MATHEMATICAL CONTROL THEORY
BANACH CENTER PUBLICATIONS, VOLUME 14
PWN – POLISH SCIENTIFIC PUBLISHERS
WARSAW 1985

AN OPTIMAL SEQUENTIAL ALGORITHM FOR THE UNIFORM APPROXIMATION OF CONVEX FUNCTIONS ON $[0,1]^2$

G. SONNEVEND

Department of Numerical Mathematics and Computer Sciences, Eötvös University
Budapest, Hungary

§ 1. Introduction. Motivations and definitions

In applications (geology, microscopy, engineering design, construction of graphic displays, image transmission (coding) systems, and in mathematical (linear and nonlinear) programming, optimal control theory, differential games) the following problem often arises.

An unknown, real-valued function x is to be uniformly approximated over a domain G_0, say $G_0 = [0,1]^p$, based on N (point) evaluations

$$x(t_i) = c_i, \quad i = 1, \ldots, N, \quad t_i \in G_0 \tag{1}$$

for other measurement specifications see below Section 4) and of course, based on some a priori knowledge about x. This knowledge is usually expressed by $x \in K^0$, where K^0 is a compact subset of the Banach space of continuous functions over G_0 denoted by $C(G_0)$.

An N step algorithm of approximation consists in

(1) specifying the choice of $t^N = (t_1, \ldots, t_N)$, which might be passive, i.e., simultaneous, or sequential, i.e., given by a sequence of functions, $A^N = (A_1, \ldots, A_N)$ and rules

$$t_1 := A_N(K^0), \quad c_i := x(t_i),$$
$$t_i := A_{N-i+1}(t_j, c_j, j = 1, \ldots, i-1, K^0), \quad i = 1, \ldots, N \tag{2}$$

and

(2) giving a pointwise approximation a of x based on the information collected, i.e., on t^N, c^N and K^0, $c^N = (c_1, \ldots, c_N)$,

$$a(t, t_i, c_i, i = 1, \ldots, N, K^0), \quad t \in G_0. \tag{3}$$

Often the main cost of the algorithm is measured by N, i.e., complicated functions (computations) (A^N, a) are allowed in order to obtain a small *global error*

$$(4)\qquad \begin{aligned} &e(A^N, K^0) := \sup\{e(A^N, x) \mid x \in K^0\}, \quad e(A^N, x) := e\big(t^N(A^N, x), x\big), \\ &e(t^N, x) := \sup\big\{x(t) - a\big(t, t_i, x(t_i),\ i = 1, \ldots, N, K^0\big) \mid\ t \in G\big\}, \end{aligned}$$

where $t^N(A^N, x) = \{t_i(A^N, x),\ i = 1, \ldots, N\}$ are defined by (2).

The global error $e(t^N, K^0)$ of a *passive algorithm* $A^N = t^N$ is defined similarly. Of course, in the construction of an algorithm, especially if we are interested only in the order of magnitude of $e(A^N, K^0)$ with respect to N, N is regarded as a free parameter. Also of special interest are those algorithms in which $A^{N_{k+1}}$ can be regarded as a "continuation" of A^{N_k} (that is $t^{N_k}(A^{N_k}, x) \subseteq t^{N_{k+1}}(A^{N_{k+1}}, x)$ for all x) at least for a subsequence $N_k \to \infty$. For a more detailed and general definition of sequential approximation problems in B-spaces see [7], [10].

If the set K^0 is convex and has a centre (i.e., is central-symmetrical or balanced), then

$$e_s(N, K^0) := \inf_{A^N} e(A^N, K^0) = \inf_{t^N} e(t^N, K^0) =: e_p(N, K^0),$$

see [1], [7]. Here we show that for natural classes K^0 of *convex* functions (which are far from being central-symmetrical, but rather are like simplices) the use of sequential N-step algorithms allows us to obtain a *significantly* (with respect to order in N) smaller global error than the use of passive ones, see Theorem 1 below.

Here we assume that G_0 is the "interval" $[0, 1]^p$, and

$$(5)\qquad \begin{aligned} K^0 &= \{x \mid x \text{ convex in } t,\ x(t_i^0) = c_i^0,\ i = 0, 1, \ldots, 2^p, \text{ fixed}\} \\ &= K^0(x, G_0), \end{aligned}$$

where the points t_i^0 are the centre and the vertices of $[0, 1]^p$.

In fact, as a corollary, a good (optimal order) algorithm results, see Remark 1 below, for the more general classes

$$(6)\qquad \begin{aligned} K^0(m, M) = \{x \mid\ &\text{Spectrum}\, D_2 x \subseteq [m, M], x(t_i^0) = c_i^0, \\ &i = 0, \ldots, 2^p \text{ fixed}\}, \end{aligned}$$

where $D_2 x(t)$ denotes the Hesse matrix (of second derivatives). (5) is the special case $m = 0$, $M = \infty$, and in (6) $K^0(m, M)$ is in fact understood as the closure in $C(G)$ of the C^2 functions with everywhere defined $D_2 x$. The corollary comes from the observation that if z belongs to $K^0(m, M)$, then $x_1 = z - \frac{1}{2} m \|t\|^2$ and $x_2 = \frac{1}{2} M \|t\|^2 - z$ both belong to $K^0(0, \infty)$ (when not stated otherwise we use the usual notation $\|\ \|$ and $\langle u, v\rangle$ for euclidean norm and scalar product in R^p).

Let us note that a similar approximation problem for convex bodies x of R^{p+1} (say for the class K^0 of bodies lying in the unit ball B_1) is reduced

to the above case where each measurement gives the value $c_i = m(d_i, x)$ of the support function $m(d, x) = \sup\{\langle d, v\rangle |\ v \in x,\ d \in R^{p+1}\}$ in some direction d_i and the Hausdorff distance $\varrho(x_1, x_2)$ is used. This distance is identical to the uniform norm in $d \in S_1(R^{p+1}) =$ the unit sphere. An equivalent norm arises if d is restricted to the surface of the unit cube, and then we have our problem (1)–(5) because $m(d, x)$ is convex in d, see [8], [9]; $\|\mathrm{grad}_d\, m(d, x)\| \leqslant 1$, $d \in S_1$, $x \subseteq B_1$.

In order to have a compact set K^0 in $C(G)$ we should assume (because otherwise there are discontinuous functions in K^0) that in (5)

$$\|\mathrm{grad}\, x(t)\| \leqslant g_0 \quad \text{for} \quad t \in G_0 \text{ (a.e. in } G_0) \tag{7}$$

for some constant g_0 (this restricted set will be denoted by $K^0(g_0)$). For reasons which will be clear later (we do not need any information about g_0 in constructing proposed algorithm), we do not include (7) in (5). Of course the finiteness of the functional $D(x, G)$ (see (16) below) which implies the optimal order $(-2/p)$ of the error, is guaranteed a priori if (7) is assumed. Note that, by the convexity in t, the functions x in K^0 are continuous and uniformly bounded on any compact subset of the open domain $(0, 1)^p$.

Instead of computing the optimal function a (defined by putting in (4) an infimum with respect to $a(\cdot)$ before the supremum) we shall — for computational simplicity and without loss (as regards the order in N of the global error) — define the function a used in (4) "locally" over the "subintervals" G_i^N, $i = 1, 2, \ldots$, into which G is decomposed (see below) to be the analogue $a(G_i^N, x)$ of the function $a(G_0, x) := a(K^0, \cdot)$ defined by

$$a(K^0, t) := \sup\{x(t) |\ x \in K^0\}, \quad t \in G_0 \tag{8}$$

for $K^0 = K^0(x, G_0)$. Of course $a(K^0, t)$ (thus each $a(G_i^N, x)$) is a convex, piecewise linear function over G_0, (resp. G_i^N), yet in general the function a thus constructed will not be convex in t. Note that for $p = 2$ the function $a(K^0, t)$ is composed from at most 4 linear pieces. The function $e(t^N, x)$ from (4) is thus giving the diameter of the set of localization $l(t^N, x)$ in the space $L(G_0)$.

Many of the constructions and methods of proof which will be presented below are valid for arbitrary values of the dimension p. The solution for the case $p = 1$, which is by no means trivial, has been given in earlier works, [8], [9], and can easily be obtained (reconstructed) by using the methods (hints) given below. Here we solve the case $p = 2$ completely. We also indicate in Section 4 those points whose generalization for $p \geqslant 3$ is not yet obtained.

PROPOSITION 1. *For the class* (5), *even when* (7) *is assumed, there exist*

constants $K_i(K^0, g_0)$, $i = 0, 1$, *such that*

$$e_p(N, K^0) \geqslant k_0 N^{-1/p} \quad \text{and} \quad e_s(N, K^0) \geqslant k_1 N^{-2/p},$$

where k_i *are positive if* K^0 *is nontrivial.*

The class K^0 is nontrivial if there exist two different functions x_1 and x_2 in it. The second inequality is then the consequence of the asymptotics (12) (see [5]) because one can construct a function x which is strongly convex and C^2 (has a constant, positive definite Hesse matrix) over a subdomain of G_0. To construct x let $\max(x_2 - x_1)$ be realized at a point t. By the gradient bound in (7) this point is interior to $G_1 = \{t |\ x_2(t) > x_1(t)\}$. The support hyperplane to the epigraph of x_2 at $\big(t, x_2(t)\big)$ is over x_1 in a convex subdomain G_2 of G_1 and a quadratic function z is chosen so that

$$z(t) = \tfrac{1}{2}\big(x_1(t) + x_2(t)\big), \quad \operatorname{grad} z(t) = \operatorname{grad} x_2(t)\,(1-\varepsilon),$$
$$D_2 z(t) \equiv d_2 I, \quad d_2 > 0.$$

Here ε, d_2 are so small that, on the boundary of G_2, z is smaller than x_1. Then $x := \max(z, x_1)$ is convex on G_2 and quadratic over a subdomain G_3 of G_2. To prove the first inequality —for simplicity for the case $p = 2$— we first choose in G_3 a square $T = (P_1, P_2, P_3, P_4)$ so that $z(P_1) = z(P_2)$, $z(P_3) = z(P_4)$, which is always possible by turning T through some angle about its centre. It is only important that the points $\big(P_i, z(P_i)\big)$, $i = 1, \ldots$ $\ldots, 4$, should lie in the same plane. Let w be the maximum of convex functions over G_3 which take at the vertices $P_1, \ldots, P_4$ and at its centre P_0 the same values as z. That is, w is linear over the four triangles $[P_0, P_i, P_{i+1}]$, and we can extend w beyond G_3 to the whole G_0, yielding in K^0 a function $\bar{w}$. If N is large then in one of the 4 triangles formed by the centre and a side of T there exists a square U similar and similarly posed to T, with side length not smaller than $r_1 N^{-1/2}$, where r_1 depends only on $K^0(g_0)$, inside of which there are no points from t^N. Let $U = (U_1, U_2, U_3, U_4)$ be a square, where $\overline{U_1 U_2}$ is parallel to $\overline{P_1 P_2}$ and nearer to $\overline{P_1 P_2}$ than to $\overline{P_3 P_4}$. Let x_3 be the maximum of the convex functions (defined over G_3) which take the same values at the points P_1, P_2, $\dfrac{U_1 + U_3}{2}$, $\dfrac{U_2 + U_4}{2}$, P_3, P_4 as the functions z, z, w, w, z, z, resp., and let x_4 be the maximum of the convex functions which take the values of x_3 at $P_1, P_2, U_1, U_2, U_3, U_4, P_3, P_4$; let both of them continue to be equal to $\bar{w}$ outside G_3. The estimation $\|x_3 - x_4\| \geqslant r_1 N^{-1/2} r_2$, r_2 depending on d_2 and G_3 (which corresponds to the one-dimensional case of the construction) is easily proved. For $p > 2$ a similar construction can be given, with simplices (rather than cubes) in the role of the sets G_0, G_3, T.

Thus an algorithm A^N or, more precisely, a sequence of algorithms $\bar{A}^\infty = (\bar{A}^1, \bar{A}^2, \ldots)$ is rightly said to be *optimal with respect to order* if, for

$x \in K^0(g_0)$,

(10) $\quad e(\bar{A}^N, x) \leqslant k_2(x) N^{-2/p}$ with $k_2(x) \leqslant k_2\big(K^0(g_0), \bar{A}^\infty\big)$,

where the constant $k_2\big(K^0(g_0), \bar{A}^\infty\big)$ is independent of N. In Theorem 1 (see (19)) we prove that for $p = 2$ a simple algorithm is optimal (modulo a small factor $\log N$).

§ 2. Description of the algorithm

The notion of individually best N point approximation, $a(\bar{t}^N, x)$, is defined by the extremal problem, for given (known) x

(11) $\quad e(N, x) := \inf_{t^N} e(t^N, x) = e(\bar{t}^N, x) \geqslant \big\|x - a_s\big(\bar{t}^N, x(t^N), K^0\big)\big\|_{C(G_0)}$.

For an analogous, essentially equivalent problem of approximation of strongly convex C^2 smooth bodies by polyhedra in the Hausdorff metrics, the asymptotics of $e(N, x)$ for $N \to \infty$ and arbitrary p has been obtained in [5]. In our case it is not difficult to prove that

(12) $\quad e(N, x) = r_p N^{-2/p} \Big(\int_G \sqrt{\det D_2 x(t)}\, dt\Big)^{2/p} + o(N^{-2/p})$,

where the constant r_p depends only on p. Roughly speaking, the best system of N nodes, for N large, must be such that their linear density, at a point t, along the principal directions (eigenvectors) of $D_2 x(t)$ is proportional to the square root of the corresponding eigenvalue. Of course our aim is to get exact upper bounds for finite N, and at present we do not know whether for the multidimensional case, $p \geqslant 2$, it is possible to construct an algorithm satisfying (10) *and* yielding, for each x with continuous and positive definite second derivative, asymptotically optimal nodes (for $p = 1$ this is true).

The nodes yielded by algorithm 1 below are not such; however, they are chosen according to the *principle of equal local errors* (see [1]) which is derived from an important property of the optimal set of nodes.

In order to describe algorithm 1 (for simplicity for the case $p = 2$), we need some preparation. Let x be an arbitrary element of K^0 and let us denote by $K(G, x)$ the set of convex functions z defined over an interval $G = [a_1, b_1] \times [a_2, b_2]$, of length $l(G) = b_1 - a_1 = b_2 - a_2$, for which the values $z(S_i^G) = x(S_i^G)$, $i = 0, 1, 2, 3, 4$, are fixed, where S_i^G are the centre $\frac{1}{2}(a_1 + a_2, b_1 + b_2)$ and the vertices of G. Thus $K^0 = K(G_0, x)$. The diameter $d(G, x)$ of $K(G, x)$ in $C(G)$ (more precisely in $L^\infty(G)$) is easily computed as the maximum of the distances of the value of x at a vertex of G from the value of the linear function (at that vertex) which is determined by the values of x at the centre and at the opposite vertex of G. By a (re-

gular) subdivision of G — for given x — we mean the generation of four "standard" sets $K(G_j, x)$, $j = 1, 2, 3, 4$, from the set $K(G, x)$, where each G_j is a half-size interval such that

$$\bigcup_{j=1}^{4} G_j = G, \quad \lambda(G_j \cap G_i) = 0 \quad \text{for} \quad i \neq j \tag{13}$$

(λ is the Lebesque measure). Let us define the indices 1, 2, 3, 4 in such a way that their growing order corresponds to the clockwise movement round G, G_1 being the down left interval, G_3 the upper right one. These sets are thus defined if we compute x, in addition to the five already fixed values of x in $K(G, x)$, at eight new points.

ALGORITHM 1. Choose a (small) positive number ε. Define by sequential subdivisions a sequence of points $t_1^n, t_2^n, \ldots, t_{8n+5}^n$, and interval subdivisions $\{G_0^n, \ldots, G_{3n}^n\}$ inductively as follows.

Let $G_0^0 = G_0 = [0, 1]^2$, $t_1^0, \ldots, t_5^0$ being given as in the definition of K^0. Suppose that $G_0^n, \ldots, G_{3n}^n$ are already defined and such that

$$\bigcup_{i=0}^{3n} G_i^n = G_0, \quad \lambda(G_i^n \cap G_j^n) = 0 \quad \text{if} \quad i \neq j. \tag{14}$$

We order the lower indices i in the following way: say that G_i^n is before G_j^n if, for an interval $G = G_k^m$, $m \leqslant n$ (see (13)) we have $G_i^n \subseteq G_\alpha$ and $G_j^n = G_\beta$ for $1 \leqslant \alpha < \beta < 4$. Now let $i = i(n)$ be the smallest, in this sense, index such that the uncertainty in the values of x over G_i^n, as computed from the five measurements (points $(s_j, x(s_j))$, $s_j \in G_i^n$), i.e., $d(G_i^n, x)$ is greater than ε. In fact, as will be clear below, we could use the following criterion as well (replacing $d(G, x)$ by $\Delta(x, G)$):

$$\Delta(x, G) := x(s_1^G) + x(s_2^G) + x(s_3^G) + x(s_4^G) - 4x(s_0^G) > \varepsilon \quad \text{for } G = G_{i(n)}^n. \tag{15}$$

Define now the next eight new measurement points, t_{8n+5+j}, $j = 1, \ldots, 8$, as those eight ones which are needed in the subdivision of $G_{i(n)}^n$. Thus are defined the new points t_j^{n+1}, $j = 1, \ldots, 8(n+1)+5$ and the new subdivision

$$G_j^{n+1}, \quad j = 0, \ldots, 3(n+1). \tag{14}'$$

The algorithm terminates at some (mega) step $N = N(\varepsilon, x, G_0)$ when after $8N$ function evaluations there exists no value $i(N)$ satisfying (15). Then obviously x is approximated by the locally constructed functions $a(K(G_j^N), t)$ (see (8)) over each G_j^N within accuracy ε.

§ 3. Error estimation for the algorithm

An important role will be played by the functional

$$D(x, G) = \int_{\partial G} \langle \operatorname{grad} x(u), n(u) \rangle du, \tag{16}$$

defined for an arbitrary interval G in G_0 and function x in K^0. Here $\langle \operatorname{grad} x(u), n(u) \rangle$ is interpreted as an arbitrary value of the subgradient of the convex function x of t in the direction of the outer normal $n(u)$. Thus the smallest value is

$$I_2 + I_1 = \int_{a_2}^{b_2} \left[\frac{\partial x}{\partial s_1}\right] ds_2 + \int_{a_1}^{b_1} \left[\frac{\partial x}{\partial s_2}\right] ds_1, \quad \text{where} \quad t = (s_1, s_2),$$

$$\left[\frac{\partial x}{\partial s_1}\right](s_2) = \frac{\partial x(b_1 - 0, s_2)}{\partial s_1} - \frac{\partial x(a_1 + 0, s_2)}{\partial s_1},$$

$$\left[\frac{\partial x}{\partial s_2}\right](s_1) = \frac{\partial x(s_1, b_2 - 0)}{\partial s_2} - \frac{\partial x(s_1, a_2 + 0)}{\partial s_2}.$$

It is well known that for C^2 smooth functions

$$D(x, G) = \int_G \Delta x(t) dt, \qquad \Delta = \frac{\partial^2}{\partial s_1^2} + \frac{\partial^2}{\partial s_2^2} \tag{17}$$

is the Laplace operator (otherwise Δx should be understood in a generalized sense; see e.g. [4]).

In fact, for a function x in $K^0(x, G)$, which might be discontinuous at the boundary of G, the value (16) can be defined (and be finite!) by a limit procedure

$$D(x, G) = \lim_{n=\infty} D(x_n, G), \quad \text{where} \quad x_n \in C(G), \quad x_n \to x \quad \text{for} \quad n \to \infty,$$

uniformly on compact subsets of the interior of G. However, without assuming an upper bound g_0 for the gradients over G, $e(A^N, K^0)$ cannot even be made to converge to zero for $N \to \infty$.

THEOREM 1. *The number of function evaluations, i.e., steps $N = N(\varepsilon, x, G)$ necessary to finish algorithm* 1 *over G, for an arbitrary convex function x with gradients less than g_0, is estimated from above by*

$$N(\varepsilon, x, G) \leqslant N_0(\varepsilon, x, G) := \left[8 \frac{D(x, G)}{\varepsilon} \log \frac{2^{5/6} l(G) g_0}{\varepsilon}\right]_+, \tag{18}$$

where $[\ ]_+$ *denotes the nonnegative entire part and* log *is of base* 2.

COROLLARY 1. *Choosing ε as the solution of* (18) *for given* N, $N_0(\varepsilon, x, G_0) = N$, *we obtain an* N*-step algorithm* $\bar{A}^N$ *whose global error is almost of order* $-l = -2/p$, (*because of* $D(x, G_0) \leqslant l(G_0) g_0$, $-D \log D \leqslant 1/2$,

$$e(\bar{A}^N, K^0) \leqslant \frac{4}{N} + \frac{32 l(G_0) g_0}{N} \log 2^{5/8} N l(G_0) g_0. \tag{19}$$

Choosing $\varepsilon = \varepsilon_k = 4^{-k}\varepsilon_0$ and performing algorithm 1 with $\varepsilon = \varepsilon_k$, $k = 1, 2, \ldots$, we obtain an infinitely continuable algorithm, for which at each step N essentially the same inequality (19) holds. For the proof of the theorem we need two lemmas.

LEMMA 1. *For arbitrary* G *and* x *the following inequalities hold*:

$$d(G, x) \leqslant \Delta(x, G) \leqslant D(x, G). \tag{20}$$

Proof. The first inequality from the left was proved already in the description of algorithm 1 (by noting that the maximum of two non-negative numbers is less than their sum). The second part can be interpreted as the statement that the simplest discrete approximation of the integral (17) (which is based on five evaluations of x to get a discrete approximation for the Laplace operator at the centre of G) is always less than the exact value of the integral thus approximated. The idea of the following proof was communicated to the author by T. Fiala.

Let v_i be the unit vector pointing from the centre s_0^G of G to the vertex s_i^G. Then

$$x(s_i^G) = x(s_0^G) + \sqrt{2} \int_0^c \operatorname{grad} x(s_0^G + t\sqrt{2}\, v_i), \quad v_i > dt, \quad c = \frac{l(G}{2}. \tag{21}$$

Since

$$\langle \operatorname{grad} x(s), v_i \rangle = \frac{1}{\sqrt{2}} \left(\pm \frac{\partial x(s)}{\partial s_1} \pm \frac{\partial x(s)}{\partial s_2} \right),$$

with $\pm$ depending on i, we obtain (20) by using the *monotonicity* of $\dfrac{\partial x(s)}{\partial s_1}$ (resp. $\dfrac{\partial x(s)}{\partial s_2}$) in s_1 (resp. s_2) and summing (18) with respect to i. It is not difficult to construct many functions x for which

$$2\Delta(x, G) = D(x, G).$$

LEMMA 2. *For arbitrary nonnegative numbers* z_1, z_2, z_3, z_4, L *such that*

$$z_1 + z_2 + z_3 + z_4 > 1 \quad \textit{and} \quad L \geqslant 1$$

we have

$$\sum_i [8z_i \log L]_+ \leqslant \Big[8\Big(\sum_i z_i\Big)\log 2L\Big]_+ - 8. \tag{22}$$

Proof. By the simple inequalities $\log L \geqslant 0$, and $\log 2 = 1$

$$\sum [u_i]_+ \leqslant \Big[\sum u_i\Big]_+ \quad \text{for} \quad u_i \geqslant 0.$$

Proof of Theorem 1. We show by induction with respect to the value of N that (18) holds. Let us assume first that $N_0(\varepsilon, x, G) = 0$. Then because of the fact that we can assume

$$l(G) g_0 \frac{1}{\varepsilon} \geqslant \sqrt{2}^{-1} \tag{23}$$

(otherwise formula (16) shows by Lemma 1 and a simple inequality that $d(x, G) < \varepsilon$ and thus $N(\varepsilon, x, G) = 0$), it follows that $D(x, G)$ must be less than ε, and thus, again by Lemma 1, $d(x, G) < \varepsilon$, i.e., $N(\varepsilon, x, G) = 0$. The function $N(\varepsilon, x, G)$ satisfies (for a subdivision (13))

$$8 + \sum_{j=1}^{4} N(\varepsilon, x, G_j) = N(\varepsilon, x, G).$$

Thus we have to prove that if $N(\varepsilon, x, G) > 1$, i.e., $d(x, G)$ (or $\Delta(x, G)$) is greater than ε, then

$$8 + \sum_{j=1}^{4} N_0(\varepsilon, x, G_j) \leqslant N_0(\varepsilon, x, G). \tag{24}$$

This, by the *additivity* of the functional $D(x, G)$ with respect to subdivisions, is a consequence of Lemma 2; put

$$z_i = D(x, G_i)\varepsilon^{-1}, \quad L = L(G_i) = \frac{g_0}{\varepsilon} \cdot 2^{5/8}(G_i) = \tfrac{1}{2} L(G).$$

Remark 1. For the class $K^0(m, M)$ (see (6)) with finite m, M the transformations $z \to x_1(z, m)$, $z \to x_2(z, M)$ lead to the following algorithm. Perform simultaneously for x_1 and x_2 algorithm 1 as described, replacing the criterion of subdivision (15) by

$$\min\big(\Delta(x_1, G),\ \Delta(x_2, G)\big) > \varepsilon \quad \text{for} \quad G = G^n_{i(n)},$$

and define the approximation $a(G, z)$ for each $G = G_i^N$ as

$$\tfrac{1}{2} m \|t\|^2 + a(G, x_1), \quad \tfrac{1}{2} M |t|^2 - a(G, x_2), \text{ resp.}$$

Remark 2. A simpler, infinitely continuable variant of algorithm 1 is the following. Instead of fixing a number ε, let us perform the sequence

of subdivisions according to the following rule: let $i(n)$ be that value of i for which $d(x, G_i^n)$ (or $\Delta(x, G_i^n)$) is maximal. The number of subdivisions (steps) $N_1(\varepsilon, x, G)$ needed to reach a global accuracy ε can be estimated by the same inductive method and turns out to be majorated by the same quantity $N_0(\varepsilon, x, G)$.

Let us note that, as the previous proof indicates, the logarithmic factor cannot be eliminated from the error estimations (18), (19). It is easy to see that for each fixed x in K^0 with a continuous and positive definite Hesse matrix, we have asymptotically for $N \to \infty$ (see (17))

$$e(t^N, x) = \frac{3}{4N} D(x, G_0) + o(N^{-1}), \qquad t^N = t^N(\bar{A}^N, x). \tag{25}$$

Indeed, because of (15), (17), for large N and continuous Δx, we have

$$\int \Delta x(t)\, dt \cong \left(3\frac{N}{8} + 1\right) 2\varepsilon. \tag{26}$$

It seems true that there exists no simple optimal order algorithm for which the asymptotic error would have the same main part as in (12); note that

$$\Delta x(t) \geqslant 2\left(\det D_2 x(t)\right)^{1/2} \quad \text{with equality iff} \quad \frac{\partial^2 x}{\partial s_1 \partial s_2} = 0, \tag{27}$$

and

$$\frac{\partial^2 x}{\partial s_1^2} = \frac{\partial^2 x}{\partial s_2^2}.$$

The reasons behind this statement and further motivation concerning our algorithm 1 (its stability properties) will be given in the next section.

§ 4. Generalizations and further comments on the algorithm

The most interesting question is of course whether or not the straightforward generalization of algorithm 1 to the $p \geqslant 3$ dimensional case yields an optimal algorithm (with respect to the order of the global error) and, if so, whether that order is equal to $(-2/p)$, the lower bound given by (12). While the answer to the first question seems to be positive, concerning the second one the following remarks ("negative results", resp. conjectures) are stated, indicating that algorithm 1 yields the optimal order $(-2/p)$ only for those classes of functions x for which $\|\Delta x\|_{L_{p/2}(G)} < \infty$ uniformly.

A generalization of inequality (20), i.e., Lemma 1, would be

$$d(x, G) \leqslant \Delta_p(x, G) \leqslant c_p l^{2-p}(G) D(x, G) \tag{28}$$

for some constant c_p, where $D(x, G)$ is given by the obvious generalization of (17) and (16),

$$\Delta_p(x, G) := \sum_1^{2^p} x(P_i^G) - 2^p x(P_0^G).$$

The first part is again easy. As regards the second part, the exact value of c_1 is 1/2 (here $d(x, G) = \Delta_1(x, G)$ and—as noted after (21) — the exact value of c_2 seems to be also 1/2). Now, in order to prove that algorithm 1 has an error of optimal order $(-2/p)$,

$$N_0(\varepsilon, x, G) := k_p \left(\frac{l^{2-p}(G) \cdot D(x, G)}{\varepsilon} \right)^{p/2} \left(1 + o(\varepsilon^{-\delta})\right), \qquad \delta > 0,$$

the following inequality should hold (generalization of Lemma 2, i.e., of the *subadditivity* property of N) (of course modulo some logarithmic factor):

$$D_p(x, G) := \left(l^{2-p}(G) D(x, G)\right)^{2/p},$$

$$\sum_{j=1}^{2^p} D_p(x, G_j) \leqslant D_p(x, G) \quad \text{if} \quad G = \bigcup_{j=1}^{2^p} G_j \tag{29}$$

for regular subdivisions and all x. Now (29) requires nothing but the norm inequality $(a_j = D_p^{2/p}(x, G_j))$, $\|a\|_{L_1}$, which is not true when $p \geqslant 3$. By Holder's inequality $D_p(x, G_j) \leqslant \|\Delta x\|_{L_{p/2}(G)}$, and the functional $\Delta^{p/2}(x, G)$ is obviously additive in G. The conjecture that the main part of N_0 is $k_p \varepsilon^{-p/2} \Delta^{p/2}(x, G)$, especially that $\Delta x \in L_{p/2}(G)$ is needed, is seen from Hadamard's inequality $\Delta^{p/2} x(t) \geqslant lp \sqrt{\det D_2 x(t)}$, and from the fact that the *Green function* for the Laplace operator having the singularity $(1/|t-s|)^{p-2}$, $p > 2$, is only in $L_{(p-\delta)/(p-2)}$, the dual of $L_{p/2+\lambda}$, $\delta, \lambda > 0$.

Let us see whether for the functional giving the main part of the individually best approximations (12)

$$A_p(x, G) := \left(\int_G \sqrt{\det D_2(x, t)}\, dt \right)^{2/p} \tag{30}$$

there exists another upper bound (see (27)) which is expressible in terms of (the values and) the gradients of x at the boundary of G. If the values of x are fixed along the boundaries of the intervals belonging to a subdivision, then these values become independent for the different subintervals; thus subdivision means decomposition, in a well-defined sense. If one applies the Cauchy–Schwarz inequality to the integral (30), then one obtains

$$A_p(x, G) \leqslant H_p(x, G) := \left(\mathrm{vol\,grad}\, x(G) \cdot \mathrm{vol}(G)\right)^{1/p}, \tag{31}$$

where vol means the Lebesgue measure and $\operatorname{grad} x(G)$ is the image of the set G in the "monotone" map $t \to \operatorname{grad} x(t)$. Notice that $H_1\big(x, (a, b)\big) = \big(x'(b) - x'(a)\big)(b - a) = D_1(x, G)$. Because of the monotonicity of $\operatorname{grad} x$,

$$\operatorname{grad} x(G) = \bigcup_{i=1}^{2^p} \operatorname{grad} x(G_i), \qquad \operatorname{vol}\big(\operatorname{grad} x(G_i) \cap \operatorname{grad} x(G_j)\big) = 0, \qquad j \neq i,$$

$H_p(x, G)$ satisfies the required subadditivity property

$$\sum_{j=1}^{2^p} H_p^{p/2}(x, G_j) \leqslant H_p^{p/2}(x, G), \tag{32}$$

because of $\|a\|_{L_q} \leqslant \|a\|_{L_1}$ for $q = 1/2$, $a \in R^{2^p}$.

Take some constant k_p for an algorithm using subdivisions for a measurement pattern in which at each step over an interval G_n we get and use information I concerning the values of x over G, $I(x, G_n) = c_n$, and accepting or subdividing the interval, depending on the value of some "barrier" functional $\Delta_p^I(x, G_n) \leqslant \varepsilon$. In order to prove the inequality

$$N(\varepsilon, x, G) \leqslant k_p \left(\frac{H_p(x, G)}{\varepsilon}\right)^{p/2}, \tag{33}$$

it would be enough to prove the inequality (the analogue of Lemma 1)

$$d_p^I(x, G_n) := d\{z \mid I(z, G_n) = I(x, G_n), z \in K^0\} \leqslant q_p \Delta_p^I(x, G_n) \leqslant r_p H_p(x, G_n), \tag{34}$$

where d stands for the diameter of a set in $L_\infty(G_n)$, and q_p, r_p are constants.

Indeed, for $p = 1$, using at each step three points

$$\Delta_1^I := \Delta_1(x, a, b) = 2d\big(x, (a, b)\big) = x(a) + x(b) = 2x\big((a+b)/2\big), \tag{35}$$

one obtains (33) with $k_1 = 1/2$. In the inductive proof for arbitrary p, $N_0(\cdot)$ is chosen in the form

$$N_0(\varepsilon, x, G) = \left[u\left(\frac{H_p(x, G)}{\varepsilon}\right)^{p/2} - v\right]_+$$

with appropriate positive constants u, v. The constant v is needed (used) in order to compensate for the number of measurements needed in a subdivision and not included in the left side of (32), see (24).

In the case $p = 2$, inequality (34) is not true, i.e., cannot be satisfied for the information pattern used in algorithm 1.

It seems true that the only information pattern I for which (34) holds is the complete one, I^c, i.e., if at step n we measure all the values $x(t)$ and subgradient $\operatorname{grad} x(t)$ along the boundary of some interval G_n, then of course (as e.g. in the case $p = 1$) we could choose Δ_p^I to be d_p^I or H_p as well.

Any discretized version of the algorithm in which the subgradients are approximated by finite differences seems (except the case $p = 1$) to be far from giving the optimal order with respect to the number of function evaluations; this will be partly explained below when we speak about the instabilities.

In the case $p = 2$, when the information I^m used at step n consists of the values of x along the boundary of G_n and along its two middle lines, Δ_2^I being taken as d_2^I, (34) is not satisfied for any r_2, yet the role of the functional H_2 can be taken up by

$$F_2(x, G) := \left(I_1(x, G) \cdot I_2(x, G)\right)^{1/2}, \tag{36}$$

where I_1, I_2 are the integrals defined earlier in (16), (17). Concerning the functional F_2 let us prove that in the case of the complete information pattern I^c inequality (34) with H_2 replaced by F_2 holds with $d_2^{I^c} = \Delta I_2^c$ and $r_2 = 2$. Indeed, for arbitrary x, z, with $I^c(x) = I^c(z)$, $x, z \in C^2(G)$, we have

$$|x(t)-z(t)| \leqslant 2\sup \iint_G \left|\frac{\partial^2 z(t)}{\partial s_1 \partial s_2}\right| dt \leqslant 2\sup \iint_G \sqrt{\frac{\partial^2 z}{\partial s_1^2}\frac{\partial^2 z}{\partial s_1^2}}\, dt \leqslant 2F_2(x, G),$$

by the convexity of z, and then the Cauchy–Schwarz inequality yields the last part by the definition of F_2.

If we compute (use) not the values of the gradients (along the boundary) but the values of x along the two middle lines then for the corresponding algorithm as in the proof of Theorem 1 (using the subadditivity of F_2) we can prove that

$$N_2^{I^m}(x, \varepsilon, G) \leqslant \left[w_1 \frac{F_2(x, G)}{\varepsilon} \log \frac{w_2 l(G) g_0}{\varepsilon}\right]_+. \tag{37}$$

Even this algorithm cannot be discretized (to give one with an optimal order) because of the following *instability*. If the values of x are known (along the middle lines and the boundary of x) only within some accuracy ε_0 (e.g., this happens if for their approximation the one-dimensional algorithm (35) is used), then the perturbation of the value of $d_2^{I^m}$ is not bounded by a quantity proportional to ε_0.

Our algorithm 1 is *stable* with respect to errors in the measurement of $x(t_i) = c_i$. More precisely: suppose that the values c_i, $i = 1, \ldots, N$, are measured within accuracy ε_0; then, by using, as in algorithm 1, five such values over the current interval G, the value of the uncertainty in $x(t)$ over G, $d(x, G, \varepsilon_0)$ can be computed easily. Let us subdivide G in such a way that (for a given ε)

$$d(x, G, \varepsilon_0) \leqslant \varepsilon + 4\varepsilon_0.$$

For this algorithm, the number of steps $N(\varepsilon, x, G)$ is estimated by the same expression as when $\varepsilon_0 = 0$. Thus in the case of measurement errors of magnitude M^{-1}, to obtain over G_0 a global accuracy $5M^{-1}$ (choose $\varepsilon = \varepsilon_0 = M^{-1}$) no more than $k(g_0, K^0)M\log M$ steps are needed (by (19)).

References

[1] N. S. Bahvalov, *Numerical Methods*, Nauka, Moscow 1973.

[2] C. de Boor and J. R. Rice, *An adaptive algorithm for multivariate approximation*, J. Approximation Theory 25, 4 (1979), 337–360.

[3] C. A. Micchelli and T. J. Rivlin, eds., *Optimal Estimation in Approximation Theory*, Plenum Press, New York 1977.

[4] A. V. Pogorelov, *Minkowski's Problem in Higher Dimensions*, Nauka, Moscow 1975.

[5] R. Schneider, *Zur Optimalen Approximation Konvexer Hyperflächen durch Polyeder*, in: Proc. Conf. on Discrete Geometry, Salzburg, 1980, preprint of the University of Freiburg.

[6] G. Sonnevend, *On the optimization of algorithms for function minimization*, J. Comput. Math. Math. Physics 17, 3 (1973), 591–609.

[7] —, *On the optimization of adaptive algorithms*, in: *Operations Research Verfahren*, vol. 31, Athenäum, 1978, 581–595.

[8] —, *Existence and computation of extremal invariant sets*, in: Lect. Notes in Control and Information Sciences, vol. 22, Springer-Verlag, 1979, 251–261.

[9] —, *Sequential, optimal order algorithms for the uniform approximation of function with convex r-th derivatives*, preprint No 2/1980, Dept. Numer. Math. and Comp. Sci., Eötvös University, Budapest, 1–56; to appear in Anal. Math.

[10] J. F. Traub, H. Woźniakowski, *A General Theory of Optimal Algorithms*, Academic Press, 1980.