{F}_0$). The operator $G = M(F)$ is, in this connection, called an *iterating function*.

It is clear that, for the iterative process $(\{M\},F)$ to have a positive radius of convergence, this process should at least be locally convergent. The (local) convergence behaviour of the iterative process $(\{M_k\},F)$ is, of course, closely related to the behaviour of $G = M(F)$.
The following expresses this relation

<u>THEOREM 2.4.1</u>. *If the iterative process* $(\{M\},F)$ *is locally convergent and* $G = M(F)$ *is continuous in a neighbourhood of* x^* *then*

$$G(x^*) = x^*.$$

This means that x^* is a fixed point of G.

Conversely, when x^* is a fixed point of G, we have

<u>THEOREM 2.4.2</u>. *Let* $(\{M\},F)$ *be an iterative process. If* x^* *is a fixed point of* $G = M(F)$ *and* $\|G'(x^*)\| < 1$ *then the iterative process* $(\{M\},F)$ *is locally convergent.*

<u>PROOF</u>. Let $\varepsilon > 0$ be such that $\|G'(x^*)\| = 1 - 2\varepsilon$. Then there exists a $\rho > 0$ such that

$$\|G(x) - G(x^*) - G'(x^*)(x-x^*)\| \leq \varepsilon\|x - x^*\|, \quad \text{for any } x \in B(x^*,\rho).$$

Hence

$$\|G(x) - x^*\| \leq \|G(x) - G(x^*) - G'(x^*)(x-x^*)\| + \|G'(x^*)(x-x^*)\| \leq$$

$$\leq (1-\varepsilon)\|x - x^*\|.$$

The conclusion is immediate. □

2.5. *Classes of operators*

In this report we restrict our attention to operators F which are members of the following subset of $\mathcal{F}^*$.

Let $\beta,\gamma > 0$ be given, then

$$(2.5.1) \qquad \mathcal{F} < \beta,\gamma > = \{F \mid F \in \mathcal{F}^*,\ D = X;$$
$$F'(x) \text{ exists and } \|F'(x)-F'(y)\| \le \gamma\ \|x-y\|$$
$$\text{for all } x,y \in X;\ \|F'(x^*)^{-1}\| \le \beta\}.$$

Let

$$(2.5.2) \qquad \mathcal{F}_1 = \bigcup_{\beta,\gamma>0} \mathcal{F} < \beta,\gamma > ,$$

then the auxiliary operator K (see (1.1.3)) is assumed to be a member of the following class of operators

$$(2.5.3) \qquad \mathcal{K}_1 = \{K \mid K: X \times X \times \mathcal{F}_1 \to X,$$

For all $F \in \mathcal{F}$, the operator $K(x,y;F)$ has the following properties

1. $K(x,x;F) = 0$, for all $x \in X$,
2. $K_1(x,y;F)$ exists for all $x,y \in X$,
3. there are $\delta_1,\delta_2 > 0$ and a neighbourhood V of x^* such that
   $\|K_1(y,x;F) - K_1(z,x;F)\| \le \delta_1\|y - z\|$,
   $\|K_1(x^*,x^*;F) - K_1(x^*,x;F)\| \le \delta_2\|x - x^*\|$, for all $x,y,z \in V\}$.

If $F \in \mathcal{F}_1$ is given, then, for ease of notation, we shall write $K(x,y)$ instead of $K(x,y;F)$ when no confusion is possible.

Examples

Given $F \in \mathcal{F}_1$,

1. $K(x,y) = F(x) - F(y)$,
2. $K(x,y) = F'(y)(x-y)$,
3. $K(x,y) = x - y$.

2.6. *Some results from analysis*

We give here three theorems that will be used subsequently.

THEOREM 2.6.1. (cf.[3]). *If* L *and* M *are bounded linear operators in* X,

$$M^{-1} \text{ exists and } \|M - L\| < \frac{1}{\|M^{-1}\|},$$

then

$$L^{-1} \text{ exists and } \|L^{-1}\| \le \frac{\|M^{-1}\|}{1-\|M^{-1}\|\,\|M-L\|}.$$

THEOREM 2.6.2. (cf[3]). *If* $F: D \to X$, $D \subset X$, D *open and convex,* $F'(x)$ *exists and* $\|F'(x)\| \le \delta$ *for all* $x \in D$, *then*

$$\|F(x) - F(y)\| \le \delta\|x - y\|, \qquad \text{for all } x,y \in D.$$

THEOREM 2.6.3. *If* $F: X \to X$ *is Fréchet-differentiable in* X *and* $\|F'(x) - F'(y)\| \le \gamma\|x - y\|$ *for all* $x,y \in X$ *and some* $\gamma > 0$, *then*

$$\|F(x) - F(y) - F'(y)(x-y)\| \le \frac{\gamma}{2}\|x - y\|^2, \qquad \text{for all } x,y \in X.$$

PROOF. This result follows from the fundamental theorem of the differential and integral calculus (cf.[3]):

$$\|F(x) - F(y) - F'(y)(x-y)\| = \left\|\int_0^1 [F'(\theta x+(1-\theta)y)-F'(y)](x-y)d\theta\right\| \le$$

$$\le \frac{\gamma}{2}\|x - y\|^2. \qquad \square$$

3. CLASS OF ITERATIVE METHODS

Before we construct the iterative methods to be dealt with in this report, we define the Runge-Kutta methods to be used.

3.1. *Runge-Kutta methods*

Let

$$(3.1.1)\qquad \begin{aligned} \dot{y}(t) &= f(t,y(t)), \qquad t \in [a,b], \\ y(0) &= y_0 \end{aligned}$$

be an initial value problem to be solved, where $f: [a,b] \times D \to X$, $D \subset X$ and $y_0 \in D$ are given.
Computational methods for solving (3.1.1) approximate the analytical solution $y(t)$ of (3.1.1) on a discrete point set $\{t_n \mid a = t_0 < t_1 < \ldots < t_N = b\}$.

Runge-Kutta methods are one-step methods, which means that, starting from y_0 and t_0, approximations y_n of $y(t_n)$, $n = 1,\ldots,N$ are obtained by

$$(3.1.2a)\qquad y_{i+1} = y_i + h_i\Phi(t_i,y_i;h_i,f), \qquad i = 0,1,\ldots,N-1;$$

where

$$(3.1.2b)\qquad h_i = t_{i+1} - t_i, \qquad i = 0,1,\ldots,N-1.$$

The function Φ is characteristic for the method. We therefore define a Runge-Kutta method in terms of Φ.

DEFINITION 3.1.1. Let $\Lambda = (\lambda_{j,\ell})$ be a strictly lower triangular $(m+1)\times(m+1)$ matrix. Then the general *m-stage Runge-Kutta method* is defined by

$$(3.1.3a)\qquad \Phi(t,y;h,f) = \sum_{\ell=1}^{m} \lambda_{m+1,\ell}k_\ell ,$$

where

$$\text{(3.1.3b)} \qquad \begin{aligned} k_1 &= f(t,y) \\ k_\ell &= f(t+\eta_\ell h, y+h \sum_{j=1}^{\ell-1} \lambda_{\ell,j} k_j), \qquad \ell = 2,\dots,m; \end{aligned}$$

and

$$\text{(3.1.3c)} \qquad \eta_\ell = \sum_{j=1}^{\ell-1} \lambda_{\ell,j}, \qquad \ell = 2,\dots,m.$$

The matrx Λ is called the *generating matrix* of the Runge-Kutta method, which, obviously, completely determines the method.

For the sake of shortness we shall use the phrase "Runge-Kutta method Λ" to mean "Runge-Kutta method with generating matrix Λ".

Moreover, given a Runge-Kutta method $\Lambda = (\lambda_{j,\ell})$, then η_ℓ is always supposed to satisfy (3.1.3c). It is usual to restrict oneself to Runge-Kutta methods for which

$$\text{(3.1.4a)} \qquad \sum_{\ell=1}^{m} \lambda_{m+1,\ell} = 1,$$

$$\text{(3.1.4b)} \qquad \eta_\ell \in (0,1], \qquad \ell = 2,\dots,m.$$

The initial value problem we want to solve is of type (1.1.11). This means that $a = 0$ and $b = 1$ in (3.1.1). In this particular case an $N \times m$-stage Runge-Kutta method $\tilde{\Phi}(t,y;h,f)$ exists with generating matrix $\tilde{\Lambda} = (\tilde{\lambda}_{j,\ell})$, such that $\tilde{y}_1 = y_N$, where

$$\tilde{y}_1 = y_0 + \tilde{\Phi}(0,y_0;1,f).$$

Moreover it is easy to see that $\sum_{\ell=1}^{N\times m} \tilde{\lambda}_{N\times m+1,\ell} = 1$, and $\tilde{\eta}_\ell \in (0,1]$, where $\tilde{\eta}_\ell = \sum_{j=1}^{\ell-1} \tilde{\lambda}_{\ell,j}$, $\ell = 2,\dots,N\times m$.

Therefore, as we are only interested in the Runge-Kutta approximation in $t = 1$, it is no restriction to assume that in (3.1.2), $N = 1$.

3.2. *Description of the iterative methods*

Let $F \in \mathcal{F}_1$, $K \in \mathcal{K}_1$, $g:[0,1] \to \mathbb{R}$ and a Runge-Kutta method Λ be given.

Let

(3.2.1) $$D = \{(t,x,y) \mid t \in [0,1];\ x,y \in X;\ [(1-t)K_1(x,y)+tF'(x)]^{-1} \text{ exists}\}.$$

For a given $x_0 \in X$, let the curve $x(t)$, defined in (1.1.6) satisfy $(t,x(t),x_0) \in D$, for all $t \in [0,1]$. Then we recall from chapter 1 that the curve $x(t)$ is a solution of the initial value problem

(3.2.2) $$\begin{aligned} \dot{x}(t) &= -[(1-t)K_1(x(t),x_0) + tF'(x(t))]^{-1} \times \\ &\quad [-K(x(t),x_0)+F(x(t))+g(t)\{(1-t)K(x(t),x_0)+tF(x(t))\}], \\ &\qquad t \in [0,1], \\ x(0) &= x_0. \end{aligned}$$

Consider $f: D \to X$,

(3.2.3) $$\begin{aligned} f(t,x,y) &= -[(1-t)K_1(x,y)+F'(x)]^{-1} \times \\ &\quad [-K(x,y)+F(x)+g(t)\{(1-t)K(x,y)+tF(x)\}], \quad (t,x,y) \in D. \end{aligned}$$

Then (3.2.2) is equivalent to

(3.2.4) $$\begin{aligned} \dot{x}(t) &= f(t,x(t),x_0), \quad t \in [0,1], \\ x(0) &= x_0. \end{aligned}$$

The Runge-Kutta approximation x_1 of $x(1) = x^*$ is given by

(3.2.5a) $$x_1 = x_0 + \sum_{\ell=1}^{m} \lambda_{m+1,\ell} k_\ell(x_0)$$

where

$$k_1(x_0) = f(0,x_0,x_0)$$

(3.2.5b)

$$k_\ell(x_0) = f(\eta_\ell, x_0 + \sum_{j=1}^{\ell-1} \lambda_{\ell,j} k_j(x_0), x_0), \quad \ell = 2,\ldots,m.$$

We have written $k_1(x_0)$ and $k_\ell(x_0)$ instead of k_1 and k_ℓ to emphasize the dependence of k_1 and k_ℓ on x_0.

It is clear that if we repeat this process in the way described in chapter 1, the generated sequence $\{x_k\}$ might be considered as being generated by x_0 and an iterative process $(\{M\},F)$ with iterating function $G = M(F)$ defined as

$$\text{(3.2.6a)} \quad G(x) = x + \sum_{\ell=1}^{m} \lambda_{m+1,\ell} k_\ell(x),$$

where

$$k_1(x) = -\, K_1(x,x)^{-1}F(x),$$

(3.2.6b)

$$k_\ell(x) = -\,[(1-\eta_\ell)K_1(x+\sum_{j=1}^{\ell-1} \lambda_{\ell,j}k_j(x),x)+\eta_\ell F'(x+\sum_{j=1}^{\ell-1}\lambda_{\ell,j}k_j(x))]^{-1} \times$$

$$[-K(x+\sum_{j=1}^{\ell-1}\lambda_{\ell,j}k_j(x),x) + F(x+\sum_{j=1}^{\ell-1}\lambda_{\ell,j}k_j(x)) +$$

$$+ g(\eta_\ell)\{(1-\eta_\ell)K(x+\sum_{j=1}^{\ell-1}\lambda_{\ell,j}k_j(x),x)+\eta_\ell F(x+\sum_{j=1}^{\ell-1}\lambda_{\ell,j}k_j(x))\}],$$

$$\ell = 2,\ldots,m.$$

We define $D(G)$ for G of type (3.2.6) as

$$\text{(3.2.7)} \quad D(G) = \{x \mid x \in X, \text{ in } x \text{ all inverses appearing in (3.2.6b) exist}\}.$$

Obviously, the operator M depends on K,g and Λ

$$\text{(3.2.8)} \quad M\colon \mathcal{F}_1 \to \mathcal{G},\ M(\cdot) \equiv M(K,g,\Lambda;\cdot).$$

From now on, for given $K \in \mathcal{K}_1$, $g\colon [0,1] \to \mathbb{R}$ and Runge-Kutta method Λ, we shall use the phrase "$M(\cdot) \equiv M(K,g,\Lambda;\cdot)$" to mean "$M(\cdot) \equiv M(K,g,\Lambda;\cdot)$, where M is of type (3.2.8)".

In the next chapters we shall investigate the convergence behaviour of iterative methods $\{M\}$ where $M(\cdot) \equiv M(K,g,\Lambda;\cdot)$ for given K,g and Λ.

4. LOCAL CONVERGENCE BEHAVIOUR OF THE ITERATIVE PROCESSES

Let $F \in \mathcal{F}_1$.

For given $K \in \mathcal{K}_1$, $g: [0,1] \to \mathbb{R}$ and Runge-Kutta method Λ, let $M(\cdot) = M(K,g,\Lambda;\cdot)$. Then $G = M(F)$ is of type (3.2.6), $G: D(G) \to X$.

It has already been observed (see Chapter 2) that the radius of convergence of the iterative process $(\{M\},F)$ is only positive when $(\{M\},F)$ is locally convergent.

In this chapter we investigate the conditions which have to be imposed on K in order that $(\{M\},F)$ is locally convergent.

Since $K \in \mathcal{K}_1$ we recall from Section 2.5 that there is a neighbourhood V of x^* and δ_1, $\delta_2 > 0$ such that

$$\|K_1(y,x;F) - K_1(z,x;F)\| \leqq \delta_1\|y-z\|,$$
(4.1)
$$\|K_1(x^*,x^*;F) - K_1(x^*,x;F) \leqq \delta_2\|x-x^*\|, \text{ for all } x,y,z \in V.$$

Moreover $F \in \mathcal{F}_1$ implies that there are $\beta,\gamma > 0$ such that

$$\|F'(x) - F'(y)\| \leqq \gamma\|x-y\|, \text{ for all } x,y \in X$$
(4.2)
$$\|F'(x^*)^{-1}\| \leqq \beta.$$

Let

(4.3) $$D_1 = \{x \mid K(x,x;F) \text{ is invertible}\}$$

and

$$(4.4)\qquad \begin{aligned} &G_1 : D_1 \to X, \\ &G_1(x) = x - K_1(x,x;F)^{-1}F(x), \text{ for all } x \in D_1. \end{aligned}$$

LEMMA 4.1. *If* $x^* \in$ *interior* (D_1) *then* $G_1'(x^*)$ *exists and* $G_1'(x^*) = I - K_1(x^*,x^*;F)^{-1}F'(x^*)$.

PROOF. For ease of notation we suppress the dependence of K on F. Since $K_1(x^*,x^*)$ is bounded and invertible, there is an $\alpha > 0$ such that $\|K_1(x^*,x^*)^{-1}\| \leq \alpha$. Now,

$$K_1(y,x) - K_1(x^*,x^*) = K_1(y,x) - K_1(x^*,x) + K_1(x^*,x) - K_1(x^*,x^*)$$

so that, using (4.1),

$$(4.5)\qquad \|K_1(y,x) - K_1(x^*,x^*)\| \leqq \delta_1\|y-x^*\| + \delta_2\|x-x^*\| \text{ for all } x,y \in V.$$

Let $\rho = \dfrac{1}{2\alpha(\delta_1+\delta_2)}$, then Theorem 2.6.1 yields that for $x \in B(x^*,\rho) \cap V$, $K_1(x,x)$ is invertible and

$$(4.6)\qquad \|K_1(x,x)^{-1}\| \leqq 2\alpha.$$

If P and Q are bounded, invertible linear operators on X, then $P^{-1} - Q^{-1} = Q^{-1}(Q-P)P^{-1}$, so

$$(4.7)\qquad \|P^{-1}-Q^{-1}\| \leqq \|P^{-1}\|\ \|Q^{-1}\|\ \|P-Q\|.$$

Let $\tau > 0$ such that $\sup\{\|F'(x)\| \mid x \in B(x^*,\rho)\} \leqq \tau$, then using Theorem 2.6.2,

$$(4.8)\qquad \|F(x)\| \leqq \tau\|x-x^*\| \text{ for all } x \in B(x^*,\rho).$$

Moreover, Theorem 2.6.3 yields

$$(4.9)\qquad \|F(x) - F'(x^*)(x-x^*)\| \leqq \frac{\gamma}{2}\|x-x^*\|^2 \text{ for all } x \in X.$$

Using (4.5), (4.6), (4.7), (4.8) and (4.9), for $x \in B(x^*,\rho) \cap V$:

$$\|G_1(x) - G_1(x^*) - [I - K_1(x^*,x^*)^{-1}F'(x^*)]\,(x-x^*)\| =$$
$$= \|-K_1(x,x)^{-1}F(x) + K_1(x^*,x^*)^{-1}F'(x^*)(x-x^*)\| =$$
$$= \|[-K_1(x,x)^{-1} + K_1(x^*,x^*)^{-1}]F(x) - K_1(x^*,x^*)^{-1}[F(x) - F'(x^*)(x-x^*)]\| \leq$$
$$\leq [2\alpha^2(\delta_1+\delta_2)\tau + \alpha\tfrac{\gamma}{2}]\,\|x^*-x\|^2. \quad \square$$

Let

(4.10) $\qquad D_2 = \{x \mid F'(x) \text{ is invertible}\}$

and

$$G_2\colon D_2 \to X,$$

(4.11) $\qquad G_2(x) = F'(x)^{-1}F(x)$, for all $x \in D_2$.

<u>LEMMA 4.2</u>. $G_2'(x^*)$ *exists and* $G_2'(x^*) = I$.

<u>PROOF</u>. According to (4.2) and Theorem 2.6.1, $B(x^*,\frac{1}{2\beta\gamma}) \subset D_2$ and

(4.12) $\qquad \|F'(x)^{-1}\| \leqq 2\beta$ for all $x \in B(x^*,\frac{1}{2\beta\gamma})$.

Now, $F(x) + F'(x)(x^*-x) = r(x)$, where $\|r(x)\| \leqq \frac{\gamma}{2}\|x-x^*\|^2$ for all $x \in X$ (Theorem 2.6.3). Therefore, for $x \in B(x^*,\frac{1}{2\beta\gamma})$:

(4.13) $\qquad \|F'(x)^{-1}F(x) - (x-x^*)\| = \|F'(x)^{-1}r(x)\| \leqq \beta\gamma\|x^*-x\|^2. \quad \square$

Let

(4.14) $\qquad D_3 = \{x \mid x \in D_1 \text{ and } F'(G_1(x)) \text{ is invertible}\}$ and

$$G_3\colon D_3 \to X$$

(4.15)

$$G_3(x) = x - F'(G_1(x))^{-1}[-K(G_1(x),x;F) + p\,F(G_1(x))], \text{ for all}$$

$x \in D_3$, where $p \in \mathbb{R}$.

LEMMA 4.3. *If* $x^* \in$ *interior* (D_3) *then* $G_3'(x^*)$ *exists and*

$$G_3'(x^*) = -p[I - K_1(x^*,x^*;F)^{-1}F'(x^*)].$$

PROOF. Again, for ease of notation, we suppress the dependence of K on F. There are $\alpha, \rho, \tau > 0$ such that (4.8) holds and for all $x \in B(x^*,\rho)$, $\|F'(x)^{-1}\| \leqq 2\beta$, $\|K_1(x,x)^{-1}\| \leqq \alpha$ and $\|G_1(x) - x^*\| \leqq \frac{1}{2\beta\gamma}$. This last inequality implies that for $x \in B(x^*,\rho)$, $\|F'(G_1(x))^{-1}\| \leqq 2\beta$. (see Theorem 2.6.1). So $B(x^*,\rho) \subset D_3$. For $x \in B(x^*,\rho)$:

$$\|F'(G_1(x))^{-1}K(G_1(x),x) + x - x^*\| =$$

$$= \|F'(G_1(x))^{-1}[K_1(x,x)(G_1(x)-x) + r_1(x)] + x - x^*\| =$$

$$= \|-F'(G_1(x))^{-1}F(x) + F'(G_1(x))^{-1}r_1(x) + x - x^*\| =$$

$$= \|-[F'(G_1(x))^{-1}-F'(x)^{-1}]F(x) + F'(G_1(x))^{-1}r_1(x) - F'(x)^{-1}F(x) + x - x^*\|,$$

where

$$\|r_1(x)\| \leqq \frac{\delta_1}{2} \|G_1(x) - x\|^2 \text{ (see Theorem 2.6.3).}$$

$$G_1(x) - x = - K_1(x,x)^{-1}F(x), \text{ so}$$

$$\|G_1(x) - x\| \leqq \alpha.\tau\|x - x^*\|, \text{ for } x \in B(x^*,\rho). \quad (4.16)$$

Using (4.7), (4.8) and (4.13):

$$\|F'(G_1(x))^{-1}K(G_1(x),x) + x - x^*\| \leqq$$

$$\leqq [(4\beta)^2\gamma.\alpha\tau^2 + 2\beta\frac{\delta_1}{2} + \beta\gamma] \|x - x^*\|^2 \text{ for all } x \in B(x^*,\rho). \quad (4.17)$$

Let $k: D_3 \to X$,

$$k(x) = G_2(G_1(x)) \text{ for all } x \in D_3.$$

Since $G_1(x^*) = x^*$ and $G_1'(x^*)$ and $G_2'(x^*)$ exist, $k'(x^*)$ exists and $k'(x^*) =$ $= G_2'(x^*)G_1'(x^*) = I - K_1(x^*,x^*)^{-1}F'(x^*)$. Then for $\varepsilon > 0$, there is a $\rho_1 > 0$ such that

$$(4.18)\qquad \|F'(G_1(x))^{-1}F(G_1(x)) - [I - K_1(x^*,x^*)^{-1}F'(x^*)](x-x^*)\| \leqq \varepsilon \|x - x^*\|, \quad \text{for all } x \in B(x^*,\rho_1).$$

Let $\rho_2 = \min\{\rho,\rho_1\}$, then for $x \in B(x^*,\rho_2)$

$$\|G_3(x) - G_3(x^*) + p[I - K_1(x^*,x^*)^{-1}F'(x^*)](x-x^*)\| =$$

$$= \|x - x^* - F'(G_1(x))^{-1}[-K(G_1(x),x) + pF(G_1(x))] +$$

$$+ p[I - K_1(x^*,x^*)^{-1}F'(x^*)](x-x^*)\| \leqq$$

$$\leqq \|F'(G_1(x))^{-1}K(G_1(x),x) + x - x^*\| +$$

$$+ |p|\ \|F'(G_1(x))^{-1}F(G_1(x)) - [I - K_1(x^*,x^*)^{-1}F'(x^*)](x-x^*)\| \leqq$$

$$\leqq \nu\|x - x^*\|^2 + |p|\ \varepsilon\ \|x - x^*\|,$$

where ν is the term between the square brackets in (4.17), ν is independent of ε, thus the conclusion of Lemma 4.3 holds. □

The next theorem shows the dependence of the local convergence behaviour of the iterative process $(\{M\},F)$ where $M(\cdot) = M(K,g,\Lambda;\cdot)$, on K. Let

$$(4.19)\qquad \Lambda_1 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \text{ and } \Lambda_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

and $g_0:[0,1] \to \mathbb{R}$ a given function.

<u>THEOREM 4.1</u>. *Let* $F \in \mathcal{F}_1$ *and* $K \in \mathcal{K}_1$, *then the following propositions* (i), (ii), (iii), *and* (iv) *are equivalent.*

(i) *The iterative process* $(\{M\},F)$, *where* $M(\cdot) = M(K,g_0,\Lambda_2;\cdot)$ *is locally quadratically convergent.*

(ii) *For any* $g\colon [0,1] \to \mathbb{R}$, *the iterative process* $(\{M\},F)$, *where* $M(\cdot) = M(K,g,\Lambda_1;\cdot)$ *is locally monotonically convergent.*

(iii) *For any Runge-Kutta method* Λ *the iterative process* $(\{M\},F)$, *where* $M(\cdot) = M(K,g_0,\Lambda;\cdot)$ *is locally quadratically convergent.*

(iv) $K_1(x^*,x^*;F) = F'(x^*)$.

PROOF. We shall prove: 1. (i) implies (iv)
2. (iv) implies (iii)
3. (iii) implies (i)
4. (iv) implies (ii)
5. (ii) implies (iv)

Of course, this is sufficient to prove that (i), (ii), (iii) and (iv) are equivalent.

1. Suppose proposition (i) holds and let $K_1(x^*,x^*;F) \neq F'(x^*)$. Let $G = M(F)$, where $M(\cdot) = M(K,g_0,\Lambda_2;\cdot)$, then

$$G(x) = x - K_1(x,x;F)^{-1}F(x) \quad \text{for all } x \in D(G).$$

Since the iterative process $(\{M\},F)$ is locally quadratically convergent, there is a neighbourhood $V \subset D(G)$ of x^* and a $\delta > 0$ such that $\|G(x) - x^*\| \leqq \delta\|x - x^*\|^2$ for all $x \in V$. Hence Lemma 4.1 applies, so $G'(x^*)$ exists and

$$G'(x^*) = I - K_1(x^*,x^*;F)^{-1}F'(x^*).$$

As $K_1(x^*,x^*;F) \neq F'(x^*)$, some $y \in X$, $y \neq 0$ exists such that $\|G'(x^*)y\| = L\|y\|$, $L > 0$. By L, a positive ρ exists such that

$$\|G(x) - G(x^*) - G'(x^*)(x-x^*) \leqq \frac{L}{2}\|x-x^*\|, \text{ for all } x \in B(x^*,\rho) \subset V. \tag{4.20}$$

Moreover, a $t_1 \in (0,\frac{L}{2\delta\|y\|})$ exists such that $x^* + ty \in B(x^*,\rho)$ for all $t \in [0,t_1]$. For $t \in [0,t_1]$:

$$\|G(x^*+ty) - G(x^*) - G'(x^*)ty\| \geqq$$
$$\geqq \|G'(x^*)ty\| - \|G(x^*+ty) - x^*\| \geqq$$
$$\geqq L\|ty\| - \delta\|ty\|^2 >$$
$$> L\|ty\| - \frac{\delta L\|y\|}{2\delta\|y\|} \cdot \|ty\| =$$
$$= \frac{L}{2}\|y\|.$$

This yields a contradiction to (4.20). So (i) implies (iv).

2. Suppose

$$K_1(x^*,x^*;F) = F'(x^*). \tag{4.21}$$

Let $\Lambda = (\lambda_{\ell,j})$ be an m-stage Runge-Kutta method. Let $G = M(F)$, where $M(\cdot) = M(K,g_0,\Lambda;\cdot)$. For ease of notation we suppress the dependence of K on F.

Now, let $D_0 = X$ and

$$\begin{aligned} &G_0 : D_0 \to X, \\ &G_0(x) = x, \quad \text{for all } x \in D_0. \end{aligned} \tag{4.22}$$

With $\eta_1 = 0$, define for $\ell = 1,\ldots,m$

$$D_\ell = \{x \mid x \in D_{\ell-1}, [(1-\eta_\ell)K_1(G_{\ell-1}(x),x) + \eta_\ell F'(G_{\ell-1}(x))] \text{ is invertible}\}, \tag{4.23}$$

and

$$\begin{aligned} &G_\ell : D_\ell \to X \\ &G_\ell(x) = x + \sum_{j=1}^{\ell} \lambda_{\ell+1,j} k_j(x), \quad \text{for all } x \in D_\ell. \end{aligned} \tag{4.24}$$

Note that $G_m = G$.

We shall prove by induction the following proposition:
For $\ell = 1,\ldots,m$ there exist $\rho_\ell, \sigma_\ell > 0$ such that

a. $B(x^*,\rho_\ell) \subset D_\ell$.

b. $k_\ell(x) = -F'(x)^{-1}F(x) + r_\ell(x)$, where $\|r_\ell(x)\| \leqq \sigma_\ell \|x-x^*\|^2$ for all $x \in B(x^*,\rho_\ell)$.

c. $k_\ell'(x^*) = -I$.

Let $\ell = 1$: $G_1(x) = x + \lambda_{2,1}k_1(x) = x - \eta_2 K_1(x,x)^{-1}F(x)$ for $x \in D_1$.

a. Since $F \in \mathcal{F}_1$ and $K \in \mathcal{K}_1$ there exists a neighbourhood V of x^* and $\delta_1,\delta_2,\beta,\gamma>0$ such that (4.1) and (4.2) hold. Now, using (4.5) and (4.21),

$$(4.25) \qquad \|K_1(x,x) - F'(x^*)\| \leqq (\delta_1+\delta_2)\ \|x-x^*\| \text{ for all } x \in V.$$

Let

$$(4.26) \qquad \rho_1 = \max\{\rho \mid B(x^*,\rho) \subset B(x^*,\frac{1}{2\beta(\delta_1+\delta_2)}) \cap V \cap B(x^*,\frac{1}{2\beta\gamma})\},$$

then from Theorem 2.6.1 it follows that both $K_1(x,x)$ and $F'(x)$ are invertible and

$$(4.27) \qquad \|K_1(x,x)^{-1}\| \leqq 2\beta \text{ and } \|F'(x)^{-1}\| \leqq 2\beta, \qquad \text{for all } B(x^*,\rho_1).$$

This implies that $B(x^*,\rho_1) \subset D_1$.

b. Let $x \in B(x^*,\rho_1)$, then

$$k_1(x) + F'(x)^{-1}F(x) = -[K_1(x,x)^{-1} - F'(x)^{-1}]F(x).$$

Using (4.21),

$$K_1(y,x) - F'(x) = K_1(y,x) - K_1(x^*,x^*) + F'(x^*) - F'(x),$$

thus (4.2) and (4.5) imply that

$$(4.28) \qquad \|K_1(y,x) - F'(x)\| \leqq (\delta_2+\gamma)\ \|x-x^*\| + \delta_1\|y-x^*\|, \quad \text{for all } x,y \in B(x^*,\rho_1).$$

Let $\tau > 0$ be such that $\sup\{\|F'(x)\| \mid x \in B(x^*,\rho_1)\} \leqq \tau$, then (4.8) holds for $\rho = \rho_1$. Then, using (4.7), (4.8), (4.27) and (4.28), (4.29) $\|k_1(x) + F'(x)^{-1}F(x)\| \leqq (2\beta)^2(\delta_1+\delta_2+\gamma)\tau_1\|x-x^*\|^2.$

So, let $\sigma_1 = (2\beta)^2(\delta_1+\delta_2+\gamma)\tau_1$.

c. Obviously, $k_1(x^*) = 0$. So using (4.26), for $x \in B(x^*,\rho_1)$

$$\|k_1(x) - k_1(x^*) + (x-x^*)\| \leqq$$
$$\leqq \|-F'(x)^{-1}F(x) + (x-x^*)\| + \sigma_1\|x-x^*\|^2.$$

Since $B(x^*,\rho_1) \subset B(x^*,\frac{1}{2\beta\gamma})$, (4.13) holds for $x \in B(x^*,\rho_1)$, so

$$\|k_1(x) - k_1(x^*) + (x-x^*)\| \leqq$$
$$\leqq (\beta\gamma+\sigma_1)\|x-x^*\|^2, \quad \text{for all } x \in B(x^*,\rho_1).$$

Therefore, $k_1'(x^*) = -I$.

So for $\ell = 1$ the proposition holds.

Now, suppose that for $j = 1,2,\ldots,\ell-1<m$ the proposition is true. Let ρ_1 satisfy (4.26) which is no restriction. Since

$$G_{\ell-1}(x) = x + \sum_{j=1}^{\ell-1} \lambda_{\ell,j}k_j(x), \quad \text{for all } x \in D_{\ell-1},$$

and

$$k_j'(x^*) = -I, \quad j = 1,\ldots,\ell-1;$$

$G_{\ell-1}'(x^*)$ exists and $\|G_{\ell-1}'(x^*)\| = 1 - \eta_\ell$. So, there is a $\tilde{\rho}_\ell$, $0 < \tilde{\rho}_\ell \leqq \min_{j=1,\ldots,\ell-1} \{\rho_j\}$, such that

$$\|G_{\ell-1}(x) - G_{\ell-1}(x^*) - G_{\ell-1}'(x^*)(x-x^*)\| \leqq \eta_\ell\|x-x^*\|,$$
$$\text{for all } x \in B(x^*,\rho_\ell).$$

Therefore,

$$\|G_{\ell-1}(x)-x^*\| =$$

$$= \|G_{\ell-1}(x) - G_{\ell-1}(x^*) - G'_{\ell-1}(x^*)(x-x^*) + G'_{\ell-1}(x^*)(x-x^*)\| \leqq$$
$$\leqq \|x-x^*\|, \quad \text{for all } x \in B(x^*,\tilde{\rho}_\ell)$$

By (4.2), (4.5) and (4.21) we have

$$\|[(1-\eta_\ell)K_1(G_{\ell-1}(x),x) + \eta_\ell F'(G_{\ell-1}(x))]-F'(x^*)\| \leqq$$
$$\leqq (1-\eta_\ell)[\delta_2\|x-x^*\|+\delta_1\|G_{\ell-1}(x)-x^*\|]+\eta_\ell\gamma\|x-x^*\|,$$
$$\text{for all } x \in B(x^*,\tilde{\rho}_\ell).$$

With $\rho_\ell = \min\{\rho_\ell, \frac{1}{2\beta[(1-\eta_\ell)(\delta_1+\delta_2)+\eta_\ell\gamma]}\}$, Theorem 2.6.1 yields that $[(1-\eta_\ell)K_1(G_{\ell-1}(x),x)+\eta_\ell F'(G_{\ell-1}(x))]$ is invertible and

(4.30) $\|[(1-\eta_\ell)K_1(G_{\ell-1}(x),x)+\eta_\ell F'(G_{\ell-1}(x))]^{-1}\| \leqq 2\beta$ for all $x \in B(x^*,\rho_\ell)$.

Therefore, $B(x^*,\rho_\ell) \subset D_\ell$.

b. For $x \in B(x^*,\rho_\ell)$:

$$k_\ell(x) + F'(x)^{-1}F(x) =$$
$$= -[(1-\eta_\ell)K_1(G_{\ell-1}(x),x)+\eta_\ell F'(G_{\ell-1}(x))]^{-1} \times$$
$$[-K(G_{\ell-1}(x),x)+F(G_{\ell-1}(x))+g_0(\eta_\ell)\{(1-\eta_\ell)K(G_{\ell-1}(x),x)+\eta_\ell F(G_{\ell-1}(x))\}]+$$
$$+ F'(x)^{-1}F(x).$$

From Theorem 2.6.3 it follows that

$$K(G_{\ell-1}(x),x) = K_1(x,x)[G_{\ell-1}(x)-x]+\tilde{s}_1(x), \quad \|\tilde{s}_1(x)\| \leqq \frac{\delta_1}{2}\|G_{\ell-1}(x)-x\|^2.$$

Thus

$$K(G_{\ell-1}(x),x) = [K_1(x,x)-F'(x)][G_{\ell-1}(x)-x] + F'(x)(G_{\ell-1}(x)-x) + \tilde{s}_1(x)$$
$$= [K_1(x,x)-F'(x)][G_{\ell-1}(x)-x] +$$
$$+ F'(x) \sum_{j=1}^{\ell-1} \lambda_{\ell,j}[-F'(x)^{-1}F(x)+r_j(x)] + \tilde{s}_1(x)$$

$$= -\eta_\ell F(x) + [K_1(x,x)-F'(x)][G_{\ell-1}(x)-x] + \sum_{j=1}^{\ell-1} \lambda_{\ell,j} r_j(x) + \tilde{s}_1(x).$$

Hence, using (4.28) and observing that

$$(4.31) \qquad \|G_{\ell-1}(x)-x\| = \|G_{\ell-1}(x)-x^*+x^*-x\| \leqq 2\|x-x^*\|,$$

we see that

$$(4.32) \qquad \begin{aligned} &K(G_{\ell-1}(x),x) = -\eta_\ell F(x) + s_1(x), \\ &\|s_1(x)\| \leqq [(\delta_1+\delta_2+\gamma)2 + \sum_{j=1}^{\ell-1} |\lambda_{\ell,j}|\sigma_j+2\delta_1]\|x-x^*\|^2. \end{aligned}$$

According to Theorem 2.6.3,

$$\begin{aligned} &F(G_{\ell-1}(x)) = F(x) + F'(x)[G_{\ell-1}(x)-x] + \tilde{s}_2(x), \\ &\|\tilde{s}_2(x)\| \leqq \frac{\gamma}{2}\|G_{\ell-1}(x)-x\|^2. \end{aligned}$$

Thus

$$F(G_{\ell-1}(x)) = F(x) + F'(x) \sum_{j=1}^{\ell-1} \lambda_{\ell,j}[-F'(x)^{-1}F(x)+r_j(x)] + \tilde{s}_2(x),$$

and therefore

$$(4.33) \qquad \begin{aligned} &F(G_{\ell-1}(x)) = (1-\eta_\ell)F(x) + s_2(x), \\ &\|s_2(x)\| \leqq [\sum_{j=1}^{\ell-1} |\lambda_{\ell,j}|\sigma_j+2\gamma] \|x-x^*\|^2. \end{aligned}$$

Now, let

$$(4.34) \qquad A(x) = [(1-\eta_\ell)K_1(G_{\ell-1}(x),x) + \eta_\ell F'(G_{\ell-1}(x))],$$

then $\|A(x)^{-1}\| \leqq 2\beta$ (see 4.30), and

$$\begin{aligned} \|A(x) - F'(x)\| \leqq{} & (1-\eta_\ell) \|K_1(G_{\ell-1}(x),x) - F'(x)\| + \\ & + \eta_\ell \|F'(G_{\ell-1}(x)) - F'(x)\|, \end{aligned}$$

So, using (4.28) and (4.30)

$$\|A(x) - F'(x)\| \leq (1-\eta_\ell)[(\delta_2+\gamma)\|x-x^*\| + \delta_1\|G_{\ell-1}(x)-x^*\|] + + \eta_\ell\gamma\|G_{\ell-1}(x)-x^*\| \leq [1-\eta_\ell)(2\delta_1+\delta_2+\gamma) + 2\eta_\ell\gamma]\|x-x^*\|. \tag{4.35}$$

(4.32), (4.33) and (4.34) then yield

$$\begin{aligned} &k_\ell(x) + F'(x)^{-1}F(x) = \\ &= -A(x)^{-1}[\eta_\ell F(x) - s_1(x) + (1-\eta_\ell)F(x) + s_2(x) + \\ &\quad + g(\eta_\ell)\{(1-\eta_\ell)(-\eta_\ell F'(x)^{-1}F(x)+s_1(x))+\eta_\ell((1-\eta_\ell)F'(x)^{-1}F(x)+s_2(x))\}]+ \\ &+F'(x)^{-1}F(x) = \\ &= [-A(x)^{-1} + F'(x)^{-1}]F(x) + \\ &\quad + A(x)^{-1}[(-1+g(\eta_\ell)(1-\eta_\ell))s_1(x)+(1+g(\eta_\ell)\eta_\ell)s_2(x)]. \end{aligned}$$

Let ν_1 and ν_2 be the terms between the square brackets in (4.32) and (4.33) respectively. We recall that a $\tau > 0$ exists such that $\|F(x)\| \leq \tau\|x-x^*\|$, for all $x \in B(x^*,\rho_1)$. Then, using (4.7), (4.27), (4.30), (4.32), (4.33) and (4.35),

$$\begin{aligned} &\|k_\ell(x) + F'(x)^{-1}F(x)\| \leq \\ &\leq \{(2\beta)^2[(1-\eta_\ell)(2\delta_1+\delta_2+\gamma) + 2\eta_\ell\gamma]\tau + \\ &\quad + (2\beta)[|-1 + g(\eta_\ell)(1-\eta_\ell)|\nu_i + |1 + g(\eta_\ell)\eta_\ell|\nu_2]\}\|x-x^*\|^2. \end{aligned}$$

Therefore, for $j = \ell$ proposition b holds.

c. It follows that for $x \in B(x^*,\rho_\ell)$ a $\sigma_\ell > 0$ exists such that

$$k_\ell(x) = -F'(x)^{-1}F(x) + r_\ell(x), \text{ where } \|r_\ell(x)\| \leq \sigma_\ell\|x-x^*\|^2$$

Obviously, $k_\ell(x^*) = 0$, so that

$$\begin{aligned} &\|k_\ell(x) - k_\ell(x^*) + (x-x^*)\| \leq \\ &\leq \|-F'(x)F(x) + (x-x^*)\| + \sigma_\ell\|x-x^*\|^2 \qquad \text{for all } x \in B(x^*,\rho_\ell). \end{aligned}$$

Since $B(x^*,\rho_\ell) \subset B(x^*,\rho_1) \subset B(x^*,\frac{1}{2\beta\gamma})$, (4.13) holds for $x \in B(x^*,\rho_\ell)$, thus

$$\|k_\ell(x) - k_\ell(x^*) + (x-x^*)\| \leqq (\beta\gamma+\sigma_1)\|x-x^*\|^2, \text{ for all } x \in B(x^*,\rho_\ell).$$

Therefore $k_\ell(x^*) = -I$ and the proposition is proved.

Now, with $\rho = \min\{\rho_1,\ldots,\rho_m,\frac{1}{2\beta\gamma}\}$, for $x \in B(x^*,\rho)$,

$$G(x) = x + \sum_{\ell=1}^{m} \lambda_{m+1,\ell}k_\ell(x) = x - F'(x)^{-1}F(x) + \sum_{\ell=1}^{m} \lambda_{m+1,\ell}r_\ell(x).$$

So, (4.13) yields

$$\begin{aligned}\|G(x) - x^*\| &= \|-F'(x)^{-1}F(x) + (x-x^*) + \sum_{\ell=1}^{m} \lambda_{m+1,\ell}r_\ell(x)\| \\ &\leqq \delta\|x-x^*\|^2,\end{aligned}$$

where $\delta = \beta\gamma + \sum_{\ell=1}^{m} |\lambda_{m+1,\ell}|\sigma_\ell$. For $x \in V_1 = B(x^*,\frac{1}{2\delta}) \cap B(x^*,\rho)$, we have

$$\|G(x) - x^*\| \leqq \tfrac{1}{2}\|x-x^*\|,$$

hence, for all $x_0 \in V_1$, the sequence $\{x_k\}$ generated by x_0 and $(\{M\},F)$, remains in V_1, converges to x^* and satisfies

$$\|x_{k+1}-x^*\| \leqq \delta\|x_k-x^*\|^2, \quad k = 0,1,\ldots$$

Therefore, (iv) implies (iii).

3. It is obvious that (iii) implies (i).

4. Suppose $K_1'(x^*,x^*;F) = F'(x^*)$.
Let $g:[0,1] \to \mathbb{R}$. Then proposition (iii) holds for $g_0 = g$, since (iv) implies (iii). Therefore the iterative process $(\{M\},F)$, where $M(\cdot) = M(K,g,\Lambda_1;\cdot)$ is locally, quadratically convergent. So, there is a neighbourhood V of x^* and $\delta > 0$ such that for all $x_0 \in V$, the sequence $\{x_k\}$ generated by x_0 and $(\{M\},F)$ satisfies $x_k \to x^*$ and $\|x_{k+1} - x^*\| \leqq \delta\|x_k - x^*\|^2$ for $k = 0,1,2,\ldots$.

Thus for $x_0 \in B:(x^*, \frac{1}{\delta}) \cap V$ the sequence $\{x_k\}$ generated by x_0 and $(\{M\},F)$ satisfies $x_k \to x^*$ and $\|x_{k+1} - x^*\| \leqq \|x_k - x^*\|$, $k = 0,1,\ldots$.

So (iv) implies (ii).

5. Suppose proposition (ii) holds and let

$$K_1(x^*,x^*) \neq F'(x^*) \tag{4.36}$$

Again, we have suppressed the dependence of K on F. Let

$$D_1 = \{x \mid K_1(x,x) \text{ is invertible}\} \tag{4.37}$$

and

$$\begin{aligned} &G_1 : D_1 \to X, \\ &G_1(x) = x - K_1(x,x)^{-1}F(x). \end{aligned} \tag{4.38}$$

For $g_1 : [0,1] \to \mathbb{R}$, $g_1(t) = 0$ for all $t \in [0,1]$, let $G_2 = M(F)$, where $M(\cdot) = M(K,g_1,\Lambda_1;\cdot)$. Then one easily verifies that $D_1 \supset D(G_2)$. Proposition (ii) implies that $x^* \in D(G_2)$, so $K_1(x^*,x^*)$ is invertible. According to (4.32) there is a $y \in X$, $y \neq 0$, such that $\|[I - K_1(x^*,x^*)^{-1}F'(x^*)]y\| = L\|y\|$, $L > 0$.

Now, let $g_2 : [0,1] \to \mathbb{R}$, such that $p = (1+g(1))L - 1 > 0$. For $G = M(F)$, where $M(\cdot) = M(K,g_2,\Lambda_1;\cdot)$,

$$G(x) = x - F'(G_1(x))^{-1}[-K(G_1(x),x) + (1+g_2(1))F(G_1(x))],$$
$$\text{for all } x \in D(G).$$

Proposition (ii) implies that $x^* \in$ interior $(D(G))$, so Lemma 4.3 applies for $D_3 = D(G)$ and $G_3 = G$, therefore $G'(x^*)$ exists and

$$G'(x^*) = -(1+g(1))[I - K_1(x^*,x^*)^{-1}F'(x^*)].$$

Now, there is a neighbourhood V of x^* such that

$$\|G(x) - x^*\| \leqq \|x - x^*\| \text{ for all } x \in V.$$

Let $t_1 > 0$ be such that for all $t \in [0,t_1]$, $x^* + ty \in V$. Then for all $t \in [0,t_1]$:

$$\|G(x^* + ty) - G(x^*) - G'(x^*)ty\| \geq$$

$$\geq (1+g(1))L\|ty\| - \|ty\| =$$

$$= p\|ty\|.$$

This yields a contradiction, so (ii) implies (iv). □

As a direct consequence of Theorem 4.1 we have

<u>THEOREM 4.2</u>. *Let* $F \in \mathcal{F}_1$ *and* $K \in \mathcal{K}_1$ *then the following propositions* (i) *and* (ii) *are equivalent:*

(i) $K_1(x^*,x^*;F) = F'(x^*)$.

(ii) For all $g:[0,1] \to \mathbb{R}$ and any Runge-Kutta method Λ, the iterative process $(\{M\},F)$, where $M(\cdot) = M(K,g,\Lambda;\cdot)$ is locally quadratically convergent.

<u>PROOF</u>. The result immediately follows from the equivalence of propositions (iii) and (iv) in Theorem 4.1. □

Examples of $K \in \mathcal{K}_1$ that satisfy the condition $K_1(x^*,x^*;F) = F'(x^*)$ for $F \in \mathcal{F}_1$ are

1. $K(x,y;F) = F(x) - F(y)$,
2. $K(x,y;F) = F'(y)(x-y)$.

5. RADIUS OF CONVERGENCE OF NEWTON'S METHOD

Let $K \in \mathcal{K}_1$ be defined by $K(x,y;\, F) = F(x) - F(y)$, where $x, y \in X$ and $F \in \mathcal{F}_1$; let $g\colon [0,1] \to \mathbb{R}$, where $g(t) = 0$ for all $t \in [0,1]$; let Λ be a Runge-Kutta method, where

$$\Lambda = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Consider the iterative method $\{M\}$, where $M(\cdot) = M(K,g,\Lambda;\cdot)$. For $F \in \mathcal{F}_1$, the iterating function $G = M(F)$ is defined by

$$(5.1)\qquad \begin{aligned} &G\colon D(G) \to X, \\ &G(x) = x - F'(x)^{-1}F(x), \text{ for all } x \in D(G). \end{aligned}$$

Therefore $\{M\}$ is Newton's method.

Now, let $\beta, \gamma > 0$. In order to calculate the radius of convergence of Newton's method, we first have to prove some lemma's.

LEMMA 5.1. *If* $F \in \mathcal{F}\langle\beta,\gamma\rangle$ *and* $x \in B(x^*,\frac{1}{\beta\gamma})$ *then* $F'(x)$ *is invertible and*

$$\|F'(x)^{-1}\| \leqq \frac{\beta}{1-\beta\gamma\|x-x^*\|}.$$

PROOF. Since $F \in \mathcal{F}\langle\beta,\gamma\rangle$, we have $\|F'(x^*)^{-1}\| \leqq \beta$ and for all $x \in B(x^*,\frac{1}{\beta\gamma})$: $\|F'(x)-F'(x^*)\| \leqq \gamma\|x-x^*\| < \frac{1}{\beta}$. Therefore, Theorem 2.6.1 applies, thus proving this lemma. □

LEMMA 5.2. *For* $F \in \mathcal{F}\langle\beta,\gamma\rangle$ *let* G *be defined by (5.1), then* $B(x^*,\frac{1}{\beta\gamma}) \subset D(G)$, *and for all* $x \in B(x^*,\frac{1}{\beta\gamma})$,

$$\|G(x)-x^*\| \leqq \frac{\beta\gamma\|x-x^*\|^2}{2(1-\beta\gamma\|x-x^*\|)}.$$

PROOF. According to (3.2.7), $D(G) = \{x \mid F'(x) \text{ is invertible}\}$. So from Lemma 5.1. it follows that $B(x^*,\frac{1}{\beta\gamma}) \subset D(G)$. Since $F \in \mathcal{F}\langle\beta,\gamma\rangle$, from Theorem 2.6.3 it follows that for $x \in X$:

$$0 = F(x^*) = F(x) + F'(x)(x^*-x) + r(x), \text{ where } \|r(x)\| \leqq \frac{\gamma}{2}\|x - x^*\|^2.$$

Thus for $x \in B(x^*, \frac{1}{\beta\gamma})$

$$\|x - F'(x)^{-1}F(x) - x^*\| = \|F'(x)^{-1}r(x)\|.$$

Now, using Lemma 5.1, we obtain $\|F'(x)^{-1}r(x)\| \leqq \dfrac{\beta\gamma\|x-x^*\|^2}{2(1-\beta\gamma\|x-x^*\|)}$.

This completes the proof. □

<u>THEOREM 5.1</u>. *The radius of convergence of Newton's method with respect to* $\mathcal{F}\langle\beta,\gamma\rangle$ *is* $\frac{2}{3\beta\gamma}$.

<u>PROOF</u>. <u>1</u>. We first prove that the radius of convergence of Newton's method with respect to $\mathcal{F}\langle\beta,\gamma\rangle$ is not less than $\frac{2}{3\beta\gamma}$.

Take an arbitrary $F \in \mathcal{F}\langle\beta,\gamma\rangle$, and let $G = M(K,g,\Lambda;F)$. G is thus defined by (5.1).

For any ε, $0 < \varepsilon \leqq \frac{2}{3}$, set $\alpha(\varepsilon) = \dfrac{\frac{2}{3} - \varepsilon}{\frac{2}{3} + 2\varepsilon}$.

Note that $0 \leqq \alpha(\varepsilon) < 1$. According to Lemma 5.2, it follows that for any $x \in \bar{B}(x^*, (\frac{2}{3} - \varepsilon)\frac{1}{\beta\gamma})$,

$$\|G(x) - x^*\| \leqq \frac{\beta\gamma\|x-x^*\|^2}{2(1-\beta\gamma\|x-x^*\|)} \leqq \frac{\frac{2}{3}-\varepsilon}{2(1-(\frac{2}{3}-\varepsilon))}\|x-x^*\| = \alpha(\varepsilon)\|x-x^*\|. \tag{5.2}$$

Now, for any $x_0 \in B(x^*, \frac{2}{3\beta\gamma})$, there is an $\varepsilon > 0$ such that $x_0 \in \bar{B}(x^*, (\frac{2}{3} - \varepsilon)\frac{1}{\beta\gamma})$. (5.2) shows that the sequence $\{x_k\}$, generated by x_0 and $(\{M\},F)$ satisfies

$$\|x_k - x^*\| \leqq [\alpha(\varepsilon)]^k \|x_0 - x^*\| \to 0 \ (k\to\infty).$$

So, the radius of convergence of Newton's method with respect to F is not less than $\frac{2}{3\beta\gamma}$.

Since $F \in \mathcal{F}\langle\beta,\gamma\rangle$ was arbitrary, it follows that the radius of convergence of Newton's method with respect to $\mathcal{F}\langle\beta,\gamma\rangle$ is not less than $\frac{2}{3\beta\gamma}$.

<u>2</u>. In order to prove that the radius of convergence of Newton's method with respect to $\mathcal{F}\langle\beta,\gamma\rangle$ is $\frac{2}{3\beta\gamma}$, we show that an $F \in \mathcal{F}\langle\beta,\gamma\rangle$ and an $x_0 \in X$ exist,

such that $\|x_0 - x^*\| = \frac{2}{3\beta\gamma}$ and the sequence $\{x_k\}$ generated by x_0 and $(\{M\},F)$ satisfies $x_0 = x_2 = x_4 = \dots$.

a. If $X = \mathbb{R}^1$ with innerproduct $(x,y) = x.y$, then define

$$\phi: \mathbb{R}^1 \to \mathbb{R}^1$$

$$(5.3) \qquad \phi(x) = \begin{cases} \frac{1}{2\beta^2\gamma} & , \text{ for } x > \frac{1}{\beta\gamma} \\ \frac{1}{\beta} x - \frac{\gamma}{2} x^2, & \text{for } 0 \leq x \leq \frac{1}{\beta\gamma} \\ \frac{1}{\beta} x + \frac{\gamma}{2} x^2, & \text{for } -\frac{1}{\beta\gamma} \leq x < 0 \\ \frac{1}{2\beta^2\gamma} & , \text{ for } x < -\frac{1}{\beta\gamma} . \end{cases}$$

$$(5.4) \qquad |\phi'(x)| \leq \frac{1}{\beta}, \text{ for all } x \in \mathbb{R}^1 .$$

It is easily verified that 0 is the unique solution of $\phi(x) = 0$, $\|\phi'(0)^{-1}\| = \beta$ and $\|\phi'(x) - \phi'(y)\| \leq \gamma\|x - y\|$ for all $x,y \in X$. Therefore, $\phi \in F\langle\beta,\gamma\rangle$.

Now take $x_0 = \frac{2}{3\beta\gamma}$, then the sequence $\{x_k\}$ generated by x_0 and $(\{M\},F)$ satisfies $x_1 = -\frac{2}{3\beta\gamma}$, $x_2 = \frac{2}{3\beta\gamma}$, etc.

b. If X is an infinitely dimensional Hilbert space, then a subset B of X exists such that the following three statements hold (cf. [1]):

For all $u,v \in B$ we have $(u,v) = 0$ if $u \neq v$, $(u,v) = 1$ if $u = v$.

For any $x \in X$, the set $B_x = \{u \mid u \in B, (u,x) \neq 0\}$ is countable.

Assuming that, for $x \in X$, B_x contains an infinite number of u (otherwise extend B_x with $u \in B$ for which $(u,x) = 0$), let $n \to u_n$ be an enumeration of the set B_x, then

$$x = \sum_{n=1}^{\infty} (x,u_n)u_n$$

and

$$\|x\|^2 = \sum_{n=1}^{\infty} (x,u_n)^2.$$

Now, for $x \in X$, let $B_x = \{u_n\}$ and set $\alpha_n = (x,u_n)$, $n = 1,2,\ldots$. Then

$$(5.5) \qquad x = \sum_{n=1}^{\infty} \alpha_n u_n .$$

Define $F_N(x) = \sum_{n=1}^{N} \phi(\alpha_n)u_n$, where ϕ is defined by (5.3). According to (5.4) and Theorem 2.6.2, $|\phi(\alpha_n)| \leqq \frac{1}{\beta}|\alpha_n|$. Thus, for $m > 0$, $\|F_{N+m}(x) - F_N(x)\|^2 = \sum_{n=N+1}^{N+m} [\phi(\alpha_n)]^2 \to 0$ $(N\to\infty)$. Therefore, $\{F_N(x)\}$ is a Cauchy sequence. Let $F(x) = \lim F_N(x)$. Then

$$(5.6) \qquad F(x) = \sum_{n=1}^{\infty} \phi(\alpha_n)u_n .$$

$F(x)$ is independent of the enumeration of B_x. We shall prove that $F \in \mathcal{F}\langle\beta,\gamma\rangle$.

i. $F(x) = 0$ has a solution $x^* = 0$. Let $y^* \in X$ be such that $F(y^*) = 0$ and $y^* \neq 0$. Let $B_{y^*} = \{v_n\}$ and $y^* = \sum_{n=1}^{\infty} \beta_n v_n$. Then $F(y^*) = \sum_{n=1}^{\infty} \phi(\beta_n)v_n$ and $0 = \|F(y^*)\|^2 = \sum_{n=1}^{\infty} [\phi(\beta_n)]^2$. Thus $\phi(\beta_n) = 0$, $n = 1,2,\ldots$. This implies that $\beta_n = 0$, $n = 1,2,\ldots$. Therefore, $y^* = 0$ which yields a contradiction.

ii. Let $x \in X$.
For $h \in X$, let $B_x \cup B_h = \{u_n\}$, $x = \sum_{n=1}^{\infty} \alpha_n u_n$ and $h = \sum_{n=1}^{\infty} h_n u_n$. Set $A_N h = \sum_{n=1}^{N} \phi'(\alpha_n)h_n u_n$, then for $m > 0$,

$$\|A_{N+m}h - A_N h\|^2 = \sum_{n=N+1}^{N+m} [\phi'(\alpha_n)]^2 h_n^2 \leqq \frac{1}{\beta^2} \sum_{n=N+1}^{N+m} h_n^2 \to 0 \ (N\to\infty).$$

Thus $\{A_n h\}$ is a Cauchy sequence. Let $Ah = \lim_{N\to\infty} A_N h = \sum_{n=1}^{\infty} \phi'(\alpha_n)h_n u_n$. A is independent of the enumeration of $B_x \cup B_h$.

We shall prove that $A = F'(x)$.

$$\|F(x+h) - F(x) - Ah\|^2 =$$

$$= \|\sum_{n=1}^{\infty} \phi(\alpha_n+h_n)u_n - \sum_{n=1}^{\infty} \phi(\alpha_n)u_n - \sum_{n=1}^{\infty} \phi'(\alpha_n)h_n u_n\|^2 =$$

$$= \|\sum_{n=1}^{\infty} \{\phi(\alpha_n+h_n) - \phi(\alpha_n) - \phi'(\alpha_n)h_n\}u_n\|^2$$

$$= \sum_{n=1}^{\infty} |\phi(\alpha_n+h_n) - \phi(\alpha_n) - \phi'(\alpha_n)h_n|^2$$

$$\leqq \sum_{n=1}^{\infty} [\frac{\gamma}{2}h_n^2]^2$$

$$= \frac{\gamma^2}{4} \sum_{n=1}^{\infty} h_n^4$$

$$\leqq \frac{\gamma^2}{4} \|h\|^4.$$

For $h \in X$,

$$\|Ah\|^2 = \sum_{n=1}^{\infty} [\phi'(\alpha_n)h_n]^2 \leqq \frac{1}{\beta^2} \sum_{n=1}^{\infty} h_n^2 = \frac{1}{\beta^2} \|h\|^2.$$

For $h_1,h_2 \in X$, let $B_{h_1} \cup B_{h_2} \cup B_x = \{v_n\}$. $x = \sum_{n=1}^{\infty} \beta_n v_n$ and $h_j = \sum_{n=1}^{\infty} h_{j,n}v_n$, $j = 1,2$. Then for real numbers θ_1, θ_2,

$$A(\theta_1 h_1+\theta_2 h_2) = \sum_{n=1}^{\infty} \phi'(\beta_n)(\theta_1 h_{1,n}+\theta_2 h_{2,n})v_n =$$

$$= \theta_1 \sum_{n=1}^{\infty} \phi'(\beta_n)h_{1,n}v_n + \theta_2 \sum_{n=1}^{\infty} \phi'(\beta_n)h_{2,n}v_n$$

$$= \theta_1 Ah_1 + \theta_2 Ah_2.$$

Therefore, A is a bounded linear operator in X and $A = F'(x)$.

iii. Let $x,y,h \in X$ and $B_x \cup B_y \cup B_h = \{u_n\}$. $x = \sum_{n=1}^{\infty} \alpha_n u_n$, $y = \sum_{n=1}^{\infty} \beta_n u_n$ and $h = \sum_{n=1}^{\infty} h_n u_n$.

Then

$$\|[F'(x) - F'(y)]h\|^2$$

$$= \|\sum_{n=1}^{\infty} \{\phi'(\alpha_n) - \phi'(\beta_n)\}h_n u_n\|^2$$

$$= \sum_{n=1}^{\infty} [\phi'(\alpha_n) - \phi'(\beta_n)]^2 h_n^2$$

$$\leqq \gamma^2 \sum_{n=1}^{\infty} (\alpha_n - \beta_n)^2 h_n^2$$

$$\leqq \gamma^2 \left\{\sum_{n=1}^{\infty} (\alpha_n - \beta_n)^2\right\} \left\{\sum_{n=1}^{\infty} h_n^2\right\}$$

$$= \gamma^2 \|x - y\|^2 \|h\|^2.$$

Therefore $\|F'(x) - F'(y)\| \leqq \gamma\|x - y\|$ for all $x,y \in X$.

iv. Let $h \in X$, and $B_h = \{u_n\}$, $h = \sum_{n=1}^{\infty} h_n u_n$. Then $F'(0)h = \sum_{n=1}^{\infty} \phi'(0)h_n u_n = \sum_{n=1}^{\infty} \frac{1}{\beta} h_n u_n = \frac{1}{\beta} h$. So $F'(0) = \frac{1}{\beta} I$, thus $F'(0)$ is invertible and $F'(0)^{-1} = \beta I$. This implies that $\|F'(0)^{-1}\| = \beta$.

Thus $F \in \mathcal{F}\langle\beta,\gamma\rangle$.

Now, for $u \in B$, let $x_0 = \frac{2}{3\beta\gamma} u$. Then the sequence $\{x_k\}$ generated by x_0 and $(\{M\},F)$ satisfies: $x_1 = -\frac{2}{3\beta\gamma} u$, $x_2 = \frac{2}{3\beta\gamma} u$, etc. Where X is infinitely dimensional, this completes the proof.

c. If X is finitely dimensional, then we can show by a similar method as part b of the proof that an $F \in \mathcal{F}\langle\beta,\gamma\rangle$ and an $x_0 \in X$, $\|x_0 - x^*\| = \frac{2}{3\beta\gamma}$, exist, such that the sequence $\{x_k\}$ generated by x_0 and $(\{M\},F)$ satisfies $x_0 = x_2 = x_4 = \dots$. □

6. ITERATIVE METHODS WITH A GREATER RADIUS OF CONVERGENCE

Let $\beta,\gamma > 0$.

In this final chapter we present a class of iterative methods (applicable to F_1) whose menbers all have a greater radius of convergence with respect to $F\langle\beta,\gamma\rangle$ than Newton's method.

Let m be an integer, $m \geq 2$. $\omega_1,\ldots,\omega_m$ are real numbers satisfying

(6.1)
$$\omega_i \in (0,1), \quad i = 1,\ldots,m-1;$$
$$\omega_m = 1.$$

Let $\Lambda = (\lambda_{\ell,j})$ be an $(m+1) \times (m+1)$ strictly lower triangular matrix, such that

(6.2)
$$\lambda_{\ell,1} = \omega_1, \quad \text{for } \ell = 2,\ldots,m+1;$$
$$\lambda_{\ell,j} = \omega_j(1 - \sum_{i=1}^{j-1} \lambda_{\ell,i}), \quad \text{for } j = 2,\ldots,m; \quad \ell = j+1,\ldots,m+1.$$

<u>LEMMA 6.1</u>. *For* $\Lambda = (\lambda_{\ell,j})$ *defined by* (6.1) *and* (6.2), *let* $\eta_\ell = \sum_{j=1}^{\ell-1} \lambda_{\ell,j}$, *for* $\ell = 2,\ldots,m+1$. *Then*

$$\eta_\ell \in (0,1), \quad \textit{for } \ell = 2,\ldots,m;$$

and

$$\eta_{m+1} = 1.$$

<u>PROOF</u>. We prove this lemma by mathematical induction.

If $\ell = 2$, then $\eta_2 = \lambda_{2,1} = \omega_1$. According to (6.1), $\eta_2 \in (0,1)$.

Suppose that for $j = 2,\ldots,\ell-1 < m+1$ the conclusion holds. According to (6.2),

(6.3)
$$\lambda_{\ell-1,j} = \lambda_{\ell,j}, \quad \text{for } j = 1,\ldots,\ell-2.$$

$$\eta_\ell = \sum_{j=1}^{\ell-1} \lambda_{\ell,j} = \lambda_{\ell,\ell-1} + \sum_{j=1}^{\ell-2} \lambda_{\ell-1,j}.$$

Thus using (6.2),

$$\eta_\ell = \omega_{\ell-1}(1-\sum_{j=1}^{\ell-1} \lambda_{\ell-1,j}) + \sum_{j=1}^{\ell-1} \lambda_{\ell-1,j} = \omega_{\ell-1} + (1-\omega_{\ell-1})\eta_{\ell-1}.$$

Now, if $\ell < m+1$, then $\omega_{\ell-1} \in (0,1)$, and since we assumed that $\eta_{\ell-1} \in (0,1)$, we obtain $\eta_\ell \in (0,1)$. If $\ell = m+1$ then according to (6.1), $\omega_{\ell-1} = 1$. This implies that $\eta_{m+1} = 1$.

This proves this lemma. □

Thus, $\Lambda = (\lambda_{\ell,j})$ satisfies (3.1.4) and it may therefore be considered as a generating matrix of an m-stage Runge-Kutta method.

Let $K \in \mathcal{K}_1$ be defined by

(6.4) $\qquad K(x,y;F) = F(x) - F(y)$, for all $F \in \mathcal{F}_1$ and $x,y \in X$.

Let

$$g: [0,1] \to \mathbb{R},$$

(6.5)
$$g(t) = \frac{1}{1-t}, \qquad \text{for } t \in [0,1),$$
$$g(t) = 1, \qquad \text{for } t = 1.$$

For $F \in \mathcal{F}\langle\beta,\gamma\rangle$, consider $G = M(F)$, where $M(\cdot) = M(K,g,\Lambda;\cdot)$. Then,

(6.6a)
$$G: D(G) \to X,$$
$$G(x) = x + \sum_{\ell=1}^{m} \lambda_{m+1,\ell}\, k_\ell(x) \qquad \text{for all } x \in D(G),$$

where

(6.6b)
$$k_1(x) = -F'(x)^{-1}F(x),$$
$$k_\ell(x) = -F'(x+\sum_{j=1}^{\ell-1} \lambda_{\ell,j}\, k_j(x))^{-1} \times$$

$$[F(x) + \frac{1}{1-\eta_\ell}\{(1-\eta_\ell)(F(x+\sum_{j=1}^{\ell-1}\lambda_{\ell,j}k_j(x)) - F(x)) + \eta_\ell F(x+\sum_{j=1}^{\ell-1}\lambda_{\ell,j}k_j(x))\}]$$

$$= -\frac{1}{1-\eta_\ell}F'(x+\sum_{j=1}^{\ell-1}\lambda_{\ell,j}k_j(x))^{-1}F(x+\sum_{j=1}^{\ell-1}\lambda_{\ell,j}k_j(x)), \quad \ell = 2,\ldots,m.$$

If we define for all $x \in D(G)$,

$$(6.7) \qquad \begin{aligned} y_1(x) &= x, \\ y_\ell(x) &= x + \sum_{j=1}^{\ell-1}\lambda_{\ell,j}k_j(x), \quad \ell = 2,\ldots,m+1; \end{aligned}$$

then, with $\eta_1 = 0$,

$$(6.8) \qquad k_\ell(x) = -\frac{1}{1-\eta_\ell}F'(y_\ell(x))^{-1}F(y_\ell(x)), \quad \ell = 1,2,\ldots,m.$$

LEMMA 6.2. *The following relations are true:*

$$(6.9a) \qquad G(x) = y_{m+1}(x), \quad \text{for all } x \in D(G),$$

where

$$(6.9b) \qquad \begin{aligned} y_1(x) &= x, \\ y_\ell(x) &= y_{\ell-1}(x) - \omega_{\ell-1}F'(y_{\ell-1}(x))^{-1}F(y_{\ell-1}(x)), \quad \ell = 2,\ldots,m+1. \end{aligned}$$

PROOF. We only have to prove that for all $x \in D(G)$,
$y_\ell(x) = y_{\ell-1}(x) - \omega_{\ell-1}F'(y_{\ell-1}(x))^{-1}F(y_{\ell-1}(x)), \quad \ell = 2,\ldots,m+1.$

For all $x \in D(G)$, according to (6.2), (6.6b) and (6.7),

$$y_2(x) = x + \lambda_{2,1}k_1(x) = x - \omega_1F'(y_1(x))^{-1}F(y_1(x)).$$

Thus, for $\ell = 2$, the relation to be proved is true.

Now, for $\ell = 3,\ldots,m+1$, according to (6.3) and (6.7), for all $x \in D(G)$

$$y_\ell(x) = x + \sum_{j=1}^{\ell-1}\lambda_{\ell,j}k_j(x)$$

$$= x + \sum_{j=1}^{\ell-2}\lambda_{\ell-1,j}k_j(x) + \lambda_{\ell,\ell-1}k_{\ell-1}(x)$$

$$= y_{\ell-1}(x) + \lambda_{\ell,\ell-1}k_{\ell-1}(x).$$

Using (6.2), (6.3) and (6.8),

$$\lambda_{\ell,\ell-1}k_{\ell-1}(x) = \frac{-\lambda_{\ell,\ell-1}}{1-\sum_{j=1}^{\ell-2}\lambda_{\ell-1,j}} F'(y_{\ell-1}(x))^{-1}F(y_{\ell-1}(x))$$

$$= \frac{-\lambda_{\ell,\ell-1}}{1-\sum_{j=1}^{\ell-2}\lambda_{\ell,j}} F'(y_{\ell-1}(x))^{-1}F(y_{\ell-1}(x))$$

$$= -\omega_{\ell-1}F'(y_{\ell-1}(x))^{-1}F(y_{\ell-1}(x)).$$

This proves the lemma. □

Thus, $G = M(F)$ might also be conceived as being defined by (6.9).

Iterative methods $\{M\}$, such that for all $F \in \mathcal{F}_1$, $G = M(F)$ is defined by (6.6) (or, equivalently, by (6.9)), will be investigated in this chapter. To that end we need some lemma's.

LEMMA 6.3. *Let* $x \in X$, *and suppose that*

1. $F \in \mathcal{F}\langle\beta,\gamma\rangle$.
2. *real numbers* κ *and* α *exist such that*
   a) $\alpha\kappa\gamma \leq \frac{1}{2}$.
   b) $F'(x)$ *is invertible and* $\|F'(x)^{-1}\| \leq \kappa$.
   c) $\|F'(x)^{-1}F(x)\| \leq \alpha$.

Then

$$\|v-x^*\| \leq \alpha, \textit{ where } v = x - F'(x)^{-1}F(x).$$

PROOF. The conclusion is a direct consequence of the well-known Newton-Kantorovich theorem (cf [2] and [3]). □

LEMMA 6.4. *If* $x,y,z,u \in X$, $z = \omega x + (1-\omega)y$ *for an* $\omega \in \mathbb{R}$, *then*

$$\|z-u\|^2 = \omega\|x-u\|^2 + (1-\omega)\|y-u\|^2 - \omega(1-\omega)\|x-y\|^2.$$

PROOF. 1. Observe that

$$\begin{aligned}\|z-u\|^2 &= \|\omega(x-u) + (1-\omega)(y-u)\|^2\\ &= (\omega(x-u)+(1-\omega)(y-u),\ \omega(x-u)+(1-\omega)(y-u))\\ &= \omega^2\ \|x-u\|^2 + (1-\omega)^2\ \|y-u\|^2 + 2\omega(1-\omega)(x-u,y-u).\end{aligned}$$

2. $$\begin{aligned}\|x-y\|^2 &= \|(x-u) - (y-u)\|^2\\ &= ((x-u)-(y-u),(x-u)-(y-u))\\ &\quad \|x-u\|^2 + \|y-u\|^2 - 2(x-u,y-u).\end{aligned}$$

Therefore,

$$2(x-u,y-u) = \|x-u\|^2 + \|y-u\|^2 - \|x-y\|^2.$$

Together with the first part of the proof, this proves the lemma. □

Let $\omega \in (0,1)$. Define

$$\zeta_\omega : [0,\tfrac{1}{\beta\gamma}) \to [0,\infty),$$

(6.10) $$\zeta_\omega(\sigma) = (1-\omega)\sigma^2 + \omega^2\Big[\frac{\beta\gamma\sigma^2}{2(1-\beta\gamma\sigma)}\Big]^2, \text{ for } \sigma \in [0,\tfrac{1}{2\beta\gamma}),$$

$$\zeta_\omega(\sigma) = (1-\omega)\sigma^2 + \omega\Big[\frac{\beta\gamma\sigma^2}{2(1-\beta\gamma\sigma)}\Big]^2 - \omega(1-\omega)\Big[\frac{1}{2\beta\gamma}(1-\beta\gamma\sigma)\Big]^2,$$

$$\text{for } \sigma \in \Big[\frac{1}{2\beta\gamma},\frac{1}{\beta\gamma}\Big)$$

Let the function ξ_ω be defined by

(6.11) $$\xi_\omega : [0,\tfrac{1}{\beta\gamma}) \to [0,\infty),\qquad \xi_\omega(\sigma) = \sqrt{\zeta_\omega(\sigma)}\ , \text{ for all } \sigma \in [0,\tfrac{1}{\beta\gamma}).$$

LEMMA 6.5. *Let* $F \in F\langle\beta,\gamma\rangle$. *For all* $y \in B(x^*,\frac{1}{\beta\gamma})$, *let*

$$z = y - \omega\ F'(y)^{-1}F(y).$$

Then the following error estimate holds:

$$\|z-x^*\| \le \xi_\omega(\|y-x^*\|).$$

Moreover, there is a μ_ω, $\frac{2}{3\beta\gamma} < \mu_\omega < \frac{1}{\beta\gamma}$, *such that for all* $y \in B(x^*,\mu_\omega)$,

$$\|z-x^*\| < \|y-x^*\|.$$

PROOF. Let $y \in B(x^*, \frac{1}{\beta\gamma})$. Set $\|y-x^*\| = \sigma$. It should be noted that according to Lemma 5.1, $F'(y)$ is invertible and

$$\|F'(y)^{-1}\| \leq \frac{\beta}{1-\beta\gamma\sigma}.$$

Let

$$v = y - F'(y)^{-1}F(y).$$

Then $z = \omega v + (1-\omega)y$. According to Lemma 6.4,

$$(6.12) \qquad \|z-x^*\|^2 = (1-\omega)\ \|y-x^*\|^2 + \omega\ \|v-x^*\|^2 - \omega(1-\omega)\alpha^2,$$

where $\alpha = \|F'(y)^{-1}F(y)\|$.

If

$$\frac{\alpha\beta\gamma}{1-\beta\gamma\sigma} \leqq \tfrac{1}{2}$$

then Lemma 6.3 applies and

$$\|v-x^*\| \leqq \alpha,\ \alpha \leqq \frac{1}{2\beta\gamma}(1-\beta\gamma\sigma).$$

Therefore,

$$(6.13a) \qquad \|z-x^*\|^2 \leqq (1-\omega)\sigma^2 + \omega^2\ \|v-x^*\|^2$$

and

$$(6.13b) \qquad \|z-x^*\|^2 \leqq (1-\omega)\sigma^2 + \omega^2[\frac{1}{2\beta\gamma}(1-\beta\gamma\sigma)]^2, \text{ if } \frac{\alpha\beta\gamma}{1-\beta\gamma\sigma} \leqq \tfrac{1}{2}.$$

If

$$\frac{\alpha\beta\gamma}{1-\beta\gamma\sigma} > \tfrac{1}{2}$$

then $\alpha > \frac{1}{2\beta\gamma}(1-\beta\gamma\sigma)$. According to Lemma 5.2,

$$\|v-x^*\| \leqq \frac{\beta\gamma\sigma^2}{2(1-\beta\gamma\sigma)}. \tag{6.14}$$

Thus

$$\|z-x^*\| \leqq (1-\omega)\sigma^2 + \omega\left[\frac{\beta\gamma\sigma^2}{2(1-\beta\gamma\sigma)}\right]^2 - \omega(1-\omega)\left[\frac{1}{2\beta\gamma}(1-\beta\gamma\sigma)\right]^2, \quad \text{if } \frac{\sigma\beta\gamma}{1-\beta\gamma\sigma} > \tfrac{1}{2}. \tag{6.15}$$

Note that

$$\frac{\beta\gamma\sigma^2}{2(1-\beta\gamma\sigma)} < \frac{1}{2\beta\gamma}(1-\beta\gamma\sigma), \text{ if } \sigma \in \left[0,\frac{1}{2\beta\gamma}\right), \tag{6.16a}$$

and

$$\frac{\beta\gamma\sigma^2}{2(1-\beta\gamma\sigma)} \geqq \frac{1}{2\beta\gamma}(1-\beta\gamma\sigma), \text{ if } \sigma \in \left[\frac{1}{2\beta\gamma},\frac{1}{\beta\gamma}\right). \tag{6.16b}$$

Using (6.13a), (6.14), (6.15) and (6.16a) we may conclude that

$$\|z-x^*\|^2 \leq \zeta_\omega(\sigma), \text{ if } \sigma \in [0,\frac{1}{2\beta\gamma}).$$

From (6.16b) it follows that

$$\begin{aligned} &(1-\omega)\sigma^2 + \omega^2\left[\frac{1}{2\beta\gamma}(1-\beta\gamma\sigma)\right]^2 \leq \\ &\leq (1-\omega)\sigma^2 + \omega\left[\frac{\beta\gamma\sigma^2}{2(1-\beta\gamma\sigma)}\right]^2 - \omega(1-\omega)\left[\frac{1}{2\beta\gamma}(1-\beta\gamma\sigma)\right]^2, \\ &\qquad \text{if } \sigma \in \left[\frac{1}{2\beta\gamma},\frac{1}{\beta\gamma}\right). \end{aligned} \tag{6.17}$$

Using (6.13b), (6.17) and (6.15) we may conclude that

$$\|z-x^*\|^2 \leq \zeta_\omega(\sigma), \text{ if } \sigma \in \left[\frac{1}{2\beta\gamma},\frac{1}{\beta\gamma}\right).$$

This proves the first part of the lemma.

It is easily verified that:

1. $\xi_\omega(\sigma)/\sigma$ is monotonically increasing on the interval $(0,\frac{1}{\beta\gamma})$.

2. $\lim_{\sigma\downarrow 0} \xi_\omega(\sigma)/\sigma < 1$.

3. $\lim_{\sigma\uparrow 1} \xi_\omega(\sigma)/\sigma = \infty$.

Consequently, there are uniquely defined constants μ_ω and η_ω such that

(6.18a) $$\xi_\omega(\eta_\omega) = \frac{1}{\beta\gamma}$$

(6.18b) $$\xi_\omega(\mu_\omega) = \mu_\omega, \qquad \frac{2}{3\beta\gamma} < \mu_\omega < \eta_\omega < \frac{1}{\beta\gamma}.$$

(6.18c) $$\xi_\omega(\sigma) < \sigma, \qquad \text{for all } \sigma \in (0,\mu_\omega).$$

Therefore, the conclusion holds. □

It should be noted that from Lemma 6.5 it follows that for any $\omega \in (0,1)$ there is a $\mu_\omega > \frac{2}{3\beta\gamma}$ such that for any $F \in F\langle\beta\ \gamma\rangle$ and any $y \in B(x^*,\mu_\omega)$, for

$$z = y - \omega F'(y)^{-1}F(y),$$

the inequality $\|z-x^*\| < \|y-x^*\|$ holds.

However, for the first Newton iterate, say v,

$$v = y - F'(y)^{-1}F(y)$$

the inequality $\|v-x^*\| \geqq \|y-x^*\|$ may hold.

THEOREM 6.1. *Let* $\omega_1,\ldots,\omega_m$, $(m\geqq 2)$, *be a sequence of real numbers satisfying* (6.1). *Let* $\Lambda = (\lambda_{\ell,j})$, K, g *be defined by* (6.2), (6.4) *and* (6.5) *respectively. Then the iterative method* $\{M\}$, *where* $M(\cdot) = M(K,g,\Lambda;\cdot)$ *has a greater radius of convergence with respect to* $F\langle\beta,\gamma\rangle$ *than Newton's method.*

PROOF. 1. For $j = 1,\ldots, m-1$, let μ_{ω_j} and η_{ω_j} be defined by (6.11) and (6.18) for $\omega = \omega_j$.

Define for $j = 1,\ldots,m-1$,

$$\phi_{\omega_j}: \left[0,\frac{1}{\beta\gamma}\right] \to \left[0,\frac{1}{\beta\gamma}\right],$$

$$(6.19) \qquad \phi_{\omega_j}(\sigma) = \xi_{\omega_j}(\sigma), \qquad \text{for } \sigma \in \left[0,\eta_{\omega_j}\right),$$

$$\phi_{\omega_j}(\sigma) = \frac{1}{\beta\gamma}, \qquad \text{for } \sigma \in \left[\eta_{\omega_j},\frac{1}{\beta\gamma}\right].$$

Now, consider the function ψ defined as follows

$$\psi: \left[0,\frac{1}{\beta\gamma}\right] \to \left[0,\frac{1}{\beta\gamma}\right],$$

$$(6.20a) \qquad \psi(\varepsilon) = \frac{\beta\gamma\sigma_m^2}{2(1-\beta\gamma\sigma_m)}, \qquad \text{for } \sigma_m \in \left[0,(-1+\sqrt{3})\frac{1}{\beta\gamma}\right),$$

$$\psi(\varepsilon) = \frac{1}{\beta\gamma}, \qquad \text{for } \sigma_m \in \left[(-1+\sqrt{3})\frac{1}{\beta\gamma},\frac{1}{\beta\gamma}\right),$$

where

$$\sigma_1 = \varepsilon,$$

$$(6.20b) \qquad \sigma_j = \phi_{\omega_{j-1}}(\sigma_{j-1}), \qquad \text{for } j = 2,\dots,m.$$

Using (6.18a) and (6.18b) it is easy to verify that there is a real $\hat{\rho}$ (depending on $\beta,\gamma,\omega_1,\dots,\omega_m$) such that

i. $\psi(\hat{\rho}) = \hat{\rho}$, $\quad \frac{2}{3\beta\gamma} < \hat{\rho} < \frac{1}{\beta\gamma}$,

ii. $\psi(\varepsilon) < \varepsilon$ for all $\varepsilon \in [o,\hat{\rho})$,

iii. For any $\rho \in [0,\hat{\rho})$ there exists an $\alpha \in (0,1)$ such that $\psi(\varepsilon) \leqq \alpha\varepsilon$ for all $\varepsilon \in [0,\rho]$.

2. Let $F \in \mathcal{F}\langle\beta,\gamma\rangle$ and $G = M(F)$.

a. Let $x \in X$ be such that $\|x-x^*\| \leqq \rho < \hat{\rho}$. Then according to the first part of the proof, an $\alpha \in (0,1)$ exists such that $\psi(\varepsilon) \leqq \alpha\varepsilon$ for all $\varepsilon \in [0,\rho]$. With $\varepsilon = \|x-x^*\|$, let σ_j, $j = 1,\dots,m$ be defined in (6.20b). Then, using Lemma 6.5, $\|y_j-x^*\| \leqq \sigma_j$, $j = 1,2,\dots,m$; where $y_1,y_2,\dots,y_m$ are defined as follows:

$$y_1 = x$$

$$(6.21) \qquad y_j = y_{j-1} - \omega_{j-1}F'(y_{j-1})^{-1}F(y_{j-1}), \qquad j = 2,\dots,m+1.$$

According to Lemma 6.2, $G(x) = y_{m+1}$ holds. We recall that $\omega_m = 1$. Using (6.21) and Lemma 5.2 it follows that

$$\|y_{m+1}-x^*\| \leqq \frac{\beta\gamma\sigma_m^2}{2(1-\beta\gamma\sigma_m)}$$

Hence (cf (6.20a)),

$$\|G(x)-x^*\| \leqq \psi(\|x-x^*\|) \leqq \alpha\|x-x^*\|.$$

b. Let $x_0 \in B(x^*\ \hat{\rho})$. Then, since $\|x_0-x^*\| \leqq \rho$ for some $\rho < \hat{\rho}$, from (a) it follows that $B(x^*,\rho) \subset S$, $S = S(\{M\},F)$ being the region of convergence of the iterative process $(\{M\},F)$.

Therefore, the radius of convergence of the iterative method $\{M\}$ with respect to F is not less than $\hat{\rho}$, where $\hat{\rho} > \frac{2}{3\beta\gamma}$.

Since $F \in \mathcal{F}\langle\beta,\gamma\rangle$ was arbitrary, it follows that the radius of convergence of $\{M\}$ with respect to $\mathcal{F}\langle\beta,\gamma\rangle$ is not less than $\hat{\rho}$, $\hat{\rho} > \frac{2}{3\beta\gamma}$.

Together with Theorem 5.1 this completes the proof. □

ACKNOWLEDGEMENTS. *I wish to thank Prof. M.N. Spijker for many helpful suggestions and constructive comments during the preparation of this paper. I am also grateful to J.C.P. Bus, Prof. P.J. van der Houwen and J.G. Verwer for their careful reading of this manuscript and their useful criticisms.*

REFERENCES

[1] BROWN, A.L. & A. PAGE, *Elements of Functional Analysis*, Van Nostrand Reinhold Company, London, 1970.

[2] GRAGG, W.B. & R.A. TAPIA, *Optimal error bounds for the Newton-Kantorovich theorem*, S.I.A.M. Journal on Numerical Analysis, 11, 1974 (10-13).

[3] KANTOROWITSCH, L.W. & G.P. AKILOV, *Funktionalanalysis in Normierten Raümen*, Akademie-Verlag, Berlin, 1964.

[4] LAMBERT, J.D., *Computational Methods in Ordinary Differential Equations*, John Wiley and Sons, London, 1973.

[5] MEYER, G.H., *On Solving Nonlinear Equations with a One-parameter Operator Imbedding*, S.I.A.M. Journal on Numerical Analysis 5, 1968 (739-752).