Asymptotic Optimality of the

Bisection Method

CUCS-92-84

K. Sikorski

Department of Computer Science
Columbia University
New York, N.Y. 10027

G.M. Trojan

Department of Physics
The University of Western Ontario
Canada N6A3K7

January 1984

AMS (MOS) Subject classifications (1970): 65H10, CR5.15.

This research was supported in part by the National Science Foundation under Grant MCS-7823676.

Abstract

The bisection method is shown to posses the asymptotically best rate of convergence for infinitely differentiable functions having zeros of arbitrary multiplicity. If the multiplicity of zeros is bounded methods are known which have asymptotically at least quadratic rate of convergence.

Summary.

We seek an approximation to a zero of an infinitely differentiable function $f: \mathbb{R} \to \mathbb{R}$ such that $f(0) \leq 0$ and $f(1) \geq 0$. It is known that the error of the bisection method using n function evaluations is $2^{-(n+1)}$. If the information used are function values, then it is known that bisection information and the bisection algorithm are optimal. Traub and Woźniakowski [6] conjectured that bisection information and the bisection algorithm are optimal even if far more general information is permitted. They permit adaptive evaluations of arbitrary linear functionals as information and arbitrary transformations of this information as algorithms. This conjecture was established in [4]. That is, for n _fixed,_ bisection information and bisection algorithm are optimal in the worst case. Thus nothing is lost by restricting oneself to function values.

One may then ask whether bisection is optimal in the _asymptotic_ worst case sense, i.e., possesses asymptotically the best rate of convergence. (Asymptotic methods are, of course, widely used in practice.) We prove that the answer to this question is positive for the class F of functions having zeros with _arbitrary multiplicity_ and continuous

functionals. Assuming that every f in F has zeros with <u>bounded multiplicity</u> there are known hybrid methods which have at least quadratic rate of convergence as n tends to infinity, see, e.g.,Brent [1], Traub [5], and Section 5.

1. Formulation of the problem.

Let $G = C^{\infty}(\mathbb{R})$ be the Fréchet space of infinitely differentiable functions on $\mathbb{R}$ with the metric ρ given by

$$\rho(f,g) = \Sigma_{i=1}^{\infty} 2^{-i}\|f-g\|_i/(1+\|f-g\|_i), \quad \forall f,g \in G,$$

where $\| \ \|_i$ is the i-th semi-norm, $i = 0,1,\ldots,$

$$\|h\|_i = \max\{|h^{(j)}(x)|, x \in [-i,i], \ j = 0,1,\ldots,i\},$$

see, e.g., Schaefer [3].

Observe that the semi-norms are monotonic, i.e., $\|h\|_{i+1} \geq \|h_i\|, \ \forall i, \ \forall h \in G$.

We seek an approximation to a zero of a function which belongs to the class F_0,

$$F_0 = \{f \in C^{\infty}[0,1], f(0) \leq 0, f(1) \geq 0, \exists\alpha: f(\alpha) = 0\}.$$

Obviously each function f in F_0 can be extended to the function $\tilde{f} \in G$. Therefore without loss of generality we consider the class F:

$$(1.1) \qquad F = \{f \in G: f(0) \leq 0, f(1) \geq 0, \exists\alpha, f(\alpha) = 0\}.$$

Define the solution operator $S: F \to [0,1]$ by

$$(1.2) \qquad S(f) = f^{-1}(0).$$

Our problem is to find an approximation to $S(f)$. To solve this problem we use an adaptive information operator (briefly information) $N: G \to \mathbb{R}^\infty$ defined as follows.

Let $f \in G$ and

$$(1.3) \qquad N(f) = [L_1(f), L_{2,f}(f), \ldots, L_{n,f}(f), \ldots]$$

where

$$L_{i,f}(\cdot) = L_i(\cdot; y_1, \ldots, y_{i-1}): G \to \mathbb{R}$$

is an arbitrary linear functional and

$$y_1 = L_1(f)$$

$$y_i = L_i(f; y_1, \ldots, y_{i-1}), \quad i = 2,3,\ldots .$$

By $N_n(f)$ we denote

$$(1.4) \qquad N_n(f) = [L_1(f), L_{2,f}(f), \ldots, L_{n,f}(f)].$$

Note that the vector $N_{n+1}(f)$ contains all components of $N_n(f)$, $N_{n+1}(f) = [N_n(f), L_{n+1,f}(f)]$. That is, increasing n we use previously computed information. It is convenient to use the notation $N = \{N_n\}$. We may assume without loss of generality that for some function f in F the functionals $L_{i,f}(\cdot)$, $i = 1,2,\ldots,n$, are linearly independent and therefore the functional L_1 is not equal to the zero functional. Let us also denote by $\mathcal{N}$ the class of all information operators of the form (1.3).

The bisection information N^{bis} is defined by

$$(1.5) \qquad L_{i,f}^{bis}(f) = f(x_i), \quad i = 1,2,\ldots,$$

where

$$x_i = (a_{i-1} + b_{i-1})/2$$

with $a_0 = 0$, $b_0 = 1$ and

$$a_i = \begin{cases} a_{i-1} & \text{if } f(x_i) > 0 \\ x_i & \text{if } f(x_i) \le 0 \end{cases}, \quad b_i = \begin{cases} b_{i-1} & \text{if } f(x_i) < 0, \\ x_i & \text{if } f(x_i) \ge 0. \end{cases}$$

Knowing $N_n(f)$ we approximate $S(f)$ by an algorithm. By the <u>algorithm</u> $\varphi = \{\varphi_n\}$ we mean a sequence of arbitrary transformations, $\varphi_n\colon N_n(G) \to \mathbb{R}$, $n = 1,2,\ldots$. Let $\Phi(N)$ denote the class of all algorithms using the information N. The n-th error of φ for an element f is defined by

$$(1.6) \qquad e_n(N,\varphi,f) = |S(f) - \varphi_n(N_n(f))|.$$

In the asymptotic setting we wish to find φ^* and N^* such that for any f in F the error $e_n(N^*,\varphi^*,f)$ goes to zero as fast as possible as n tends to infinity. The information N^* and algorithm φ^* are called optimal iff

$$\forall N \in \mathcal{N}, \quad \forall \varphi \in \Phi(N), \quad \exists f^* \in F \text{ such that}$$

$$(1.7) \qquad \limsup_{n\to\infty} \frac{e_n(N,\varphi,f^*)}{e_n(N^*,\varphi^*,f)} > 0, \qquad \forall f \in F.$$

The bisection algorithm $\varphi^{bis} = \{\varphi_n^{bis}\}$ is defined by

$$\varphi_n^{bis}(N_n^{bis}(f)) = (a_n + b_n)/2.$$

It is known that for every f in F,

$$e_n(N^{bis},\varphi^{bis},f) \le 2^{-(n+1)}$$

and

$$e_n(N^{bis},\varphi^{bis},f^*) = \begin{cases} 2^{-(n+1)}/3 & n\text{-even}, \\ 2^{-(n+1)}/6 & n\text{-odd}, \end{cases}$$

for $f^*(x) = x - 1/6$.

It was shown in [4] that for a <u>fixed</u> n:

$$\sup_{f\in F}|S(f) - \varphi_n(N_n(f))| \ge \sup_{f\in F}|S(f) - \varphi_n^{bis}(N_n^{bis}(f))|$$
$$= 2^{-(n+1)},$$

for every $N \in \mathcal{N}$ and $\varphi \in \Phi(N)$, i.e., that bisection information and algorithm are optimal for the worst case model with a <u>fixed</u> number n of functional evaluations.

Here we show that bisection information and algorithm are <u>nearly</u> optimal for the asymptotic worst case model.

More precisely, we show that for every <u>continuous</u> N, i.e., $L_{i,f}$ in (1.3) are continuous, $L_{i,f}(g_k) \xrightarrow[k\to\infty]{} L_{i,f}(g)$ whenever $\rho(g_k,g) \xrightarrow[k\to\infty]{} 0$, every algorithm $\varphi \in \Phi(N)$ and every sequence $\{\delta_n\}_{n=1}^{\infty}, \delta_n \searrow 0$, there exists a function f^* in F such that the upper limit

$$\limsup_{n\to\infty} \frac{e_n(N,\varphi,f^*)}{\delta_n\, 2^{-n}} > 0,$$

and obviously for every f in F

$$\limsup_{n\to\infty} \frac{e_n(N,\varphi,f^*)}{\delta_n e_n(N^{bis},\varphi^{bis},f)} \geq \limsup_{n\to\infty} \frac{e_n(N,\varphi,f^*)}{\delta_n\, 2^{-n}},$$

since $\delta_n 2^{-n} \geq 2\delta_n e_n(N^{bis},\varphi^{bis},f)$. The sequence $\{\delta_n\}$ may converge to zero arbitrarily slowly. Therefore we say that the bisection information and algorithm are <u>nearly</u> optimal for the asymptotic worst case model; compare also to [6, p. 199] and [8]. We formulate this result in

Theorem 1.1: For every continuous information $N \in \mathcal{N}$, every algorithm $\varphi \in \Phi(N)$ and every sequence $\{\delta_n\}_{n=1}^{\infty}$, $\delta_n \searrow 0$, there exists a function f^* in F such that

$$(1.8) \qquad \limsup_{n\to\infty} e_n(N,\varphi,f^*)/(\delta_n 2^{-n}) > 0. \qquad \square$$

2. Sketch of the proof.

First we give a sketch of the proof of Theorem 1.1. The proof is by contradiction. Suppose that there exists an operator $N^* = \{N_n^*\}$ and an algorithm $\varphi^* = \{\varphi_n^*\}$ such that for every function f in F

$$(2.1) \qquad |\varphi_n^*(N_n^*(f)) - S(f)| = o(\delta_n 2^{-n}).$$

We construct a Cauchy sequence of functions $\{g_r\}_{r=1}^{\infty}$ in $C^{\infty}(\mathbb{R})$ such that $g = \lim_{r\to\infty} g_r$ is in F and does not satisfy (2.1).

Let $\{\delta_n'\}_{n=1}^{\infty}$ be any sequence of positive numbers, $\delta_n' \searrow 0$, $\delta_1' < \frac{1}{2}$, such that $\delta_n = o(\delta_n')$. Then we construct a sequence of functions $\{f_n^1\}_{n=1}^{\infty}$, $f_n^1 \in G$ such that $f_n^1(x) = 0$ for $x \in I_n^1 = [\alpha_n^1, \beta_n^1]$ and $\delta_n' 2^{-n} \le \operatorname{diam}(I_n^1) \le 2^{-n}$, $f_n^1(x) < 0$ (resp. > 0) for $x \in [-\infty, \alpha_n^1)$ (resp. $x \in (\beta_n^1, +\infty]$) and $N_n^*(f_{n+q}^1) = N_n^*(f_n^1)$, $q = 0,1,\ldots$. Moreover $\{f_n^1\}$ is a Cauchy sequence and $I_{n+1}^1 \subset I_n^1$, $\forall n$. We prove that the limit function $f^1 = \lim_{n\to\infty} f_n^1$ is in F. By the continuity of N_n^* we get

$$N_n^*(f^1) = N_n^*(f_n^1), \quad \forall n.$$

Let $\{m_j\}_{j=1}^{\infty}$ be an increasing sequence of integers such that

$$(2.2) \qquad \delta_n < 6^{-j}\delta_n' \quad \text{for } \forall n > m_j.$$

Equation (2.1) implies that there exist $n_1 > m_1$ such that

$$(2.3) \qquad |\varphi_n^*(N_n^*(f^1)) - S(f^1)| < 6^{-1}\delta_n' 2^{-n}, \quad \forall n \geq n_1.$$

Then we define $g_1 = f_{n_1}^1$ and construct the next Cauchy sequence of functions $\{f_n^2\}_{n=1}^{\infty}$ by setting $f_n^2 = f_n^1$ for $n \leq n_1$ and constructing f_n^2, $n \geq n_1 + 1$ in such a way that $f_n^2 \in G$, $f_n^2(x) = 0$ for $x \in I_n^2 = [\alpha_n^2, \beta_n^2]$ and $6^{-1}\delta_n' 2^{-n} \leq \mathrm{diam}(I_n^2) \leq 6^{-1}2^{-n}$, $n > n_1$, $f_n^2(x) < 0$ (resp. > 0) for $x \in [-\infty, \alpha_n^2)$ (resp. $x \in (\beta_n^2, +\infty]$), $N_n^*(f_{n+q}^2) = N_n^*(f_n^2)$, $q = 0,1,\ldots,$ $I_{n+1}^2 \subset I_n^2$, $\forall n$ and

$$\mathrm{dist}(I_{n+1}^2, \varphi_n^*(N_n^*(f^1))) \geq 6^{-1}\delta_{n_1}' 2^{-n_1}, \quad \forall n \geq n_1.$$

We prove that the limit function $f^2 = \lim_{n\to\infty} f_n^2$ is in F and by the continuity of N_n^* get

$$N_n^*(f^2) = N_n^*(f_n^2), \quad \forall n.$$

Then we choose $n_2 > \max(n_1, m_2)$ such that

$$|\varphi_n^*(N_n^*(f^2)) - S(f^2)| < 6^{-2}\delta_n' 2^{-n}, \quad \forall n \geq n_2,$$

define $g_2 = f_{n_2}^2$ and repeat our construction. In this way we obtain the sequence $\{n_r\}_{r=1}^{\infty}$, and $\{f_n^r\}_{n=1}^{\infty}$, $r = 1,2,\ldots,$ such that

(2.4) $\{f_n^r\}_{n=1}^{\infty}$ is a Cauchy sequence for every r,

$$f_n^r = \begin{cases} < 0 & x \in (-\infty, \alpha_n^r) , \\ = 0 & x \in I_n^r = [\alpha_n^r, \beta_n^r] , \\ > 0 & x \in (\beta_n^r, +\infty) , \end{cases}$$

$$6^{-r}\delta_n' 2^{-n} \leq \operatorname{diam}(I_n^r) \leq 6^{-r}2^{-n} \text{ for } n > n_{r-1},$$

$$n_r > \max(n_{r-1}, m_r), \qquad n_0 = 0$$

and

$$|\varphi_n^*(N_n^*(f^r)) - S(f^r)| < 6^{-r}\delta_n' 2^{-n}, \quad \forall n \geq n_r \tag{2.5}$$

where

$$f^r = \lim_{n\to\infty} f_n^r \quad \text{belongs to } F,$$

$$N_n^*(f^r) = N_n^*(f_n^r), \qquad \forall n \tag{2.6}$$

$$N_n^*(f_{n_r}^r) = N_n^*(f_{n_{r+q}}^{r+q}), \quad q = 0,1,\ldots, \qquad n \leq n_r \tag{2.7}$$

and

$$\operatorname{dist}(I_{n+1}^{r+1}, \varphi_n^*(N_n^*(f^r))) \geq 6^{-r}\delta_{n_r}' 2^{-n_r} \geq \delta_{n_r} 2^{-n_r}, \tag{2.8}$$

$$\forall n \geq n_r.$$

We define $g_r = f_{n_r}^r$, show that $\{g_r\}_{r=1}^{\infty}$ is a Cauchy sequence and that the limit function $g = \lim_{r\to\infty} g_r$ belongs to F. By the continuity of N_n^* and (2.7) we get

$$N_n^*(g) = N_n^*(g_r) \quad \forall r, \quad n \leq n_r. \tag{2.9}$$

Construction of g_r implies that $S(g)$ belongs to the interval I_{n+1}^{r+1}, $\forall r$. Therefore (2.4), (2.5) and (2.6) yield

$$\begin{aligned} |\varphi_{n_r}^*(N_{n_r}^*(g))-S(g)| &\geq \operatorname{dist}(\varphi_{n_r}^*(N_{n_r}^*(f^r)), I_{n_r+1}^{r+1}) \\ &\geq \delta_{n_r} 2^{-n_r}, \quad \forall n_r, \end{aligned}$$

which contradicts (2.1) and completes the proof.

3. Auxiliary lemmas.

In order to construct the functions in Section 2 we need a few auxiliary lemmas.

Let $U_1,\ldots,U_m$ be linearly independent continuous linear functionals on G, and $E_1,\ldots,E_m$ closed subintervals of $\mathbb{R}$. Denote $A_m = \cup_{j=1}^{m} E_j$ and

$$C(A_m) = \{f \in G: \operatorname{supp}(f) \subset A_m\}.$$

Lemma 3.1: For every positive ϵ and every family of nondegenerated intervals E_i, $i = 1,2,\ldots,m-1$, such that $U_1,\ldots,U_{m-1}$ are linearly independent on $C(A_{m-1})$, there exists an interval $E_m \subset \mathbb{R}$, with $\operatorname{diam}(E_m) = \epsilon$, such that $U_1,\ldots,U_m$ are linearly independent on A_m. □

Proof: This is the same as the proof of Lemma 2.1 of [4] with $C^\infty[a,b]$ replaced by $C^\infty(\mathbb{R})$, and therefore the proof is omitted. □

Proposition 3.1: For every n, ϵ, $'\epsilon > 0$, and continuous information $N \in \eta$ there exist a function $f_n \in G$, interval $I_n = [\alpha_n,\beta_n]$, $\operatorname{diam}(I_n) \in (0,2^{-n}]$ and intervals E_j, $j = 1,2,\ldots,k_n$, where k_n is the maximal number of linearly independent functionals on G among $L_{1,f_n},\ldots,L_{n,f_n}$

(denote them by $L^*_{1,n},\dots,L^*_{k_n,n}$) such that:

(i) $L^*_{1,n},\dots,L^*_{k_n,n}$ are linearly independent on $C(\cup_{j=1}^{k_n} E_j)$,

(ii) $$f_n(x) = \begin{cases} < 0 & x \in (-\infty,\alpha_n), \\ = 0 & x \in [\alpha_n,\beta_n], \\ > 0 & x \in (\beta_n,+\infty), \end{cases}$$

and $\operatorname{dist}(E_j,I_n) \geq \frac{1}{2}\operatorname{diam}(E_j)$, $\forall j$. □

Proof: The proof is similar to the proof of Lemma 2.2 of [4]. We prove below a more general Lemma 3.3 which combined with Lemma 3.2 yields the proposition by induction. The formulation of Lemma 3.3 enables easy varying the interval I_n which is needed in the proof of Theorem 1.1. □

Lemma 3.2: The proposition 3.1 holds for $n = 1$ with f_1 such that $\operatorname{diam}(I_1) \geq \delta_1'/2$, where δ_1' is the first element of the sequence $\{\delta_n'\}$ from Section 2. □.

Proof: Since $L_1 \neq 0$ on G then as in the proof of Lemma 2.1 of [3] we conclude that there exists an interval E_1, $\operatorname{diam}(E_1) \leq 1/4$, such that $L_1 \neq 0$ on $C(E_1)$. Let

$$v_1(x) = \begin{cases} -\exp(-x^2) & x \in (-\infty,0), \\ 0 & x \in [0,1/4], \\ \exp(-(x-1/4)^{-2}) & x \in [1/4,+\infty), \end{cases}$$

and

$$v_2(x) = v_1(x-3/4).$$

Note that $v_i \in G$. Define f_1 by

$$f_1 = \begin{cases} v_1 & \text{if } E_1 \subset [3/8,+\infty), \\ v_2 & \text{otherwise.} \end{cases}$$

Then $\operatorname{diam}(I_1) = 1/4$, so $\delta_1'/2 \le \operatorname{diam}(I_1) \le 1/2$ and $\operatorname{dist}(I_1,E_1) \ge 1/8 = 1/2 \operatorname{diam}(E_1)$ which shows that f_1 satisfies Lemma 3.2. $\square$

<u>Lemma 3.3</u>: Suppose that Proposition 3.1 holds for n, and let Z_n be an arbitrary interval $Z_n \subset I_n$. Then Proposition 3.1 holds for $n+1$ with $I_{n+1} \subset Z_n$ such that

$$\operatorname{diam}(I_{n+1}) \ge (\delta_{n+1}'/\delta_n')\operatorname{diam}(Z_n)/2,$$

where $\{\delta_n'\}$ is as in Section 2, $k_n \le k_{n+1} \le k_n + 1$ and $L_{i,n+1}^* = L_{i,n}^*$, $i = 1,2,\ldots,k_n$. $\square$

<u>Proof</u>: Let $Z_n = [e_1,e_2]$, $d_n = \min(2^{-n}, \min_{j=1,\ldots,k_n} \operatorname{diam}(E_j)/2)$

$I_n = [\alpha_n,\beta_n]$, $b_n = \operatorname{diam}(Z_n)(1 - \frac{\delta_{n+1}'}{\delta_n'})/2$ and $M = (e_1+e_2)/2$.

Define the functions, see Fig. 3.1,

$$H_{n,1}(x) = \begin{cases} -\exp(-(x-\alpha_n+d_n)^{-2}(x-M-b_n)^{-2}) & x \in [\alpha_n-d_n, M+b_n], \\ \exp(-(x-e_2)^{-2}(x-\beta_n-d_n)^{-2}) & x \in [e_2,\alpha_n+d_n], \\ 0 & \text{otherwise}, \end{cases}$$

$$H_{n,2}(x) = \begin{cases} \exp(-(x-M+b_n)^{-2}(x-\beta_n-d_n)^{-2}) & x \in [M-b_n,\beta_n+d_n], \\ -\exp(-(x-\alpha_n+d_n)^{-2}(x-e_1)^{-2}) & x \in [\alpha_n-d_n,e_1], \\ 0 & \text{otherwise.} \end{cases}$$

Consider the functions $G_{n,1}$, $G_{n,2} \in C(\cup_{i=1}^{k_n} E_i)$ such that

$$L^*_{i,n}(H_{n,j}+G_{n,j}) = 0, \qquad i = 1,2,\ldots,k_n, \; j = 1,2.$$

Such functions exist since $L^*_{i,n}$ are linearly independent on $C(\cup_{j=1}^{k_n} E_j)$. Let

$$h_{n,j} = H_{n,j} + G_{n,j}. \tag{3.1}$$

Choose a positive constant c_n so small that

$$c_n \max_{x\in\mathbb{R}}|h_{n,j}(x)| \leq \min \begin{cases} \min\limits_{x\in(-\infty,\alpha_n-d_n]\cup[\beta_n+d_n,+\infty)} |f_n(x)|/4 \\ \max\limits_{j=1,2} \max\limits_{x\in\mathbb{R}} |h_{n-1,j}(x)|\cdot c_{n-1}/2 \end{cases} \tag{3.2}$$

and

$$c_n\|h_{n,j}\|_n \leq 2^{-n} \qquad \text{for } j = 1,2 \tag{3.3}$$

where $h_{0,j}(x) \equiv 1$ and $c_0 = 1$.

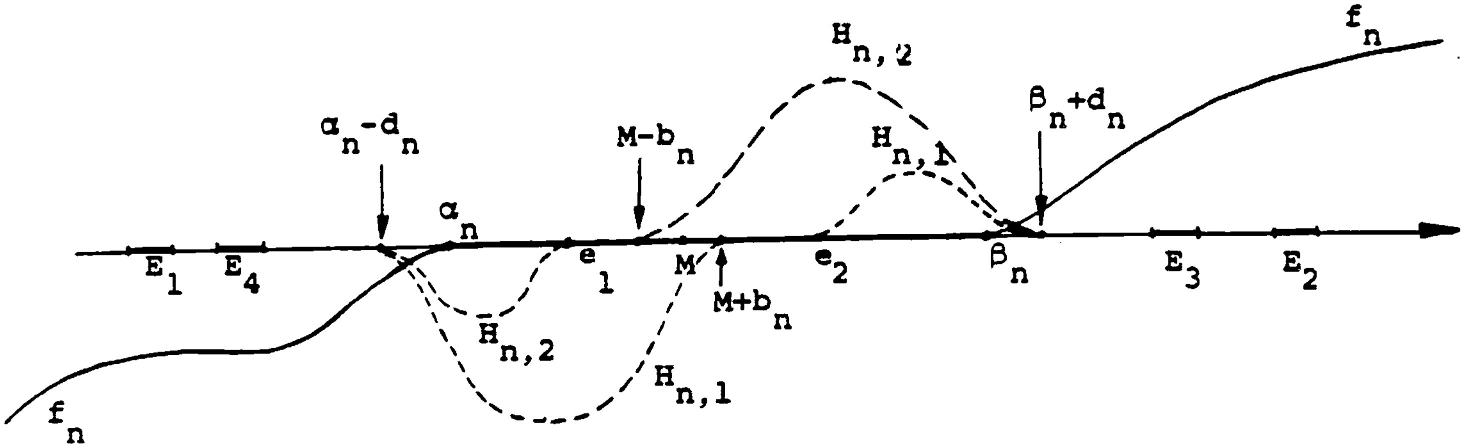

Fig 3.1

Let $f_{n,j} = f_n + c_n h_{n,j}$, $j = 1,2$. Then $f_{n,j} \in G$ and

(3.4) $$N_n(f_{n,j}) = N_n(f_n),$$

since $L^*_{i,n}(h_{n,j}) = 0$, $i = 1,2,\ldots,k_n$ and $L_{i,f_n}(h_{n,j}) = 0$, $i = 1,2,\ldots,n$. The information operator N_{n+1} yields a functional L_{n+1,f_n}. If $L^*_{1,n},\ldots,L^*_{k_n,n},L_{n+1,f_n}$ are linearly dependent on G then Lemma 3.3 holds for $f_{n+1} = f_{n,1}$ (also for $f_{n+1} = f_{n,2}$) and $k_{n+1} = k_n$. Of course in both cases

$$\operatorname{dist}(E_i,I_{n+1}) \geq \operatorname{diam}(E_i)/2, \; i = 1,2,\ldots,k_n$$

and $\operatorname{diam}(I_{n+1}) = \operatorname{diam}(Z_n)/2 - b_n = \operatorname{diam}(Z_n)\delta'_{n+1}/(2\delta'_n)$. If $L^*_{1,n},\ldots,L^*_{k_n,n},L_{n+1,f_n}$ are linearly independent on G then Lemma 2.1 yields that there exists an interval

$E_{k_{n+1}} = [e_1, e_1+b_n]$, $k_{n+1} = k_n + 1$, such that they are linearly independent on $C(\cup_{j=1}^{k_n} E_j)$. Then define f_{n+1} by

$$(3.5) \qquad f_{n+1} = \begin{cases} f_{n,1} & \text{if } E_{k_{n+1}} \subset (-\infty, M+b_n/2], \\ f_{n,2} & \text{otherwise.} \end{cases}$$

Then

$$I_{n+1} = \begin{cases} [M+b_n, e_2] & \text{if } E_{k_{n+1}} \subset (-\infty, M+b_n/2], \\ [e_1, M-b_n] & \text{otherwise.} \end{cases}$$

Obviously $\mathrm{dist}(E_i, I_{n+1}) \geq \mathrm{diam}(E_i)/2$, $i = 1,2,\ldots,k_{n+1}$ and $\mathrm{diam}(I_{n+1}) = \mathrm{diam}(Z_n)/2 - b_n = \mathrm{diam}(Z_n)\cdot\delta'_{n+1}/(2\delta'_n)$. Thus the function f_{n+1} satisfies Lemma 3.3. □

Lemma 3.4: Let $\{f_n\}$ be a sequence of functions constructed by applying Lemma 3.3 to the function f_1 from Lemma 3.2. Then $\{f_n\}$ is convergent in G and $f = \lim_{n\to\infty} f_n$ belongs to F. □

Proof: Observe that each of the functions f_n, $n \geq 2$ is of the form

$$(3.6) \qquad f_n = f_1 + \Sigma_{i=1}^{n-1} h_i$$

where $h_i = c_i h_{i,j}$ for $j = 1$ or $j = 2$, see (3.1), (3.2) and (3.3), $h_i \in G$ and

$$(3.7) \qquad \max_{x\in\mathbb{R}}|h_i(x)| \le \min\begin{cases}\min|f_i(x)|/4 , \\ x \in (-\infty,\alpha_i-d_i]\cup[\beta_i+d_i,+\infty) , \\ \max|h_{i-1}(x)|/2 ,\end{cases}$$

and

$$(3.8) \qquad \|h_i\|_i \le 2^{-i}.$$

We first prove that $\{f_n\}$ is a Cauchy sequence, which combined with completeness of G implies convergence.

Assume without loss of generality that $n > m$. Then (3.6), (3.8) and monotonicity of semi-norms imply that

$$\begin{aligned}\rho(f_n,f_m) &\le \Sigma_{i=1}^{m-1}\ 2^{-i}\|f_n-f_m\|_i + \Sigma_{i=m}^{\infty}\ 2^{-i} \\ &= \Sigma_{i=1}^{m-1}\ 2^{-i}\|\Sigma_{j=m}^{n-1}\ h_j\|_i + 2^{-(m-1)} \\ &\le \Sigma_{i=1}^{m-1}\ 2^{-i}\ \Sigma_{j=m}^{n-1}\ \|h_j\|_j + 2^{-(m-1)} \\ &\le 2^{-(m-1)}1 + 2^{-(m-1)} = 4\cdot 2^{-m},\end{aligned}$$

which yields convergence. Let $f = \lim_{n\to\infty} f_n$. Now we prove that f is in F, i.e., that it has exactly one zero. Recall that $I_{n+1} \subset I_n$ and $\operatorname{diam}(I_n) \le 2^{-n}$. Therefore $f(\alpha) = 0$, $\alpha = \cap_{n=1}^{\infty} I_n$, and $f^{(j)}(\alpha) = 0$, $j = 1,2,\ldots$, since convergence in metric ρ implies uniform convergence with all derivatives on every closed interval in $\mathbb{R}$. Now we show that α is the only zero of f. Namely take arbitrary

$x \neq \alpha$ and assume without loss of generality that $x > \alpha$. Then we show that $f(x) > 0$. Since $I_{n+1} \subset I_n$ and $\mathrm{diam}(I_n) \leq 2^{-n}$ then there exists an index j^* such that

$$x \notin [\alpha_n - 2^{-n}, \beta_n + 2^{-n}] \quad \forall n \geq j^*.$$

Using (3.6), (3.7), the fact that $x \in [\beta_{j^*} + 2^{-j^*}, +\infty)$ and $d_n \leq 2^{-n}$ we get

$$\begin{aligned} f(x) &= f_1(x) + \Sigma_{j=1}^{j^*-1} h_j(x) + \Sigma_{j=j^*}^{\infty} h_j(x) = f_{j^*}(x) + \Sigma_{j=j^*}^{\infty} h_j(x) \\ &\geq f_{j^*}(x) - \Sigma_{j=j^*}^{\infty} |h_j(x)| \geq f_{j^*}(x) - \max_{t \in \mathbb{R}} |h_{j^*}(t)| \Sigma_{j=j^*}^{\infty} 2^{j^*-j} \\ &\geq f_{j^*}(x) - 2 \max_{t \in \mathbb{R}} |h_{j^*}(t)| \\ &\geq \frac{1}{2}\left(\min_{t \in (-\infty, \alpha_{j^*} - 2^{-j^*}] \cup [\beta_{j^*} + 2^{-j^*}, +\infty)} |f_{j^*}(t)| \right) > 0, \end{aligned}$$

$\square$

which completes the proof.

4. Constructions needed in the proof of Theorem 1.1

In order to complete the proof of Theorem 1.1 we construct the sequences $\{f_n^r\}_{n=1}^{\infty}$, $r = 1,2,\ldots$, by use of Lemmas 3.3 and 3.2 for the information N^*.

Namely let $f_1^1 = f_1$ from Lemma 3.2 and let $\{f_n^1\}_{n=2}^{\infty}$ be the sequence of functions from Lemma 3.3 with the intervals Z_n equal to I_n for every n. Lemma 3.4 yields that $f^1 = \lim_{n\to\infty} f_n^1$ exists and belongs to F. Moreover (3.4) implies that $N_n(f^1) = N_n(f_n^1), \forall n$. Constructions in the proof of Lemma 3.3 imply that f_n^1 has all the properties from Section 2.

Now suppose we have constructed the sequence $\{f_n^r\}_{n=1}^{\infty}$, $r \geq 1$, by applying Lemma 3.3 to the function f_1 from Lemma 3.2, such that (2.4), (2.5) and (2.6) are satisfied, where $f^r = \lim_{n\to\infty} f_n^r$ exists and belongs to F by Lemma 3.4. We set $g_r = f_{n_r}^r$ and define the next sequence $\{f_n^{r+1}\}_{n=1}^{\infty}$ as follows:

Set $f_n^{r+1} = f_n^r$ for $n \leq n_r$ and let $f_{n_r+1}^{r+1}$ be the function from Lemma 3.3 applied to the function $f_{n_r}^r$, with the interval Z_{n_r} given by

$$Z_{n_r} = \begin{cases} [\alpha^r_{n_r}, \alpha^r_{n_r} + \mathrm{diam}(I^r_{n_r})/6] & \text{if } S(f^r) \geq (\alpha^r_{n_r} + \beta^r_{n_r})/2, \\ [\beta^r_{n_r} - \mathrm{diam}(I^r_{n_r})/6, \beta^r_{n_r}] & \text{otherwise}. \end{cases}$$

Then f^{r+1}_{n+1} are constructed by Lemma 3.3 with $Z_n = I^{r+1}_n$ for every $n \geq n_r+1$. Lemma 3.4 implies that $f^{r+1} = \lim_{n\to\infty} f^{r+1}_n$ exists and belongs to F. Moreover the above definition of Z_r combined with (2.5) yields (2.8), since $\mathrm{diam}(I^r_{n_r}) \geq 6^{-r} \delta'_{n_r} 2^{-n_r}$. Lemma 3.3 yields also that (2.4) and (2.6) hold for the sequence $\{f^{r+1}_n\}$. The construction of f^{r+1}_n implies that $N^*_n(f^r_{n_r}) = N^*_n(f^{r+1}_{n_{r+1}})\ \forall\, n \leq n_r$ since $n_{r+1} > n_r$. Therefore by induction (2.7) holds.

Let $\{p_n\}^\infty_{n=1}$ be the sequence of functions

$$p_n = f^r_n \quad \text{for} \quad n_{r-1} < n \leq n_r, \quad r = 1,2,\ldots, \quad r_0 = 0.$$

Then Lemma 3.4 yields that $g = \lim_{n\to\infty} p_n$ exists and belongs to F. Moreover $g = \lim_{r\to\infty} g_r$, since $\{g_r\}$ is a subsequence of $\{p_n\}$.

Observe that $S(g) = \cap^\infty_{r=1} I^r_{n_r}$, since $I^{r+1}_{n_{r+1}} \subset I^{r+1}_{n_r+1} \subset I^r_{n_r}$ and $\mathrm{diam}(I^r_{n_r}) \leq 6^{-r} 2^{-n_r}$. Therefore $S(g) \in I^{r+1}_{n_r+1}$ which finally completes our constructions.

<u>Remark 4.1</u>. Observe that $g^{(j)}_r(x) = 0$, $j = 0,1,\ldots$, $x \in I^r_{n_r}$. Therefore, as in the proof of Lemma 3.4 we conclude that

$g^{(j)}(S(g)) = 0$, $j = 0,1,\ldots$, i.e., that g has a zero with infinite multiplicity. $\square$

5. Final remarks.

(i) Remark 4.1 indicates that bisection is nearly optimal in the subclass of F consisting of functions having zeros with arbitrary multiplicity.

(ii) The idea of the proof is based on the "remanance" property introduced by Delahaye and Germain-Bonne in [2]. For other applications, see also Trojan [7].

(iii) If the multiplicity of zeros of functions in F is bounded it is possible to construct information N and algorithms φ which guarantee asymptotically quadratic convergence. If the multiplicity of a zero is known, say m, it is enough to use a combination of bisection and modified Newton's method: $x_{i+1} = x_i - mf(x_i)/f'(x_i)$, which converges quadratically for $i \to \infty$, see [5, p. 127]. If the multiplicity m of a zero is unknown we can calculate it by using a combination of bisection and Newton's method and applying Aitken's δ^2 formula, see [5, p. 129, Appendix D]. Then knowing m we proceed as above.

Acknowledgements.

We are greatly indebted to K. Rygielski, J.F. Traub, G.W. Wasilkowski, A. Werschulz and H. Woźniakowski for the remarks on the manuscript.

References.

[1] Brent, R.P., Algorithms for Minimization without Derivatives, Prentice Hall, Englewood Cliffs, New Jersey, 1973.

[2] Delahaye, J.P., Germain-Bonne, B., The Set of Logarithmically Convergent Sequences Cannot be Accelerated, SIAM. J. Num. Anal., August 1982, v. 19, No. 4, pp. 840-844.

[3] Schaefer, H.H., Topological Vector Spaces, the Mac-Millan Company, New York, 1967.

[4] Sikorski, K., Bisection is Optimal, Num. Math., 40, 1982, pp. 111-117.

[5] Traub, J.F., Iterative Methods for the Solution of Equations, Prentice-Hall, Englewood Cliffs, 1964. Reissued by Chelsea, New York, 1982.

[6] Traub, J.F., Woźniakowski, H., A General Theory of Optimal Algorithms, Acad. Press, New York, 1980.

[7] Trojan, G.M., An Upper Bound on the Acceleration of Convergence, submitted for publication.

[8] Trojan, G.M., Asymptotic Model for Linear Problems, in progress.