ASYMPTOTIC NEAR OPTIMALITY OF THE BISECTION METHOD

K. Sikorski
Department of Computer Science
University of Utah
Salt Lake City
UT 84112

&

G. M. Trojan
Department of Computer Science
University of Waterloo
Waterloo, Ontario N2L3G1

UUCS-88-007

AMS (MOS) Subject Classifications : 65 H 10, CR 5.15

Abstract

The bisection method is shown to possess the nearly 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 : [0,1] → $\Re$ 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 conjectured in [4] that the bisection information and algorithm are optimal even if far more general information is permitted. They permit adaptive (sequential) evaluations of arbitrary linear functionals and arbitrary transformations of this information as algorithms. This conjecture was established in [2]. That is for n fixed, the bisection information and algorithm are optimal in the worst case setting. Thus nothing is lost by restricting oneself to function values.

One may then ask whether bisection is nearly optimal in the asymptotic worst case sense, that is, possesses asymptotically nearly the best rate of convergence. Methods converging fast asymptotically, like Newton or secant type, are of course, widely used in scientific computation. We prove that the answer to this question is positive for the class F of functions having zeros of infinite multiplicity and information consisting of evaluations of continuous linear functionals. Assuming that every f in F has zeroes with bounded multiplicity, there are known hybrid methods which have at least quadratic rate of convergence as n tends to infinity, see. e.g., Brent [1], Traub [3] and Section 1.

1. Formulation of the Problem.

Let $G = C^\infty[0,1]$ be the space of infinitely differentiable real-valued functions on the interval $I = [0,1]$ with the metric ρ given by

$$\rho(f,\, g) = \sum_{i=1}^{\infty} 2^{-i} \parallel f - g \parallel_i /(1 + \parallel f - g \parallel_i),\ \forall f, g \epsilon G$$

where

$$\parallel f \parallel_i = \max_{0 \leq j \leq i} \sup_{x \epsilon I} \mid f^{(j)}(x) \mid \ .$$

Let $S(f) = f^{-1}(0)$ denote the set of all zeros of the function f. We seek an approximation to a zero of a function which belongs to the class F:

(1.1)

$$F = \{f \epsilon G :\ f(0) \leq 0, f(1) \geq 0 \textit{ and } S(f) \textit{ is a singleton}\};$$

i.e. every function in F has exactly one zero. To solve this problem, we use an adaptive information operator (briefly information) N : G $\rightarrow \Re^{\infty}$ defined as follows:

Let $f \epsilon G$ **and**

(1.2)

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

where

$$L_{i,f}(\cdot) = L_i(\ \cdot\ ; y_1, \cdots y_{i-1}) \quad : G \rightarrow \Re$$

is an arbitrary linear functional and

$$y_1 = L_1(f), \qquad y_j = L_j(f;\ y_1, \cdots y_{j-1}) \quad , j = 2, 3, \ldots i - 1.$$

Observe that $L_{i,f}(\cdot)$ depends on the previously computed values y_j , $j = 1, .. i - 1$.

By $N_n(f)$ we denote

(1.3)

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

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 do use previously computed information. We may assume without loss of generality that the functionals in N($\cdot$) are linearly independent, i.e.,

(1.4) $L_1, L_{2,f}, \ldots \quad L_{n,f}$ are linearly independent for every $f \ \epsilon \ G$, $n = 1, 2, \ldots$.

Let us denote by $\mathcal{N}$ the class of all information operators of the form (1.3). Knowing $N_n(f)$ we approximate $S(f)$ by an algorithm. By the algorithm $\phi = \{\phi_{\ n}\}$ we mean a sequence of arbitrary transformations, $\phi_n : \ N_n(G) \rightarrow I$, $n = 1, 2 \ldots$. Let $\phi(N)$ be the class of all algorithms using information N. The n- th error of ϕ for an element f is defined by

(1.5)

$$e_n(N, \phi, f) = | \ S(f) - \phi_n(N_n(f))|.$$

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

The information N^* and algorithm ϕ^* are called nearly optimal iff $\forall N \in \mathcal{N}$, $\forall \phi \in \phi(N)$ *and* $\forall$ *sequence* δ_n,

$$\delta_n \searrow 0 \ (\delta_n \ \textit{strictly decreasing}),$$

$$\exists f^* \in F \quad \textit{such that} \ \forall f \in F :$$

(1.6)

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

This means that an arbitrary algorithm ϕ does not converge essentially faster for the function f^* than the algorithm ϕ^* for any function f.

The bisection information N^{bis} is defined by

(1.7)

$$L_{i,f}^{bis}(f) = f(x_i) \ , \ i = 1, 2, \dots ,$$

where

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

with

$$a_0 = 0, \quad b_0 = 1 \quad and$$

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

The bisection algorithm $\phi^{bis} = \{\phi_n^{bis}\}$ is given by

$$\phi_n^{bis}(N_n^{bis}(f)) = \begin{cases} (a_n + b_n)/2 & if \quad f(a_n) \cdot f(b_n) < 0, \\ a_n & if \quad f(a_n) = 0, \\ b_n & if \quad f(b_n) = 0. \end{cases}$$

It is known that for every f in F

(1.8)

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

and that there exists functions f in F such that

(1.9)

$$e_n(N^{bis}, \phi^{bis}, f) \geq c\, 2^{-(n+1)},$$

for some c>0, like for example $f_i(t) = t - \frac{i}{6}$, $i = 1, 2, 4, 5$. In fact there exists an infinite number of such functions.

It was shown in [2] that for a fixed n

$$\sup_{f \epsilon F} e_n(N, \phi, f) \geq \sup_{f \epsilon F} e_n(N^{bis}, \phi^{bis}, f) \geq 2^{-(n+1)},$$

for every $N \epsilon \mathcal{N}$ *and* $\phi \epsilon \phi\ (N)$, i.e., that the bisection information and algorithm are optimal for the worst case model with a fixed number of functional evaluations.

Here we show that the bisection information and algorithm are nearly optimal for the asymptotic worst case setting. More precisely, assume that the information N is continuous, i.e.,

$$L_{i,f}(g_k) \rightarrow L_{i,f}(g) \text{ whenever } \rho(g_k, g) \underset{k \rightarrow \infty.}{\rightarrow 0}$$

For an arbitrary sequence δ_n, $\delta_n \searrow 0$, any $N \epsilon \mathcal{N}$ and any $\phi \epsilon \phi\ (N)$ define the set $B = B(N, \phi, \delta_n)$ of functions from F such that the error $e_n(N, \phi, f)$ is essentially at least $\delta_n \cdot 2^{-n}$, i.e.,

(1.10)

$$B = \{f \epsilon F : \limsup_{n \rightarrow \infty} \frac{e_n(N, \phi, f)}{\delta_n \cdot 2^{-n}} > 0.\}$$

To prove near optimality of the bisection method, it is enough to show that the set B is not empty for any δ_n, $N and\ \phi$. Indeed, taking any $f^* \epsilon B$ and any f from F we have

$$\limsup_{n \rightarrow \infty} \frac{e_n(N, \phi, f^*)}{\delta_n e_n(N^{bis}, \phi^{bis}, f)} \geq \limsup_{n \rightarrow \infty} \frac{e_n(N, \phi, f^*)}{\delta_n \cdot 2^{-n}} > 0.$$

We will show more by proving that the Lebesgue measure of the set S(B) of zeros of all functions from B is unity. This in particular implies that the set B is uncountable. Precisely, define the set S(B) by

(1.11)

$$S(B) = \{x \epsilon\ [0, 1] : \exists f \epsilon B : x \epsilon S(f)\}$$

We prove

Theorem 1.1

For every continuous information $N \in \mathcal{N}$, every algorithm $\phi \in \phi\ (N)$ and any sequence $\delta_n,\ \delta_n \searrow 0$, the Lebesgue measure μ of the set S(B) is unity, i.e.,

$$\mu(S(B)) = 1.$$

We remark that if the multiplicity m of a zero of f is finite, then it is possible to construct information N and algorithm ϕ which guarantee asymptotically quadratic convergence, see [1] and [3]. We can calculate m by using a combination of bisection and Newton's methods and applying Aitken's δ^2 formula, see [3, p.129, Appendix D]. Knowing m we may use the modified Newton's method [3, p. 127] $x_{i+1} = x_i - m\ f(x_i)/f'(x_i)$ which converges quadratically for $i \rightarrow \infty$. For such information and algorithm, the set B contains functions with zeros of infinite multiplicity. Therefore, we can not essentially beat the bisection only for functions having infinite multiplicity zeros.

In the next section we present auxiliary lemmas and the proof of Theorem 1.1.

2. Auxiliary Lemmas.

In this section, we prove a few auxiliary lemmas needed in the proof of Theorem 1.1. The first lemma 2.1, was proved in [2]. Namely, let I_i , $i = 1, , \ldots k$, be closed intervals in [0,1] and

$$G(\bigcup_{i=1}^{k} I_i) = \{f \epsilon G :\ supp(f) \subset \bigcup_{i=1}^{k} I_i\}.$$

Lemma 2.1

Let $L_i : G \rightarrow \Re$, $i = 1, \ldots k$ be linearly independent linear funtionals. Then for every positive α and any family of closed intervals $I_i \subset [0,1], i = 1, \ldots, k-1$ such that $L_1, \ldots, L_{k-1}$ are linearly independent on $G\bigcup_{i=1}^{k-1} I_i)$ there exists a closed interval $I_k \subset [0,1]$ of length α, such that $L_1,\ldots,L_k$ are linearly independent on $G(\bigcup_{i=1}^{k} I_i)$.

In the next lemma, we construct a family of functions from G needed in the proof of Theorem 1.1.

Lemma 2.2

For every $\varepsilon > 0, 0 < \varepsilon < \frac{1}{14}$, there exists a family of functions F^{ε} ,

$$F^{\varepsilon} = \{f_{i,n} \in G \; n = 0, 1, ..; i \epsilon [1, 3^n]\}$$

with the following properties:

(2.1) F^{ε} is tree structured, where $f_{1,0}$ is the root of the tree,

$$f_{1,0} = \begin{cases} -exp(-(x-\varepsilon)^{-2}) & x \epsilon [0, \varepsilon], \\ 0 & x \epsilon (\varepsilon, 1-\epsilon), \\ exp(-(x-1+\varepsilon)^{-2}) & x \epsilon [1-\varepsilon, 1], \end{cases}$$

and the functions on the n- th level, n = 1,2,.. , are constructed inductively in what follows:

(2.2) Every function $f \epsilon F^{\varepsilon}$ satisfies

$$f(x) = \begin{cases} < 0 & x \epsilon [0, \alpha^{*}_{f}) \, , \\ = 0 & x \epsilon [\alpha^{*}_{f}, \alpha^{**}_{f}] \, , \\ > 0 & x \epsilon (\alpha^{**}_{f}, 1] \, , \end{cases}$$

for some $[\alpha^{*}_{f} \alpha^{**}_{f}] \subset [0, 1]$.

(2.3)

For every $f = f_{i,n}$ there exists closed intervals $I_1, I_2, \dots I_n \subset [0, 1]$ such that the functionals $L_1, L_{2,f}, ..L_{n,f}$ see (1.3), are linearly independent on $G(\bigcup_{i=1}^{n} I_i)$, and the distance dist $(I_i, [\alpha^{*}_{f}, \alpha^{**}_{f}]) \geq \varepsilon_n, \; i = 1, \dots n, \; where \; \varepsilon_n = \varepsilon \cdot 2^{-2n}$,

$$\text{and dist } (X, Y) = \min_{\substack{x \epsilon X \\ y \epsilon Y}} | x - y |$$

with the convention dist $(\Phi, Y) = +\infty$.

Proof. (Construction).
Let ε be a small positive number, $0 < \varepsilon < \frac{1}{14}, \varepsilon_n = \varepsilon \cdot 2^{-2n}$ and let $\{\delta_n\}$ be an arbitrary sequence monotonically decreasing to $0, \delta_n \searrow 0$. Define a sequence of indices $\{n_k\}, k = 1, 2, \dots$ such that

(2.4)

$$\delta_{n_k} = o((\frac{2}{3})^k), as\ k \to \infty.$$

Let

(2.5)

$$H(x;a,b) = \begin{cases} exp(-(x-a)^{-2}(x-b)^{-2}) & , \quad x\epsilon[a,b], \\ 0 & , \quad otherwise, \end{cases}$$

for any interval [a,b] $\subset [0,1]$.

The family F^{ϵ} is tree structured. Namely, at the root we have the function $f_{1,0}$ defined in (2.1). At the n- th level of the tree we have c_n functions, where $c_n \leq 3 \cdot c_{n-1}$ if $n = n_k + 1$ for some k or $c_n \leq 2c_{n-1}$ otherwise. Thus, the number $c_n \leq 3^n$. We define the functions $f_{i,n}$ inductively, i.e., construct $f_{i,n}$ assuming that all $f_{j,k}$, $k = 0, \ldots n-1$, have already been constructed.

When n = 0 then $f_{1,0}$ is defined in (2.1). Obviously, $\alpha^*_{f_{1,0}} = \varepsilon\ and\ \alpha^{**}_{f_{1,0}} = 1 - \varepsilon$. Since there exists no intervals I_j in this case then dist $(\Phi, [\varepsilon, 1-\varepsilon]) = +\infty > \varepsilon_0 = \varepsilon$.Thus $f_{1,0}$ satisfies the induction basis.

Assume now that all $f_{j,k}, k = 0, \ldots n-1$, have been constructed. Let $k = n-1$ and let $f = f_{i,n-1}$ be any function on the n-1 -st level. The information operator N_n yields the functional $L_{n,f}$, see (1.3). Due to assumption (1.4), the functionals $L_1, L_{2,f}, \ldots L_{n,f}$ are linearly independent on G. Thus, Lemma 2.1 with $\alpha = 2\varepsilon_n$ and $k = n-1$ yields an interval $I_n, I_n = [m - \varepsilon_n, m + \varepsilon_n]$ such that $L_1, L_{2,f}, .L_{n,f}$ are linearly independent on $G(\bigcup_{i=1}^n I_i)$, where $I_1, \ldots I_{n-1}$ are the intervals from (2.3) for the function f. Now we construct the functions on the n- th level which are successors in F^{ϵ} to f. Let $\alpha^* = \alpha^*_f\ and\ \alpha^{**} = \alpha^{**}_f$.

If $\alpha^{**} - \alpha^* \leq 6\varepsilon_n$ then we let f be a leaf of the tree and therefore the successors are not defined.

If $\alpha^{**} - \alpha^* > 6\varepsilon_n$ then we define the successors $f_{i_j,n}, j\epsilon\{1,2,3\}$, $i_j\epsilon[1,3^n]$ depending on whether $n = n_k + 1$ for some k or not.

Let M = $(\alpha^* + \alpha^{**})/2$ and define the auxiliary functions $H_j, j\epsilon\{1,2,3\}$ by:

(2.6)

$$If\ n \neq n_k + 1\ for\ any\ k\ then:$$

(i)

$$if\ \alpha^* + 3\varepsilon_n \leq m \leq \alpha^{**} - 3\varepsilon_n\ then$$

$$H_1(x) = -H(x; \alpha^* - \varepsilon_n, m + 2\varepsilon_n),$$

$$H_2(x) = H(x; m - 2\varepsilon_{n,} \alpha^{**} + \varepsilon_n);$$

(ii)

$$if \quad m < \alpha^* + 3\varepsilon_n \; then$$

$$H_1(x) = -H(x; \alpha^* - \varepsilon_n, \; max(M, m + 2\varepsilon_n))$$

$$H_2(x) = -H(x; \alpha^* - \varepsilon_{n,} max(\alpha^* - \varepsilon_n, m + 2\varepsilon_n)) + H(x; max(M, m + 2\varepsilon_n), \alpha^{**} + \varepsilon_n))$$

(iii) if $m > \alpha^{**} - 3\varepsilon$ then

$H_1(x) = H(x; min(M, m - 2\varepsilon_n), \alpha^{**} + \varepsilon_n)$,

$H_2(x) = -H(x; \alpha^* - \varepsilon_{n,} \; min(M, m - 2\varepsilon_n)) \; + H(x; min(m - 2\varepsilon_{n,} \alpha^{**} + \varepsilon_n), \alpha^{**} + \varepsilon_n))$.

(2.7)

If $n = n_k + 1$ for some k then suppose first that $\alpha^{**} - \alpha^* \leq 10\varepsilon_n$. Then we define the functions $H_{j,} j = 1, 2$ as in (2.6).

If $\alpha^{**} - \alpha^* > 10\varepsilon_n$ then we have three cases:

(i)

$$m \leq \alpha^* + 3\varepsilon_n \; or \; m \geq \alpha^{**} - 3\varepsilon_n.$$

In both of these we define $H_{j,} j = 1, 2$ as in (2.6) (ii) and (iii) respectively.

(ii)

$$M - 2\varepsilon_n \leq m \leq M + 2\varepsilon_n.$$

In this case $H_{j,} j = 1, 2$ are defined as in (2.6) (i).

(iii)

a) $\alpha^* + 3\varepsilon_n < m < M - 2\varepsilon_n$ or
b) $M + 2\varepsilon_n < m < \alpha^{**} - 3\varepsilon_n$.

In this case we define three functions Hj, $j = 1, 2, 3$.

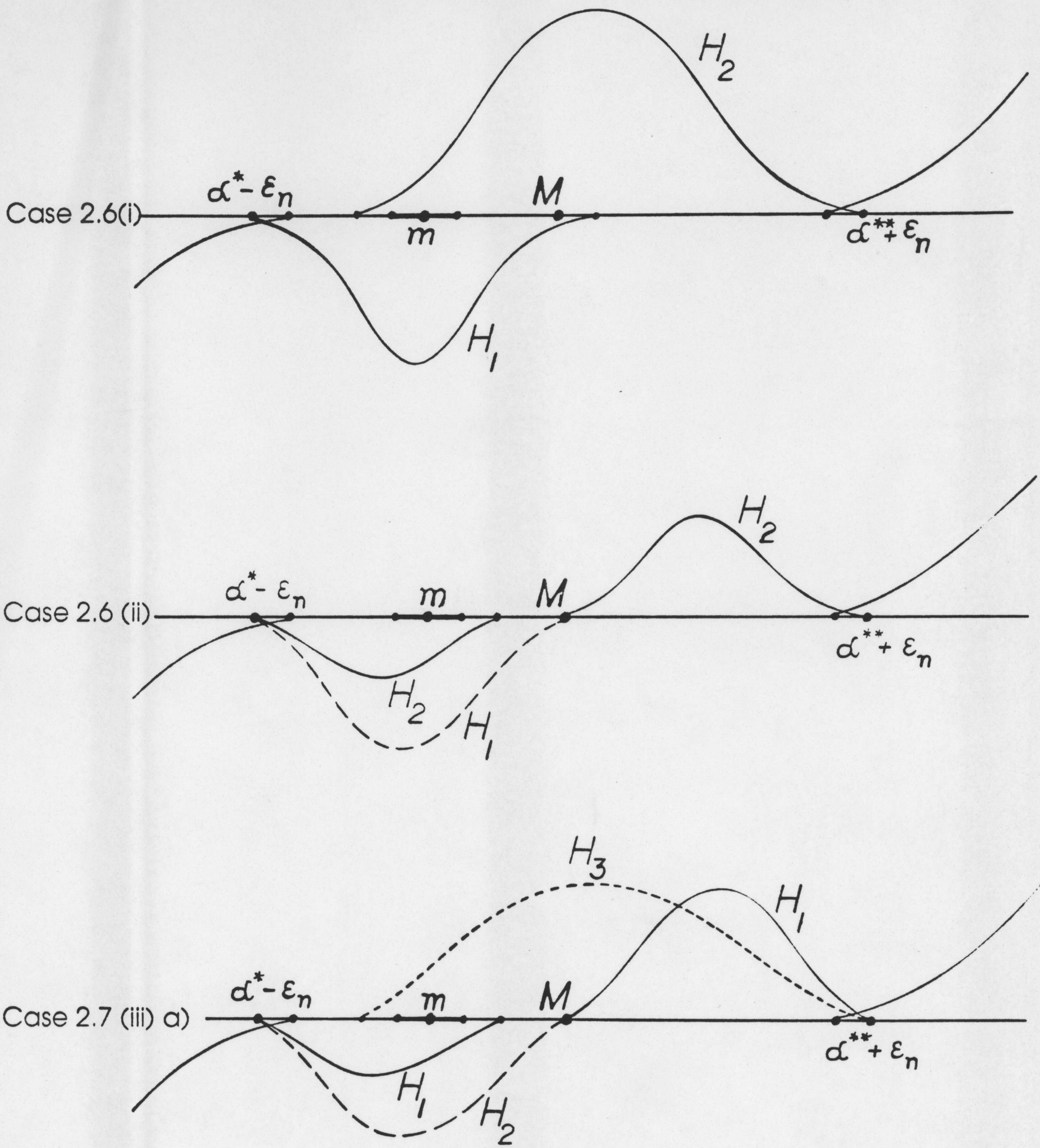
In the case a) we have

$$\begin{aligned} H_1(x) &= -H(x; \alpha^* - \varepsilon_n, m + 2\varepsilon_n) + H(x; M, \alpha^{**} + \varepsilon_n), \\ H_2(x) &= -H(x; \alpha^* - \varepsilon_n, M) \;, \\ H_3(x) &= H(x; m - 2\varepsilon_n, \alpha^{**} + \varepsilon_n) \;; \end{aligned}$$

In the case b) we have

$$\begin{aligned}
H_1,(x) &= -H(x;\alpha^* - \varepsilon_n, M) + H(x; m - 2\varepsilon_{n,}\alpha^{**} + \varepsilon_n),\\
H_2(x) &= H(x; M, \alpha^{**} + \varepsilon_n),\\
H_3(x) &= H(x; \alpha^* - \varepsilon_n, m + 2\varepsilon_n)
\end{aligned}$$

The functions H_1, H_2 and H_3 are illustrated in Figure 2.1



H_2
Case 2.6(i)
$\alpha^* - \varepsilon_n$
M
m
$\alpha^{**} + \varepsilon_n$
H_1
H_2
Case 2.6 (ii)
$\alpha^* - \varepsilon_n$
m
M
$\alpha^{**} + \varepsilon_n$
H_2
H_1
H_3
H_1
Case 2.7 (iii) a)
$\alpha^* - \varepsilon_n$
m
M
$\alpha^{**} + \varepsilon_n$
H_1
H_2

Figure 2.1

For any of the cases (2.6) or (2.7) let $\bar{H}_j \epsilon G(\bigcup_{i=1}^{n-1} I_i)$ be the solutions of

(2.8)

$$L_{i,f}(H_j + \bar{H}_j) = 0 \quad , i = 1, \ldots n-1, j\epsilon[1,2,3].$$

Such functions exist since the functionals $L_{i,f}$, $i = 1, \ldots n-1$, are linearly independent on G($\bigcup_{i=1}^{n-1} I_i$). Let

(2.9)

$$h_{j,n}(x) = c(H_j + \bar{H}_j),$$

where c is a positive constant so small that

(2.10)

$$\| h_{j,n} \|_n \leq 2^{-n} \ ,$$

and

(2.11)

$$\| h_{j,n} \|_0 \leq 2^{-n} \cdot \min_{x\epsilon[0,\alpha^* - \varepsilon_n] \cup [\alpha^{**} + \varepsilon_n, 1]} | f(x) |$$

We define

(2.12)

$$f_{i_j,n} = f + h_{j,n} \ , \quad \begin{array}{l} j\epsilon[1,2,3], \\ i_j\epsilon[1,3^n]. \end{array}$$

Note that $S(f_{i_j,n}) = S(H_j) \bigcap [\alpha^*, \alpha^{**}]$. This, (2.11) and the choice of H_j imply that (2.2) and (2.3) are satisfied.

The next lemma 2.3 characterizes more properties of the functions in F^ϵ :

Lemma 2.3
Let $f_{i,n}$ be an arbitrary function in F^ϵ as constructed in Lemma 2.2. Then:

(i) The length of the interval of zeros of $f_{i,n}$ is at most $(\frac{1}{2})^k$ for $n_k \leq n < n_{k+1}$, *i.e.*,

(2.13)

$$\mu(S(f_{i,n})) \leq (\frac{1}{2})^k \ for \ n_k \leq n < n_{k+1} ;$$

(ii) For every n the Lebesgue measure of the set $\bigcup_i S(f_{i,n})$ is at least $1 - D_n \varepsilon$, where

$$D_o = 2 \text{ and } D_n = D_{n-1} + 2\cdot(\frac{3}{4})^n ,$$

i.e.

2.14

$$\mu(\bigcup_i S(f_{i,n})) \geq 1 - D_n \cdot \varepsilon \ , \ \forall n.$$

This in particular implies that $D_n \leq 8$, *i.e.*,

that
(2.15)

$$\mu(\bigcup_i S(f_{i,n})) \geq 1 - 8 \cdot \varepsilon.$$

(iii) There exist infinite branches in the tree F^ε.

(iv) The functions in every infinite branch in F^ε form a Cauchy sequence in G.

(v) If $n > m$ and $f_{j,m}$ is a predecessor of $f_{i,n}$ in F^ε , *then* $N_m(f_{j,m}) = N_m(f_{i,n})$.

Proof:

The construction in Lemma 2.2 implies that for all n, if $f_{i,n}$ is a successor in F^ε to $f_{j,n-1}$ then $\mu(S(f_{i,n})) \leq \mu(S(f_{j,n-1}))$ and for $n = n_k + 1$ $\mu(S(f_{i,n})) \leq \frac{1}{2}\ \mu\ (S(f_{j,n-1}))$. This yields (2.13).

Now we prove (2.14). Let $f_{i_j,n}$ be the successors of $f_{l,n-1}$ in F^ε , $i_j \epsilon [1, 3^n], j \epsilon \{1, 2, 3\}$. Recall that any intersection of the sets $S(f_{i_j,n})$ has measure zero and that $S(f_{i_j,n}) \subset S(f_{l,n-1})$. The construction of F^ε yields

2.16

$$\mu(\bigcup_{j\epsilon\{1,2,3\}} S(f_{i_j,n})) = \sum_{j\epsilon\{1,2,3\}} \mu(S(f_{i_j,n})) \geq \mu(S(f_{l,n-1})) - 6\varepsilon_n, \forall_n.$$

We show (2.14) by simple induction. If n=0 then $S(f_{o,o}) = [\varepsilon, 1 - \varepsilon]$ and $\mu(S(f_{o,o})) = 1 - 2\varepsilon \geq 1 - D_o\varepsilon$.

Assume now that (2.14) holds for $n - 1, n \geq 1$.
Then

$$\mu(\bigcup_i S(f_{i,n})) = \sum_i \mu(S(f_{i,n})) \geq \sum_l (\mu(S(f_{l,n-1})) \ - 6\varepsilon_n$$

$$\geq \mu(\bigcup_l S(f_{l,n-1})) - 3^{n-1} \cdot 6\varepsilon_n \geq 1 - D_{n-1}.\varepsilon - 3^{n-1} 6\varepsilon_n$$

$$= 1 - (D_{n-1} + 2 \cdot (\frac{3}{4})^n) \cdot \varepsilon = 1 - D_n \cdot \varepsilon \,,$$

since $\varepsilon_n = \varepsilon \cdot 4^{-n}$ and the total number of functions on the n-1 -st level is at most 3^{n-1}.

This completes the proof.

By solving the recurrence relation for D_n one obtains $D_n = 8 - 6 \cdot (\frac{3}{4})^n$, i.e., $D_n \leq 8, \forall n$.

Now we show (iii). Suppose by contrary that all branches in F^ϵ are finite, i.e., that F^ϵ has at most n levels for some n. Then the tree F^ϵ has all leaves on at most n-th level. Recall that $f_{i,n}$ is a leaf, iff $\mu(S(f_{i,n})) \leq 6\varepsilon_n$.

This yields that

$$\mu(\bigcup_i S(f_{i,n})) \leq 3^n \cdot 6\varepsilon_n = 6 \cdot \varepsilon(\frac{3}{4})^n \leq 6\varepsilon < \frac{6}{14}.$$

But (2.15) implies $\mu(\bigcup_i S(f_{i,n})) \geq 1 - 8\varepsilon > 1 - \frac{8}{14} = \frac{6}{14}$ which contradicts our assumption.

Now we show (iv). Note first that if $f_{i,n}$ is a successor to $f_{j,m}$, $n > m$, *in* F^ϵ then the construction in Lemma 2.2 implies

(2.17)
$$f_{i,n} = f_{j,m} + \sum_{k=m+1}^{n} h_{\cdot,k} \,.$$

where $h_{\cdot,k}$ are the functions defined in (2.9) and the summation is taken along the branch of F^ϵ connecting $f_{j,m}$ to $f_{i,n}$.

Observe that (2.10) implies that $\| h_{\cdot,k} \|_l \leq 2^{-k}$ for any $0 < l \leq k$. Therefore

$$\| f_{i,n} - f_{j,m} \|_l \leq \sum_{k=m+1}^{M} \| h_{\cdot,k} \|_l \leq \sum_{k=m+1}^{M} 2^{-k} < 2^{-m},$$

for any $0 < l \leq m$.

Consequently

$$\rho(f_{i,n}, f_{j,m}) \leq \sum_{l=1}^{m} 2^{-l} \parallel f_{i,n} - f_{j,m} \parallel_l + \sum_{k=m+1}^{\infty} 2^{-k} \leq$$

$$\leq 2^{-m} \sum_{l=1}^{m} 2^{-l} + 2^{-m} \leq 2^{-m} + 2^{-m} = 2^{-(m-1)}.$$

Since $2^{-(m-1)}$ is arbitrarily small for large m, the proof is completed.

The point (v) of Lemma is an immediate consequence of (2.8) and (2.17). Indeed, letting $f = f_{j,m}$ in the construction of Lemma 2.2, the formula (2.8) implies that

$$L_{l,f}(h_{\cdot,k}) = 0 \ , l = 1, \ldots, m, k = m+1, \ldots n,$$

where $f_{i,n} = f_{j,m} + \sum_{k=m+1}^{n} h_{\cdot,k}$, as in (2.17).

Thus

$$L_{l,f}(f_{i,n}) = L_{l,f}(f_{j,m}) \ ,$$

i.e.,

$$N_m(f_{i,n}) = N_m(f_{j,m}).$$

This finally completes the proof of Lemma 2.3

Since G is a Frêchet space, then every Cauchy sequence in G is convergent. Therefore Lemma 2.3 implies that the following class of functions F_o is well defined:

(2.18)

$$F_o = \{f \epsilon G : \ f = \lim_n f_{\cdot,n}, f_{\cdot,n} \epsilon F^{\epsilon}\},$$

where $f_{\cdot,n}$ constitute the infinite branches in F^{ϵ}, and the limit is taken with respect to the ρ - metric in G.

In the next Lemma 2.4 we show that every function f *in* F_o has exactly one zero, and that $f(0) \leq 0$ and $f(1) \geq 0$. Moreover, we show that the set of zeros of all f from F_o has Lebesgue measure arbitrarily close to 1.

Lemma 2.4

The set F_o is a subset of F, i.e.,

(i) $F_o \subset F$;

The set of zeros of all functions f from F_0 has almost full measure for $\varepsilon \to 0$. More precisely,

(ii)

$$\mu(\bigcup_{f\epsilon F_0} S(f)) \geq 1 - 8 \cdot \varepsilon .$$

Proof:

We first show (i). Note that if $f \epsilon F_0$, i.e., $f = \lim_n f_{\cdot,n}$, then

$$\alpha_f := \bigcap_{n=0}^{\infty} S(f_{\cdot,n}) \subset S(f).$$

We will show that α_f is the only zero of f. This, combined with $f_{\cdot,n}(0) \leq 0$, and $f_{\cdot,n}(1) \geq 0$ implies (i).

Indeed, take any $\alpha \neq \alpha_f, \alpha\epsilon[0,1]$. *Since* $\mu(S(f_{\cdot,n})) \to 0$ as $n \to \infty$, see (2.13), then there exists an index $m \geq 1$, such that $\alpha \notin [\alpha^*_{f\cdot,n} - \varepsilon_n, \alpha^{**}_{f\cdot,n} + \varepsilon_n]$, for $n \geq m$.

Using (2.11) and (2.17) with $f_{i,n} = f$, *and* $f_{j,m} = f_{\cdot,m}(n = +\infty)$ we get

$$| f(\alpha) | = | f_{\cdot,m}(\alpha) + \sum_{k=m+1}^{\infty} h_{\cdot,k}(\alpha) | \geq$$

$$\geq | f_{\cdot,m}(\alpha) | - \sum_{k=m+1}^{\infty} | h_{\cdot,k}(\alpha) | \geq$$

$$\geq | f_{\cdot,m}(\alpha) | - \quad | f_{\cdot,m}(\alpha) | \cdot \sum_{k=m+1}^{\infty} 2^{-k} \geq$$

$$\geq | f_{\cdot,m}(\alpha) | (1 - 2^{-m}) > 0,$$

which completes the proof of (i).

Now we show (ii). Define

$$S_n := \bigcup_j S(f_{j,n}).$$

Then the set of zeros of all functions from F_0 *is* :

$$\bigcup_{f \epsilon F_\circ} S(f) = \bigcap_{n=0}^{\infty} S_n.$$

Observe that $S_{n+1} \subset S_n$.

This and 2.15 yield:

$$\mu(\bigcup_{f \epsilon F_\circ} S(f)) = \mu(\bigcap_{n=0}^{\infty} S_n) \geq \liminf_{n \to \infty} \mu(S_n) \geq 1 - 8 \cdot \varepsilon ,$$

which proves (ii).

<u>**Proof of the Theorem 1.1**</u>

To complete the proof of Theorem 1.1 we will show that for every $\varepsilon, 0 < \varepsilon < \frac{1}{14}$, **every sequence $\delta_n \searrow 0$, any** $N \epsilon \mathcal{N}$ and any $\phi \epsilon \phi(N)$ the measure $\mu(S(B) \cap S(F_\circ)) \geq 1 - 8\varepsilon$, *i.e.*,

(2.19)

$$1 \geq \mu(S(B)) \geq \mu(S(B) \cap S(F_\circ)) \geq 1 - 8\varepsilon,$$

where S(A) denotes the set of zeros of all functions from A and B is defined as in (1.10).

The proof is completed by taking $\varepsilon \to 0$ in (2.19).

To show (2.19) we need only to prove that $\mu(T) = 0, where\ T = \bigcup(S(f) : f \epsilon F_\circ\ and\ e_n(N, \phi, f) = 0(\delta_n \cdot 2^{-n}))$.

Indeed: μ(T) = 0 **and Lemma 2.4 (ii) imply that** μ (S(B) $\cap S(F_\circ)) = \mu(\bigcup(S(f) : f \epsilon F_\circ$ and $\limsup_{n \to \infty} \frac{e_n(N,\phi,f)}{\delta_n \cdot 2^{-n}} > 0)) = \mu(\bigcup(S(f) :\ f \epsilon F_\circ)) \geq 1 - 8\varepsilon$.

Now we concentrate on the proof of:

(2.20)

$$\mu(T) = 0.$$

Let $x_{i,n} = \phi_n(N_n(f_{i,n})), i = 1, \cdots c_n$, **for any function** $f_{i,n}$ **on the n-th level of** F^ε **from Lemma 2.2. Since the functionals in** N_n **are continuous, then Lemma 2.3 implies that** $x_{i,n} = \phi_n(N_n(f))$, **for any** $f \epsilon F_\circ$ **such that** $f_{i,n}$ **belongs to the branch** $\{f_{\cdot,n}\}$ **of** F^ε **with** $f = \lim_n f_{\cdot,n}$.

Let $\delta_n^1 = 2^{-n} c_n \cdot \delta_n$. **Observe that the definition (2.4) of the sequence** n_k **implies that for** $n_k < n \leq n_{k+1}$ we have

$$\delta_n^1 \leq 2^{-n} (\frac{3}{2})^k \cdot 2^n \cdot \delta_n = o(1),$$

i.e., δ_n^1 converges to zero, $\delta_n^1 \to 0$.

Let M be a positive interger and

$$V^M(x_{i,n}) = \{x\epsilon[0,1] : | x - x_{i,n} | \leq M \cdot \frac{\delta_n^1}{c_n} = M \cdot \delta_n \cdot 2^{-n}\}.$$

Define

$$V_n^M = \bigcup_{i=1}^{c_n} V^M(x_{i,n}) ,$$

$$T_m^M = \bigcap_{n=m}^{\infty} V_n^M \quad ,$$

and

$$T_\delta = \bigcup_{M=1}^{\infty} \bigcup_{m=1}^{\infty} T_m^M .$$

Observe that

(2.21)

$$T_\delta = \{x\epsilon[0,1] :| x - x_{k_n,n} |= 0(\delta_n 2^{-n})\}$$

where

$$x_{k_n,n} = \phi_n(N_n(f_{k_n,n})) \; and \; f_{k_n,n} \; forms$$

an arbitrary infinite branch in F^ϵ.

Indeed, let

$$A = \{x\epsilon[0,1] :| x - x_{k_n,n} |= 0(\delta_n 2^{-n})\}.$$

Take any $x\epsilon T_\delta$. Then $\exists M \; and \; m$ such that $x\epsilon T_m^M$. Thus, $x\epsilon \bigcap_{n=m}^{\infty} V_n^M; i.e. \; \forall \; n \geq m, | x - x_{k_n,n} | \leq M \cdot \delta_n 2^{-n}$, for some sequence $x_{k_n,n}$ along a branch of F^ϵ. Thus $x\epsilon A$. Conversely, if $x\epsilon A$ then $\exists \; m, M$, such that $| x - x_{k_n,n} | \leq M\delta_n 2^{-n} \; \forall n \geq$ m and some $x_{k_n,n}$. This implies that $x\epsilon V^M(x_{k_n,n}) \; for \; \forall n \geq m$, i.e., $x\epsilon \bigcap_{n=m}^{\infty} V_n^M$ which yields $x\epsilon T_m^M$ and $x\epsilon T_\delta$, and completes the proof of (2.21).

Observe now that

$$\mu(V^M(x_{i,n})) = 2M\delta_n^1/c_n.$$

Therefore

$$\mu(V_n^M) \leq \sum_{i=1}^{c_n} \mu(V^M(x_{i,n})) = 2 \cdot c_n \cdot M\delta_n^1 \cdot /c_n = 2M\delta_n^1,$$

and

$$\mu(T_m^M) \leq \inf_{n \geq m} \mu(V_n^M) = 0, \; since \; \delta_n^1 \rightarrow 0.$$

This yields $\mu(T_\delta) = 0$, since T_δ is a countable union of sets of measure zero.

Finally, since $T = \cup(S(f) : f \epsilon F_o \; and \; | x_{k_n,n} - S(f) | = 0(\delta_n 2^{-n}))$, where

$$x_{k_n,n} = \phi_n(N_n(f_{k_n,n}))$$

for some infinite branch $f_{k_n,n}$ of F^ϵ , then T is a subset of T_δ. Thus $\mu(T) = 0$, which completes the proof of (2.20).

Acknowledgements

We would like to thank the referee for the remarks on the manuscript.

References:

[1] Brent, R.P., *Algorithms for Minimization Without Derivatives*, Prentice Hall, Englewood Cliffs, 1973

[2] Sikorski, K., *Bisection is Optimal, Numer. Math. 40,* pp. 111-117, 1982

[3] Traub, J.F., *Iterative Methods for the Solution of Equations*, Prentice Hall, Englewood Cliffs, 1964

[4] Traub, J.F. and Wozniakowski, H., *A General Theory of Optimal Algorithms*, Academic Press, New York, 1980

[5] Traub, J.F., Wasilkowski, G. and Wozniakowski, H., *Information Based Complexity*, Academic Press, 1988

[6] Trojan, G.M., *Lower Bounds and Fast Algorithms for Sequence Acceleration*, J.ACM. 31, pp. 329-335, 1984