JOURNAL OF COMPLEXITY **15**, 239–281 (1999)
Article ID jcom.1999.0501, available online at http://www.idealibrary.com on IDEAL®

Finding Zeros of Analytic Functions: α Theory for Secant Type Methods

Jean-Claude Yakoubsohn

Laboratoire Approximation et Optimisation, Université Paul Sabatier, 31062 Toulouse Cedex, France
E-mail: yak@cict.fr

Received May 25, 1998

We present new results concerning the convergence of secant type methods with only conditions at a point. The radius of robustness of these methods is given, and we apply it to the study of the complexity of homotopy methods for approximating roots. In particular, we say how to use the secant type method to get an approximate zero relative to the Newton method. © 1999 Academic Press

Key Words: Regula Falsi; secant method; Newton method; approximate zero; α-theory; homotopy method; complexity.

Contents.

0. INTRODUCTION AND MAIN RESULTS

We consider here the zero-finding problem for an analytic function $f: E \to F$ between two real or complex Banach spaces. A classical algorithm to solve this problem is Newton's sequence defined by

$$x_{k+1} = x_k - Df(x_k)^{-1} f(x_k),$$

0885-064X/99 $30.00

with x_0 given. Under certain hypotheses this sequence converges quadratically towards a zero ζ of f.

To analyse these convergence properties two approaches are possible: Kantorovitch-type theorems that require estimates of the first and second derivatives of f over a certain domain containing the starting point x_0, see [5, 6], and Smale's α-theory in terms of data computed at the point x_0 alone [9].

In this paper we study Regula Falsi (RF) and secant (S) methods to compute approximately the zeros of f and we prove α-theory-like theorems for these iterations. When $f\colon \mathbb{R} \to \mathbb{R}$, the Regula Falsi sequence is defined by

$$r_0, r_1 \text{ given,} \qquad r_{k+1} = r_k - ([r_k, r_0]\, f)^{-1}\, f(r_k), \quad k \geqslant 1,$$

and the secant sequence

$$s_0, s_1 \text{ given,} \qquad s_{k+1} = s_k - ([s_k, s_{k-1}]\, f)^{-1}\, f(s_k), \quad k \geqslant 1,$$

where the starting points r_0, r_1, s_0, s_1 are given and

$$[t, r]\, f = \frac{f(t) - f(r)}{t - r}$$

is the first divided difference of f at points t and r. A study of the convergence of these sequences can be found in [7]. Geometrically, the point r_{k+1} is the intersection point of the straight line through the points $(r_k, f(r_k))$ and $(r_0, f(r_0))$ with the x-axis. The fundamental notion which defines these methods is the quantity $[t, r]\, f$. It can be seen as a linear operator from $\mathbb{R}$ into $\mathbb{R}$ verifying the functional equation $f(t) - f(r) = [t, r]\, f(t - r)$.

When the function f is a polynomial or an analytic function, we have obviously

$$[t, r]\, f = f'(r) + \sum_{k \geqslant 2} \frac{f^{(k)}(r)}{k!} (t - r)^{k-1}.$$

When E and F are finite dimensional Banach spaces (with the same dimension) one can define in the same way a secant method, known as the $n+1$-point secant Steffensen method, see [6]. When E and F are Banach spaces and f is a Lipschitz function, secant type methods can be defined if for $x, y \in E$ there exists a bounded linear operator $A(x, y)\colon E \to F$ which satisfies the functional equation

$$f(x) - f(y) = A(x, y)(x - y).$$

From an algebraic point of view, $A(x, y)$ is a Bezoutian, see [3] for an overview.

In this paper, since we deal with analytic functions, the divided difference operator will be defined by

DEFINITION OF THE DIVIDED DIFFERENCE OPERATOR. Let f be an analytic function from E into F. The divided difference operator at the points x and y lying in E is the linear operator defined from E into F by

$$[y, x] f = Df(x) + \sum_{k \geqslant 2} \frac{D^k f(x)}{k!} (y-x)^{k-1}.$$

Notice that this definition makes sense: the series is converging in a ball centered at x, and the functional equation

$$f(y) - f(x) = [y, x] f(x-y)$$

holds.

We are now able to define the RF and S sequences, similarly as in the one dimensional case. The Regula Falsi method computes the RF-sequence

$$x_0, x_1 \text{ given in } E, \qquad x_{k+1} = x_k - ([x_k, x_0] f)^{-1} f(x_k). \qquad \text{(RF)}$$

The secant method computes the S-sequence

$$x_0, x_1 \text{ given in } E, \qquad x_{k+1} = x_k - ([x_k, x_{k-1}] f)^{-1} f(x_k). \qquad \text{(S)}$$

$[y, x] f$ involves the derivatives $D^k f(x)$, $k \geqslant 1$, when $E = F = \mathbb{R}^n$ or $\mathbb{C}^n$, we are able to compute them without any knowledge of these derivatives. This aspect will be made precise in Subsection 8.3. The main interest of Regula Falsi and secant methods lies in this remark since it will allow us to compute approximate zeros of f even if the computation of the derivative $Df(x)$ is difficult or impossible. This is the case when f is given by some kind of black-box allowing us to evaluate f at a point but no additional knowledge of f is available.

A good way to study the convergence of the previous sequences is to deal with the invariants

$$\beta(f, A, x) = \|A^{-1} f(x)\|, \qquad \gamma(f, A, x) = \sup_{k \geqslant 2} \left(\frac{\|A^{-1} D^k f(x)\|}{k!} \right)^{1/(k-1)},$$

where $x \in E$ and $A\colon E \to F$ is an invertible bounded linear operator. Recall from the closed graph theorem that such an operator has a bounded inverse.

Smale's α-theory for Newton's method involves three invariants:

$$\beta(f, x)=\beta(f, Df(x), x), \quad \gamma(f, x)=\gamma(f, Df(x), x), \quad \alpha(f, x)=\beta(f, x)\,\gamma(f, x).$$

Our main results on RF and S methods are of two types, gamma theorems and alpha theorems. Gamma theorems give an estimate of the size of a disc of convergence about a zero. Alpha theorems give criteria for convergence on a point from the value of alpha at that point. For comparison's sake we state versions of α-theorems and γ-theorems for Newton's method.

N-α-THEOREM [4]. *Let $x_0 \in E$ such that*

$$\alpha=\alpha(f, x_0) \leqslant \frac{13-3\sqrt{17}}{4}.$$

Let

$$t_1=\frac{\alpha+1-\sqrt{\alpha^2-6\alpha+1}}{4\gamma},$$

with $\gamma=\gamma(f, x_0)$. Then

(1) *There exists a zero ζ of the analytic function f in the open ball $B(x_0, t_1)$ and $t_1 \leqslant (5-\sqrt{17})/4\gamma$.*

(2) *The Newton sequence*

$$x_{k+1}=x_k-Df(x_k)^{-1} f(x_k), \qquad k \geqslant 0, \tag{N}$$

is well defined, and we have for $k \geqslant 0$

$$\|x_{k+1}-x_k\| \leqslant (\tfrac{1}{2})^{2^k-1} \beta(f, x_0), \qquad \text{and} \qquad \|x_k-\zeta\| \leqslant (\tfrac{1}{2})^{2^k-1} t_1.$$

In order to avoid complicated notations we now introduce

$$\beta_i=\beta(f, [x_1, x_0] f, x_i), \qquad \gamma_i=\gamma(f, [x_1, x_0] f, x_i), \quad i=0, 1.$$

We will show in Section 3 the following RF-α-theorem.

RF-α-THEOREM 0.1. *Let x_0 and x_1 be two points given in E such that $[x_0, x_1] f$ is invertible. Let*

$$b=\frac{1-2\gamma_0(\beta_0-\beta_1)}{1-\gamma_0(\beta_0-\beta_1)}\,\beta_0, \qquad t_1=\frac{b\gamma_0+1-\sqrt{(b\gamma_0)^2-6b\gamma_0+1}}{4\gamma_0}.$$

Let us suppose that the following conditions hold:

- $\gamma_0 \|x_1 - x_0\| t_1 \leqslant \beta_0 - \beta_1 \leqslant t_1$.
- $0 \leqslant b\gamma_0 \leqslant (13 - 3\sqrt{17})/4$.

Then

(1) *There exists a zero* ζ *of the analytic function* f *in the open ball* $B(x_0, t_1)$ *and,* $t_1 \leqslant (5 - \sqrt{17})/4\gamma_0$.

(2) *The RF-sequence is well defined, and we have, for* $k \geqslant 1$,

$$\|x_{k+1} - x_k\| \leqslant (\tfrac{1}{2})^{k-1}\, b, \qquad \text{and} \qquad \|x_k - \zeta\| \leqslant (\tfrac{1}{2})^{k-1}\,(t_1 - \beta_0 + \beta_1).$$

We will see in Section 3 that this previous theorem is a corollary of a more general RF α-theorem. We also will state a RF-α-theorem involving the quantity $\gamma(f, Df(x_0), x_0)$.

In Section 4 we will give a S-α-theorem. In particular

S-α-THEOREM 0.2. *Let* x_0 *and* x_1 *be two points given in* E *such that* $Df(x_0)$ *is invertible. Let*

$$b = \frac{1 - 2\gamma(\beta_0 - \beta_1)}{1 - \gamma(\beta_0 - \beta_1)}\,\beta_0, \qquad t_1 = \frac{b\gamma + 1 - \sqrt{(b\gamma)^2 - 6b\gamma + 1}}{4\gamma}.$$

with $\gamma = \gamma(f, Df(x_0), x_0)$. *Under the assumptions*

- $0 \leqslant \beta_0 - \beta_1 \leqslant t_1$,
- $0 \leqslant b\gamma \leqslant (13 - 3\sqrt{17})/4$,

it follows:

(1) *The analytic function* f *possesses a zero* $\zeta \in B(x_0, t_1)$, *and,* $t_1 \leqslant (5 - \sqrt{17})/4\gamma$.

(2) *The S-sequence is well defined and converges towards* ζ. *Moreover, for* $k \geqslant 1$,

$$\|x_{k+1} - x_k\| \leqslant (\tfrac{1}{2})^{i_k - 1}\, b, \qquad \text{and} \qquad \|x_k - \zeta\| \leqslant (\tfrac{1}{2})^{i_k - 1}\,(t_1 - \beta_0 + \beta_1),$$

where $i_0 = i_1 = 1$ *and* $i_{k+1} = i_k + i_{k-1}$, *is the Fibonacci sequence.*

When $x_0 = x_1$, the hypotheses of RF and S α are reduced at the one of N α-theorem, i.e., $\alpha \leqslant (13 - 3\sqrt{17})/4$.

In the RF-α and S-α-theorems, the points x_0 and x_1 are supposed to be known. In Subsection 8.1 we explain a possible choice of the point x_1 when we only know x_0. When the computation of the derivative $Df(x_0)$ is easy, we will establish that the points x_0 and x_1 defined by

$$x_1 = x_0 - \lambda\, Df(x_0)^{-1} f(x_0), \qquad 0 \leqslant \lambda \leqslant 1,$$

are good points to test the assumptions of the RF-α-theorem and S-α-theorem.

The second part of this paper is devoted to state γ-theorems of the secant type methods and their application to the homotopy method. In fact, to follow numerically a homotopy path we need a result which details under which conditions we can say that a point is close to a zero. The notion of an approximate zero relative to the Newton method has been introduced by S. Smale [9]:

N Approximate Zero [9]. We say that x_0 is an N approximate zero of f with associated zero ζ, if the sequence $x_{k+1} = x_k - Df(x_k)^{-1} f(x_k)$ is well defined for all $k \leqslant 0$, with

$$\|x_k - \zeta\| \leqslant (\tfrac{1}{2})^{2^k - 1} \|x_0 - \zeta\|, \qquad k \geqslant 0.$$

In [1] we find a sufficient condition to get an approximate zero in terms of the invariant $\gamma(f, Df(\zeta), \zeta)$.

N-γ-THEOREM [1]. *Let ζ be a zero of the analytic function f. Let $x_0 \in E$ be such that $u = \gamma(f, Df(\zeta), \zeta)\, \|\zeta - x_0\| \leqslant (3 - \sqrt{7})/2$. The Newton sequence (N) is well defined, and converges towards ζ, with*

$$\|\zeta - x_k\| \leqslant \left(\frac{u}{1 - 4u + 2u^2}\right)^{2^k - 1} \|\zeta - x_0\| \leqslant \left(\frac{1}{2}\right)^{2^k - 1} \|\zeta - x_0\|, \qquad k \geqslant 0.$$

Hence x_0 is an N approximate zero of f with associated zero ζ.

In the sequel, we define a notion of approximate zero relative to RF and S methods. We next give RF and S γ-theorems using respectively the invariants $\gamma(f, [\zeta, x_0]\, f, \zeta)$, and $\gamma(f, Df(\zeta), \zeta)$ where ζ is a root of f.

We first give the definition of an approximate zero relative to the RF metthod:

DEFINITION: RF APPROXIMATE ZERO. Let x_0 be given. We say that x_0 is a RF approximate zero of f with associated zero ζ if there exists $x_1 \in E$ such that the RF-sequence, $x_{k+1} = x_k - ([x_k, x_0] f)^{-1} f(x_k)$, is well defined for all $k \geqslant 1$, and converges towards ζ, with

$$\|x_k - \zeta\| \leqslant (\tfrac{1}{2})^{k-1} \|x_1 - \zeta\|, \qquad k \geqslant 1.$$

The first RF-γ-theorem gives a sufficient condition of the RF approximate zero with the quantity $\gamma(f, [\zeta, x_0] f, \zeta)$. In a certain sense, it is natural to consider this quantity rather than $\gamma(f, Df(\zeta), \zeta)$. In fact, when the sequence $(x_k)_k$ converges towards ζ, the operator sequence $([x_k, x_0] f)_k$ converges towards the operator $[\zeta, x_0] f$.

RF-γ-THEOREM 0.3. *Let ζ be a zero of f and x_0, x_1 be given. Let us consider the quantities $u = \gamma(f, [\zeta, x_0] f, \zeta) \|\zeta - x_0\|$ and $v = \gamma(f, [\zeta, x_0] f, \zeta) \|\zeta - x_1\|$, and suppose*

$$\frac{u}{1 - u - 2v + uv} \leqslant \frac{1}{2}.$$

Then x_0 is a RF approximate zero of f with associated zero ζ.

Consequently, we can determine a neighbourhood of x_0 which satisfies the notion of RF approximate zero.

COROLLARY 0.1. *Let $u = \gamma(f, [\zeta, x_0] f, \zeta) \|x_0 - \zeta\| \leqslant (5 - \sqrt{21})/2$. Let x_1 be such that*

$$\gamma(f, [\zeta, x_0] f, \zeta) \|x_1 - x_0\| \leqslant \frac{1 - 5u + 2u^2}{2 - u}.$$

Then x_0 is a RF approximate zero of f with associated zero ζ.

There is equally a criterion of RF approximate zero using $\gamma(f, Df(\zeta), \zeta)$.

RF-γ-THEOREM 0.4. *Let ζ be a zero of f. Let x_0 and x_1 be such that $u = \gamma(f, Df(\zeta), \zeta) \|x_0 - \zeta\|$ and $v = \gamma(f, Df(\zeta), \zeta) \|x_1 - \zeta\|$ verify*

$$\frac{u}{1 - 2u - 2v + 2uv} \leqslant \frac{1}{2}.$$

Then x_0 is a RF approximate zero of f with associated zero ζ.

COROLLARY 0.2. *Let* $u=\gamma(f, Df(\zeta), \zeta)\, \|x_0-\zeta\| \leqslant (3-\sqrt{7})/2$. *Let* x_1 *be such that*

$$\gamma(f, Df(\zeta), \zeta)\, \|x_1-x_0\| \leqslant \frac{1-6u+2u^2}{2(1-u)}.$$

Then x_0 *is a RF approximate zero if* f *with associated zero* ζ.

In the same way we introduce the notion of approximate zero relative to the S method:

DEFINITION: S APPROXIMATE ZERO. Let x_0 be given in E. We say that x_0 is a S approximate zero of f with associated zero ζ if there exists x_1 such that the S-sequence $x_{k+1}=x_k-([x_k, x_{k-1}]\, f)^{-1} f(x_k)$ is well defined for all $k \geqslant 1$, with

$$\|\zeta - x_k\| \leqslant (\tfrac{1}{2})^{i_k-1}\, \|\zeta - x_1\|, \qquad k \geqslant 1,$$

where $(i_k)_k$ is the Fibonacci sequence with respect to $i_0=i_1=1$.

In Section 6, we give a sufficient condition to get an S approximate zero relatively to the S method using the quantity $\gamma(f, Df(\zeta), \zeta)$.

S-γ-THEOREM 0.5. *Let* ζ *a zero of* f *and* x_0, x_1 *be such that the quantities* $u=\gamma(f, Df(\zeta), \zeta)\, \|x_0-\zeta\|$ *and* $v=\gamma(f, Df(\zeta), \zeta)\, \|x_1-\zeta\|$ *verify*

$$\frac{u}{1-2u-2v+2uv} \leqslant \frac{1}{2} \qquad \text{and} \qquad \frac{v}{1-2u-2v+2uv} \leqslant \frac{1}{2}.$$

Then x_0 *is a S approximate zero of* f *with associated zero* ζ.

COROLLARY 0.3. *Let* $u=\gamma(f, Df(\zeta), \zeta)\, \|x_0-\zeta\| \leqslant (3-\sqrt{7})/2$. *Let* x_1 *be such that*

$$\gamma(f, Df(\zeta), \zeta)\, \|x_1-x_0\| \leqslant \frac{1-6u+2u^2}{2(2-u)}.$$

Then x_0 *is an S approximate zero of* f *with associated zero* ζ.

An application of these gamma theorems is the study of the complexity of path-following methods. Let h: $[0, 1] \times E \rightarrow F$ be a continuous function such that $h_t = h(t, \cdot)$ is an analytic function. The set $\{h_t \colon t \in [0, 1]\}$ is named homotopy. A regular associated path the curve in E is given by

$$\{\zeta_t \colon t \in [0, 1],\, h_t(\zeta_t)=0 \text{ and } Dh_t(\zeta_t)^{-1} \text{ exists}\}.$$

A homotopy path is the set $\{(h_t, \zeta_t) : t \in [0, 1]\}$. The numerical path-following consists in defining a subdivision $0 = t_0 < t_1 < \cdots < t_k = 1$, and in computing a sequence $(z_i)_i$ such that each point $z_i, 0 \leqslant i \leqslant k$, is an approximate zero of $h_i = h_{t_i}$ relative to a numerical method.

We study here two numerical path-following methods. The first method is defined by

$$z_0, z_1 \text{ given}, \qquad z_{i+1} = z_i - ([z_i, z_{i-1}]\, h_{i+1})^{-1}\, h_{i+1}(z_i), \quad 1 \leqslant i \leqslant k-1. \tag{H1}$$

Let $r \geqslant 0$ be given. The second method is defined by

$$z_0,\ y_0 \in \bar{B}(z_0, r) \text{ given},$$

$$z_{i+1} = y_i - ([y_i, z_i]\, h_{i+1})^{-1}\, h_{i+1}(y_i), \quad y_{i+1} \in \bar{B}(z_{i+1}, r), \quad 1 \leqslant i \leqslant k-1, \tag{H2}$$

where $\bar{B}(x, r)$ denotes the closed ball. When we take $y_i = z_i$ for each i, the sequence (H2) generalizes

$$z_0 \text{ given}, \qquad z_{i+1} = z_i - Dh_{i+1}(z_i)^{-1}\, h_{i+1}(z_i), \quad 1 \leqslant i \leqslant k-1. \tag{H0}$$

In the following we compute a sufficient k for z_i to be an N approximate zero of h_i with associated zero $\zeta_i := \zeta_{t_i}$, see the definition above. The N path following theorem gives an answer for the sequence (H0) using the three quantities:

- $\gamma(h) = \max_{0 \leqslant t \leqslant 1} \{\gamma(h_t, Dh_t(x_t), x_t)\}$,
- $C(h) = \max_{0 \leqslant t \leqslant 1} \|Dh_t(x_t)^{-1}\|$,
- $L(h)$ the length of the curve $t \in [0, 1] \to h_t$.

N Path Following Theorem [2]. Let $\varepsilon = (3 - \sqrt{7})/4 \sim 0.0885621722$. There exists a subdivision, $0 = t_0 < t_1 < \cdots < t_k = 1$, such that $z_0 = \xi_0$ and each z_i defined by (H0) is an N approximate zero of h_i with associated zero ζ_i. Moreover, we can take

$$k = \left\lceil \frac{C(h)\, \gamma(h)\, L(h)}{\varepsilon} \right\rceil.$$

The properties of the sequences (H1) are given by the following result.

THEOREM 0.6. *Let* $\varepsilon=(4-\sqrt{13})/6 \sim 0.0657414541$. *There exists a subdivision* $0=t_0<t_1<\cdots<t_k=1$ *such that the following conditions hold*:

— $z_0=\zeta_0$.
— $\gamma(h)\,\|\zeta_1-\zeta_0\|\leqslant\varepsilon/2$, *and* $\gamma(h)\,\|z_1-\zeta_0\|\leqslant\varepsilon/2$.
— *the* z_i*'s of the sequence* (H1) *satisfy*

$$\gamma(h)\,\|z_i-\zeta_i\|\leqslant\varepsilon, \qquad 0\leqslant i\leqslant k.$$

Moreover, we can take

$$k=\left\lceil\frac{\gamma(h)\,C(h)\,L(h)}{\varepsilon}\right\rceil+1.$$

Hence for all i, $0\leqslant i\leqslant k$, *each point* z_i *is a N approximate zero of* h_i *with associated zero* ζ_i.

The next theorem concerns the sequence (H2).

THEOREM 0.7. *Let* $\lambda\geqslant 0$ *be such that* $\varepsilon=(3-\sqrt{7})/4-\lambda/2>0$. *Let* $r_\lambda=2\lambda(\lambda+\sqrt{7})/(2\lambda+\sqrt{7}-1)$. *There exists a subdivision* $0=t_0<t_1<\cdots<t_k=1$ *such that the following conditions hold*:

— $z_0=\zeta_0$ *and* y_0 *lies in the closed ball* $\bar{B}(0,r_\lambda)$.
— *the* z_i*'s of the sequence* (H2) *satisfy*

$$\gamma(h)\,\|z_i-\zeta_i\|\leqslant\varepsilon, \qquad 0\leqslant i\leqslant k.$$

Moreover, we can take

$$k=\left\lceil\frac{\gamma(h)\,C(h)\,L(h)}{\varepsilon}\right\rceil.$$

Hence for all i, $0\leqslant i\leqslant k$, *each point* z_i *is a N approximate zero of* h_i *with associated zero* ζ_i.

These theorems provide a sufficient condition for path-following algorithms with secant method to obtain N approximate zeros. In particular Theorem 0.7 becomes N path-following with $\lambda=0$.

In Section 7 we also give result of complexity of the homotopy using the quantities $C(h)$, $L(h)$, and

$$\gamma_\varepsilon(h)=\max_{0\leqslant t\leqslant 1}\max_{z}\{\gamma(h_t,[\zeta_t,z]\,h_t,\zeta_t):\gamma(h_t,[\zeta_t,z]\,h_t,\zeta_t)\,\|z-\zeta_t\|\leqslant\varepsilon\}.$$

In conclusion, we have chosen to separate α-theory and γ-theory as in the first studies concerning α-theory (see [8]). This way is different from the one given in [1], in which the α-theorem is a consequence of the classical fixed point theorem and of the γ-theorem. The approach given here permits us to compute the best values for the universal constants α and γ relative to the secant type methods.

On the other hand, we have studied the gamma theorems of these methods, and we have used numerical path-following with secant type methods to get an approximate zero relative to the Newton method.

1. REGULA FALSI METHOD AND SECANT METHOD FOR UNIVERSAL FUNCTION

We use Kantorovitch analysis [5] to establish results of convergence of RF-sequence and S-sequences. The Kantorovitch analysis establishes that the convergence of the RF-sequence or S-sequence is conditioned by the convergence of the real sequences $(r_k)_{k\geqslant 0}$ or $(s_k)_{k\geqslant 0}$ respectively associated to the universal function $h(t)$ defined by

$$h(t)=b-2t+\frac{t}{1-gt},$$

where b and g are real positive numbers defined in Sections 3 and 4. In this section we study precisely the rates of convergence of the RF-sequence and S-sequence associated to the function $h(t)$. For this, it is more convenient to deal with the function $h^*(\tau)$ defined by

$$h^*(\tau)=gh\left(\frac{\tau}{g}\right)=a-2\tau+\frac{\tau}{1-\tau},$$

where $a=bg$. This function $h^*(\tau)$ is a convex function on the interval $[0, 1[$. Let us suppose that the inequality

$$a<3-2\sqrt{2}$$

holds in this section. Under this condition, the function $h^*(\tau)$ has two roots $\tau_1<\tau_2$ in the interval $[0, 1[$, given by

$$\tau_i=\frac{a+1\pm\sqrt{(a+1)^2-8a}}{4}, \qquad i=1, 2.$$

PROPOSITION 1.1. *Consider the following RF sequence* $(\rho_k)_{k\geqslant 0}$ *associated with the function* $h^*(\tau)$:

$$\rho_0 = 0 < \rho_1 \leqslant \tau_1, \qquad \rho_{k+1} = -\frac{h^*(0)}{[\rho_k, 0]\, h^*}, \quad k \geqslant 1.$$

Denote $q = \tau_1(1-\tau_2)/\tau_2(1-\tau_1)$ *and* $u_k = (\tau_1 - \rho_k)/(\tau_2 - \rho_k)$. *Then the sequence* $(\rho_k)_{k\geqslant 0}$ *is strictly increasing and converges towards* τ_1. *The convergence rate is given by*:

(1) $\tau_1 - \rho_{k+1} = ((\tau_2 - \tau_1)/(1 - q^k u_1))\, y_1 q^k \leqslant (\tau_1 - \rho_1)\, q^k$, $k \geqslant 0$.

(2) $\rho_{k+1} - \rho_k = ((\tau_2 - \tau_1)/(1 - q^{k-1}u_1)(1 - q^k u_1))(1-q)\, q^{k-1} u_1 \leqslant a q^{k-1}$, $k \geqslant 1$.

(3) $r_k = \rho_k / g$.

Proof. The convexity of the function $h^*(\tau)$ ensures that the sequence $(\rho_k)_{k\geqslant 0}$ is strictly increasing and converges to τ_1. Let us study the rate of convergence of this sequence. From the definition of the sequence $(\rho_k)_{k\geqslant 0}$, it follows that

$$\rho_{k+1} = \frac{a(1-\rho_k)}{1-2\rho_k}.$$

Next, using the fact τ_1 is a root of the function $h^*(\tau)$, we can write $\tau_1(2\tau_1 - 1) + a(1-\tau_1) = 0$. A direct computation gives

$$\begin{aligned}\tau_1 - \rho_{k+1} &= \frac{\tau_1(1-2\rho_k) - a(1-\rho_k)}{1-2\rho_k}\\ &= \frac{\tau_1(1-2\rho_k) - a(1-\rho_k) + \tau_1(2\tau_1 - 1) + a(1-\tau_1)}{1-2\rho_k}\\ &= \frac{(\tau_1 - \rho_k)(2\tau_1 - a)}{1-2\rho_k}.\end{aligned}$$

Similarly we also have

$$\tau_2 - \rho_{k+1} = \frac{(\tau_2 - \rho_k)(2\tau_2 - a)}{1-2\rho_k}.$$

Since $2\tau_1\tau_2 = a$, we find

$$\frac{2\tau_1 - a}{2\tau_2 - a} = \frac{a(1-\tau_2)}{\tau_2} \frac{\tau_1}{a(1-\tau_1)} = q.$$

Hence

$$\frac{\tau_1 - \rho_{k+1}}{\tau_2 - \rho_{k+1}} = q \frac{\tau_1 - \rho_k}{\tau_2 - \rho_k} = \cdots = q^k \frac{\tau_1 - \rho_1}{\tau_2 - \rho_1}.$$

The inequality $\tau_1 < \tau_2$ implies $q < 1$. Some direct computations prove

$$\tau_1 - \rho_{k+1} = \frac{\tau_2 - \tau_1}{1 - q^k u_1} u_1 q^k \leqslant \frac{\tau_2 - \tau_1}{1 - u_1} u_1 q^k = (\tau_1 - \rho_1)\, q^k.$$

Part (2) is proved similarly. Let us prove the inequality in part (2). We have $u_1 < \tau_1/\tau_2 = u_0$ and a straightforward computation shows

$$\frac{\tau_2 - \tau_1}{(1 - q^{k-1}u_1)(1 - q^k u_1)} (1-q)\, u_1 < \frac{\tau_2 - \tau_1}{(1-u_0)(1-qu_0)} (1-q)\, u_0 = a,$$

and this lemma follows. ▌

PROPOSITION 1.2. *Let us consider the secant sequence* $(\sigma_k)_{k \geqslant 0}$ *associated to the function* $h^*(\tau)$:

$$\sigma_0 = 0, \qquad 0 < \sigma_1 < \tau_1, \qquad \sigma_{k+1} = \sigma_k - \frac{h^*(\sigma_k)}{[\sigma_k, \sigma_{k-1}]\, h^*}.$$

We also define for $k \geqslant 0$, *the sequence* $u_k = (\sigma_k - \tau_1)/(\sigma_k - \tau_2)$, *and* $q = (1-\tau_2)\,\tau_1/(1-\tau_1)\,\tau_2$. *Moreover we consider the Fibonacci sequence* $i_0 = i_1 = 1$, *and* $u_{k+1} = i_k + i_{k-1}$ *for* $k \geqslant 1$.

The sequence $(\sigma_k)_{k \geqslant 0}$ *is strictly increasing and converges towards* τ_1. *The rate of convergence of this sequence is given by*

(1) $u_{k+1} = (q/u_0)\, u_k u_{k-1} \leqslant (u_0/q)\, q^{i_{k+1}}$, $k \geqslant 1$.

(2) $\tau_1 - \sigma_{k+1} = ((\tau_2 - \tau_1)/(1 - u_{k+1}))\, u_{k+1}$
$\leqslant ((\tau_2 - \tau_1)/(1 - u_1))\, u_1 q^{i_{k+1} - 1} = (\tau_1 - \sigma_1)\, q^{i_{k+1} - 1}$, $k \geqslant 0$.

(3) $\sigma_{k+1} - \sigma_k = ((\tau_2 - \tau_1)/(1 - u_k)(1 - u_{k+1}))(1 - (q/u_0)\, u_{k-1})\, u_k$
$\leqslant a q^{i_k - 1}$, $k \geqslant 1$.

(4) $s_k = \sigma_k / g$.

Proof. Since $h^*(\tau)=2(\tau-\tau_1)(\tau-\tau_2)/(1-\tau)$, we obtain from a straightforward computation

$$\sigma_{k+1}-\tau_1=\frac{(\sigma_k-\tau_1)(\sigma_{k-1}-\tau_1)(1-\tau_2)}{\sigma_k+\sigma_{k-1}-\sigma_k\sigma_{k-1}+\tau_1\tau_2-\tau_1-\tau_2}, \qquad k\geqslant 1.$$

A similar formula holds for $\sigma_{k+1}-\tau_2$. Hence

$$u_{k+1}=cu_ku_{k-1},$$

where $c=q/u_0$. Since $u_1<u_0$, it follows that

$$u_{k+1}=\frac{1}{c}(cu_1)^{i_k}(cu_0)^{i_{k-1}}\leqslant\frac{1}{c}(cu_0)^{i_{k+1}},$$

where (i_k) is the Fibbonacci sequence

$$i_0=i_1=1, \qquad i_{k+1}=i_k+i_{k-1}.$$

This proves the first part of this lemma follows easily. The second and third parts follow from straightforward computations. Let us prove the inequality of the third part.

The sequence (u_k) decreases and, for $c=q/u_0$, the function $u\to(1-cu)/(1-u)$ increases. Consequently

$$\begin{aligned}\frac{\tau_2-\tau_1}{(1-u_k)(1-u_{k+1})}(1-cu_{k-1})&\leqslant\frac{\tau_2-\tau_1}{(1-u_k)(1-u_{k+1})}(1-cu_k)\\&\leqslant\frac{\tau_2-\tau_1}{(1-u_0)(1-qu_0)}(1-q)=a.\end{aligned}$$

The proposition follows. ▮

The growth of the Fibonacci numbers is given by the following

LEMMA 1.1. *Let* $n_1=(1+\sqrt{5})/2$ *and* $n_2=(1-\sqrt{5})/2$ *be the roots of* $t^2-t-1=0$.

(1) $i_k=(1/\sqrt{5})(n_1^{k+1}-n_2^{k+1})$.

(2) *For all* $k\geqslant 1$, *we have* $(1/\sqrt{5})(n_1^{k+1}-n_2^{k+1})\geqslant n_1^{k-1}$.

Proof. The first part is well known. We observe $n_1^2 - n1 - 1 = 0$ and $\sqrt{5} = n_1 - n_2$. Consequently, we have for all $k \geqslant 1$

$$n_1^{k+1} - n_2^{k+1} - (n_1 - n_2)\, n_1^{k-1} = n_2^2(n_1^{k-1} - n_2^{k-1}) \geqslant 0. \quad \blacksquare$$

We end this section with some obvious lemmas.

LEMMA 1.2. $[t, r]\, h + 1 = 1/(1 - gt)(1 - gr) - 1$.

LEMMA 1.3. *Let us suppose* $bg < 3 - 2\sqrt{2}$. *For all* $0 < t \leqslant t_1$, *we have* $-1 = h'(0) < [t, 0]\, h < 0$. *Hence* $0 < [t, 0]\, h + 1 < 1$.

LEMMA 1.4. *Let* $t = s - h(s)/[s, r]\, h$. *Then we have*

$$h(t) = \frac{g(t - r)(t - s)}{(1 - gr)(1 - gs)(1 - gt)}.$$

Proof. We have $h(t) = h(s) + [s, t]\, h(t - s) + [t, s, r]\, h(t - s)(t - r) = [t, s, r]\, h(t - s)(t - r)$, where $[t, s, r]\, h$ is the second divided difference of h at points r, s, t. Since $h(s) + [s, t]\, h(t - s) = 0$ and $[t, s, r]\, h = g/(1 - gr)(1 - gs)(1 - gt)$, the lemma follows. $\blacksquare$

LEMMA 1.5. *We have* $h'((1 - \sqrt{2}/2)/g) = 0$. *Consequently, the inequalities* $h(t) > 0$ *and* $gt \leqslant 1 - \sqrt{2}/2$ *imply* $t < t_1$.

2. POINT ESTIMATES AND BASIC LEMMAS

This section gives some technical lemmas which are used in the following sections. We will use frequently

LEMMA 2.1. *For all* $0 \leqslant t < 1$, *and* $k \geqslant 0$, *we have* $\sum_{i \geqslant 0} \binom{k+i}{i} t^i = 1/(1 - t)^{k+1}$.

In this section A denotes a bounded linear map from E into F. We first recall von Neumann's perturbation lemma which is used in all point estimates:

LEMMA 2.2. *Let A be a bounded linear map from E into F. If* $\|I - A\| < 1$ *then A is invertible and*

$$\|A^{-1}\| \leqslant \frac{1}{1 - \|I - A\|}.$$

We now justify the existence of the operator $[y, x_0]\, f^{-1}$ for some $y \in E$.

LEMMA 2.3. *Let x_0 be given in E.*

(1) *For some $x_1 \in E$, let us suppose that $A = [x_1, x_0] f$ is invertible. Define for $i \in \{0, 1\}$, $\gamma_i = \gamma(f, [x_1, x_0] f, x_i)$, and suppose $1 - \gamma_i \|x_0 - x_1\| > 0$ and $1 - \gamma_i \|y - x_i\| > 0$ for some y in E. Then*

$$\|[x_1, x_0] f^{-1} [y, x_0] f - I\| \leqslant \frac{\gamma_i \|y - x_1\|}{(1 - \gamma_i \|x_0 - x_1\|)(1 - \gamma_i \|y - x_i\|)}.$$

In addition if

$$(1 - \gamma_i \|x_0 - x_1\|)(1 - \gamma_i \|y - x_i\|) > \gamma_i \|y - x_1\|$$

then $[y, x_0] f$ is invertible, and the point estimate

$$\|[y, x_0] f^{-1}[x_1, x_0] f\| \leqslant \frac{(1 - \gamma_i \|x_0 - x_1\|)(1 - \gamma_i \|y - x_i\|)}{(1 - \gamma_i \|x_0 - x_1\|)(1 - \gamma_i \|y - x_i\|) - \gamma_i \|y - x_1\|}$$

holds.

(2) *Let $A = Df(x_i)$ be invertible for $i = 0$ or $i = 1$. Define $\gamma = \gamma(f, Df(x_i), x_i)$. For some x, y in E, we assume $1 - \gamma \|x - x_i\| > 0$, and $1 - \gamma \|y - x_i\| > 0$. Then we have*

$$\|Df(x_i)^{-1} [y, x] f - I\| \leqslant \frac{1}{(1 - \gamma \|x - x_i\|)(1 - \gamma \|y - x_i\|)} - 1.$$

In addition if

$$2(1 - \gamma \|x - x_i\|)(1 - \gamma \|y - x_i\|) - 1 > 0,$$

the operator $[y, x] f$ is invertible with the point estimate

$$\|[y, x] f^{-1} Df(x_i)\| \leqslant \frac{(1 - \gamma \|x - x_i\|)(1 - \gamma \|y - x_i\|)}{2(1 - \gamma \|x - x_i\|)(1 - \gamma \|y - x_i\|) - 11}.$$

(3) *In particular, if $Df(x_0)$ is invertible then the operator $[y, x_0] f$ is invertible for all y such that $2\gamma \|y - x_0\| < 1$, with $\gamma = \gamma(f, Df(x_0), x_0)$. Moreover we have*

$$\|[y, x_0] f^{-1} Df(x_0)\| \leqslant \frac{1 - \gamma \|y - x_0\|}{1 - 2\gamma \|y - x_0\|}.$$

(4) *Reciprocally, if the operator* $[x_1, x_0]f$ *is invertible and if* $2\gamma_0 \|x_0 - x_1\| < 1$ *with* $\gamma_0 = \gamma(f, [x_1, x_0]f, x_0)$, *then the operator* $Df(x_0)$ *is invertible. Moreover, we have*

$$\|Df(x_0)^{-1}[x_1, x_0]f\| \leqslant \frac{1 - \gamma_0 \|x_1 - x_0\|}{1 - 2\gamma_0 \|x_1 - x_0\|}.$$

Proof. First, we write with $A = [x_1, x_0]f$

$$[y, x_0]f - A = \int_0^1 \int_0^t D^2 f(sx_0 + (t-s)x_1 + (1-t)y)(y - x_1)\,ds\,dt.$$

Next, using Taylor's formula at x_i, the quantity $A^{-1}[y, x_0]f - I$ is equal to

$$\int_0^1 \int_0^t 2 \sum_{j \geqslant 0} \binom{j+2}{2} \frac{A^{-1} D^{j+2} f(x_i)}{(j+2)!} \times (s(x_0 - x_i) + (t-s)(x_1 - x_i) + (1-t)(y - x_i))^j (y - x_1)\,ds\,dt.$$

The definition of $\gamma_i = \gamma(f, [x_1, x_0]f, x_i)$, and the assumptions imply

$$\begin{aligned}\|A^{-1}[y, x_0]f - I\| &\leqslant \int_0^1 \int_0^t 2\gamma_i \sum_{j \geqslant 0} \binom{j+2}{2} \\ &\quad \times \gamma_i^j (s\|x_0 - x_i\| + (t-s)\|x_1 - x_i\| + (1-t)\|y - x_i\|)^j \|y - x_1\|\,ds\,dt \\ &\leqslant \int_0^1 \int_0^t \frac{2\gamma_i \|y - x_1\|}{(1 - \gamma_i(s\|x_0 - x_i\| + (t-s)\|x_1 - x_i\| + (1-t)\|y - x_i\|))^3}\,ds\,dt.\end{aligned}$$

This previous integral is equal to

$$\frac{\gamma_i \|y - x_1\|}{(1 - \gamma_i \|x_0 - x_1\|)(1 - \gamma_i \|y - x_i\|)},$$

and part (1) of this lemma follows.

To prove part (2), we write

$$[y, x] f = \int_0^1 Df(tx + (1-t) y)\, dt$$
$$= Df(x_i) + \int_0^1 \sum_{j \geqslant 1} (j+1) \frac{D^{j+1} f(x_i)}{(j+1)!} \times (t(x - x_i) + (1-t)(y - x_i))^j\, dt.$$

Since $1 - \gamma \|x - x_i\| > 0$, and $1 - \gamma \|y - x_i\| > 0$ with $\gamma = \gamma(f, Df(x_i), x_i)$, it follows

$$\|Df(x_i)^{-1} [y, x] f - I\| \leqslant \int_0^1 \sum_{j \geqslant 1} (j+1) \gamma^j (t \|x - x_i\| + (1-t) \|y - x_i\|)^j\, dt$$
$$\leqslant \int_0^1 \frac{1}{(1 - \gamma(t \|x - x_i\| + (1-t) \|y - x_i\|))^2}\, dt - 1$$
$$\leqslant \frac{1}{(1 - \gamma \|x - x_i\|)(1 - \gamma \|y - x_i\|)} - 1,$$

and from Lemma 2.2, part (2) is established.

Part (3) (respectively part (4)) is a direct consequence of part (1), with $x_1 = x_0$ (respectively $y = x_0$). ∎

LEMMA 2.4. *Let A be an invertible linear bounded operator from E into F, and $\gamma_i = \gamma(f, A, x_i)$ for $i \in \{0, 1\}$. Let us consider $x_2 = x_1 - ([x_1, x_0] f)^{-1} f(x_1)$. Then we have*

$$\|A^{-1} f(x_2)\| \leqslant \frac{\gamma_i \|x_2 - x_1\| \, \|x_2 - x_0\|}{(1 - \gamma_i \|x_2 - x_i\|)(1 - \gamma_i \|x_1 - x_i\|)(1 - \gamma_i \|x_0 - x_i\|)}, \quad i = 0, 1.$$

Proof. Since $f(x_1) + ([x_1, x_0] f)(x_2 - x_1) = 0$, we can write for $i = 0, 1$:

$$f(x_2) = f(x_1) + ([x_1, x_0] f)(x_2 - x_1)$$
$$+ \int_0^1 \int_0^t D^2 f(s x_0 + (t-s) x_1 + (1-t) x_2)(x_2 - x_1)(x_2 - x_0)\, ds\, dt$$
$$= \int_0^1 \int_0^t 2 \sum_{j \geqslant 0} \binom{j+2}{2} \frac{D^{j+2} f(x_i)}{(j+2)!} \times (s(x_0 - x_i) + (t-s)(x_1 - x_i) + (1-t)(x_2 - x_i))^j (x_2 - x_1)(x_2 - x_0)\, ds\, dt.$$

Hence, we have successively

$$\frac{\|A^{-1}f(x_2)\|}{\|x_2-x_1\|\ \|x_2-x_0\|}$$

$$\leqslant \int_0^1 \int_0^t 2\gamma_i \sum_{j\geqslant 0} \binom{j+2}{2}$$

$$\times \gamma_i^j (s\ \|x_0-x_i\| + (t-s)\ \|x_1-x_i\| + (1-t)\ \|x_2-x_i\|)^j\, ds\, dt$$

$$\leqslant \int_0^1 \int_0^t \frac{2\gamma_i}{(1-\gamma_i(s\ \|x_0-x_i\| + (t-s)\ \|x_1-x_i\| + (1-t)\ \|x_2-x_i\|))^3}\, ds\, dt$$

$$\leqslant \frac{\gamma_i}{(1-\gamma_i\ \|x_0-x_i\|)(1-\gamma_i\ \|x_1-x_i\|)(1-\gamma_i\ \|x_2-x_i\|)}, \qquad i=0, 1,$$

and the lemma is proved. ∎

3. RF-α-THEOREM

We state a general RF-α-theorem. Let A be an invertible bounded linear operator from E into F. Let us consider the universal function $h(t)$ studied in Section 1 with

$$b = \frac{1-2g(\beta_0-\beta_1)}{1-g(\beta_0-\beta_1)}\, \beta_0, \qquad g = \gamma(f, A, x_0).$$

The quantities β_0, β_1, γ, and γ_0 are those defined in the Introduction. Let us denote by t_1 and t_2, the roots of $h(t)$ when $bg \leqslant 3-2\sqrt{2}$, and by (r_k) is the Regula Falsi sequence associated to $h(t)$ with $r_0=0$ and $r_1=\beta_0-\beta_1$.

RF-α-THEOREM 3.1. *Let x_0 and x_1 be two points given in E such that $Df(x_0)$ is invertible. Let A be an invertible bounded linear operator from E into F. Let us suppose that the following conditions hold*:

- $2\gamma\ \|x_1-x_0\| < 1.$
- $\forall y \in E$, $\forall r \geqslant 0$, $(\|y-x_0\| \leqslant r \leqslant t_1$ *and* $\|y-x_1\| \leqslant r-r_1) \Rightarrow \|A^{-1}[y, x_0]f - I\| \leqslant [r, 0]h+1.$ (1)
- $0 \leqslant \beta_0-\beta_1 \leqslant t_1.$
- $0 \leqslant bg < 3-2\sqrt{2}.$

Then

(1) *The function* $h(t)$ *has two positive roots,* $t_1 < t_2$, *and the sequence* $(r_k)_{k \geqslant 0}$ *converges towards* t_1, *as is stated in Proposition* 1.1.

(2) *There exists a zero* ζ *of the analytic function* f *in the open ball* $B(x_0, t_1)$.

(3) *The RF-sequence is well defined, and we have, for* $k \geqslant 1$,

$$\|x_{k+1} - x_k\| \leqslant r_{k+1} - r_k, \qquad \text{and} \qquad \|x_k - \zeta\| \leqslant t_1 - r_k.$$

Let us give the signification of the assumptions in the previous theorem. The condition $2\gamma \, \|x_1 - x_0\| < 1$ ensures that the divided difference operator $[x_1, x_0] f$ is invertible (Lemma 2.3). The condition (1) is a technical inequality in order to apply Von Neumann's perturbation lemma: in this section, we will see that the operators $A = Df(x_0)$ and $A = [x_1, x_0] f$ satisfy this condition. The inequality $bg < 3 - 2\sqrt{2}$ implies that the function $h(t)$ has two positive roots and hence Kantorovitch's analysis can be done. The inequality $0 \leqslant \beta_0 - \beta_1 \leqslant t_1$ shows that $(r_k)_{k \geqslant 0}$ is an increasing sequence and converges towards t_1. Hence, the rate of convergence of the sequence $(x_k)_{k \geqslant 0}$ is given by the one of the sequence $(r_k)_{k \geqslant 0}$ which has been studied in Proposition 1.1.

Proof. To establish $\|x_{k+1} - x_k\| \leqslant r_{k+1} - r_k$, we proceed by induction. From assumption the operator $Df(x_0)$ is invertible. From Lemma 2.3 part (3), and the inequality $1 - 2\gamma \, \|x_1 - x_0\| > 0$, the operator $[x_1, x_0] f$ is invertible. The point x_2 is well defined. The function $h(t)$ and the sequence $(r_k)_{k \geqslant 0}$ have been constructed in order to get

$$\|x_2 - x_0\| = \beta_0 = r_2 - r_0 = r_2 \qquad \text{and} \qquad \|x_2 - x_1\| = \beta_1 = r_2 - r_1.$$

From Section 1, the inequality $bg < 3 - 2\sqrt{2}$ ensures that the function $h(t)$ has two roots $t_1 < t_2$. Since $0 \leqslant r_1 = \beta_0 - \beta_1 < t_1$, the sequence $(r_k)_{k \geqslant 0}$ increases and converges towards t_1.

Let us suppose now that the x_j's exist for all j, $2 \leqslant j \leqslant k$, and satisfy

$$\|x_j - x_{j-1}\| \leqslant r_j - r_{j-1}.$$

We first prove x_{k+1} is well defined, i.e., the operator $[x_k, x_0] f$ is invertible. We have

$$\begin{aligned}
\|x_k - x_0\| &\leqslant \sum_{j=3}^{k} \|x_j - x_{j-1}\| + \|x_2 - x_0\| \\
&\leqslant \sum_{j=3}^{k} r_j - r_{j-1} + r_2 - r_0 = r_k.
\end{aligned}$$

For the same reasons, we also have

$$\|x_k - x_1\| \leqslant r_k - r_1.$$

We now give an upper bound for r_k. Since $bg < 3 - 2\sqrt{2}$, we have

$$r_k < t_1 \leqslant \frac{bg+1}{4g} \leqslant \frac{1}{g}\left(1 - \frac{\sqrt{2}}{2}\right) < \frac{1}{2g}.$$

Consequently, we have simultaneously $\|x_k - x_0\| \leqslant r_k \leqslant t_1$ and $\|x_k - x_0\| \leqslant r_k - r_1$.

From condition (1) on the operator A, we obtain the point estimate

$$\|A^{-1}[x_k, x_0]\, f - I\| \leqslant [r_k, 0]\, h + 1.$$

In other hand, Lemma 1.2 and the inequality $1 - 2gr_k > 0$ imply $0 < [r_k, 0]\, h + 1 = 1/(1 - gr_k) - 1 < 1$. From Lemma 2.2, the operator $[x_k, x_0]\, f$ is invertible with the point estimate

$$\|[x_k, x_0]\, f^{-1} A\| \leqslant \frac{1}{[r_k, 0]\, h}. \tag{2}$$

Hence x_{k+1} is well defined.

Let us now prove the inequality $\|x_{k+1} - x_k\| \leqslant r_{k+1} - r_k$. For that, we write

$$\begin{aligned} \|x_{k+1} - x_k\| &= \|[x_k, x_0]\, f^{-1} f(x_k)\| \\ &\leqslant \|[x_k, x_0]\, f^{-1} A\| \; \|A^{-1} f(x_k)\|. \end{aligned}$$

The inequality (2) gives an upper bound for $\|[x_k, x_0]\, f^{-1} A\|$. Lemma 2.4 gives an upper bound for $\|A^{-1} f(x_k)\|$. More precisely

$$\begin{aligned} \|A^{-1} f(x_k)\| &\leqslant \frac{g\, \|x_k - x_{k-1}\|\; \|x_k - x_0\|}{(1 - g\, \|x_k - x_0\|)(1 - g\, \|x_{k-1} - x_0\|)} \\ &\leqslant \frac{g(r_k - r_{k-1})(r_k - r_0)}{(1 - gr_k)(1 - gr_{k-1})}. \end{aligned}$$

Since $r_0 = 0$, and from Lemma 1.3, we obtain

$$\|A^{-1} f(x_k)\| \leqslant \frac{g(r_k - r_{k-1})\, r_k}{(1 - gr_k)(1 - gr_{k-1})} = h(r_k).$$

Finally

$$\|x_{k+1}-x_k\| \leqslant \frac{-1}{[r_k, 0]h} h(r_k) = r_{k+1}-r_k.$$

Hence the conclusions of this theorem follow easily. ∎

The two propositions below give examples of operators A which satisfy condition (1) of Theorem 3.1.

PROPOSITION 3.1. *The operator* $A = Df(x_0)$ *satisfies condition* (1) *of the RF-α-theorem.*

Proof. Here $g = \gamma = \gamma(f, Df(x_0), x_0)$. From Lemma 2.3, we have for all y such that $1-\gamma \|y-x_0\| > 0$

$$\|A^{-1}[y, x_0]f - I\| \leqslant \frac{\gamma \|y-x_0\|}{1-\gamma \|y-x_0\|}.$$

Hence for all r, such that $\|y-x_0\| \leqslant r \leqslant t_1$, it follows

$$\|A^{-1}[y, x_0]f - I\| \leqslant \frac{\gamma r}{1-\gamma r} = [r, 0]h + 1 < 1. \quad ∎$$

PROPOSITION 3.2. *If the inequality*

$$\|x_1 - x_0\| \gamma_0 t_1 \leqslant \beta_0 - \beta_1 = r_1,$$

holds, then the operator $A = [x_1, x_0]f$ *satisfies condition* (1) *of the RF-α-theorem.*

Proof. Here $g = \gamma_0 = \gamma(f, [x_1, x_0]f, x_0)$. Lemma 2.3 establishes the following inequality for $A = [x_1, x_0]f$: for all y such that $1-\gamma_0 \|y-x_0\| > 0$, we have, with $\delta = 1/(1-\gamma_0 \|x_0 - x_1\|)$,

$$\|[x_1, x_0]f^{-1}[y, x_0]f - I\| \leqslant \frac{\delta\gamma_0 \|y-x_1\|}{1-\gamma_0 \|y-x_0\|}.$$

Under the assumptions $\|y-x_0\| \leqslant r \leqslant t_1$, and $\|y-x_1\| \leqslant r - r_1$, it follows

$$\|[x_1, x_0]f^{-1}[y, x_0]f - I\| \leqslant \frac{\delta\gamma_0(r-r_1)}{1-\gamma_0 r}.$$

Let us show now $\delta\gamma_0(r-r_1)/(1-\gamma_0 r) \leqslant -1+1/(1-\gamma_0 r) = [r, 0]h+1$. In fact

$$\frac{\delta\gamma_0(r-r_1)}{1-\gamma_0 r} - \frac{\gamma_0 r}{1-\gamma_0 r} = \frac{(\delta-1)\,\gamma_0 r - \delta\gamma_0 r_1}{1-\gamma_0 r}.$$

This previous quantity is negative under the condition $r < t_1 \leqslant \delta r_1/(\delta-1) = (\beta_0-\beta_1)/\gamma_0\,\|x_1-x_0\|$, and the proposition follows. ∎

Proof of RF-α-Theorem 0.1. Let us verify the assumptions of Theorem 3.1. The assumption $1-2\gamma\,\|x_1-x_0\|$ is not necessary in this case since the operator $[x_1, x_0]\,f$ is assumed invertible. The operator $[x_1, x_0]\,f$ is invertible and the inequality $1-2\gamma_0\,\|x_0-x_1\|>0$ holds. From Proposition 3.2, the condition (1) is satisfied. The inequality $b\gamma_0 \leqslant (13-3\sqrt{17})/4$ implies in Proposition 1.1,

$$\tau_1 = \gamma_0 t_1 \leqslant \frac{13-3\sqrt{17}}{4}, \qquad q = \frac{\tau_1(1-\tau_2)}{\tau_2(1-\tau_1)} \leqslant \frac{1}{2}.$$

Hence

$$r_{k+1}-r_k = \frac{1}{\gamma_0}\,(\rho_{k+1}-\rho_k) \leqslant \left(\frac{1}{2}\right)^{k-1} b = \left(\frac{1}{2}\right)^{k-1} \frac{1-2\gamma_0(\beta_0-\beta_1)}{1-\gamma_0(\beta_0-\beta_1)}\,\beta_0,$$

and we obtain the point estimate on $\|x_{k+1}-x_k\|$. The point estimate on $\|x_k-x\|$ is obtained in a similar way. ∎

Remark 3.1. Proposition 3.1 implies a RF-α-theorem using the Smale's invariant $\gamma(f, x_0)$. The next section proves under the same assumptions that the S-sequence also converges.

4. S-α-THEOREM

In this section we consider $A = Df(x_0)$ and the universal function $h(t)$ with

$$b = \frac{1-2\gamma(\beta_0-\beta_1)}{1-\gamma(\beta_0-\beta_1)}\,\beta_0, \qquad g = \gamma.$$

Let us consider the sequence $(s_k)_{k\geqslant 0}$ with $s_0 = 0$ and $s_1 = \beta_0-\beta_1$.

S-α-THEOREM 4.1. *Let x_0 and x_1 be two points given in E such that $Df(x_0)$ is invertible. Under the assumptions*

- $2\gamma \|x_1 - x_0\| < 1$
- $b\gamma < 3 - 2\sqrt{2}$
- $0 \leqslant \beta_0 - \beta_1 \leqslant t_1$,

it follows:

(1) *The function $h(t)$ has two positive roots, $t_1 < t_2$, and the sequence $(s_k)_{k \geqslant 0}$ converges towards t_1, as is stated in Proposition* 1.2.

(2) *The analytic function f possesses a zero $\zeta \in B(x_0, t_1)$.*

(3) *The S-sequence is well defined and converges towards ζ. Moreover, for $k \geqslant 1$,*

$$\|x_{k+1} - x_k\| \leqslant s_{k+1} - s_k \qquad \textit{and} \qquad \|x_k - \zeta\| \leqslant t_1 - s_k.$$

Proof. The structure of the proof is similar to the proof of the RF-α-theorem. We proceed by induction. The point x_2 is well defined as in the proof of the RF-α-theorem with

$$\|x_2 - x_0\| = \beta_0 = s_2 - s_0 \qquad \text{and} \qquad \|x_2 - x_1\| = \beta_1 = s_2 - s_1.$$

Since $\beta_0 - \beta_1 \leqslant t_1$ the sequence $(s_k)_{k \geqslant 0}$ converges towards t_1 as described in Proposition 2.2.

Let us suppose now the x_j's exist for all j, $2 \leqslant j \leqslant k$ and satisfy

$$\|x_j - x_{j-1}\| \leqslant s_j - s_{j-1}.$$

We first prove that x_{k+1} is well defined and verifies $\|x_{k+1} - x_k\| \leqslant s_{k+1} - s_k$.

We have

$$\|x_k - x_0\| \leqslant \sum_{j=3}^{k} \|x_j - x_{j-1}\| + \|x_2 - x_0\| \leqslant \sum_{j=3}^{k} s_j - s_{j-1} + s_2 - s_0 = s_k.$$

This implies $\|x_k - x_0\| < t_1 < (b\gamma + 1)/4\gamma$. Since $b\gamma < 3 - 2\sqrt{2}$, it follows $\gamma \|x_k - x_0\| < 1 - \sqrt{2}/2$. This inequality also holds for $k - 1$. Consequently,

$$\gamma \|x_k - x_0\| < 1 - \frac{\sqrt{2}}{2} < 1 - \frac{1}{2(1 - \gamma \|x_{k-1} - x_0\|)}.$$

Hence the inequality

$$2(1-\gamma\,\|x_k-x_0\|)(1-\gamma\,\|x_{k-1}-x_0\|)-1>0$$

holds. By Lemma 2.3, part (2) applied at the points x_k, x_{k-1}, x_0, the operator $[x_k, x_{k-1}]\,f$ is invertible. Hence the point x_{k+1} is well defined. Moreover, we have the following point estimate:

$$\begin{aligned}\|[x_k, x_{k-1}]\,f^{-1}\,Df(x_0)\| &\leqslant \frac{(1-\gamma\,\|x_k-x_0\|)(1-\gamma\,\|x_{k-1}-x_0\|)}{2(1-\gamma\,\|x_k-x_0\|)(1-\gamma\,\|x_{k-1}-x_0\|)-1}\\ &\leqslant \frac{(1-\gamma s_k)(1-\gamma s_{k-1})}{2(1-\gamma s_k)(1-\gamma s_{k-1})-1}\\ &\leqslant \frac{-1}{[s_k, s_{k-1}]\,h}.\end{aligned}$$

We now prove $\|x_{k+1}-x_k\| \leqslant s_{k+1}-s_k$. We have

$$\begin{aligned}\|x_{k+1}-x_k\| &= \|[x_k, x_{k-1}]\,f^{-1}f(x_k)\|\\ &\leqslant \|[x_k, x_{k-1}]\,f^{-1}\,Df(x_0)\|\;\|Df(x_0)^{-1}\,f(x_k)\|.\end{aligned}$$

We have previously obtained an upper bound for $\|[x_k, x_{k-1}]\,f^{-1}\,Df(x_0)\|$. Lemma 2.4 gives an upper bound for $\|Df(x_0)^{-1}\,f(x_k)\|$. More precisely, for $k \geqslant 2$

$$\begin{aligned}\|Df(x_0)^{-1}\,f(x_k)\| &\leqslant \frac{\gamma\,\|x_k-x_{k-1}\|\;\|x_k-x_{k-2}\|}{(1-\gamma\,\|x_k-x_0\|)(1-\gamma\,\|x_{k-1}-x_0\|)(1-\gamma\,\|x_{k-2}-x_0\|)}\\ &\leqslant \frac{\gamma(s_k-s_{k-1})(s_k-s_{k-2})}{(1-\gamma s_k)(1-\gamma s_{k-1})(1-\gamma s_{k-2})}.\end{aligned}$$

From Lemma 1.3, this last quantity is equal to $h(s_k)$. Finally the previous point estimates imply

$$\|x_{k+1}-x_k\| \leqslant \frac{-h(s_k)}{[s_k, s_{k-1}]\,h} = s_{k+1}-s_k.$$

Hence the theorem follows easily. ∎

Proof of S-α-Theorem 0.2. It is a corollary of Theorem 4.1. Under the assumption $b\gamma \leqslant (13-3\sqrt{17})/4$, the quantity q of Proposition 1.2 is bounded by 1/2. Moreover, always by Proposition 1.2, we have

$$s_{k+1}-s_k \leqslant \frac{1}{\gamma}(\sigma_{k+1}-\sigma_k) \leqslant \left(\frac{1}{2}\right)^{i_k-1} b.$$

We end this proof by straightforward computation. ∎

5. RF-γ-THEOREM

In this section we prove a general RF-γ-theorem. Here ζ is a zero of f. The goal is in estimating the radius of a ball centered in ζ containing RF approximate zeros with respectively the quantities $\gamma(f, [\zeta, x_0] f, \zeta)$ and $\gamma(f, Df(\zeta), \zeta)$.

We first state a RF-γ-theorem relative to the quantity $\gamma(f, [\zeta, x_0] f, \zeta)$.

THEOREM 5.1. *Let x_0, x_1 be given in E and, ζ be a zero of the analytic function f. Let us consider the quantities $u=\gamma(f, [\zeta, x_0] f, \zeta)\, \|\zeta-x_0\|$ and $v=\gamma(f, [\zeta, x_0] f, \zeta)\, \|\zeta-x_1\|$ such that*

$$R(u, v)=\frac{u}{1-u-2v+uv}<1.$$

Then the RF-sequence $(x_k)_k$ is well defined and converges towards x with

$$\|\zeta-x_k\| \leqslant R(u, v)^{k-1}\, \|\zeta-x_1\|, \qquad k \geqslant 1.$$

The RF-γ-theorem relative to the quantity $\gamma(f, Df(\zeta), \zeta)$ is

THEOREM 5.2. *Let x_0, x_1 be given in E and ζ be a zero of f. Let us consider the quantities $u=\gamma(f, Df(\zeta), \zeta)\, \|\zeta-x_0\|$ and $v=\gamma(f, Df(\zeta), \zeta)$ $\|\zeta-x_1\|$ such that*

$$S(u, v)=\frac{u}{1-2u-2v+2uv}<1.$$

Then the RF-sequence $(x_k)_k$ is well defined, and converges towards ζ with

$$\|\zeta-x_k\| \leqslant S(u, v)^{k-1}\, \|\zeta-x_1\|, \qquad k \geqslant 1.$$

The proof of these RF-γ-theorems needs some preliminary results.

LEMMA 5.1. *Let L be a k-multilinear symmetric application from E^k into $\mathscr{L}(E, F)$. For all z_1 and z_2 given in E, we have*

$$\int_0^1 ((k+1)\, L(tz_1 + (1-t)\, z_2)^k - Lz_1^k)\, dt$$

$$= \int_0^1 (k+1) \sum_{i=0}^{k-1} \binom{k}{i} t^i (1-t)^{k-i}\, Lz_1^i z_2^{k-i}\, dt.$$

Proof. Since L is a k-multilinear symmetric, we have

$$(k+1)\, L(tz_1 + (1-t)\, z_2)^k - Lz_1^k)$$

$$= (k+1) \sum_{i=0}^{k-1} \binom{k}{i} t^i (1-t)^{k-i}\, Lz_1^i z_2^{k-i} + ((k+1)\, t^k - 1)\, Lz_t^k.$$

Hence the lemma follows easily. ∎

LEMMA 5.2. *Let A be an invertible bounded linear operator from E into F. Let ζ be a zero of f. For x_0 and y given $\in E$, we introduce the quantities $u = \gamma(f, A, \zeta)\, \|\zeta = x_0\|$ and $v = \gamma(f, A, \zeta)\, \|\zeta - y\|$. Let us suppose $1 - u > 0$ and $1 - v > 0$. Then we have*

$$\|A^{-1}([y, x_0]\, f(y - \zeta) - f(y))\| \leqslant \frac{u}{(1-u)(1-v)}\, \|y - \zeta\|.$$

Proof. We have simultaneously

$$[y, x_0]\, f = \int_0^1 Df(ty + (1-t)\, x_0)\, dt$$

$$= Df(\zeta) + \int_0^1 \sum_{k \geqslant 1} \frac{D^{k+1} f(\zeta)}{k!} (t(y - \zeta) + (1-t)(x_0 - \zeta))^k\, dt,$$

and

$$f(y) = \left(Df(\zeta) + \sum_{k \geqslant 1} \frac{D^{k+1} f(\zeta)}{(k+1)!} (y - \zeta)^k \right) (y - \zeta).$$

From Lemma 5.2 with $L = D^{k+1}f(\zeta)/(k+1)!$, $z_1 = y-\zeta$, and $z_2 = x_0-\zeta$, the quantity $A^{-1}([y, x_0]\, f(y-\zeta)-f(y))$ is equal to

$$\int_0^1 \left(\sum_{k\geqslant 1} (k+1) \sum_{i=0}^{k-1} \binom{k}{i} t^i(1-t)^{k-i} \right.$$
$$\left. \times \frac{A^{-1}D^{k+1}f(\zeta)}{(k+1)!} (y-\zeta)^i\, (x_0-\zeta)^{k-i} \right) (y-\zeta)\, dt.$$

We obtain successively with $\gamma(f, A, \zeta)$,

$$\begin{aligned}
&\|A^{-1}([y, x_0]\, f(y-\zeta)-f(y))\| \\
&\quad\leqslant \int_0^1 \sum_{k\geqslant 1} (k+1)\, \gamma^k \sum_{i=0}^{k-1} \binom{k}{i} \\
&\qquad \times t^i(1-t)^{k-i}\, \|y-\zeta\|^i\, \|x_0-\zeta\|^{k-i}\, dt\, \|y-\zeta\| \\
&\quad\leqslant \int_0^1 \sum_{k\geqslant 1} (k+1) \\
&\qquad \times \gamma^k((t\,\|y-\zeta\| + (1-t)\,\|x_0-\zeta\|)^k - \|y-\zeta\|^k\, t^k)\, dt\, \|y-\zeta\| \\
&\quad\leqslant \int_0^1 \left(\frac{1}{(1-tv-(1-t)u)^2} - \frac{1}{(1-tv)^2} \right) \|y-\zeta\| \\
&\quad\leqslant \frac{u}{(1-u)(1-v)}\, \|y-\zeta\|.
\end{aligned}$$

Hence the lemma is proved. ∎

We next define the following functions of two real variables:

$$\phi(u, v) = 1-u-2v+uv, \qquad R(u, v) = \frac{u}{\phi(u, v)}.$$

We state obvious properties of the function $R(u, v)$.

LEMMA 5.3. (1) *For all* $u \in [0, 1/3]$ *and* $0 \leqslant v \leqslant (1-3u)/(2-u)$, *we have* $R(u, v) \leqslant 1/2$.

(2) *In particular, for all* $u \in [0, (5-\sqrt{21}/2]$ *and* $0 \leqslant w \leqslant (1-5u+2u^2)/(2-u)$, *we have* $R(u, u+w) \leqslant 1/2$.

(3) *For all* u *and* v *such that* $\phi(u, v) > 0$, *the function* $v \to 1/\phi(u, v)$ *is a continuous strictly increasing function.*

Proof. The proof is obvious and left to the reader.

Proposition 5.1. *Let ζ be a zero of f. For x_0 and y given $\in E$, let us consider the quantities*

$$u=\gamma(f,[\zeta,x_0]f,\zeta)\,\|\zeta-x_0\|, \qquad v=\gamma(f,[\zeta,x_0]f,\zeta)\,\|\zeta-y\|.$$

Let us suppose $\phi(u,v)>0$. Then the point $z=y-([y,x_0]f)^{-1}f(y)$ is well defined, and we have

$$\|z-\zeta\|\leqslant R(u,v)\,\|y-\zeta\|.$$

Proof. Let us consider $A=[\zeta,x_0]f$. Let us show that $(1-u)(1-v)/\phi(u,v)$ is an upper bound for $\|[y,x_0]f^{-1}[\zeta,x_0]f\|$. The inequality $\phi(u,v)>0$ is equivalent to $(1-u)(1-v)>0$. From Lemma 2.3, part (1), the operator $[y,x_0]f$ is invertible, and the point z is well defined. Moreover

$$\|[y,x_0]f^{-1}[\zeta,x_0]f\|\leqslant\frac{1}{1-v/(1-u)(1-v)}=\frac{(1-u)(1-v)}{\phi(u,v)}.$$

We now write

$$\begin{aligned}\|z-\zeta\|&=\|y-\zeta-([y,x_0]f)^{-1}f(y)\|\\&\leqslant\|[y,x_0]f^{-1}[\zeta,x_0]f\|\,\|([\zeta,x_0]f)^{-1}([y,x_0]f(y-\zeta)-f(y))\|.\end{aligned}$$

Using Lemma 5.2 which gives an upper bound for $\|([\zeta,x_0]f)^{-1}([y,x_0]f(y-\zeta)-f(y))\|$, the result follows easily. ∎

Proof of the RF-γ-Theorem 5.1. We proceed by induction. For $k=1$, it is obvious. Let us suppose x_k is well defined and $\|x_k-\zeta\|\leqslant R(u,v)^{k-1}\,\|x_1-\zeta\|$. Consequently $\|x_k-\zeta\|\leqslant\|x_1-\zeta\|$. Hence

$$\phi(u,\gamma(f,[x_0,\zeta]f,\zeta)\,\|x_k-\zeta\|)>\phi(u,v)>0.$$

From Proposition 5.1 the point $x_{k+1}=x_k-([x_k,x_0]f)^{-1}f(x_k)$ is well defined and satisfies

$$\|x_{k+1}-\zeta\|\leqslant R(u,\gamma(f,[x_0,\zeta]f,\zeta)\,\|x_k-\zeta\|)\,\|x_k-\zeta\|.$$

From Lemma 5.3, the function $v\to R(u,v)$ is strictly increasing, and we have

$$\|x_{k+1}-\zeta\|\leqslant R(u,v)\,\|x_k-\zeta\|.$$

Hence the RF-γ-theorem follows. ∎

Proof of RF-γ-Theorem 0.3. The condition $u/(1-u-2v+uv) \leqslant 1/2$ is $R(u, v) \leqslant 1/2$. From Theorem 5.1, the result follows easily. ∎

Proof of Corollary 0.1. Let $w = \gamma(f, [\zeta, x_0] f, \zeta) \|x_1 - x_0\|$. We have $v \leqslant u + w$. This implies $R(u, v) \leqslant R(u, u+w)$. From Lemma 5.3, the conditions $u \leqslant (5-\sqrt{21})/2$ and $0 \leqslant w \leqslant (1-5u+2u^2)/(2-u)$ give $R(u, u+w) \leqslant 1/2$. From Theorem 5.1 the result follows. ∎

To prove Theorem 5.2, we introduce the functions

$$\psi(u, v) = 1 - 2u - 2v + 2uv, \qquad S(u, v) = \frac{u}{\psi(u, v)}.$$

We first have the obvious lemma:

LEMMA 5.4. (1) *If* $1 - 4u - 2v + 2uv \geqslant 0$, $S(u, v) \leqslant 1/2$.

(2) *If* $1 - 2u - 4v + 2uv \geqslant 0$, $S(v, u) \leqslant 1/2$.

(3) *For all* $u \in [0, (3-\sqrt{7})/2]$ *and* $0 \leqslant w \leqslant (1 - 6u + 2u^2)/2(2-u)$, *we have* $S(u, u+w) \leqslant S(u+w, u) \leqslant 1/2$.

(4) *For all* $u > 0$ *and* $v > 0$ *such that* $\psi(u, v) > 0$, *the functions* $u \to 1/\psi(u, v)$ *and* $v \to 1/\psi(u, v)$ *are continuous strictly increasing functions.*

PROPOSITION 5.2. *Let* ζ *be a zero of* f. *For* y *and* z *given* $\in E$, *let us consider the quantities*

$$u = \gamma(f, Df(\zeta), \zeta) \|\zeta - y\|, \qquad v = \gamma(f, Df(\zeta), \zeta) \|\zeta - z\|.$$

Let us suppose $\psi(u, v) > 0$. *Then the point* $z' = z - ([z, y] f)^{-1} f(z)$ *is well defined, and we have*

$$\|z' - \zeta\| \leqslant S(u, v) \|z - \zeta\|.$$

Proof. Let us consider $A = Df(\zeta)$. From Lemma 2.3 part (2), the inequality

$$2(1 - \gamma(f, Df(\zeta), \zeta) \|\zeta - y\|)(1 - \gamma(f, Df(\zeta), \zeta) \|\zeta - z\|) - 1 = \psi(u, v) > 0$$

implies that the operator $[z, y] f$ is invertible. The point z' is well defined, and we have the point estimate $\|[z, y] f^{-1} Df(\zeta)\| \leqslant (1-u)(1-v)/\psi(u, v)$.

On the other hand, we know, from Lemma 5.2 with $A = Df(\zeta)$, that the quantity

$$\frac{u}{(1-u)(1-v)} \|z-\zeta\|$$

is an upper estimate $\|Df(\zeta)^{-1}([z, y] f(z-\zeta) - f(z)\|$. Writing

$$\begin{aligned}\|z'-\zeta\| &= \|z-\zeta-([z, y] f)^{-1} f(z)\| \\ &\leqslant \|[z, y] f^{-1} Df(\zeta)\| \, \|Df(\zeta)^{-1}([z, y] f(z-\zeta) - f(z)\|,\end{aligned}$$

this proposition follows easily. ∎

Proof of Theorem 5.2. Substituting ϕ by ψ and R by S, and using the previous proposition, this proof is similar to the one of Theorem 5.1. ∎

Proof of the RF-γ-Theorem 0.4. The condition $u/(1-2u-2v+2uv) \leqslant 1/2$ is $S(u, v) \leqslant 1/2$. From Theorem 5.2, the result follows easily. ∎

Proof of Corollary 0.2. Let $w = \gamma(f, Df(\zeta), \zeta) \|x_1 - x_0\|$. We have $v \leqslant u + w$. This implies $S(u, v) \leqslant S(u, u+w)$. From Lemma 5.4, part (3), the conditions $u \leqslant (3-\sqrt{7})/2$ and $0 \leqslant w \leqslant (1-6u+2u^2)/2(1-u)$ give $S(u, u+w) \leqslant 1/2$. From Theorem 5.2, the result follows. ∎

6. S-γ-THEOREM

In this section we deal with the quantities

$$u = \gamma(f, Df(\zeta), \zeta) \|x_0 - \zeta\|, \qquad v = \gamma(f, Df(\zeta), \zeta) \|x_1 - \zeta\|,$$

where ζ is a zero of f, and x_0, x_1 are two points in E. We also use the function $S(u, v) = u/(1-2u-2v+2uv)$. We will use the Fibonacci sequence $i_{-1} = 0$, $i_0 = 1$, $i_{k+1} = i_k + i_{k-1}$, $k \geqslant 0$.

We first have the following

THEOREM 6.1. *Let x_0 and x_1 be given in E such that $u = \gamma(f, Df(\zeta), \zeta) \|x_0 - \zeta\|$ and $v = \gamma(f, Df(\zeta), \zeta) \|x_1 - \zeta\|$ verify*

$$1 - 3u - 2v + 2uv > 0, \qquad \textit{and} \qquad 1 - 2u - 3v + 2uv > 0.$$

Then the S-sequence $x_{k+1}=x_k-([x_k,x_{k-1}]f)^{-1}f(x_k)$, $k\geqslant 1$, *is well defined and satisfies for all* $k\geqslant 2$

$$\begin{aligned}\|x_k-\zeta\| &\leqslant S(v,u)^{i_{k-1}-1}\,S(u,v)^{i_{k-2}}\,\|x_1-\zeta\|\\ &\leqslant S(u,v)^{i_{k-1}}\,S(u,v)^{i_{k-2}-1}\,\|x_0-\zeta\|.\end{aligned}$$

Proof. We proceed by induction. The inequality $1-3u-2v+2uv>0$ implies $S(u,v)<1$ and $\psi(u,v)>0$. Similarly, the inequality $1-2u-3v+2uv>0$ implies $S(v,u)<1$.

We have

$$\|x_1-\zeta\|=S(v,u)^{i_0-1}\,S(u,v)^{i_{-1}}\,\|x_1-\zeta\|=S(u,v)^{i_0}\,S(u,v)^{i_{-1}-1}\,\|x_0-\zeta\|.$$

Hence the case $k=1$ holds. Let us suppose that the inequalities hold for all i, $2\leqslant i\leqslant k$. Consequently $\|x_i-\zeta\|\leqslant\|x_1-\zeta\|$ and $\|x_i-\zeta\|\leqslant\|x_0-\zeta\|$. Hence $\psi(\gamma(f,Df(\zeta),\zeta)\,\|x_{i-1}-\zeta\|,\quad \gamma(f,Df(\zeta),\zeta)\,\|x_i-\zeta\|)\geqslant\psi(u,v)>0$. From Proposition 5.2, the point x_{k+1} is well defined and satisfies

$$\|x_{k+1}-\zeta\|\leqslant S(\gamma(f,Df(\zeta),\zeta)\,\|x_{k-1}-\zeta\|,\gamma(f,Df(\zeta),\zeta)\,\|x_k-\zeta\|)\,\|x_k-\zeta\|.$$

Using the induction assumption and the definition of Fibonacci numbers, it follows

$$\begin{aligned}\|x_{k+1}-\zeta\| &\leqslant\frac{\gamma(f,Df(\zeta),\zeta)}{\psi(\gamma(f,Df(\zeta),\zeta)\,\|x_{k-1}-\zeta\|,\gamma(f,Df(\zeta),\zeta)\,\|x_k-\zeta\|)}\\ &\quad\times\|x_{k-1}-\zeta\|\,\|x_k-\zeta\|\\ &\leqslant\frac{\gamma}{\psi(u,v)}\,S(v,u)^{i_{k-1}-1}\,S(u,v)^{i_{k-2}}\\ &\quad\times\|x_1-\zeta\|\,S(v,u)^{i_{k-2}-1}\,S(u,v)^{i_{k-3}}\,\|x_1-\zeta\|\\ &\leqslant S(v,u)^{i_k-1}\,S(u,v)^{i_{k-1}}\,\|x_1-\zeta\|.\end{aligned}$$

The inequality $\|x_{k+1}-\zeta\|\leqslant S(v,u)^{i_k}\,S(u,v)^{i_{k-1}-1}\,\|x_0-\zeta\|$ is obtained in the same way. ∎

Proof of the S-γ-Theorem 0.5. The case $k=1$ is obvious. The inequalities $1-4u-2v+uv\geqslant 0$, and $1-2u-4v+2uv\geqslant 0$ imply respectively $S(u,v)\leqslant 1/2$ and $S(v,u)<1/2$. From Theorem 6.1, it follows for all $k\geqslant 2$,

$$\begin{aligned}\|x_k-\zeta\| &\leqslant(\tfrac{1}{2})^{i_k-1}\,\|x_1-\zeta\|\\ &\leqslant(\tfrac{1}{2})^{i_k-1}\,\|x_0-\zeta\|.\quad ∎\end{aligned}$$

Proof of Corollary 0.3. Let $w = \gamma(f, Df(\zeta), \zeta)\,\|x_1 - x_0\|$. We have $v < u + w$. This implies $S(u, v) \leqslant S(u, u+w)$ and $S(v, u) \leqslant S(u+w, u)$. From Lemma 5.4, part (3), the conditions $u \leqslant (3-\sqrt{7})/2$ and $0 \leqslant w \leqslant (1-6u+2u^2)/2(2-u)$ give $S(u, u+w) \leqslant 1/2$ and $S(u+w, u) \leqslant 1/2$. From Theorem 6.1, the result follows. ∎

7. PATH-FOLLOWING

We consider the homotopy defined in the Introduction. Remember that the first path-following method computes the sequence (H1)

$$z_0, z_1 \text{ given}, \qquad z_{i+1} = z_i - ([z_i, z_{i-1}]\,h_{i+1})^{-1}\,h_{i+1}(z_i), \quad 1 \leqslant i \leqslant k-1.$$

THEOREM 7.1. *Let us suppose that there exists* $g(h) > 0$ *and* $\varepsilon > 0$ *such that for all* $t \in [0, 1]$ *with* $h_t(\zeta_t) = 0$, *we have*

$$\forall y, z,\ g(h)\,\|y - \zeta_t\| \leqslant 3\varepsilon, \quad \text{and} \quad g(h)\,\|z - \zeta_t\| \leqslant 2\varepsilon \Rightarrow g(h)\,\|z' - \zeta_t\| \leqslant \varepsilon,$$

where $z' = z - ([z, y]\,h_t)^{-1}\,h_t(z)$.

There exists a subdivision $0 = t_0 < t_1 < \cdots < t_k = 1$ *such that the following conditions hold with* $\zeta_i := \zeta_{t_i}$:

— $z_0 = \zeta_0$.

— $g(h)\,\|\zeta_1 - \zeta_0\| \leqslant \varepsilon/2$, *and* $g(h)\,\|z_1 - \zeta_0\| \leqslant \varepsilon/2$.

— *the* z_i*'s of the sequence* (H1) *satisfy*

$$g(h)\,\|z_i - \zeta_i\| \leqslant \varepsilon, \qquad 0 \leqslant i \leqslant k.$$

Moreover, we can take

$$k = \left\lceil \frac{g(h)\,C(h)\,L(h)}{\varepsilon} \right\rceil + 1.$$

We need the following lemma to prove this theorem:

LEMMA 7.1. *Let a subdivision* $0 = t_0 < t_1 < \cdots < t_k = 1$ *be given. Then*

$$\sum_{i=0}^{k-1} \|\zeta_{i+1} - \zeta_i\| \leqslant C(h)\,L(h).$$

Proof. The identity $h_t(\zeta_t)=0$ implies $\dot{\zeta}_t = -Dh_t(\zeta_t)^{-1}\, \dot{h}_t(\zeta_t)$ where $\dot{\zeta}_t$ is the derivative by respect to t. Consequently

$$\begin{aligned}\sum_{i=0}^{k-1} \|\zeta_{i+1}-\zeta_i\| &\leqslant \sum_{i=0}^{k-1} \int_{t_i}^{t_{i+1}} \|\dot{\zeta}_t\|\, dt = \int_0^1 \|\dot{\zeta}_t\|\, dt \\ &\leqslant \int_0^1 \|Dh_t(\zeta_t)^{-1}\, \dot{h}_t(\zeta_t)\|\, dt \\ &\leqslant \int_0^1 \|Dh_t(\zeta_t)^{-1}\|\, \|\dot{h}_t(\zeta_t)\|\, dt \\ &\leqslant C(h)\, L(h). \quad \blacksquare\end{aligned}$$

Proof of Theorem 7.1. From Lemma 7.1 we have

$$\sum_{i=0}^{k-1} \|\zeta_{i+1}-\zeta_i\| \leqslant C(h)\, L(h).$$

There exists a subdivision, $0=t_0<t_1<\cdots<t_k=1$, such that the quantities $\|\zeta_{i+1}-\zeta_i\|$ are equal for $i\geqslant 1$ and $2\,\|\zeta_0-\zeta_1\| \leqslant \|\zeta_{i+1}-\zeta_i\|$. Hence

$$\|\zeta_{i+1}-\zeta_i\| \leqslant \frac{C(h)\, L(h)}{k-\frac{1}{2}} \leqslant \frac{C(h)\, L(h)}{k-1}, \qquad i\geqslant 1.$$

For the value $k=\lceil g(h)\, C(h)\, K(h)/\varepsilon \rceil + 1$, it follows for $i\geqslant 1$,

$$g(h)\, \|\zeta_{i+1}-\zeta_i\| \leqslant \varepsilon.$$

Hence $g(h)\, \|\zeta_1-\zeta_0\| \leqslant \frac{1}{2}\, \|\zeta_{i+1}-\zeta_i\| \leqslant \varepsilon/2$. Choose z_1 such that $g(h)\, \|z_1-\zeta_0\| \leqslant \varepsilon/2$. At this step, we proceed by induction to prove that the z_i's defined by (H1) satisfy

$$g(h)\, \|z_i-\zeta_i\| \leqslant \varepsilon, \qquad i\geqslant 1. \tag{3}$$

This inequality holds for $i=1$. In fact,

$$g(h)\, \|z_1-\zeta_1\| \leqslant g(h)\, \|z_1-\zeta_0\| + g(h)\, \|\zeta_0-\zeta_1\| \leqslant \varepsilon.$$

Let us suppose now the inequality (3), for $i\geqslant 1$, be given and prove it for $i+1$. We first have

$$g(h)\, \|z_i-\zeta_{i+1}\| \leqslant g(h)(\|z_i-\zeta_i\| + \|\zeta_i-\zeta_{i+1}\|) \leqslant \varepsilon+\varepsilon = 2\varepsilon.$$

On the other hand,

$$g(h)\,\|z_{i-1}-\zeta_{i+1}\| \leqslant g(h)(\|z_{i-1}-\zeta_{i-1}\| + \|\zeta_{i-1}-\zeta_{i+1}\|) \leqslant \varepsilon + 2\varepsilon = 3\varepsilon.$$

By the assumption, it follows $g(h)\,\|z_{i+1}-\zeta_{i+1}\| \leqslant \varepsilon$. The inequality (3) holds for $i+1$, and the theorem follows. ▮

Proof of Theorem 0.6. Let us verify the assumption of Theorem 7.1. Here $g(h)=\gamma(h)$. Let us consider y and z such that $\gamma(h)\,\|y-\zeta_t\| \leqslant 3\varepsilon$, and $\gamma(h)\,\|z-\zeta_t\| \leqslant 2\varepsilon$. Then $S(\gamma(h)\,\|y-\zeta_t\|,\ \gamma(h)\,\|z-\zeta_t\| \leqslant S(3\varepsilon, 2\varepsilon)$. This quantity is equal to $\frac{1}{2}$ for $\varepsilon=(4-\sqrt{13})/6$. From Proposition 5.2, the point $z'=z-([z, y]\,h_t)^{-1}\,h_t(z)$ is well defined and satisfies $\|z'-\zeta_t\| \leqslant S(3\varepsilon, 2\varepsilon)\,\|z-\zeta_t\| = \frac{1}{2}\,\|z-\zeta_t\|$. Hence the result follows Theorem 7.1. ▮

Let us recall that the second method of path-following is defined by the sequence (H2)

$$z_0,\ y_0 \in \bar{B}(z_0, r) \textit{ given},$$
$$z_{i+1} = y_i - ([y_i, z_i]\,h_{i+1})^{-1}\,h_{i+1}(y_i), \quad y_{i+1} \in \bar{B}(z_{i+1}, r),$$

where $r>0$ is given.

THEOREM 7.2. *Let us suppose that there exists* $g(h)>0$, $\varepsilon>0$, *and* $r>0$ *such that for all* $t\in[0,1]$ *with* $h_t(\zeta_t)=0$, *we have*

$$\forall z, \forall y,\ (g(h)\,\|z-\zeta_t\| \leqslant 2\varepsilon, \quad \textit{and} \quad g(h)\,\|y-z\| \leqslant r) \Rightarrow g(h)\,\|z'-\zeta_t\| \leqslant \varepsilon,$$

where $z'=y-([y, z]\,h_t)^{-1}\,h_t(z)$.

There exists a subdivision $0=t_0<t_1<\cdots<t_k=1$ *such that the* z_i*'s of the sequence* (H2) *satisfy*

$$z_0=\delta_0, \qquad g(h)\,\|z_i-\zeta_i\| \leqslant \varepsilon, \quad 0 \leqslant i \leqslant k.$$

Moreover, we can take

$$k = \left\lceil \frac{g(h)\,C(h)\,L(h)}{\varepsilon} \right\rceil.$$

Proof. The proof is similar to the proof of Theorem 7.1.

Proof of Theorem 0.7. Let us recall that $\lambda \geqslant 0$, $\varepsilon = (3-\sqrt{7})/4 - \lambda/2$ and $r_\lambda = 2\lambda(\lambda+\sqrt{7})/(2\lambda+\sqrt{7}-1)$. Here $g(h) = \gamma(h)$. Let us verify the assumptions of Theorem 7.2. For this, we consider y and z satisfying $\gamma(h)\,\|z-\zeta_t\| \leqslant 2\varepsilon$ and $\gamma(h)\,\|y-z\| \leqslant r_\lambda$. Then $S(\gamma(h)\,\|z-\zeta_t\|, \gamma(h)\,\|y-\zeta_t\|) \leqslant S(2\varepsilon, 2\varepsilon + r_\lambda)$. It is easy to see that this quantity is equal to $\frac{1}{2}$. From Proposition 5.2, the point $z' = z - ([z, y]\,h_t)^{-1}\,h_t(z)$ is well defined and satisfies $\|z'-\zeta_t\| \leqslant S(2\varepsilon, 2\varepsilon + r_\varepsilon)\,\|z-\zeta_t\| = \frac{1}{2}\,\|z-\zeta_t\|$. Hence the theorem follows. ∎

Let us define the quantity

$$\gamma_\varepsilon(h) = \max_{0 \leqslant t \leqslant 1} \max_{z} \{\gamma(h_t, [\zeta_t, z]\,h_t, \zeta_t) : \gamma(h_t, [\zeta_t, z]\,h_t, \zeta_t)\,\|z-\zeta_t\| \leqslant \varepsilon\}.$$

We state two results of complexity of path following using $\gamma_\varepsilon(h)$.

Corollary 7.1. *Let* $\varepsilon = (13-\sqrt{145})/12 \sim 0.079867118$ *and* $g(h) = \gamma_{3\varepsilon}(h)$. *Then Theorem* 7.1 *holds.*

Proof. Let us consider y and z be such that $\gamma_{3\varepsilon}(h)\,\|y-\zeta_t\| \leqslant 3\varepsilon$ and $\gamma_{3\varepsilon}(h)\,\|z-\zeta_t\| \leqslant 2\varepsilon$.

Then $R(\gamma_{3\varepsilon}(h)\,\|y-\zeta_t\|, \gamma_{3\varepsilon}(h)\,\|z-\zeta_t\|) \leqslant R(3\varepsilon, 2\varepsilon)$. We see that this quantity is equal to $\frac{1}{2}$. From Proposition 5.1, the point $z' = z - ([z, y]\,h_t)^{-1}\,h_t(z)$ is well defined and satisfies $\|z'-\zeta_t\| \leqslant R(3\varepsilon, 2\varepsilon)\,\|z-\zeta_t\| = \frac{1}{2}\,\|z-\zeta_t\|$. Hence the corollary follows. ∎

Corollary 7.2. *Let* $\varepsilon = (5-\sqrt{21})/4 - \lambda > 0$, $g(h) = \gamma_{2\varepsilon}(h)$, *and* $r_\lambda = 2\lambda(\lambda+\sqrt{21})/(2\lambda+\sqrt{21}-1)$. *Then Theorem* 7.2 *holds.*

Proof. Let us consider $\gamma_{2\varepsilon}(h)\,\|z-\zeta_t\| \leqslant 2\varepsilon$ and $y \in \bar{B}(z, r_\lambda)$. Then $R(\gamma_{2\varepsilon}(h)\,\|y-\zeta_t\|, \gamma_{2\varepsilon}(h)\,\|y-\zeta_t\|) \leqslant R(2\varepsilon, 2\varepsilon + r_\varepsilon)$. We see that this quantity is equal to $\frac{1}{2}$. From Proposition 5.1, the point $z' = z - ([z, y]\,h_t)^{-1}\,h_t(z)$ is well defined and satisfies $\|z'-\zeta_t\| \leqslant R(2\varepsilon, 2\varepsilon + r_\varepsilon)\,\|z-\zeta_t\| = \frac{1}{2}\,\|z-\zeta_t\|$. Hence the corollary follows. ∎

8. MISCELLANEOUS

8.1. *Computing a Good Point* x_1. The goal is to show why

$$x_1 = x_0 - \lambda\, Df(x_0)^{-1}\, f(x_0), \qquad 0 \leqslant \lambda \leqslant 1,$$

is a good point x_1 for the Regula Falsi and secant methods.

In all of this section we will suppose $Df(x_0)$ invertible. We recall that

$$\beta = \beta(f, Df(x_0), x_0), \qquad \gamma = \gamma(f, Df(x_0), x_0), \qquad \gamma_0 = \gamma(f, [x_1, x_0]\, f, x_0),$$
$$\alpha = \beta\gamma, \qquad \beta_i = \beta(f, [x_1, x_0]\, f, x_i), \quad i = 0,1.$$

The two following corollaries are applications of the RF-α-Theorem 3.1 and S-α-Theorem 4.1 in this choice of point x_1. The case $A = Df(x_0)$ in the RF-α-Theorem is studied in the following corollary.

COROLLARY 8.1. *Let* $\lambda \in [0, 1]$, *and suppose that the point* $x_1 = x_0 - \lambda\, Df(x_0)^{-1} f(x_0)$ *is well defined. Let* $b = ((1 - 2\gamma(\beta_0 - \beta_1))/(1 - \gamma(\beta_0 - \beta_1)))\, \beta_0$.

(1) *If* $2\lambda\alpha \leqslant 1 - \sqrt{\alpha}$, *then* $b\gamma \leqslant \alpha$.

(2) *If the inequalities*

$$b\gamma < 3 - 2\sqrt{2}, \qquad \text{and} \qquad \lambda\alpha \leqslant \frac{5 - \sqrt{13}}{6} \sim 0.2324081207$$

hold, then the points x_0 *and* x_1 *satisfy the assumptions of the RF-α-Theorem and S-α-Theorem.*

Corollary 8.2 studies the case $A = [x_1, x_0]\, f$.

COROLLARY 8.2. *Let* $\lambda \in [0, 1]$ *and the point* $x_1 = x_0 - \lambda\, Df(x_0)^{-1} f(x_0)$ *is well defined. Let* $A = [x_1, x_0]\, f$ *and* $b = ((1 - 2\gamma_0(\beta_0 - \beta_1))/(1 - \gamma_0(\beta_0 - \beta_1)))\, \beta_0$.

(1) *If* $2\lambda\beta\gamma_0 \leqslant 1 - \sqrt{\beta\gamma_0}$ *then we have*

$$b\gamma_0 \leqslant \frac{\beta\gamma_0(1 - 2\lambda\beta\gamma_0)}{(1 - \lambda\beta\gamma_0)^2} \leqslant \beta\gamma_0.$$

(2) *If the inequalities*

$$b\gamma_0 < 3 - 2\sqrt{2}, \qquad \text{and} \qquad \lambda\beta\gamma_0 \leqslant \frac{\sqrt{2}}{4 + \sqrt{2}} \sim 0.2612038749$$

hold, then the points x_0 *and* x_1 *satisfy the assumptions of the RF-α-Theorem.*

(3) *In particular if*

$$\beta\gamma_0 \leqslant \tfrac{19}{34} - \tfrac{3}{17}\sqrt{2} - \tfrac{1}{34}\sqrt{44\sqrt{2} - 43} \sim 0.1802953273$$

the points x_0 *and* $x_1 = x_0 - Df(x_0)^{-1} f(x_0)$ *satisfy the assumptions of the RF-α-Theorem.*

We first state

LEMMA 8.1. *Let us suppose* $1-2\alpha>0$. *For* $x_1=x_0-\lambda\, Df(x_0)^{-1} f(x_0)$, $0\leqslant\lambda\leqslant 1$, *we have the point estimates*

$$(1)\quad \beta_0\leqslant\frac{1-\lambda\alpha}{1-2\lambda\alpha}\beta \qquad \textit{and} \qquad \beta\leqslant\frac{1-\lambda\beta\gamma_0}{1-2\lambda\beta\gamma_0}\beta_0.$$

$$(2)\quad \beta_1\leqslant\left(1-\lambda+\frac{\lambda^2\alpha}{1-\lambda\alpha}\right)\left(\frac{1-\lambda\alpha}{1-2\lambda\alpha}\right)\beta.$$

$$(3)\quad \beta_1\leqslant(1-\lambda)\,\beta_0+\frac{\lambda^2\beta\gamma_0}{1-\lambda\beta\gamma_0}\beta.$$

$$(4)\quad \gamma_0\leqslant\frac{1-\lambda\alpha}{1-2\lambda\alpha}\gamma.$$

$$(5)\quad \beta_0-\beta_1\leqslant\lambda\beta.$$

Proof. Part (1) is a direct consequence of Lemma 2.3, part (3) and part (4) with $\|x_1-x_0\|=\lambda\beta$. Parts (2) and (3) provide the point estimate obtained with the Taylor formula

$$\|A^{-1}f(x_1)\|\leqslant(1-\lambda)\,\beta(f, A, x_0)+\frac{\lambda^2\beta\gamma(f, A, x_0)}{1-\lambda\beta\gamma(f, A, x_0)}\beta,$$

where $A=Df(x_0)$ or $A=[x_1, x_0]\, f^{-1}$. Part (4) is easy and left to the reader. Let us prove part (5). Under the condition $1-2\alpha>0$, and from Lemma 2.3, the operator $[x_1, x_0]\, f$ is invertible. The definition of $[x_1, x_0]\, f$ implies

$$\begin{aligned}\beta_0-\beta_1&=\|([x_1, x_0]\, f)^{-1} f(x_0)\|-\|([x_{11}, x_0]\, f)^{-1} f(x_1)\|\\&\leqslant\|[x_1, x_0]\, f^{-1}(fx_0)-f(x_1))\|=\|x_0-x_1\|=\lambda\beta. \quad \blacksquare\end{aligned}$$

PROPOSITION 8.1. *If* $1-3\lambda\beta\gamma_0\geqslant 0$ *or* $\lambda\alpha\leqslant(5-\sqrt{13})/6$, *then* $\beta_1\leqslant\beta_0$.

Proof. From Lemma 8.1, parts (1) and (3), we have

$$\beta_1\leqslant\left(1-\lambda+\frac{\lambda^2\beta\gamma_0}{1-2\lambda\beta\gamma_0}\right)\beta_0.$$

Hence it is sufficient to have

$$1-\lambda+\frac{\lambda^2\beta\gamma_0}{1-2\lambda\beta\gamma_0}\leqslant 1.$$

This inequality is trivially satisfied if $1-3\lambda\beta\gamma_0\geqslant 0$. Moreover, using the point estimate (4) of Lemma 8.1, it is sufficient to have

$$3\,\frac{1-\lambda\alpha}{1-2\lambda\alpha}\,\lambda\alpha\leqslant 1,$$

i.e., $1-5\lambda\alpha+3(\lambda\alpha)^2\geqslant 0$. For $\lambda\alpha\leqslant(5-\sqrt{13})/6$ this previous inequality holds. ∎

PROPOSITION 8.2. *If* $2\lambda\alpha\leqslant 1-\sqrt{\alpha}$, *the point* $x_1=x_0-\lambda\, Df(x_0)^{-1}\,f(x_0)$ *satisfies the inequality*

$$b\gamma=\beta_0\gamma\left(2-\frac{1}{1-\gamma(\beta_0-\beta_1)}\right)\leqslant(\gamma\beta_1+\lambda\alpha)\left(2-\frac{1}{1-\lambda\alpha}\right)\leqslant\alpha.$$

Proof. A straightforward computation shows that the function

$$t\to t\left(2-\frac{1}{1-t+\gamma\beta_1}\right)$$

is an increasing function for $t\in[0,1+\gamma\beta_1-\frac{1}{2}\sqrt{2(1+\gamma\beta_1)}]$. From Lemma 8.1, part (5), we have $\beta_0\leqslant\beta_1+\lambda\beta$. Let us show that the condition $2\lambda\alpha\leqslant 1-\sqrt{\alpha}$ implies $\gamma(\beta_1+\lambda\beta_1)\leqslant 1+\gamma\beta_1-\frac{1}{2}\sqrt{2(1+\gamma\beta_1)}$.

In fact, the inequality $\sqrt{2(1+\gamma\beta_1)}\leqslant 2(1-\lambda\alpha)$ is equivalent to $1+\gamma\beta_1\leqslant 2(1-\lambda\alpha)^2$. From Proposition 8.1, part (2), we have $\gamma\beta_1\leqslant(1-\lambda)((1-\lambda\alpha)/(1-2\lambda\alpha))\,\alpha+\lambda^2\alpha^2/(1-2\lambda\alpha)$. Hence it is sufficient to verify $1+(1-\lambda)((1-\lambda\alpha)/(1-2\lambda\alpha))\,\alpha+\lambda^2\alpha^2/(1-2\lambda\alpha)\leqslant 2(1-\lambda\alpha)^2$, i.e., $-(1-\lambda\alpha)(4\lambda^2\alpha^2-4\lambda\alpha+1-\alpha)\leqslant 0$. Hence, we find $2\lambda\alpha\leqslant 1-\sqrt{\alpha}$.

Then, we can write, with $\beta_0\gamma\leqslant(\beta_1+\lambda\beta)\gamma\leqslant(1-\lambda+\lambda^2\alpha/(1-\lambda\alpha))((1-\lambda\alpha)/(1-2\lambda\alpha))\,\alpha+\alpha=((1-\lambda\alpha)/(1-2\lambda\alpha))\,\alpha$,

$$\begin{aligned}\gamma\beta_0\left(2-\frac{1}{1-\gamma(\beta_0-\beta_1)}\right)&\leqslant\gamma(\beta_1+\lambda\beta)\left(2-\frac{1}{1-\lambda\alpha}\right)\\&\leqslant\frac{1-\lambda\alpha}{1-2\lambda\alpha}\,\alpha\left(2-\frac{1}{1-\lambda\alpha}\right)\\&\leqslant\alpha.\end{aligned}$$ ∎

Proof of Corollary 8.1. From Proposition 8.2, the condition $2\lambda\alpha \leqslant 1-\sqrt{\alpha}$ implies $b\gamma \leqslant \alpha$.

Let us prove now part (2), and verify the assumptions of the RF-α-Theorem 3.1 with $A = Df(x_0)$. From Proposition 3.1, the operator $Df(x_0)$ satisfies the condition (1). From Proposition 8.1, the condition $\lambda\alpha \leqslant (5-\sqrt{13})/6$ ensures both $1-2\lambda\alpha > 0$, and $\beta_1 \leqslant \beta_0$. On the other hand, we know $(\beta_0 - \beta_1)\gamma \leqslant \lambda\alpha \leqslant (5-\sqrt{13})/6 < 1-\sqrt{2}/2$. Moreover, we have

$$h(\beta_0 - \beta_1) = \frac{\beta_1(1-2(\beta_0-\beta_1)\gamma)}{1-(\beta_0-\beta_1)\gamma} > 0.$$

From Lemma 1.4, we obtain $\beta_0 - \beta_1 < t_1$.

The assumption $b\gamma_0 < 3-2\sqrt{2}$ achieves the proof of the corollary. ∎

PROPOSITION 8.3. *If* $2\lambda\beta\gamma_0 \leqslant 1-\sqrt{\beta\gamma_0}$, *the point* $x_1 = x_0 - \lambda\, Df(x_0)^{-1} f(x_0)$ *satisfies the inequality*

$$\begin{aligned} b\gamma_0 &= \beta_0\gamma_0\left(2 - \frac{1}{1-\gamma_0(\beta_0-\beta_1)}\right) \\ &\leqslant (\beta_1 + \lambda\beta)\,\gamma_0\left(2-\frac{1}{1-\lambda\beta\gamma_0}\right) \leqslant \frac{\beta\gamma_0(1-2\lambda\beta\gamma_0)}{(1-\lambda\beta\gamma_0)^2}. \end{aligned}$$

Proof. From Lemma 8.1, $\beta_1 \leqslant (1-\lambda+\lambda^2\beta\gamma_0/(1-2\lambda\beta\gamma_0))\,\beta_0 \leqslant ((1-\lambda)((1-\lambda\alpha)/(1-2\lambda\alpha)) + \lambda^2\beta\gamma_0/(1-2\lambda\beta\gamma_0))\,\beta_0$. As in the proof of Proposition 8.2, we can write under the condition $2\lambda\beta\gamma_0 \leqslant 1-\sqrt{\beta\gamma_0}$

$$\gamma_0\beta_0\left(2-\frac{1}{1-\gamma_0(\beta_0-\beta_1)}\right) \leqslant \gamma_0(\beta_1+\beta)\left(2-\frac{1}{1-\lambda\beta\gamma_0}\right).$$

Now from Lemma 8.1, part (3) and part (5), we get $\beta_1 \leqslant (1-\lambda)(\beta_1+\lambda\beta) + (\lambda^2\beta\gamma_0/(1-\lambda\beta\gamma_0)\beta$. Hence after computation

$$\beta_1 + \lambda\beta \leqslant \frac{\beta}{1-\lambda\beta\gamma_0}.$$

It follows that

$$b\gamma_0 \leqslant \frac{\beta\gamma_0}{1-\lambda\beta\gamma_0}\left(2-\frac{1}{1-\lambda\beta\gamma_0}\right) = \frac{\beta\gamma_0(1-2\lambda\beta\gamma_0)}{(1-\lambda\beta\gamma_0)^2}. \quad ∎$$

Proof of Corollary 8.2. The first part follows from Proposition 8.3. In order to prove part (2), we verify now the assumptions of the RF-α-Theorem 3.1 with $A=[x_1, x_0] f$.

We first prove that the operator $[x_1, x_0] f$ verifies the condition (1). According to Proposition 3.2, it is sufficient to prove $\|x_1-x_0\| \gamma_0 t_1 \leqslant \beta_0-\beta_1$. Since $\gamma_0 t_1<1-\sqrt{2}/2$ we show $\lambda\beta(1-\sqrt{2}/2)+\beta_1 \leqslant \beta_0$. In fact using Lemma 8.1 and $\lambda\beta\gamma_0 \leqslant \sqrt{2}/(4+\sqrt{2})$, we have

$$\lambda\beta\left(1-\frac{\sqrt{2}}{2}\right)+\beta_1 \leqslant \lambda\left(1-\frac{\sqrt{2}}{2}\right)\frac{1-\lambda\beta\gamma_0}{1-2\lambda\beta\gamma_0}\,\beta_0+(1-\lambda)\,\beta_0+\frac{\lambda^2\beta\gamma_0}{1-2\lambda\beta\gamma_0}\,\beta_0$$

$$\leqslant\left(1+\frac{\lambda((4+\sqrt{2})\,\lambda\beta\gamma_0-\sqrt{2})}{2(1-2\lambda\beta\gamma_0)}\right)\beta_0 \leqslant \beta_0.$$

Hence $\lambda\beta(1-\sqrt{2}/2)+\beta_1 \leqslant \beta_0$, and the operator $[x_1, x_0] f$ verifies the condition (1).

From Proposition 8.1, the condition $\lambda\beta\gamma_0 \leqslant \sqrt{2}/(4+\sqrt{2})<1/3$ ensures $\beta_1 \leqslant \beta_0$, and $\beta_0-\beta_1 \leqslant \lambda\beta \leqslant (1/\gamma_0)(1-\sqrt{2}/2)$.

On the other hand, since $(\beta_0-\beta_1)\,\gamma_0 \leqslant \lambda\beta\gamma_0 \leqslant \sqrt{2}/(4+\sqrt{2})$, we have

$$h(\beta_0-\beta_1)=\frac{\beta_1(1-2(\beta_0-\beta_1)\,\gamma_0)}{1-(\beta_0-\beta_1)\,\gamma_0} \geqslant \beta_1\left(1-\frac{\sqrt{2}}{4}\right)>0.$$

From Lemma 1.4, we obtain $\beta_0-\beta_1<t_1$.

Finally, since $b\gamma_0<3-2\sqrt{2}$, all the assumptions of the RF-α-Theorem are satisfied.

Let us show part (3). Let $\lambda=1$. In order that $b\gamma_0<3-2\sqrt{2}$, it is sufficient to have

$$\beta\gamma_0\,\frac{1-2\beta\gamma_0}{(1-\beta\gamma_0)^2}<3<2\sqrt{2} \qquad \text{and} \qquad 2\beta\gamma_0 \leqslant 1-\sqrt{\beta\gamma_0}.$$

A straightforward computation gives the result announced. The corollary is proved. ∎

8.2. *Case of Polynomial Systems of Degree Two.* Many applications need to solve polynomial systems of degree two. We make precise α-theorems and γ-theorems for these systems. The universal function is in this case

$$h(t)=\frac{\beta_0}{1-(\beta_0-\beta_1)}-t+gt^2$$

with $g=\gamma(f, [x_1, x_0] f, x_0)$ or $g=\gamma(f, Df(x_0), x_0)$. We obtain the three following theorems.

THEOREM 8.1. *The N alpha theorem holds with* $\alpha<\frac{1}{4}$. *Moreover* $t_1\leqslant\frac{1}{3}$.

THEOREM 8.2. *The RF alpha Theorem* 0.1 *holds with* $0\leqslant\beta_0-\beta_1\leqslant t_1$ *and* $b\gamma_0\leqslant\frac{2}{9}$. *Moreover* $t_1\leqslant\frac{1}{3}$.

THEOREM 8.3. *The S alpha Theorem* 0.2 *holds with* $0\leqslant\beta_0-\beta_1\leqslant t_1$ *and* $b\gamma\leqslant\frac{2}{9}$. *Moreover* $t_1\leqslant\frac{1}{3}$.

Let ζ be a zero of the polynomial system f of degree two. The γ-theorems become

THEOREM 8.4. *The N-γ-theorem holds for all* x_0 *such that* $\gamma(f, Df(\zeta), \zeta)\,\|x_0-\zeta\|\leqslant\frac{1}{4}$. *Moreover, the Newton sequence N verifies*

$$\|\zeta-x_k\|\leqslant\left(\frac{u}{1-2u}\right)^{2^k-1}\|\zeta-x_0\|\leqslant\left(\frac{1}{2}\right)^{2^{k-1}}\|\zeta-x_0\|, \qquad k\geqslant 0.$$

THEOREM 8.5. *The RF-γ-Theorem* 0.3 *holds with* $u/(1-v)\leqslant\frac{1}{2}$.

THEOREM 8.6. *The RF-γ-Theorem* 0.4 *holds with* $u/(1-u-v)\leqslant\frac{1}{2}$.

THEOREM 8.7. *The S-γ-Theorem* 0.5 *holds with* $u/(1-u-v)\leqslant\frac{1}{2}$ *and* $v/(1-u-v)\leqslant\frac{1}{2}$.

8.3. *Computing the Divided Difference Operator.* Let us consider $x\in\mathbb{C}^n$, and let $f(x)=(f_1(x), ..., f_n(x))=0$ be an analytic system. The computation of the operator $[y, x] f$ does not need the knowledge of the derivatives $D^k f(x)$. In fact, by definition

$$[y, x] f=([y, x]_j f_i)_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}},$$

with the convention

$$[y, x]_j f_i = \begin{cases} \dfrac{f_i(y_1, ..., y_{j-1}, y_j, x_{j+1}, ..., x_n)-f_i(y_1, ..., y_{j-1}, x_j, x_{j+1}, ..., x_n)}{y_j-x_j} \\ \qquad \text{if} \quad y_j\neq x_j, \\ \dfrac{\partial f_i}{\partial x_j}(y_1, ..., y_{j-1}, x_j, x_{j+1}, ..., x_n) \qquad \text{else.} \end{cases}$$

The interesting case is when for all j, we have $y_j \neq x_j$. In fact, each iteration of the secant type method requires the evaluation of the functions f_i's at the points

$$(x_1, ..., x_n), (y_1, x_2, ..., x_n), (y_1, y_2, x_3, ..., x_n), ..., (y_1, ..., y_n).$$

But, if we consider that the functions f_i's have been evaluated in $(x_1, ..., x_n)$ at the step $k-1$ of the iteration, then the step k needs exactly n^2 evaluations at the points

$$(y_1, x_2, ..., x_n), (y_1, y_2, x_3, ..., x_n), ..., (y_1, ..., y_n),$$

to get the operator $[y, x] f$.

Comparatively the Newton method requires the evaluation of n functions and of n^2 partial derivatives at each step. In conclusion the secant type methods may be used when the evaluation of the derivative is difficult, or when the functions are not known explicitly, but are, for example, evaluated by straight line programs.

REFERENCES

1. L. Blum, F. Cucker, M. Shub, and S. Smale, "Complexity and Real Computation," Springer-Verlag, New York/Berlin, 1998.
2. J. P. Dedieu, Condition number analysis for sparse polynomial systems, *in* "Foundations of Computational Mathematics" (F. Cucker and M. Shub, Eds.), Springer-Verlag, New York/Berlin, 1997.
3. M. ElKadi and B. Mourrain, "Approche effective des résidus algébriques," Rapport de Recherche 2834, INRIA.
4. D.-F. Han and X.-H. Wang, On dominating sequence in the point estimate and smale theorem, *Sci. China Ser. A* **33**, No. 2 (1990), 135–144.
5. L. Kantorovitch and G. Akilov, "Analyse fonctionnelle," Vol. 2, Éditions MIR, 1981.
6. J. M. Ortega and W. C. Rheinboldt, "Iterative Solution of Nonlinear Equations in Several Variables," Academic Press, San Diego, 1970.
7. A. M. Ostrowski, "Solutions of Equations and Systems of Equations," 3rd ed., Academic Press, San Diego, 1973.
8. M. Shub and S. Smale, Complexity of Bezout's theorem. 1. Geometric aspects, *J. Amer. Math. Soc.* **6** (1993), 459–501.
9. S. Smale, Newton's method estimates from data at one point, *in* "The Merging of Disciplines: New Directions in Pure, Applied, and Computational Mathematics" (R. Erwing, K. Gross, and C. Martin, Eds.), Springer-Verlag, New York/Berlin, 1986.
10. J. C. Yakoubsohn, A universal constant for the convergence of Newton's method and an application to the classical homotopy method, *Numer. Algorithms* **9** (1995), 223–244.