Journal of Applied
and Numerical Analysis

2023, Vol. 1, pp. 72–79
doi: 10.30970/ana.2023.1.72

In the search for all zeros of smooth functions

L. Plaskota

*Faculty of Mathematics, Informatics and Mechanics, University of Warsaw,
2, S. Banacha str., 02-097, Warsaw, Poland*

Анотація. Ми розглядаємо *багатозначну* задачу знаходження *усіх* розв'язків рівняння $f(x) = 0$ у просторі функцій $f : [0,1] \to \mathbb{R}$ таких, що похідна $f^{(r)}$ з $r \in \{0,1,2,\ldots\}$ існує і є неперервною за Гьольдером з показником $\varrho \in (0,1]$. Доступні алгоритми використовують інформацію про значення f та/або її похідних у адаптивно вибраних n точках і похибка між істинним розв'язком $Z(f)$ і наближеним розв'язком $Z_n(f)$ вимірюється відстанню Хаусдорфа $d_H\big(Z(f), Z_n(f)\big)$.

Ми показуємо, що незважаючи на те, що *найгірша похибка* будь-якого алгоритму є нескінченна, можна побудувати неадаптивні наближення Z_n^* так, що похибка $d_H\big(Z(f), Z_n^*(f)\big)$ збігається до нуля при $n \to +\infty$. Однак збіжність може бути як завгодно повільною. Зокрема, для довільної послідовності наближень $\{Z_n\}_{n\ge1}$, які використовують адаптивно вибрані n значення функції та/або її похідні, та для довільної додатньої послідовності $\{\tau_n\}_{n\ge1}$, що збігається до нуля, у нашому просторі існують такі функції f, що $\sup_{n\ge1} \tau_n^{-1} d_H\big(Z(f), Z_n(f)\big) = +\infty$.

Ми припускаємо, що та сама нижня межа має місце, якщо ми надаємо інформацію про значення n довільних та адаптивно вибраних лінійних функціоналів на f.

Abstract. We consider the *multi-valued* problem of finding *all* solutions of the equation $f(x) = 0$ in the space of functions $f : [0,1] \to \mathbb{R}$ such that the derivative $f^{(r)}$ with $r \in \{0,1,2,\ldots\}$ exists and is Hölder continuous with exponent $\varrho \in (0,1]$. Available algorithms use information about values of f and/or its derivatives at adaptively selected n points, and the error between the true solution $Z(f)$ and approximate solution $Z_n(f)$ is measured by the Hausdorff distance $d_H\big(Z(f), Z_n(f)\big)$.

We show that, despite the fact that the *worst case error* of any algorithm is infinite, it is possible to construct nonadaptive approximations Z_n^* such that the error $d_H\big(Z(f), Z_n^*(f)\big)$ converges to zero as $n \to +\infty$. However, the convergence can be arbitrarily slow. Specifically, for arbitrary sequence of approximations $\{Z_n\}_{n\ge1}$ that use n adaptively chosen function values and/or its derivatives, and for arbitrary positive sequence $\{\tau_n\}_{n\ge1}$ converging to zero there are functions f in our space such that $\sup_{n\ge1} \tau_n^{-1} d_H\big(Z(f), Z_n(f)\big) = +\infty$.

We conjecture that the same lower bound holds if we allow information about values of n arbitrary and adaptively selected linear functionals at f.

1 Preliminaries

Let

$$D := [0,1].$$

For $\gamma = r + \varrho$, where $r \in \{0,1,2,\ldots\}$ and $0 < \varrho \le 1$, let H_γ be the linear space of functions $f : D \to \mathbb{R}$ such that $f \in C^r(D)$ and $f^{(r)}$ is Hölder continuous with exponent ϱ, i.e.,

$$[f]_\gamma := \sup_{0\le x<y\le1} \frac{|f^{(r)}(x) - f^{(r)}(y)|}{|x-y|^\varrho} < +\infty.$$

(Note that $[f]_\gamma$ is a seminorm in H_γ.) For $f \in H_\gamma$, denote by $Z(f)$ the set of all zeros of f,

$$Z(f) := \big\{ z \in D \;\big|\; f(z) = 0 \big\}.$$

2020 *Mathematics Subject Classification*: 65H05, 65H20.
Key words: zero finding, Hausdorff distance, convergence rate.

We formally treat Z as a multi-valued mapping acting from H_γ to the power set of D, i.e., $Z : H_\gamma \to 2^D$. Our problem is to recover $Z(f)$ based on (exact) evaluations of f and/or its derivatives,

$$N_n f = \big(f^{(s_1)}(x_1), f^{(s_2)}(x_2), \dots, f^{(s_n)}(x_n)\big),$$

where the points x_i and the degrees s_i of the derivatives are selected adaptively. That is, for $i = 1, 2, \dots, n$ we have

$$x_i = \alpha_i\big(f^{(s_1)}(x_1), \dots, f^{(s_{i-1})}(x_{i-1})\big) \quad \text{and} \quad s_i = \beta_i\big(f^{(s_1)}(x_1), \dots, f^{(s_{i-1})}(x_{i-1})\big),$$

for arbitrary functions $\alpha_i : \mathbb{R}^{i-1} \to [0,1]$ and $\beta_i : \mathbb{R}^{i-1} \to \{0, 1, \dots, r\}$. Then the nth approximation of $Z(f)$ is given as

$$Z_n(f) := \Phi_n(N_n f), \tag{1.1}$$

where

$$\Phi_n : \mathbb{R}^n \to 2^D$$

is an arbitrary multi-valued mapping. The error of approximating $Z(f)$ by $Z_n(f)$ equals

$$d_H\big(Z(f), Z_n(f)\big),$$

where d_H is the *Hausdorff distance* of two sets in a metric space (X, d),

$$d_H(W, Z) := \max\left\{ \sup_{w \in W} \inf_{z \in Z} d(z, w),\ \sup_{z \in Z} \inf_{w \in W} d(w, z) \right\}.$$

For our problem we obviously have $X = D$ and $d(x, y) = |x - y|$. Here we use the convention that $d_H(\emptyset, \emptyset) = 0$, and $d_H(W, \emptyset) = +\infty$ for a nonempty W.

In the space H_γ, we distinguish functions with exactly one zero,

$$\widehat{H}_\gamma := \big\{ f \in H_\gamma \,\big|\, \#Z(f) = 1 \big\}.$$

Remark 1.1. *Observe that if we know that an underlying function f has exactly one zero, i.e., $f \in \widehat{H}_\gamma$, then we can restrict our considerations to single-valued approximations. Indeed, denote by $z(f)$ the only zero of f. For a given multi-valued approximation Z_n, define a single-valued approximation Z'_n as $Z'_n(f) = \{z_n(f)\}$, where $z_n(f)$ is any element of $Z_n(f)$. Then*

$$d_H\big(Z_n(f), Z(f)\big) = d_H\big(Z_n(f), \{z(f)\}\big) \geq |z_n(f) - z(f)| = d_H\big(Z'_n(f), Z(f)\big).$$

We first show a negative result concerning a worst case setting. (See, e.g., [6–8] for a general introduction to the worst case setting.) Recall that the worst case error of an approximation Z_n with respect to a class $\mathcal{H} \subset H_\gamma$, is given as

$$e^{\mathrm{wor}}(\mathcal{H}, Z_n) := \sup_{f \in \mathcal{H}} d_H\big(Z(f), Z_n(f)\big).$$

For $M > 0$, we correspondingly define the classes

$$\begin{aligned} \mathcal{H}_\gamma(M) &:= \big\{ f \in H_\gamma \,\big|\, [f]_\gamma \leq M \big\}, \\ \widehat{\mathcal{H}}_\gamma(M) &:= \big\{ f \in \widehat{H}_\gamma \,\big|\, [f]_\gamma \leq M \big\}. \end{aligned}$$

Observe that for all $0 < \varrho_1 \leq \varrho_2 \leq 1$ we have

$$\mathcal{H}_{r+\varrho_1}(M) \supseteq \mathcal{H}_{r+\varrho_2}(M) \supseteq \mathcal{H}_{r+1}(M), \tag{1.2}$$

and we have analogous inequalities for $\widehat{\mathcal{H}}_{r+\varrho}(M)$.

Proposition 1.1. *(cf. [5]). For any approximation Z_n of the form* (1.1) *we have*

$$e^{\mathrm{wor}}\big(\mathcal{H}_\gamma(M), Z_n\big) = +\infty \qquad \text{and} \qquad e^{\mathrm{wor}}\big(\widehat{\mathcal{H}}_\gamma(M), Z_n\big) \ge 1/2.$$

Proof. In view of (1.2), it suffices to consider the case $\varrho = 1$. Let us fix a function

$$\Psi \in C^r(\mathbb{R}) \tag{1.3}$$

that is supported on $(0,1)$ and unimodal in this interval with the maximum at $1/2$, and $\Psi|_{[0,1]} \in \mathcal{H}_{r+1}(M)$. (This function will also be used later in the proof of Theorem 2.3.)

Consider the class $\mathcal{H}_{r+1}(M)$. Let $\delta = 1/(n+1)$ and $\epsilon = \Psi(1/2)\delta^{r+1}$. Suppose that for given (adaptive in general) information consisting of evaluations at n points we have $f(x_i) = \epsilon$ and $f^{(s)}(x_i) = 0$ for all $1 \le s \le n$ and $i = 1, 2, \ldots n$. (It will be clear in a moment that such information is possible for some functions $f \in \mathcal{H}_{r+1}(M)$.) We select an interval $(a, a+\delta)$ that does not contain any of the points x_i and define two functions: f_1 is the constant function equal to ϵ, and

$$f_2(x) = \epsilon - \psi_{a,\delta}(x) \quad \text{with} \quad \psi_{a,\delta}(x) = \Psi\left(\frac{x-a}{\delta}\right)\delta^{r+1}.$$

We have that

$$\left|f_2^{(r)}(x) - f_2^{(r)}(y)\right| = \delta\left|\Psi^{(r)}\left(\frac{x-a}{\delta}\right) - \Psi^{(r)}\left(\frac{y-a}{\delta}\right)\right| \le M\,|x-y|,$$

so that $f_1, f_2 \in \mathcal{H}_{r+1}(M)$. Moreover, these functions share the assumed information, i.e., $f_1(x_i) = \epsilon = f_2(x_i)$ and $f_1^{(s)}(x_i) = 0 = f_2^{(s)}(x_i)$ for $1 \le s \le r$. But $Z(f_1)$ is the empty set, while $Z(f_2) = \{a + \delta/2\} \ne \emptyset$. This means that either $d_H(Z(f_1), Z_n(f_1)) = +\infty$ or $d_H(Z(f_2), Z_n(f_2)) = +\infty$, as claimed.

For the class $\widehat{\mathcal{H}}_{r+1}(M)$, we proceed similarly. We let $\delta = h/(n+1)$ and $\epsilon = \Psi(1/2)\delta^{r+1}$ with some $0 < h < 1/2$, and select two intervals, $(a_1, a_1+\delta) \subset [0,h]$ and $(a_2, a_2+\delta) \subset [1-h, 1]$, that do not contain any of the points x_i. Then the two functions $f_1(x) = \epsilon - \psi_{a_1,\delta}(x)$ and $f_2(x) = \epsilon - \psi_{a_2,\delta}(x)$ share the same information, both are in $\widehat{\mathcal{H}}_{r+1}(M)$, and the distance between their zeros equals $(a_2 + \delta/2) - (a_1 + \delta/2) = a_2 - a_1 \ge 1 - 2h$. This means that the error of approximation is at least $1/2 - h$. Since h can be arbitrarily small, the proposition follows. □

We add that the lower bound $1/2$ for the class $\widehat{\mathcal{H}}_\gamma$ in Proposition 1.1 is achieved by the constant approximation $Z_n(f) = \{1/2\}$ for all f.

Remark 1.2. *Proposition 1.1 shows that smoothness only is not enough to get convergence to the solution, and some additional assumptions on the function class are necessary. For instance, it was shown in [3] that the bisection method is optimal in the worst case setting for the class of continuous functions f having exactly one zero and satisfying $f(0) \le 0 \le f(1)$, where the minimum error equals $2^{-(n+1)}$. Moreover, this result was proven under the assumption that evaluations of n arbitrary and adaptively selected linear functionals at f are allowed.*

2 Asymptotic setting

Because of the negative result of Proposition 1.1 concerning the worst case setting, we switch to an asymptotic setting. That is, we are interested in the behavior of the errors $d_H\big(Z(f), Z_n(f)\big)$ as $n \to +\infty$, where $Z_n(f) = \Phi_n(N_n f)$ is as before an nth approximation to $Z(f)$. In this case, we assume that information is nested, which means that

$$N_{n+1}f = \big(N_n f, f^{(s_n)}(x_n)\big) \tag{2.1}$$

with x_n and s_n selected adaptively depending on the value of $N_n f$.

We separately consider the upper and the lower bounds.

2.1 Upper bound

We first construct approximations Z_n^M for functions in the class $\mathcal{H}_\varrho(M)$, i.e., for $r=0$. Let

$$x_1^*,\, x_2^*,\, x_3^*,\, x_4^*,\, \dots \tag{2.2}$$

be an infinite sequence of pairwise different points that are dense in D, such as, for instance,

$$0,\, 1,\, \tfrac{1}{2},\, \tfrac{1}{4},\, \tfrac{3}{4},\, \tfrac{1}{8},\, \tfrac{3}{8},\, \tfrac{5}{8},\, \tfrac{7}{8},\, \tfrac{1}{16},\, \dots .$$

The approximations $Z_n^M(f)$ use evaluations of f at the n first points of that sequence, i.e.,

$$N_n f = \big(f(x_1^*), f(x_2^*), \dots, f(x_n^*)\big),$$

and are defined as follows. For given information $y_i = f(x_i^*)$, $1 \le i \le n$, denote by f_-^M and f_+^M respectively the lower envelope and upper envelope corresponding to this information, i.e.,

$$f_-^M(x) := \max_{1\le i\le n} \left\{ f(x_i^*) - M|x - x_i^*|^\varrho \right\}, \qquad f_+^M(x) := \min_{1\le i\le n} \left\{ f(x_i^*) + M|x - x_i^*|^\varrho \right\}.$$

Then

$$Z_n^M(f) := \left\{ x \in D \,\middle|\, f_-^M(x) \le 0 \le f_+^M(x) \right\}.$$

We stress that the approximations Z_n^M use nonadaptive information.

We have that $f_-^M \le f \le f_+^M$ and therefore

$$Z(f) \subseteq Z_n^M(f)$$

for all $f \in \mathcal{H}_\varrho(M)$. Furthermore, the sets $Z_n^M(f)$ are compact and

$$Z_n^M(f) \supseteq Z_{n+1}^M(f).$$

Lemma 2.1. *For all $f \in \mathcal{H}_\varrho(M)$ we have*

$$\lim_{n\to+\infty} d_H\big(Z(f), Z_n^M(f)\big) \;=\; 0.$$

Proof. If for some n is $Z_n^M(f) = \emptyset$ then also $Z(f) = \emptyset$ and $d_H(Z(f), Z_n^M(f)) = d_H(\emptyset,\emptyset) = 0$. Suppose that $Z_n^M(f)$ is nonempty for all n. Since $Z_n^M(f)$ is closed, the limiting set

$$Z^M(f) := \bigcap_{n=1}^{+\infty} Z_n^M(f)$$

is nonempty. Due to Lemma 2.4 of the Appendix, it suffices to show that $Z^M(f) = Z(f)$. Indeed, suppose that $z \in Z^M(f) \setminus Z(f)$. Without loss of generality we can assume that $f(z)$ is positive. For sufficiently small positive δ,

$$0 < \delta < \left(\frac{f(z)}{M}\right)^{1/\varrho},$$

the function f is positive in $[z-\delta, z+\delta]$. Then, due to density of the sampling points (2.2) in D, for all sufficiently large n, the lower envelopes f_-^M are also positive in that interval, which means that $z \notin Z^M(f)$. □

We now modify the approximations $Z_n^M(f)$ to obtain approximations $Z_n^*(f)$ that converge to $Z(f)$ for all $f \in H_\gamma$ with $\gamma > 0$. These are defined for $n \ge 2$ as follows. Let

$$M_n(f) := \max_{1\le i<j\le n} \frac{|f(x_i^*) - f(x_j^*)|}{|x_i^* - x_j^*|^\varrho} \qquad \text{if} \quad \gamma < 1,$$

and

$$M_n(f) := \max_{2\le i\le n} \frac{|f(x_i^*) - f(x_{i-1}^*)|}{|x_i^* - x_{i-1}^*|} \qquad \text{if} \quad \gamma \ge 1.$$

Then

$$Z_n^*(f) := Z_n^{2M_n(f)}(f).$$

Theorem 2.2. *For all* $f \in H_\gamma$ *we have*

$$\lim_{n\to+\infty} d_H\big(Z(f), Z_n^*(f)\big) \;=\; 0.$$

Proof. Consider first the case of $r = 0$, i.e., $\gamma = \varrho$. By continuity of f, for all sufficiently large n we have $[f]_\varrho \le 2M_n(f) \le 2[f]_\varrho$. This means that

$$Z(f) \subseteq Z_n^{[f]_\varrho} \subseteq Z_n^*(f) \subseteq Z_n^{2[f]_\varrho}$$

and

$$d_H\big(Z(f), Z_n^*(f)\big) \le d_H\big(Z(f), Z_n^{2[f]_\varrho}\big).$$

Since $f \in \mathcal{H}_\varrho\big(2[f]_\varrho\big)$, in view of Lemma 2.1 we conclude that $d_H\big(Z(f), Z_n^*(f)\big)$ goes to zero as $n \to +\infty$.

In the case of $r \ge 1$ the theorem follows from that fact that H_γ is a subset of H_1. □

2.2 Lower bound

To prove a lower bound in the asymptotic setting for a given sequence of approximations, one has to construct functions for which the error converges to the solution 'sufficiently slowly'. The first result of this kind was presented in [9] where a kind of equivalence of the worst case and asymptotic settings for solving linear problems was established. Those results were then generalized in various directions, see, e.g., [1, 2, 10].

We now show a lower bound in the spirit of [9] for our problem of zeros finding. Specifically, we show that the nonadaptive approximations Z_n^* constructed in Section 2.1 are optimal in the asymptotic setting, meaning that for any other approximations the convergence is arbitrarily slow.

Theorem 2.3. *Let* $\{Z_n\}_{n\ge1}$ *be an arbitrary sequence of approximations of the form* (1.1) *that use nested information* (2.1)*. Then for any* $M > 0$ *and any positive real sequence* $\{\tau_n\}_{n\ge1}$ *converging to zero there exists* $f^* \in \widehat{\mathcal{H}}_\gamma(M)$ *such that*

$$\sup_{n\ge1} \tau_n^{-1}\, d_H\big(Z(f^*), Z_n(f^*)\big) \;=\; +\infty. \tag{2.3}$$

Proof. By (1.2) it suffices to find $f^* \in \widehat{H}_{r+1}(M)$ satisfying (2.3). Moreover, in view of Remark 1.1, we can restrict our considerations to single-valued approximations only. Then we denote by z_n the single-valued mapping corresponding to Z_n, i.e., $Z_n(f) = \{z_n(f)\}$.

Suppose that the approximations z_n use (adaptive) information N_n. We shall construct an infinite sequences of indexes $0 = n_0 < n_1 < n_2 < \cdots$, intervals

$$[0,1] = [a_0, b_0] \supseteq [a_1, b_1] \supseteq [a_2, b_2] \supseteq \cdots \quad \text{with} \quad 0 < b_k - a_k \le 2^{-k},$$

and functions

$$0 \;=\; f_0 \;\le\; f_1 \;\le\; f_2 \;\le\; \cdots$$

that possess the following properties for all $k \ge 1$:

(i) $N_{n_k}(f_{k+1}) = N_{n_k}(f_k)$,
(ii) $Z(f_k) = [a_k, b_k]$, and this set has empty intersection with the closed ball $B\big(z_{n_k}(f_k), \sqrt{\tau_{n_k}}\big)$.

In the construction we use induction on k. The base case $k=0$ is done. For the induction step, suppose that we have constructed n_k, $[a_k,b_k]$ and f_k satisfying (i) and (ii) for some $k\ge 0$. Then n_{k+1} is such that

$$\sqrt{\tau_{n_k+1}}\le \tfrac{1}{6}(b_k-a_k).$$

The function f_{k+1} is constructed by successively adding to f_k finitely many 'bump' functions of small support as follows. Using another induction, we first construct (locally for given k) functions $g_0,g_1,\dots,g_{n_{k+1}-n_k}$. We let $g_0=f_k$. For the induction step, let $x_1,x_2,\dots,x_{n_k+s}$ be the points at which values of g_s or its derivatives are evaluated, and $x_{n_k+s+1}=\alpha_{n_k+s+1}(N_{n_k+s}g_s)$. If $g_s(x_{n_k+s+1})>0$ then $g_{s+1}=g_s$. Otherwise g_{s+1} is produced by adding to g_s a 'bump' function

$$\psi_{x_{n_k+s+1},\delta}(x)=\frac{1}{2}\Psi\left(\frac{x-x_{n_k+s+1}}{\delta_s}\right)\delta_s^{r+1},$$

where Ψ is as in (1.3) and δ_s is such that $0<\delta_s\le(b_k-a_k)/4^s$ and the closed ball $B(x_{n_k+s+1},\delta_s)$ does not contain any of the points x_i for $1\le i\le n_k+s$. (The last assumption assures that the values of g_s and g_{s+1} coincide at all such points.)

Denote $g:=g_{n_k+1}-g_{n_k}$. Then $[a_{k+1},b_{k+1}]$ is defined as an arbitrary nontrivial interval of length at most $2^{-(k+1)}$ that is contained in $Z(g)$ and has empty intersection with the ball $B\big(z_{n_k+1}(f),\sqrt{\tau_{n_k+1}}\big)$. Such an interval exists since the diameter of the ball is at most $\frac{1}{3}(b_k-a_k)$ and $Z(g)$ is a sum of finitely many mutually disjoint closed intervals of total length at least $\frac{2}{3}(b_k-a_k)$.

The set $Z(g)\setminus(a_{k+1},b_{k+1})$ still consists of a sum of finitely many and mutually disjoint intervals, say $[c_j,d_j]$ for $1\le j\le m$. To complete the construction, we let

$$f_{k+1}=g+\sum_{j=1}^{m}\varphi_j,$$

where

$$\varphi_j(x)=\frac{1}{2}\Psi\left(\frac{x-y_j}{h_j}\right)h_j^{r+1},\quad 1\le j\le m,$$

with $y_j=(c_j+d_j)/2$ and h_j such that each ball $B(y_j,h_j)$ does not contain any of the points x_i for $1\le i\le n_{k+1}$, and their sum provides a covering of the set $Z(g)\setminus[a_{k+1},b_{k+1}]$ in such a way that $Z(f_{k+1})=[a_{k+1},b_{k+1}]$.

We claim that the desired function f^* is given as

$$f^*(x)=\lim_{k\to+\infty}f_k(x).$$

To show this, we first note that f^* is well defined as it is an infinite sum of 'bump' functions, but for each x only at most two of these functions do not vanish at x. Since all the 'bump' functions are in $\mathcal{H}_{r+1}(M/2)$, we have that $f^*\in\mathcal{H}_{r+1}(M)$. Furthermore, f^* has the only zero at $z(f^*)$ such that

$$\big\{z(f^*)\big\}=\bigcap_{k=0}^{+\infty}[a_k,b_k],$$

and all the derivatives of f^* up to rth also nullify at $z(f^*)$. Hence $f^*\in\widehat{\mathcal{H}}_{r+1}(M)$.

Now, by (i) we have $N_{n_k}(f^*)=N_{n_k}(f_{n_k})$, and therefore $Z_{n_k}(f^*)=Z_{n_k}(f_k)$. Since $z(f^*)\in[a_k,b_k]$, using (ii) we obtain that for $k\ge 1$

$$d_H\big(Z(f^*),Z_{n_k}(f^*)\big)=|z(f^*)-z_{n_k}(f^*)|=|z(f^*)-z_{n_k}(f_k)|\ge\sqrt{\tau_{n_k}}.$$

This implies (2.3) and completes the proof. □

Remark 2.1. *The lower bound of Theorem 2.3 shows that smoothness only is not enough to get satisfactory asymptotic convergence to the solution for all functions in the class. It is worthwhile to mention the result of [4] where it was shown that, similarly to the worst case setting, the convergence rate* $2^{-(n+1)}$ *given by the bisection method cannot be essentially beaten in the asymptotic setting for the class of continuous functions having exactly one zero and satisfying* $f(0) \le 0 \le f(x)$*. This was proven under the assumption that evaluations of* n *arbitrary and adaptively selected linear functionals at* f *are allowed. We conjecture a similar result for our problem, that the lower bound of Theorem 2.3 holds true even when information about values of* n *arbitrary and adaptively selected linear functionals at* f *is allowed.*

Appendix

Lemma 2.4. *Let* Z_n *for* $n \ge 1$ *be nonempty compact sets in a normed space,*

$$Z_1 \supseteq Z_2 \supseteq \cdots \supseteq Z_n \supseteq \cdots .$$

Let $Z = \bigcap_{n=1}^{+\infty} Z_n$*. Then for any set* W *we have*

$$\lim_{n\to+\infty} d_H(W, Z_n) = d_H(W, Z).$$

Proof. We first show that

$$\lim_{n\to+\infty} d_H(Z, Z_n) = 0.$$

Indeed, if this limit equals $\epsilon > 0$ then for all n we have $d_H(Z, Z_n) \ge \epsilon$. Equivalently, there are $x_n \in Z_n$ such that $\|x_n - z\| \ge \epsilon$ for all $z \in Z$. Let x^* be an accumulation point of the sequence $\{x_n\}_{n\ge1}$ (it exists due to compactness of Z_1), and $x^* = \lim_{k\to+\infty} x_{n_k}$. Then, by continuity, $\|x^* - z\| \ge \epsilon$ for all $z \in Z$, so that $x^* \notin Z$. On the other hand, since $x_n \in Z_j$ for all $n \ge j$, we have that x^* is a member of all the sets Z_n, and consequently $x^* \in Z$.

Now the lemma follows from the inequalities

$$d_H(W, Z_n) - d_H(Z, Z_n) \le d_H(W, Z) \le d_H(Z, Z_n) + d_H(W, Z_n)$$

by letting $n \to +\infty$. □

References

1. Kacewicz B.Z.: Asymptotic error of algorithms for solving nonlinear problems. Journal of Complexity. **3**, 41-56 (1987)
2. Kacewicz, B.Z., Plaskota, L.: Noisy information for linear problems in the asymptotic setting. Journal of Complexity. **7**, 35-57 (1991)
3. Sikorski, K.: Bisection is optimal. Numerische Mathematik. **40**, 111-117 (1982)
4. Sikorski, K., Trojan, G.M.: Asymptotic near optimality of the bisection method. Numerische Mathematik. **57**, 421-433 (1990)
5. Sikorski, K., Woźniakowski, H.: For which error criteria can we solve nonlinear equations? Journal of Complexity. **2**, 163-178 (1986)
6. Traub, J.F., Wasilkowski, G.W., Woźniakowski, H.: Information, Uncertainty, Complexity, Academic Press (1983)
7. Traub, J.F., Wasilkowski, G.W., Woźniakowski, H.: Information-Based Complexity, Academic Press (1988)
8. Traub, J.F., Woźniakowski, H.: A General Theory of Optimal Algorithms. Academic Press (1980)

9. Trojan, G.M.: Asymptotic setting for linear problems. Unpublished manuscript (1983). See also pp. 338-400 Traub et al. (1988)
10. Wasilkowski, G.W., Woźniakowski, H.: On optimal algorithms in an asymptotic model with Gaussian measure. SIAM Journal on Mathematical Analysis. **19**, 632-647 (1988)

Received 17.08.2023
Revised 22.09.2023