

Algorithms for Multi-Extremal Mathematical Programming Problems Employing the Set of Joint Space-Filling Curves

ROMAN G. STRONGIN

Software Department, University of Gorky, Gagarin Ave., 23, Gorky 603600, Russia

(Received: 15 October 1991)

Abstract. Some powerful algorithms for multi-extremal non-convex-constrained optimization problems are based on reducing these multi-dimensional problems to those of one dimension by applying Peano-type space-filling curves mapping a unit interval on the real axis onto a multi-dimensional hypercube. Here is presented and substantiated a new scheme simultaneously employing several joint Peano-type scannings which conducts the property of nearness of points in many dimensions to a property of nearness of pre-images of these points in one dimension significantly better than in the case of a scheme with a single space-filling curve. Sufficient conditions of global convergence for the new scheme are investigated.

Key words. Non-convex programming, space-filling curves.

1. Introduction

One of the recent approaches to global optimization (see, e.g., Butz (1968), Strongin (1973)) is based on the idea of reducing the initial multi-dimensional multi-extremal optimization problems to equivalent one-dimensional problems using classical Peano-space-filling curves (some alternative approaches are viewed, for example, in Evtushenko (1985), Horst (1990), Korotkich (1989), Pinter (1988), Rinnooy Kan and Timmer (1989)). Numerical methods for the arising problem of computing this type of map are described in Butz (1971), Strongin (1978, 1990) and in Gergel *et al.* (1987). Optimization schemes for one-dimensional problems generated by this approach are suggested and investigated in Strongin (1973, 1978, 1984, 1989, 1990) and in Strongin and Markin (1986, 1987). These algorithms are substantially global and do not make use of penalties in treating constraints (each constraint is provided with a separate account).

Some disadvantage of this approach is in the fact that one-dimensional problem obtained by the above reduction leaks some information on the closeness of iteration points in the initial multi-dimensional space (two close images in the multi-dimensional hypercube may have substantially non-close pre-images in the

*This paper was presented at the II. IIASA Workshop on Global Optimization, Sopron (Hungary), December 9–14, 1990.

real axis interval mapped onto the hypercube by the space-filling curve). To overcome this obstacle, it was suggested to store all pre-images for each iteration point (close images will always have some close pre-images). But, in this case, the amount of pre-images to be stored grows exponentially with the increase of problem dimensionality (see Strongin (1978)).

Here is introduced a new scheme which allows us to reflect, in the case of the reduced one-dimensional problem, some essential information on the nearness of iteration points in the initial multi-dimensional domain. In this scheme, the amount of additional information for each iteration point does not depend on the dimensionality. Primarily, the ideas of this scheme were tested in pattern-recognition problems (see Gergel *et al.* (1987)).

2. Algorithm with a Single Scanning

PROBLEM. Let us consider the N -dimensional problem

$$\min\{\varphi(y) : y \in D, \quad g_i(y) \leq 0, \quad 1 \leq i \leq m\}, \quad (2.1)$$

where the *domain of search*

$$D = \{y \in \mathbb{R}^N : -2^{-1} \leq y_j \leq 2^{-1}, \quad 1 \leq j \leq N\}, \quad (2.2)$$

\mathbb{R}^N is the N -dimensional Euclidean space and the *objective* function $\varphi(y)$ to be minimized (henceforth denoted by $g_{m+1}(y)$) and the left-hand sides $g_i(y)$, $1 \leq i \leq m$, of the constraints satisfy Lipschitz conditions (with respective constants L_i , $1 \leq i \leq m+1$) and may be multi-extremal. It is admitted that functions $g_i(y)$ are defined and computable only at the points $y \in D$ satisfying the conditions

$$g_k(y) \leq 0, \quad 1 \leq k < i. \quad (2.3)$$

This property will be referred to as the *partial computability* of problem functionals.

The *feasible domain* associated with the problem (2.1) may be presented in a more unified form by introducing the additional constraint

$$g_0(y) \leq 0 \quad (2.4)$$

with the left-hand side

$$g_0(y) = \max\{|y_j| - 2^{-1} : 1 \leq j \leq N\}. \quad (2.5)$$

This unification is a provision for simplicity of notation in the following definition.

DEFINITION: A point y_ε is said to be an ε -reserved solution to the problem (2.1) if

$$\varphi(y_\varepsilon) = \min\{\varphi(y) : y \in \mathbb{R}^N, \quad g_i(y) \leq -\varepsilon_i, \quad 0 \leq i \leq m\}, \quad (2.6)$$

where $g_0(y)$ is from (2.5) and $\varepsilon = (\varepsilon_0, \dots, \varepsilon_m)$ is a vector with positive coordinates ("reserves" for each of the corresponding constraints). We also introduce a

set

$$Y_\varepsilon = \{y \in \mathbb{R}^N : g_i(y) \leq 0, \quad 0 \leq i \leq m, \quad \varphi(y) \leq \varphi(y_\varepsilon)\} \tag{2.7}$$

of admissible points for the problem (2.1) which are no worse (in the value of the objective function φ) than the ε -reserved solution. This set contains a *point of attraction* for the algorithm we suggest below.

The existence of an ε -reserved solution to the problem (2.1) plays (see Section 4) the role of the *regularity condition* somewhat similar to other forms of regularity requirements in classical nonlinear programming problems (see, e.g., Kuhn and Tucker (1951)).

SPACE-FILLING CURVES. To introduce some necessary concepts and notation, we provide a brief description of Peano-type maps. Without loss of generality, we shall conduct this consideration following the scheme suggested by Hilbert.

Let us divide the hypercube D from (2.2) into 2^N equal hypercubes of "first-partition" by cutting D with the set of N mutually orthogonal hyperplanes (each plane is parallel to one of the coordinate ones and passes through the middle points of D edges orthogonal to this hyperplane). Then we divide (in the above manner) each of the obtained first-partition hypercubes into 2^N second-partition hypercubes. Continuing this process, i.e., consequently cutting each hypercube of a current partition into 2^N hypercubes of the subsequent partition (with twice shorter edge length), we shall obtain hypercubes of any M th partition with the edge-length equal to 2^{-M} . The total number of subcubes of the M th partition is equal to 2^{MN} .

Next, we cut the interval $[0, 1]$ on the x -axis into 2^N equal parts. Then, once again, we cut each of these parts into 2^N smaller (equal) parts, etc. The subinterval of the M th partition is designated $d(M, v)$, where v is the coordinate of the left end-point of this subinterval. The length of $d(M, v)$ is equal to 2^{-MN} . We assume that $v \in d(M, v)$, but the right end-point of this subinterval does not belong to it (with the exception: if the right end-point is equal to 1, then it belongs to this subinterval).

Now, we establish a mutually single-valued correspondence between all subintervals of any particular M th partition and all subcubes of the same M th partition and, henceforth, notation $D(M, v)$ will stay for the subcube corresponding to the subinterval $d(M, v)$ and vice versa. We demand this correspondence to satisfy the following requirements:

CONDITION 1. $D(M + 1, v') \in D(M, v'')$ if and only if $d(M + 1, v') \in d(M, v'')$.

CONDITION 2. Two subintervals $d(M, v')$ and $d(M, v'')$ have a common end-point (this point may only be either v' or v'') if and only if the corresponding subcubes $D(M, v')$ and $D(M, v'')$ have a common face (i.e., these subcubes must be contiguous).

A constructive way of establishing such a correspondence is described and substantiated in Strongin (1978, 1990) and in Gergel and Strongin (1987), but the details of this scheme are not essential for the consideration we carry out here (as far as it is already proved that both of the above conditions could somehow be met). For the accepted correspondence of subintervals and subcubes we stipulate the third requirement:

CONDITION 3. *If $x \in d(M, v)$ then $y(x) \in D(M, v)$, $M \geq 1$.*

This last condition defines a single-valued continuous map $y(x)$ (note, that for any given integer $M \geq 1$ and for any given point $x \in [0, 1]$ there is just a single subinterval of the M th partition containing x ; the continuity is a consequence of Conditions 1 and 2). The center $y^c(x)$ of the subcube of the M th partition containing the point $y(x)$ may be interpreted as a discrete approximation to $y(x)$ and, in this case, the inequality

$$\max\{|y_j^c(x) - y_j(x)| : 1 \leq j \leq N\} \leq 2^{-(M+1)} \tag{2.8}$$

indicates the approximation accuracy, which may be governed by the choice of the number, M , of the partition employed. The function $y^c(x)$ corresponding to the particular value of M maps the uniform grid having the step-size equal to 2^{-MN} in the interval $[0, 1]$ onto the uniform grid in the hypercube (2.2) having the step-size equal to 2^{-M} (in each coordinate).

REDUCTION TO ONE DIMENSION. Employing the continuous single-valued Peano curve $y(x)$ mapping the unit interval $[0, 1]$ on the x -axis onto the N -dimensional domain (2.2), it is possible to find the minimum in (2.1) by solving the one-dimensional problem

$$\varphi(y(x^*)) = \min\{\varphi(y(x)) : x \in [0, 1], g_i(y(x)) \leq 0, 1 \leq i \leq m\}, \tag{2.9}$$

where (due to (2.3)) the functions $g_i(y(x))$ are defined and computable in the domains:

$$Q_1 = [0, 1], Q_{i+1} = \{x \in Q_i : g_i(y(x)) \leq 0\}, 1 \leq i \leq m, \tag{2.10}$$

and

$$[0, 1] = Q_1 \supset Q_2 \supset \dots \supset Q_{m+1} \supset Q_{m+2} = \emptyset \tag{2.11}$$

(the empty set Q_{m+2} is introduced to simplify the subsequent notation).

ITERATIONS. Each iteration of the proposed method starts by determining an *iteration point* $x \in (0, 1)$ (in accordance with the decision rules given below) and prolongs by computing $\nu = \nu(x)$, where $\nu - 1$ is the number of constraints which are satisfied at this point (this number is referred below also as an *index*), and the value

$$f(x) = g_\nu(y(x)) \tag{2.12}$$

corresponding to x and $\nu(x)$. Hence, the subscript $\nu = \nu(x)$ from (2.12) is given by the conditions:

$$x \in Q_\nu \text{ and } x \notin Q_{\nu+1}, \tag{2.13}$$

and for it we have the inequalities $1 \leq \nu(x) \leq m + 1, x \in (0, 1)$. The existence and uniqueness of the index follow from the inclusions (2.11). Function $f(x)$ is either the value of the left-hand side of the first constraint violated at the point x (this is if $\nu \leq m$), or $f(x)$ is the value of the function to be minimized (this is if $\nu = m + 1$). The above *response* (i.e., the pair $\nu(x), f(x)$) is to be obtained by successive computing values of functions $g_i(y(x)), 1 \leq i \leq \nu$, at the point x . The calculations are terminated when either the inequality $g_\nu(y(x)) > 0$ or the equality $\nu = m + 1$ is satisfied. Figure 1 illustrates the above concepts and notation.

REDUCTION TO AN UNCONSTRAINED CASE. The main idea the proposed technique is based upon is to reduce the conditional problem (2.9) to the unconditional problem:

$$\psi(x^*) = \min\{\psi(x) : x \in [0, 1]\}, \tag{2.14}$$

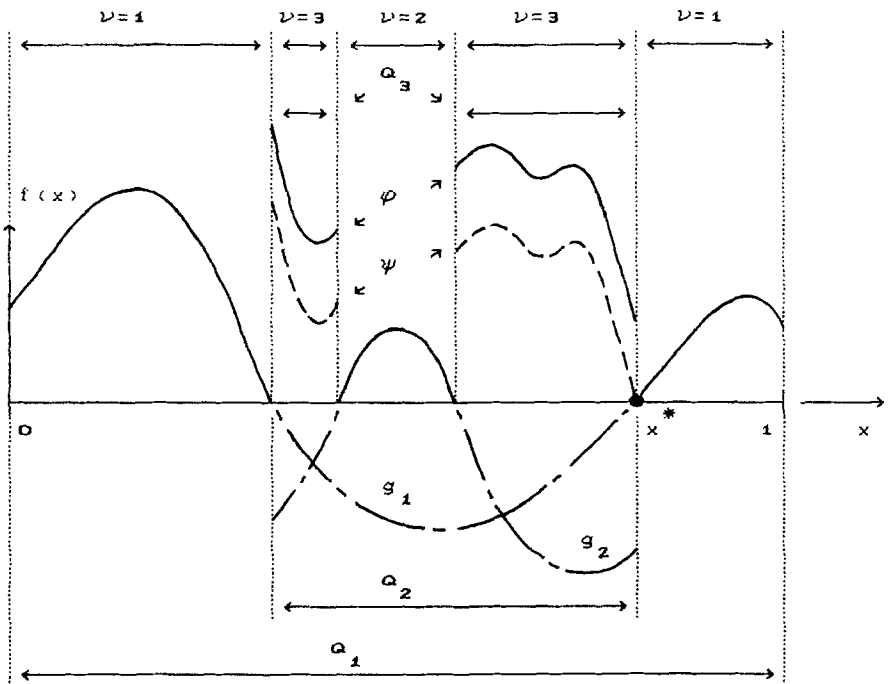


Fig. 1. Behavior of $f(x)$ (the solid line) and its relations with partially computable functionals g_1, g_2 and φ . Function $\psi(x)$ coinciding with $f(x)$, while $x \in Q_1/Q_3$, and not coinciding (the dotted line) with $f(x)$, while $x \in Q_2$, is from the connected auxiliary non-linear problem (2.14).

$$\psi(x) = f(x) - \begin{cases} 0, & \nu(x) \leq m, \\ \varphi(y(x^*)), & \nu(x) = m + 1, \end{cases} \tag{2.15}$$

where x^* is from (2.9) and $\nu = \nu(x)$ is from (2.13). The sets of solutions of problems (2.9) and (2.14) are the same, since the function ψ is strictly positive at all points which are not a solution of (2.9), while it vanishes at points which are a solution.

Functions $g_i(y(x))$ corresponding to Lipschitz functions $g_i(y)$ (with respective constants $L_i, 1 \leq i \leq m + 1$) satisfy (see, for instance, Strongin (1978)) uniform Hölder conditions:

$$|g_i(y(x')) - g_i(y(x''))| \leq K_i(|x' - x''|)^{1/N}, \quad (x', x'') \in Q_i,$$

with the respective coefficients

$$K_i = 4L_i\sqrt{N}, \quad 1 \leq i \leq m + 1. \tag{2.16}$$

By substituting the divided functions $g_i(y(x))/K_i, 1 \leq i \leq m + 1$, into (2.12) and (2.15) instead of the initial ones, we obtain the function $\psi(x)$ which in any subset Q_i from (2.10) satisfies Hölder conditions with unique coefficient $K = 1$.

If functions $g_i(y(x)), 1 \leq i \leq m + 1$, admit Hölder continuations in all the interval $(0, 1)$, then for any point $x^0 \in (0, 1)$ with the given value $\psi(x^0)$ it is true that

$$\psi(x) \geq \psi(x^0) - (|x - x^0|)^{1/N}, \quad x \in (0, 1).$$

These inequalities are sufficient for estimating a location of x^* after a finite number of function evaluations, though the function ψ may have discontinuities of the first kind at boundary points of Q_i . After k function evaluations at x^1, \dots, x^k :

$$x^* \in \{x \in [0, 1]: (|x - x^i|)^{1/N} \geq \psi(x^i) - \psi^+, 1 \leq i \leq k\}, \tag{2.17}$$

where ψ^+ is the best known estimate, i.e.,

$$\psi^+ = \psi(x^+) = \min\{\psi(x^i): 1 \leq i \leq k\} \tag{2.18}$$

(see Figure 2). Thus, to minimize the univariate function $\psi(x)$ we may apply some modification of the generalized algorithm of global search from Strongin (1973) (a particular modified version is presented below) or to develop some similar generalizations of *saw-tooth covers* (e.g., from Hansen *et al.* (1991), Piyavskii (1972) or Sukharev (1971)).

The particular technique we are to implement (the *index method*) is based on the above considerations and proposed in Strongin (1984) and in Strongin and Markin (1986, 1987). According to this method, the unknowns $\varphi(x^*)$ and K_i appearing in (2.15) and in the divided functions $g_i(y(x))/K_i, 1 \leq i \leq m + 1$, are replaced by running estimates.

ALGORITHM. The first iteration is carried out at an arbitrary point $x \in (0, 1)$.

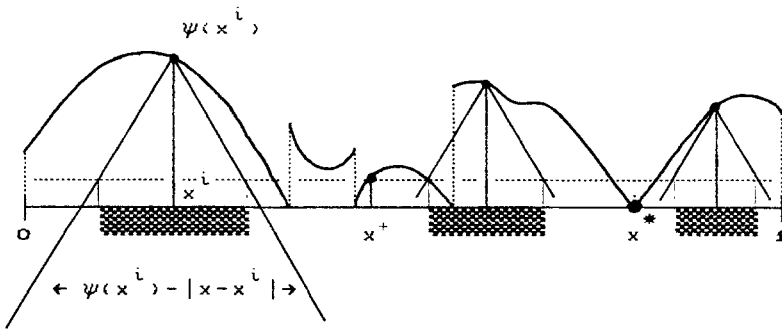


Fig. 2. Case $N = 1, k = 4$. The union of subintervals shaded below the x -axis is an estimated complement of a set (2.17) and it does not contain the optimal point x^* . Point x^+ is the best known estimate from (2.18).

The choice of any subsequent point $x^{k+1}, k \geq 1$, is determined by the following rules:

Rule 1. The points x^1, \dots, x^k from the previous iterations are renumbered by subscripts in increasing order of the coordinate, i.e.,

$$0 = x_0 < x_1 < \dots < x_i < \dots < x_k < x_{k+1} = 1 \tag{2.19}$$

and associated with the values $z_i = f(x_i), 1 \leq i \leq k$, computed at these points (f is from (2.12)). The points $x_0 = 0$ and $x_{k+1} = 1$ are additionally introduced into the series (2.19) (the values z_0 and z_{k+1} are undefined).

Rule 2. By constructing the sets:

$$I_\nu = \{i: 1 \leq i \leq k, \nu = \nu(x_i)\}, \quad 1 \leq \nu \leq m + 1,$$

all subscripts $i, 1 \leq i \leq k$, in the series (2.19) are classified with respect to the number of constraints met at the corresponding points x_i . These sets are united as

$$S_\nu = I_0 \cup \dots \cup I_{\nu-1}, \quad T_\nu = I_{\nu+1} \cup \dots \cup I_{m+2}, \quad 1 \leq \nu \leq m + 1,$$

where $I_0 = \{0, k + 1\}, I_{m+2} = \emptyset$ and $i \in S_\nu (i \in T_\nu)$ if $\nu(x_i) < \nu (\nu(x_i) > \nu)$.

Rule 3. Running lower bounds:

$$\mu_\nu = \max\{|z_j - z_i|(x_j - x_i)^{-1/N} : i, j \in I_\nu, i < j\} \tag{2.20}$$

for respective Hölder's coefficients (2.16) of functions $g_\nu(y(x)), 1 \leq \nu \leq m + 1$, are calculated. If the set I_ν contains less than two elements or if μ_ν from (2.20) is equal to zero, it is assumed that $\mu_\nu = 1$. Formula (2.20) directly implies that estimates μ_ν are non-decreasing while iterating.

Rule 4. For all non-empty sets $I_\nu, 1 \leq \nu \leq m + 1$, values

$$z_\nu^* = \begin{cases} -\varepsilon_\nu, & \text{if } T_\nu \neq \emptyset, \\ \min\{z_i: i \in I_\nu\}, & \text{if } T_\nu = \emptyset, \end{cases} \tag{2.21}$$

are determined (real numbers $\varepsilon_\nu, 1 \leq \nu \leq m$, are positive parameters correspond-

ing to coordinates of a pre-assigned vector of reserves from (2.6)). The running value $\alpha = \min\{\nu: T_\nu = \emptyset\}$ indicates the “running objective function” $g_\alpha(y(x))$ the algorithm is minimizing and z_α^* is a best non-increasing (and positive if $\alpha \leq m$) upper bound for the minimal value of this function. So that some different functions $g_i(y(x))$ may play this role until α attains the value $m + 1$. In the last case, z_{m+1}^* is a best running estimate for $\varphi(y(x^*))$ from (2.9).

Rule 5. For each interval (x_{i-1}, x_i) , $1 \leq i \leq k + 1$, the value $R(i)$ (called the *characteristic*) is computed, where

$$R(i) = \Delta_i + (z_i - z_{i-1})^2 / \mu_\nu^2 \Delta_i - 2(z_i + z_{i-1} - 2z_\nu^*) / r\mu_\nu, \quad i - 1, \quad i \in I_\nu, \tag{2.22}$$

$$R(i) = \Delta_i - 4(z_i - z_\nu^*) / r\mu_\nu, \quad i \in I_\nu, \quad i - 1 \in S_\nu, \tag{2.23}$$

$$R(i) = \Delta_i - 4(z_{i-1} - z_\nu^*) / r\mu_\nu, \quad i - 1 \in I_\nu, \quad i \in S_\nu, \tag{2.24}$$

$$\Delta_i = (x_i - x_{i-1})^{1/N}, \tag{2.25}$$

(here r is a *parameter*, with a value greater than 1, which allows us, by proper choice of r , to use the product $r\mu_\nu$ as an upper bound for Hölder’s coefficient K_ν from (2.16); this consideration will be continued in Section 4).

Rule 6. The interval (x_{t-1}, x_t) with the maximal characteristic

$$R(t) = \max\{R(i): 1 \leq i \leq k + 1\} \tag{2.26}$$

is determined. If $\nu = \nu(x_t) = \nu(x_{t-1})$, then the next iteration is carried out at a point

$$x^{k+1} = (x_t + x_{t-1}) / 2 - \text{sign}(z_t - z_{t-1}) [|z_t - z_{t-1}| / \mu_\nu]^N / 2r. \tag{2.27}$$

Otherwise, i.e., if $\nu(x_t) \neq \nu(x_{t-1})$, the second term in (2.27) is omitted.

The above algorithm may be supplied with the *termination rule* (see Strongin (1973)), which stops iterating if Δ_t from (2.25) is less than the given tolerance (t is from (2.26)).

Global convergence conditions for this algorithm (for the case when all ε_ν from (2.21) are zero-valued) are considered in Strongin and Markin (1986, 1987). This consideration employs the assumption that all domains Q_i from (2.11) are the unions of a finite number of intervals of positive length. Here we use a less rigorous requirement of the existence of an ε -reserved solution and this convergence study deals with a more general algorithm from Section 4.

COMMENT 1. The algorithm described above may be interpreted as some kind of saw-tooth cover. But the origin of this technique is based on a stochastic model representing the functions $g_i(y(x))$ (see Strongin (1978, 1989)). Within the framework of this model, a Gaussian distribution is assumed for the increments of these functions depending on the location of global optimizers. Maximum likelihood estimates for global optimizers (based on respective posterior densities obtained after a number of function evaluations) are treated as points of

iterations. This idea (after some transformations caused by the fact that the function (2.15) is piecewise presented by arcs of functions $g_v(y(x))$) leads to the formula (2.27), which may be viewed as a point estimate for a global optimizer. Some alternative stochastic decision rules for unconstrained univariate global optimization are viewed, for example, in Betro (1991) and Žilinskas (1981).

COMMENT 2. It is possible to generalize this algorithm for solving multiple criteria problems (see Strongin *et al.* (1988)) and for implementation on multi-processor systems (see Sergeev and Strongin (1987, 1989)).

3. Set of Joint Scannings

Let us consider the interval $[0, L + 1]$ on the x -axis and the family of hypercubes

$$D_l = \{y \in \mathbb{R}^N : -2^{-1} \leq y_j + 2^{-l} \leq 3 \cdot 2^{-1}, 1 \leq j \leq N\}, \quad 0 \leq l \leq L, \quad (3.1)$$

where the hypercubes D_l are obtainable by translation of hypercubes D_{l-1} along the main diagonal with the displacements equal to 2^{-l} (in each coordinate). Figure 3 presents the case $L = N = 2$.

Suppose that $y^0(x)$ maps the interval $[0, 1]$ on the x -axis onto the hypercube D_0 from (3.1), i.e.,

$$D_0 = \{y^0(x) : x \in [0, 1]\}, \quad (3.2)$$

and this map is defined following the scheme presented in Section 2, i.e., $y^0(x)$ is a space-filling curve. Any subcube of the M th partition of the hypercube D_0 generated by this definition has an edge-length equal to $2^{-(M-1)}$ and, henceforth, is designated $D_0(M, v)$, where v has a unique value for each particular subcube (v is the left end-point of the subinterval $d(M, v)$ linked with $D_0(M, v)$ by the definition). The map $y^0(x)$ and the equalities

$$y_j^l = y_j^{l-1}(x) + 2^{-l}, \quad 1 \leq j \leq N, \quad 1 \leq l \leq L, \quad (3.3)$$

define the curves $y^l(x)$ mapping the interval $[0, 1]$ onto respective hypercubes D_l , $1 \leq l \leq L$, from (3.1). Hence, the curve

$$Y(x) = y^{[x]}(x - [x]), \quad x \in [0, L + 1), \quad (3.4)$$

where $[x]$ is the integer part of x , maps the interval $[l, l + 1)$ into the hypercube D_l , $0 \leq l \leq L$. From (3.3), for each particular element $D_0(M, v)$ of the M th partition of the hypercube D_0 , there exists the element $D_l(M, v)$ of the M th partition of the hypercube D_l , and this element may be obtained by translating $D_0(M, v)$ along the main diagonal at the distance $2^{-1} + 2^{-2} + \dots + 2^{-l}$ (in each coordinate). Subcubes $D_0(2, 5/8)$ and $D_1(2, 5/8)$ are depicted in Figure 3. As follows from (3.3) and (3.4), if $d(M, v) \subset [0, 1)$ is juxtaposed to $D_0(M, v)$, then there exist a set of subcubes

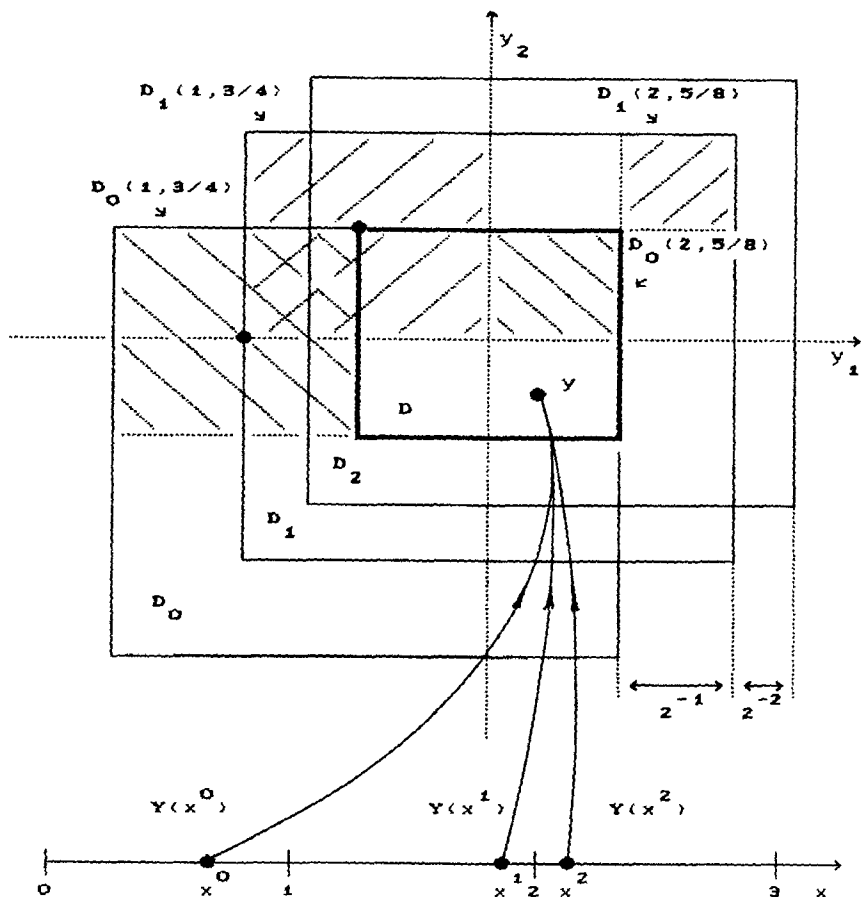


Fig. 3.

$$D_l(M, v) = y^l(d(M, v)), \quad 1 \leq l \leq L,$$

linked with the subintervals

$$d_l(M, v_l) \subset [l, l+1), \quad v = v_l - [v_l], \quad 0 \leq l \leq L,$$

where $d(M, v) = d_0(M, v_0)$, $v = v_0$, and

$$D_l(M, v) = Y(d_l(M, v_l)), \quad 0 \leq l \leq L. \tag{3.5}$$

Taking account of the obvious inclusion $D \subset D_l$, $0 \leq l \leq L$, it is possible to present the initial hypercube (2.2) in any of the following forms (see Figure 3)

$$D = \{Y(x): x \in [l, l+1), g_0(Y(x)) \leq 0\}, \quad 0 \leq l \leq L, \tag{3.6}$$

where $g_0(y)$ is from (2.5). Hence, the scanning (3.4) covers hypercube (2.2) $L + 1$ times, while x varies from 0 up to $L + 1$. This is the reason to refer to the curve $Y(x)$ from (3.4) as the *multiple scanning*. As a consequence, any point

$y \in D$ has its pre-image x^l in each interval $[l, l + 1)$, i.e.,

$$y = y^{\lfloor x^l \rfloor}(x - \lfloor x^l \rfloor), \quad x^l \in [l, l + 1), \quad 0 \leq l \leq L, \tag{3.7}$$

(see Figure 3). The next statement exhibits the property of curve (3.4) we aimed to achieve by introducing the above construction.

THEOREM 1. *Let a point y^* from the domain of search (2.2) be contained in a line segment with end-points $y', y'' \in D$ meeting the requirements:*

$$|y'_j - y''_j| \leq 2^{-p}; \quad y'_i = y''_i = y^*_i, \quad 1 \leq i \leq N, \quad i \neq j, \tag{3.8}$$

where p is an integer and $1 \leq p \leq L$ (i.e., the segment is collinear with the j th axis in \mathbb{R}^N). Then the curve (3.4) assures the existence of an integer number l , $0 \leq l \leq L$, and of pre-images $x^*, x', x'' \in [l, l + 1)$ satisfying the following conditions:

$$y^* = Y(x^*), \quad y' = Y(x'), \quad y'' = Y(x'') \tag{3.9}$$

and

$$\max\{|x' - x^*|, |x'' - x^*|, |x' - x''|\} \leq 2^{-pN}. \tag{3.10}$$

COMMENT 3. The conditions (3.8) introduce a particular type of 2^{-p} -vicinity of the point y^* . This vicinity comprises all the “neighbours” of the given point y^* differing from this point in the j th coordinate, but with differences not exceeding 2^{-p} (the range for p depends on the number L of the scannings employed). By changing j , $1 \leq j \leq N$, in (3.8), it is possible to obtain the neighbours in any N coordinate directions.

As is stated in the theorem, any two neighbours from the j th 2^{-p} -vicinity of y^* will have at least two pre-images in 2^{-pN} -vicinity of some pre-image x^* on the x -axis. This is the way the inverse multivalued mapping $Y^{-1}(y)$ reflects the property of closeness in \mathbb{R}^N in any N direction.

Proof. (1) The edge-length for both $D_l(l, v)$ and $D_{l-1}(l, v)$ is equal to $2^{-(l-1)}$, and $D_l(l, v)$ may be obtained by translation of $D_{l-1}(l, v)$ with displacements equal to 2^{-l} in each coordinate. As a consequence, the center of $D_l(l, v)$ is simultaneously a vertex of $D_{l-1}(l, v)$ and, vice versa, the center of $D_{l-1}(l, v)$ is a vertex of $D_l(l, v)$ (as an illustration of this fact, the subcubes $D_0(1, 3/4)$ and $D_1(1, 3/4)$ are singled out by the hatching in Figure 3). Thus, the union of all vertices of all subcubes $D_l(l, v)$ and $D_{l-1}(l, v)$, which are contained in the hypercube D from (2.2) (l is supposed to be fixed), forms a 2^{-l} -net (in each coordinate) in the domain of search.

(2) By assumption (3.8), points y', y'' are either both in the same subcube $D_p(p + 1, v)$ (see Case 1 in Figure 4) or they are in the two different contiguous subcubes:

$$D' = D_p(p + 1, v'), \quad D'' = D_p(p + 1, v'')$$

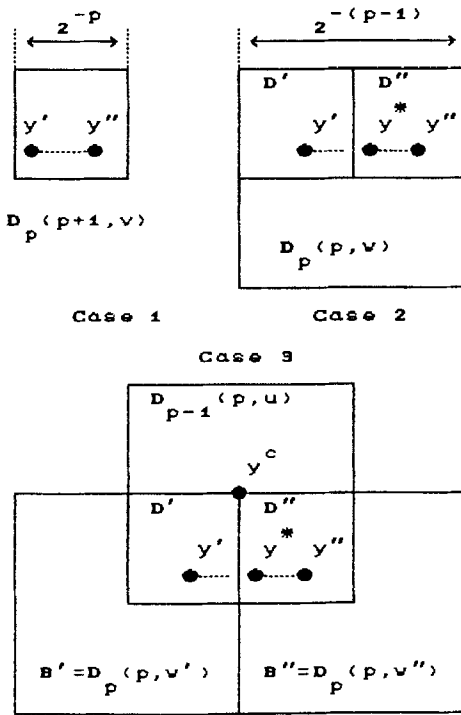


Fig. 4.

having a common $(N - 1)$ -dimensional *face* orthogonal to the j th coordinate axis. In the last case, the subcubes D' , D'' are either the elements of a next partition of some subcube $D_p(p, w)$ (see Case 2 in Figure 4) or

$$D' \subset B' = D_p(p, w') \quad \text{and} \quad D'' \subset B'' = D_p(p, w''),$$

where the subcubes B' and B'' are contiguous and have the common face containing the common face of the subcubes D' and D'' (and thus orthogonal to the j th coordinate axis). In this final case (see Case 3 in Figure 4), D' and B' have the only common vertex y^c which is also the only common vertex of the subcubes D'' and B'' .

Hence, taking account of part 1 of this proof, we may state that the point y^c as a vertex of the subcube $D_p(p, w')$ is simultaneously the center of some subcube $D_{p-1}(p, u)$ containing $D' \cup D''$, because the center y^c is the common vertex of D' and D'' , i.e., in any case, there exists some subcube $D_l(p, u)$, $l \leq p \leq L$, containing the points y' , y'' and the point y^* in the line segment between y' and y'' . The subinterval $d_l(p, u_l) \subset [l, l + 1)$ corresponding to the above subcube contains the pre-images x^* , x' , x'' of the points from (3.9). The length of $d_l(p, u_l)$ is equal to 2^{-pN} and, thus, the statement (3.10) holds true. ■

4. Algorithm with a Multiple Scanning

REDUCTION TO ONE DIMENSION. With the use of $Y(x)$ from (3.4) and taking into account (3.6), the initial problem (2.1) may be reduced to the following one-dimensional problem:

$$\min\{\varphi(Y(x)): x \in [0, L + 1), g_i(Y(x)) \leq 0, 0 \leq i \leq m\}, \tag{4.1}$$

where the functions $g_i(Y(x))$ are defined and computable in the domains Q_i :

$$Q_0 = [0, L + 1), \tag{4.2}$$

$$Q_{i+1} = \{x \in Q_i: g_i(Y(x)) \leq 0\}, 0 \leq i \leq m. \tag{4.3}$$

The algorithm from Section 2, being modified, is applicable for solving this new problem.

ITERATIONS. An iteration of the modified method at a (non-integer) point x^q includes:

- (i) calculation of an index $\nu = \nu(x^q)$ meeting the conditions (2.13) (note that ν is in the range from 0 up to $m + 1$ and the sets Q_ν are from (4.2), (4.3));
- (ii) computation of the value

$$f(x^q) = g_\nu(Y(x^q)); \tag{4.4}$$

(iii) determination of the *pre-images* $x^{ql}, 0 \leq l \leq L$, from (3.7) for the corresponding point $y^q = Y(x^q)$. Hence, an iteration at a point x^q (note, that $x^q \in \{x^{q0}, \dots, x^{qL}\}$) may be *interpreted* as $L + 1$ iterations at the points x^{q0}, \dots, x^{qL} with responses:

$$\begin{aligned} \nu(x^{q0}) &= \dots = \nu(x^{qL}) = \nu(x^q), \\ f(x^{q0}) &= \dots = f(x^{qL}) = f(x^q), \end{aligned} \tag{4.5}$$

If $\nu(x^q) = 0$ (i.e., if $y^q \notin D$), then pre-images, which are different from x^q , are not computed, and the set $\{x^{q0}, \dots, x^{qL}\}$ is presented by $\{x^q\}$ (i.e., in this case, the step (iii) is not performed).

MODIFICATIONS. Now we shall consider some modifications of the rules of the algorithm from Section 2 that have to be done to adapt this algorithm for solving the problem (4.1).

Modification 1. In *Rule 1* of the algorithm the set (2.19) has to be constructed by sequencing the points

$$\{x^{10}, \dots, x^{1L}, x^{20}, \dots, x^{2L}, x^{30}, \dots, x^{qL}\} \cup \{0, 1, \dots, L + 1\} \tag{4.6}$$

(points that are not existing due to the rules of step (iii) are omitted in the set (4.6)). Hence, $k \leq (q + 1)(L + 1) - 1$ and it is assumed that values z_i corresponding to the integer points $x_i \in \{0, 1, \dots, L + 1\}$ are undefined. To distinguish the number q of the *last iteration* being performed from the number k characterizing

the ordered set (2.19), the point x^{k+1} from (2.27) in *Rule 6* is henceforth denoted by x^{q+1} .

Modification 2. Implementation of the additional constraint (2.4) widens the range of change for indices (a zero-value becomes attainable) which causes the introduction of modified sets:

$$\begin{aligned} I_{-1} &= \{i: 0 \leq i \leq k + 1, \quad x_i \in \{0, 1, \dots, L + 1\}\}, \\ I_\nu &= \{i: 1 \leq i \leq k, \quad i \notin I_{-1}, \quad \nu = \nu(x_i)\}, \\ S_\nu &= I_{-1} \cup \dots \cup I_{\nu-1}, \quad T_\nu = I_{\nu+1} \cup \dots \cup I_{m+2}, \quad 0 \leq \nu \leq m + 1, \end{aligned}$$

for the classification of subscripts in *Rule 2*.

Modification 3. The formula

$$\begin{aligned} \mu_\nu &= \max_{0 \leq l \leq L} \max\{|z_j - z_i|(x_j - x_i)^{-1/N} : \\ & i, j \in I_\nu, \quad i < j, \quad (x_i, x_j) \subset [l, l + 1)\} \end{aligned} \tag{4.7}$$

has to be used instead of (2.20) in *Rule 3*.

Finally, the *first iteration* is to be carried out at any arbitrary point $y^1 \in D$ (which must have pre-images x^{1l} in all the subintervals $[l, l + 1)$, $0 \leq l \leq L$, to ensure the equalities (4.5)). All the other rules hold true.

THEOREM 2. *Assume that the following conditions are satisfied:*

- (1) *the problem (2.1) has an ϵ -reserved solution y_ϵ from (2.6);*
- (2) *functions $g_i(y)$, $1 \leq i \leq m + 1$, admit Lipschitz extensions $G_i(y)$ (with respective constants L_i) over each domain D_i from (3.1), i.e.*

$$g_i(y^l(x)) = G_i(y^l(x)), \quad x \in Q_i \cap [l, l + 1), \quad 0 \leq l \leq L$$

(note, that $g_0(y)$ from (2.5) is Lipschitz in any domain D_l , $0 \leq l \leq L$ and, thus, we may admit $G_0 \equiv g_0$);

- (3) *parameters ϵ_i , $0 \leq i \leq m$, in the formula (2.21) have the same values as the corresponding coordinates of the vector ϵ in (2.6);*
- (4) *since some step (i.e., if k in (2.19) is sufficiently large), values μ_ν , $0 \leq \nu \leq m + 1$, from (4.7) are meeting the inequalities:*

$$r\mu_\nu > 16L_\nu \sqrt{N}, \quad 0 \leq \nu \leq m + 1. \tag{4.8}$$

Then any accumulation point \bar{y} of the sequence of iteration points $\{y^n\}_{n=1}^\infty$ generated by the described modified algorithm satisfies the conditions:

$$\varphi(\bar{y}) = \inf\{\varphi(y^n): n \in \mathbb{N}_1, \quad g_i(y^n) \leq 0, \quad 0 \leq i \leq m\} \leq \varphi(y_\epsilon), \tag{4.9}$$

where \mathbb{N}_1 is the set of integers ≥ 1 .

Proof. To be more observable, the proof is separated into successive steps. First three steps are given by lemmas presenting some intermediate statements under the conditions (1)–(4) of Theorem 2.

LEMMA 1. For any accumulation point \bar{y} of the sequence $\{y^n\}_{n=1}^\infty$, the algorithm generates an infinite nested sequence of intervals:

$$\{(x_{t-1}, x_t) : t = t(q^p)\}_{p=1}^\infty \tag{4.10}$$

satisfying the conditions:

$$\bar{x} \in \bigcap_{p=1}^\infty [x_{t-1}, x_t], \tag{4.11}$$

$$\lim_{p \rightarrow \infty} \Delta_t = 0, \tag{4.12}$$

$$\lim_{p \rightarrow \infty} R(t(q^p)) \leq 0, \tag{4.13}$$

where \bar{x} is some pre-image of \bar{y} , q is the step-number from (i)–(iii) and $q^1 < q^2 < \dots$; $R(t)$, Δ_t and t are from (2.22)–(2.24), (2.25) and (2.26), respectively.

Proof. At least for one subinterval $[s, s + 1)$, $0 \leq s \leq L$, the sequence $\{x^q\}_{q=1}^\infty$ produced by a repetitive use of (2.27) (recall the concluding remark from Modification 1) generates the subsequence of pre-images:

$$\{x^{qs} : q \in \{q^p\}_{p=1}^\infty\} \tag{4.14}$$

from (iii) converging to the accumulation point $\bar{x}^s \in [s, s + 1)$ (i.e., $\bar{y} = Y(\bar{x}^s)$). Without loss of generality, we may assume that for each point x^{qs} from (4.14) there is an interval (x_{t-1}, x_t) from (2.19) with t from (2.26) satisfying the conditions:

$$\bar{x}^s, x^{qs} \in (x_{t-1}, x_t), \quad t = t(q^p). \tag{4.15}$$

These intervals form the set (4.10) and ensure the validity of (4.11).

From (2.27), (4.7) and (4.15),

$$\max\{(x_t - x^{qs}), (x^{qs} - x_{t-1})\} \leq \gamma(x_t - x_{t-1}), \tag{4.16}$$

where $\gamma = (r + 1)/2r < 1$ (because of the constraint $r > 1$). As a consequence of (2.25), (4.16), we obtain the *contraction property* (4.12). If there are two intervals simultaneously (i.e., for the same integer k in (2.19)) containing \bar{x}^s (which means that \bar{x}^s is an end-point of both intervals), then the number $t = t(q)$ is assigned to the one contracting in accordance with (4.12).

From (2.21) and (4.4),

$$z_j = g_\nu(Y(x_j)) \geq z_\nu^*, \quad \nu = \nu(x_j), \quad 1 \leq j \leq k. \tag{4.17}$$

Thus, from (2.22)–(2.24), (4.12), taking into account condition (2), we derive the statement (4.13). ■

LEMMA 2. There exists a step number $h \geq 1$ satisfying the condition:

$$\nu(x^h) = m + 1. \tag{4.18}$$

Proof. Let $x_\varepsilon^l \in [l, l + 1)$, $0 \leq l \leq L$, be a pre-image of y_ε from the condition (1), and let $j = j(q)$ be the number of the interval $[x_{j-1}, x_j]$ containing x_ε^l at the q th iteration. From condition (2), extensions $G_\nu(Y(x))$, $x \in [l, l + 1)$, $0 \leq \nu \leq m + 1$, must satisfy uniform Hölder conditions with the fractional exponent $1/N$ and with respective coefficients K_ν from (2.16). Hence,

$$z_j = g_\nu(Y(x_j)) \leq g_\nu(Y(x_\varepsilon^l)) + K_\nu \Delta_j, \tag{4.19}$$

where $\nu = \nu(x_j)$ and Δ_j is from (2.25). Similar inequality holds for z_{j-1} . From (2.6) and (2.21),

$$g_\nu(Y(x_\varepsilon^l)) = g_\nu(y_\varepsilon) \leq -\varepsilon_\nu \leq z_\nu^*, \quad 0 \leq \nu \leq m, \tag{4.20}$$

and if we admit that for any integer $h \geq 1$ (4.18) is not true, then from (2.22)–(2.24) and (2.16), (4.8), (4.19), (4.20), we derive the inequality:

$$R(j(q)) \geq \Delta_j(r\mu_\nu - 4K_\nu)/r\mu_\nu > 0. \tag{4.21}$$

In accordance with (2.26) from *Rule 6* and (4.13), (4.21), the interval $[x_{j-1}, x_j]$ containing x_ε^l must be subdivided by some point (2.27). Repetition of this partitioning with $q \rightarrow \infty$ yields to an infinite nested sequence of intervals contracting to x_ε^l . Then y_ε , which is an internal point of the admissible domain (cf. (2.1) and (2.6)), is an accumulation point of $\{y^n\}_{n=1}^\infty$ and, thus, in its vicinity must be some feasible point $y^h = Y(x^h)$ with x^h satisfying (4.18). ■

LEMMA 3. *Any accumulation point \bar{x} has a feasible image $\bar{y} = Y(\bar{x})$ and a nested sequence of intervals (4.10) satisfying the conditions (4.11), (4.12) and*

$$R(t(q^p)) > 0, \quad p \in \mathbb{N}_1. \tag{4.22}$$

Proof. From (2.21) and (4.18), for any $q \geq h$

$$z_\nu^* = -\varepsilon_\nu, \quad 0 \leq \nu \leq m, \quad z_{m+1}^* = \min\{z_i; i \in I_{m+1}\}, \tag{4.23}$$

and there exists an interval (x_{t-1}, x_t) , $t = t(q)$, satisfying the conditions:

$$\max\{\nu(x_{t-1}), \nu(x_t)\} = m + 1, \tag{4.24}$$

and either $z_{t-1} = z_{m+1}^*$ or $z_t = z_{m+1}^*$. If $\nu(x_{t-1}) \neq \nu(x_t)$ then from (2.23), (2.24), $R(t(q)) = \Delta_t > 0$. If $\nu(x_{t-1}) = \nu(x_t)$ then by taking account of Hölder conditions, we obtain inequality:

$$\max\{z_{t-1}, z_t\} \leq z_{m+1}^* + K_{m+1} \Delta_t,$$

which in conjunction with (2.22) and (2.16), (4.8) implies $R(t(q)) > \Delta_t/2 > 0$. As was already mentioned in the proof of Lemma 2, *Rule 6* and inequalities (4.13) and $R(t(q)) > 0$ ensure the existence of the sequence (4.10) contracting to some point \bar{x} (cf. (4.11)) and satisfying (4.12) and (4.22).

Suppose that some accumulation point \bar{y} is external to the closed feasible domain of the problem (2.1). From (4.11) and the continuity of $Y(x)$, $x \in [l, l + 1)$, $0 \leq l \leq L$,

$$\bar{y} = Y(\bar{x}) = \lim_{p \rightarrow \infty} Y(x_{t-1}) = \lim_{p \rightarrow \infty} Y(x_t).$$

Then, all the points $Y(x_{t-1}), Y(x_t), t = t(q^p), p \geq p',$ are also not feasible if p' is sufficiently large. Hence,

$$\nu = \max\{\nu(x_{t-1}), \nu(x_t)\} < m + 1, \quad t = t(q^p), \quad p \geq p', \tag{4.25}$$

and from (2.22)–(2.24), (4.12), (4.17), (4.23), (4.25)

$$R(t(q^p)) < -2\varepsilon_\nu / r\mu_\nu < 0, \quad p \geq p' \tag{4.26}$$

(with p' large enough). But from *Rule 6*, (4.22) and (4.26), the interval (x_{t-1}, x_t) could not be subdivided by iterations if p' is large enough. Therefore, our assumption that \bar{y} is not feasible must be false. ■

Proof of Theorem 2. (1) We shall prove the left equality in (4.9) by validating the relation:

$$\varphi(\bar{y}) \leq z_{m+1}^* \tag{4.27}$$

for the feasible accumulation point $\bar{y} = Y(\bar{x})$ of the sequence $\{y^n\}_{n=1}^\infty$ which exists by Lemma 3. By this lemma, there is a nested sequence of intervals (4.10) contracting to \bar{x} (cf. (4.11), (4.12)) and satisfying the conditions (4.24). Supposition that (4.27) is not true implies the inequalities:

$$z_i > z_{m+1}^* + \delta \quad \text{if} \quad \nu(x_i) = m + 1, \quad i = t - 1, t, \tag{4.28}$$

where $t = t(q^p), p \in \mathbb{N}_1,$ and δ is some positive real number. From (2.22)–(2.24), (4.24), (4.28) with account of (4.12),

$$R(t(q^p)) < 2\Delta_t - 4\delta / r\mu_{m+1} < -2\delta / r\mu_{m+1} < 0, \tag{4.29}$$

where $p \geq p'$ and p' is sufficiently large. But (4.29) contradicts the inequality (4.22). Hence, (4.27) must be true.

(2) Now, we shall prove the right inequality in (4.9) by demonstrating that for any real $\beta > 0$ the relation:

$$\varphi(y_\varepsilon) \leq z_{m+1}^* - \beta \tag{4.30}$$

must not be true for sufficiently large values of n . Suppose the contrary, i.e., that (4.30) holds true at any step q for some fixed value $\beta > 0$.

From (4.18), the interval $[x_{j-1}, x_j], j = j(h),$ containing the point $x_\varepsilon^j,$ which was introduced in the proof of Lemma 2, must satisfy the equality (4.24). Then, from (4.19), (4.24) and (4.30),

$$z_i \leq z_{m+1}^* - \beta + K_{m+1} \Delta_j \quad \text{if} \quad \nu(x_i) = m + 1, \quad i = j - 1, j,$$

which in conjunction with (2.22)–(2.24), (2.16) and (4.8) implies:

$$R(j(q)) > 4\beta / r\mu_{m+1} > 0. \tag{4.31}$$

But (4.31) contradicts (4.13): on the one hand, in accordance with *Rule 6* (cf.

(4.13), (4.31)) the interval $[x_{j-1}, x_j]$ should contract (cf. (4.12)), but on the other hand, by Lemma 1, the interval contracting to a pre-image of an accumulation point must satisfy (4.13). Hence, the infimum of z_{m+1}^* , which is the middle part of (4.9), does not exceed $\varphi(y_\varepsilon)$. ■

THEOREM 3. *Assume that under the conditions of Theorem 2 the point \bar{y} , which is not locally optimal to the problem (2.1), has a vicinity*

$$U_p(\bar{y}) = \{y \in \mathbb{R}^N : |y_j - \bar{y}_j| \leq 2^{-p}, \quad 1 \leq j \leq N\}$$

comprised of feasible points, i.e.,

$$g_i(y) \leq 0, \quad 0 \leq i \leq m, \quad y \in U_p(\bar{y}). \tag{4.32}$$

Then \bar{y} is not an accumulation point of the sequence $\{y^n\}_{n=1}^\infty$ generated by the above algorithm if

$$L \geq p + 2, \tag{4.33}$$

where L is from (3.4).

COMMENT 4. As follows from Lemma 3 and (4.9), any accumulation point \bar{y} of $\{y^n\}_{n=1}^\infty$ belongs to \mathbb{Y}_ε from (2.7). Theorem 3 singles out the subset of \mathbb{Y}_ε which could not contain accumulation points. This subset depends on the number of scannings employed in the algorithm.

Proof. First we prove some preliminary statements.

LEMMA 4. *Under the conditions of Theorem 3*

$$\bar{y} \in \text{int } D^0 \subset U_p(\bar{y}), \tag{4.34}$$

where D^0 containing \bar{y} among its internal points is an element of the l th partition of either D_{l-1} or D_l from (3.1), $l \leq L$ (see Section 3).

Proof. From $\bar{y} \in D \subset D_l$, $0 \leq l \leq L$, there exists $D_{p+1}(p+1, v)$ with the edge-length equal to 2^{-p} and

$$\bar{y} \in D_{p+1}(p+1, v) \subset U_p(\bar{y}),$$

where \bar{y} is either an internal point or a border point of $D_{p+1}(p+1, v)$. The last case we split into two subcases. The first subcase, is when \bar{y} is not contained in any hyperplane orthogonal to some edge of $D_{p+1}(p+1, v)$ and passing through a middle point of this edge. The second case is when \bar{y} is contained in such a hyperplane.

In the first subcase, \bar{y} is an internal point of some subcube $D_p(p+1, v')$. The center of this subcube is a vertex of $D_{p+1}(p+1, v)$ (see Figure 5, Case 1, and part 1 of the proof of Theorem 1).

In the second subcase, \bar{y} is in the edge of some element $D_{p+1}(p+2, w) \subset D_{p+1}(p+1, v)$ (see Figure 5, Case 2, where \bar{y} is a vertex of $D_{p+1}(p+2, w)$).

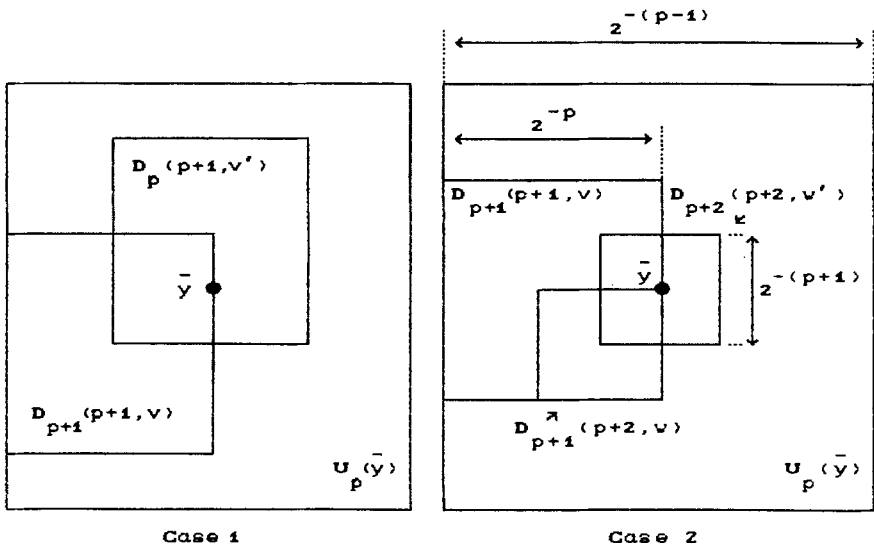


Fig. 5.

Hence, \bar{y} is an internal point of some subcube $D_{p+2}(p+2, w')$. This concludes the proof of (4.34). ■

Proof of Theorem 3. Let y^0 be defined by the relation:

$$\varphi(y^0) = \min\{\varphi(y) : y \in D^0\} < \varphi(\bar{y}), \tag{4.35}$$

where the strict inequality is a consequence of the local non-optimality of \bar{y} . From (3.5) and by Lemma 4, $D^0 = Y(d^0)$, where $Y(\cdot)$ is from (3.4) and d^0 is either $d_{l-1}(l, v_{l-1})$ or $d_l(l, v_l)$, $l \leq L$, and let x^0 designate the pre-image of y^0 in the interval d^0 .

Now, suppose that \bar{y} , which is an internal point of feasible set, is an accumulation point of $\{y^n\}_{n=1}^\infty$. Then there should be some points of this sequence in D^0 (satisfying (4.32)) and some pre-images of these points in d^0 . Hence, there exists an interval (x_{t-1}, x_t) satisfying the following conditions:

$$x^0 \in [x_{t-1}, x_t] \subset d^0, \quad t = t(q^p), \quad p \geq p', \tag{4.36}$$

where p' is large enough, and the end-points of this interval meet (4.24). If $\nu(x_t) = m + 1$ then from (4.9), (4.27), (4.35) and with account of Hölder conditions,

$$z_t = \varphi(Y(x_t)) \leq \varphi(y^0) + K_{m+1} \Delta_t < z_{m+1}^* + K_{m+1} \Delta_t.$$

Similar inequalities hold for z_{t-1} if $\nu(x_{t-1}) = m + 1$. Thus from (2.22)–(2.24) and (2.16), (4.8),

$$R(t(q^p)) > \Delta_t(r\mu_{m+1} - 4K_{m+1})/r\mu_{m+1} > 0, \quad p \geq p'. \tag{4.37}$$

Rule 6, (4.13) and (4.37) force the interval (x_{i-1}, x_i) to be subdivided by some point (2.27), and one of two subintervals obtained by this subdivision will satisfy (4.36), (4.37). As a consequence, y^0 has to be an accumulation point of $\{y^n\}_{n=1}^\infty$ and, from (4.35), some point $y^q \in D^0 \subset U_p(\bar{y})$ will inevitably meet the inequality $\varphi(y^q) < \varphi(\bar{y})$ which is a contradiction to the left equality in (4.9). Therefore, \bar{y} must not be an accumulation point of $\{y^n\}_{n=1}^\infty$ under the conditions (4.32). ■

NUMERICAL EXAMPLE. Figure 6 presents the implementation of the above technique for solving the *test problem*:

$$\begin{aligned} \text{minimize } \varphi(y) &= 1.5y_1^2 \exp[1 - y_1^2 - 20.25(y_1 - y_2)^2] + (0.5y_1 - 0.5y_2)^4 \\ &\quad (y_2 - 1)^4 \exp[2 - (0.5y_1 - 0.5)^4 - (y_2 - 1)^4], \\ \text{subject to } &1.21 \leq (y_1 - 2.2)^2 + (y_2 - 1.2)^2 \leq 2.25, \\ &0 \leq y_1 \leq 4, \quad -1 \leq y_2 \leq 3 \end{aligned}$$

(the objective function is from Uosaki *et al.* (1970)). The search was terminated by the rule $\Delta_i \leq 10^{-6}$. The maps employed ($L = 8$) were approximated by

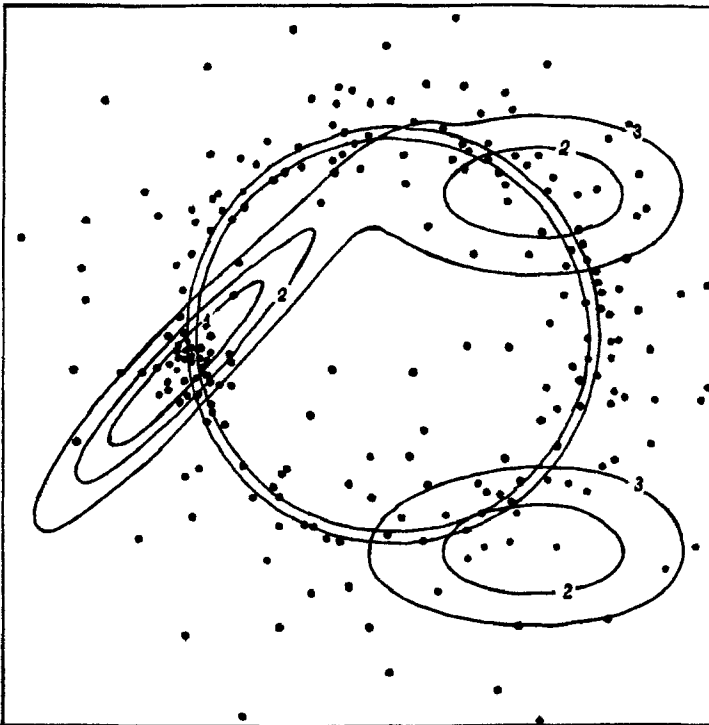


Fig. 6. The depicted square presents the domain of search, the narrow belt (lying inside the big circle and outside the small one) presents the admissible set. Level sets for the objective function are also plotted. Iteration points for the case $L = 8$ are marked by dark dots. The run with $L = 1$ (iteration points are not plotted) missed the global solution (cf. Theorem 3).

respective functions from (2.8) with $M = 10$. Reserves were given by $\varepsilon_i = 10^{-6}$, $0 \leq i \leq 2$, and $r = 2$.

After 150 iterations, the mixed scheme was employed to combine global iterations with local search (see Markin and Strongin (1987), and Strongin *et al.* (1988)). According to this scheme, each odd iteration is governed by the rules we have described. When performing the choice of even iterations, values $R(i)$ in (2.26) are replaced with

$$R^*(i) = R(i) / [(T(i))^{1/N} + 10^{-6} \mu_\nu],$$

where

$$T(i) = \begin{cases} (z_i - z_\nu^*)^2, & \text{if } \nu = \nu(x_i) > \nu(x_{i-1}), \\ (z_i - z_\nu^*)(z_{i-1} - z_\nu^*), & \text{if } \nu = \nu(x_i) = \nu(x_{i-1}), \\ (z_{i-1} - z_\nu^*)^2, & \text{if } \nu = \nu(x_{i-1}) > \nu(x_i), \end{cases}$$

and $R(i)$ is from Rule 5.

Acknowledgements

The author thanks the referees for valuable comments and D. Markin for programming and computing the test problem.

References

- Betro, B. (1991), Bayesian Methods in Global Optimization, *Journal of Global Optimization* **1**(1), 1–14.
- Butz, A. R. (1968), Space Filling Curves and Mathematical Programming, *Inform. Control* **12**(4), 314–330.
- Butz, A. R. (1971), Alternative Algorithm for Hilbert's Space Filling Curve, *IEEE Trans. Computers* **C-20**(4), 424–426.
- Evtushenko, Y. G. (1985), *Numerical Optimization Techniques*, Translation Series in Mathematics and Engineering, Optimization Software Inc., Publication Division, New York.
- Gergel, V. P., Strongin, L. G., and Strongin, R. G. (1987), The Vicinity Method in Pattern Recognition, *Engineering Cybernetics* (Transl. from *Izv. Acad. Nauk USSR, Techn. Kibernetika* **4**, 14–22).
- Hansen, P., Jaumard, B., and Lu, S. H. (1991), On Timonov's Algorithm for Global Optimization of Univariate Lipschitz Functions, *Journal of Global Optimization* **1**(1), 37–46.
- Horst, R. (1990), Deterministic Methods in Constrained Global Optimization: Some Recent Advances and New Fields of Application, *Naval Research Logistics Quarterly* **37**, 433–471.
- Korotkich, V. V. (1989), Multilevel Dichotomy Algorithm in Global Optimization, in *System Modelling and Optimization*, 161–169. Proceedings of the 14th IFIP Conference, Leipzig. Springer-Verlag.
- Kuhn, H. W. and Tucker, A. W. (1951), Nonlinear Programming, in Neyman, J., ed., *Proceedings of the Second Berkeley Symposium on Math. Stat. and Probab.* Univ. of Calif. Press, Berkeley and Los Angeles, Calif., 481–492.
- Markin, D. L. and Strongin, R. G. (1987), A Method for Solving Multi-Extremal Problems with Non-Convex Constraints, That Uses a Priori Information about Estimates of the Optimum, *U.S.S.R. Comput. Maths. Math. Phys.* **27**(1), 33–39. Pergamon Press 1988 (*Zh. vychisl. Mat. mat. Fiz.* **27**(1), 52–62, 1987).

- Pinter, J. (1988), Branch and Bound Algorithms for Solving Global Optimization Problems with Lipschitzian Structure, *Optimization* **19**, 101–110.
- Piyavskii, S. A. (1972), An Algorithm for Finding the Absolute Extremum of a Function, *U.S.S.R. Comput. Maths. Math. Phys.* **12**, 57–67 (*Zh. vychisl. Mat. mat. Fiz.* **12** (1972) 888–896).
- Rinnooy Kan, A. H. G. and Timmer, G. T. (1989), Global Optimization, in Nemhauser G. L. *et al.*, eds., *Handbooks in OR&MS* **1**, 631–662, North-Holland, Elsevier Science Publishers.
- Sergeev, Ya. D. and Strongin, R. G. (1987), Effective Algorithm for Global Optimization with Parallel Computations, in Kurzhanski, A., Neumann, K., and Pallaschke, D., eds., *Lecture Notes in Economics and Mathematical Systems* **304**, 158–162. Proceedings 1987. Springer-Verlag. Berlin, Heidelberg 1988.
- Sergeev, Ya. D. and Strongin, R. G. (1989), A Global Minimization Algorithm with Parallel Iterations, *U.S.S.R. Comput. Maths. Math. Phys.* **29**(2), 7–15. Pergamon Press 1990 (*Zh. vychisl. Mat. mat. Fiz.* **29**(3), 332–345, 1989).
- Strongin, R. G. (1973), On the Convergence of an Algorithm for Finding a Global Extremum, *Engineering Cybernetics* **11**, 549–555.
- Strongin, R. G. (1978), *Numerical Methods for Multiextremal Problems*, Nauka, Moscow (in Russian).
- Strongin, R. G. (1984), Numerical Methods for Multiextremal Nonlinear Programming Problems with Nonconvex Constraints, in Demyanov, V. F. and Pallaschke, D., eds., *Lecture Notes in Economics and Mathematical Systems* **255**, 278–282. Proceedings 1984. Springer-Verlag. IIASA, Laxenburg/Austria 1985.
- Strongin, R. G. (1989), The Information Approach to Multiextremal Optimization Problems, *Stochastics and Stochastics Reports* **27**, 65–82. Gordon and Breach Science Publishers, Inc.
- Strongin, R. G. (1990), *Search for Global Optimum*, Series of Mathematics and Cybernetics 2, Znanie, Moscow (in Russian).
- Strongin, R. G., Gergel, V. P., and Markin, D. L. (1988), Multicriterion Multiextreme Optimization with Nonlinear Constraints, in Lewandovski, A. and Volkovich, V., eds., *Lecture Notes in Economics and Mathematical Systems* **351**, 120–127. Proceedings 1988. Springer-Verlag. IIASA, Laxenburg/Austria 1991.
- Strongin, R. G. and Markin, D. L. (1986), Minimization of Multiextremal Functions with Nonconvex Constraints, *Cybernetics* **22**(4), 486–493. Translated from Russian. Consultant Bureau. New York.
- Sukharev, A. G. (1971), Best Sequential Search Strategies for Finding an Extremum, *U.S.S.R. Comput. Math. Math. Phys.* **12**(1), 39–59. Pergamon Press 1972 (*Zh. vychisl. Mat. mat. Fiz.* **12**(1), 35–50, 1971).
- Uosaki, K., Imamura, H., Tasaka, M., and Sugiyama, H. (1970), A Heuristic Method for Maxima Searching in Case of Multimodal Surfaces, *Technol. Repts., Osaka Univ.* **20**, 337–344.
- Žilinskas, A. (1981), On Multimodal Minimization Algorithm Constructed Axiomatically, *Methods of Operations Research* **40**, 197–200.