

On Global Optimization Using Interval Arithmetic

H. Munack, Tübingen

Received November 8, 1991; revised April 27, 1992

Abstract — Zusammenfassung

On Global Optimization Using Interval Arithmetic. A method for finding all global minimizers of a real-valued objective function of several variables is presented. For this purpose a problem-oriented type of number is used: the set of real compact intervals. The range of the objective function over a rectangular set is estimated by natural interval extension of a suitable modelling function. An algorithm for interpolation and approximation in multidimensional spaces is developed. This optimization method can be applied successfully to conventionally, e.g. with real arithmetic, programmed functions.

Key words: Global optimization, interval arithmetic, inclusion function

Über globale Optimierung mittels Intervallarithmetik. Vorgestellt wird eine Methode zum Auffinden aller globalen Minimierer einer reellwertigen Zielfunktion mehrerer Veränderlicher. Dabei wird ein problemorientierter Zahlentyp verwendet: der Bereich der reellen kompakten Intervalle. Der Wertebereich der Zielfunktion über einem rechteckigen Gebiet wird durch natürliche Intervallerweiterung einer geeigneten Modellfunktion abgeschätzt. Es wird ein Algorithmus zur Interpolation und Approximation in mehrdimensionalen Räumen entwickelt. Diese Optimierungsmethode kann auf herkömmliche, das heißt mit reeller Arithmetik programmierte, Funktionen erfolgreich angewendet werden.

1. Introduction

Computer simulations have become a convenient tool for the design of complex systems in various fields. Since capabilities of modern computer systems are increasing rapidly, the modelling techniques become more complex, too. The designer not only wants to simulate a single device but also demands an optimal result. This intention requires a systematic variation of all variables used in the objective function simulating the real world system. In the following we are treating minimization problems, which means that the variations have to produce decreasing values of the objective function. These variations can be done manually [10] or by an automatic minimization procedure [7]. Manual optimization assumes a throughout knowledge of the behaviour of the objective function. Nevertheless, in cases of two or more variables this method would generally produce unsatisfactory results. Therefore, many automatic minimization algorithms have been developed which tend to be stable and fast [5]. These so-called *local* methods proceed descending

from a starting point to a single end point by using classical algorithms like gradient methods, Newton methods, etc. The end point—the local minimizer—is to be reached using a minimum number of function evaluations considering the high computational expenses for each intermediate system calculation. The disadvantage of local minimization, however, is the uncertainty whether there are other local minimizers which have been missed due to an unfavourable choice of the starting point.

Global minimization, on the other hand, means the common tradition to isolate *all global* minimizers. Nevertheless, the investigation of additional (local) minimizers is neglected although they can be of interest in the treatment of real world problems, e.g. if the technical realization of a suitable system corresponding to a local minimizer can be carried out with less effort. The importance of global optimization, however, has not been accepted commonly. One reason is probably the so-called ‘curse of dimensionality’. The few existing global methods require a huge amount of computer memory and computer time which is increasing extremely with the number of dimensions of the configuration space. A second reason is that these methods cannot guarantee the reliable solution of the global problem. Since the whole area of interest has to be checked for local minimizers, most of the deterministic methods, such as dynamic programming [21], use a mesh whose nodes represent points in the variable space. Depending on the coarseness of this mesh, global minimizers can be lost. The same holds for stochastic algorithms like Monte-Carlo-methods, even those which are combined with local deterministic procedures, since they are limited by the finite number of random searches.

The method presented in this article avoids these defects by using a problem oriented kind of number: the real compact intervals. The reader is assumed to be familiar with the principles of the corresponding interval mathematics. The concept of interval inclusion functions and the fundamental principle of interval mathematics will be sufficient for a throughout understanding of the global minimization method.

The original algorithm of E. R. Hansen [8], [9] only has been applied to analytical objective functions such as polynomials and rational functions. In fact, objective functions used for modelling chemical, physical or technical systems cannot be stated in such simple terms but are represented by large computer programs using real arithmetic. Porting these programs to interval arithmetic is still very cumbersome since tools like interval pre-compilers [22] are not commonly available, yet.

Therefore, an algorithm for the treatment of arbitrary real valued objective functions has been developed. Due to the applied approximation method, however, the absolute reliability of Hansen’s method to enclose all global minimizers is lost. Nevertheless, this seems to be no drawback in the treatment of smooth objective functions where the approximation error can be kept very small. This has been shown in the successful application of our method to the global optimization of an electron optical device [14], [15].

2. Basic Concepts of Interval Mathematics

Interval mathematics still is a growing branch of pure and applied mathematics. The foundations have been carried out by Ramon E. Moore [12], [13]. The original intention was the treatment of rounding errors in digital computers. Besides this so-called ‘machine interval arithmetic’ many other interval-based applications have been developed. Nevertheless, interval mathematics is not commonly known, yet.

The following sections will give a short summary of the terminology used in this paper. The notation is that of the two standard books on interval analysis cited above and of Alefeld/Herzberger [1]:

The symbol \mathbb{I} denotes the set of real compact intervals $A = [a_1, a_2] = \{a \in \mathbb{R} | a_1 \leq a \leq a_2\}$ containing a subset of the real numbers \mathbb{R} between the lower bound $a_1 = \underline{A}$ and the upper bound $a_2 = \bar{A}$. In the following text both notations for interval bounds will be used, if appropriate. For example, an interval vector $A = (A_1, \dots, A_n)^T \in \mathbb{I}^n$ denotes a rectangular region $A_1 \times \dots \times A_n$ in the n -dimensional space \mathbb{R}^n and the notation $A_i = [\underline{A}_i, \bar{A}_i]$ for its i -th component avoids the introduction of additional indices. The real numbers $r \equiv [r, r]$ will be called *degenerated* or *point intervals* and therefore can be considered a subset of \mathbb{I} . The components of point vector quantities are denoted by corresponding lower case greek letters in cases where ambiguities with samples $\{x_i\}$ may occur, for example $x = (\xi_1, \dots, \xi_n)^T \in \mathbb{R}^n$. To distinguish point matrices from interval numbers $A - Z$ the notation $\mathcal{A} - \mathcal{Z}$ of Alefeld/Herzberger [1] is used.

2.1 Interval Metrics

Two intervals $A, B \in \mathbb{I}$ are *equal* if, and only if $a_1 = b_1$ and $a_2 = b_2$ holds. The *center* of an interval is denoted by $\text{mid}(A) := (a_1 + a_2)/2$. The center of an interval vector is a real vector and is calculated by application of the mid-function to every component: $\text{mid}(A) := (\text{mid}(A_1), \dots, \text{mid}(A_n))^T$. The positive number $\text{span}(A) := a_2 - a_1$ is called the *span* of A . $\text{span}(A) := \max_{i=1}^n \{\text{span}(A_i)\}$ defines the span of an interval vector as its longest side. The *distance* between two intervals A, B is determined by $\text{dist}(A, B) := \max(|a_1 - b_1|, |a_2 - b_2|)$. This definition satisfies the axioms of a topological distance function and implies the norm function $|A| := \text{dist}(A, 0) = \max(|a_1|, |a_2|)$. It can be shown that, according to this distance function, \mathbb{I} is a complete set, that means every Cauchy sequence converges in \mathbb{I} .

2.2 Interval Arithmetic

The following arithmetic rules only work with interval boundaries and can be implemented easily in a digital computer:

$$A + B = [a_1 + b_1, a_2 + b_2]$$

$$A - B = [a_1 - b_2, a_2 - b_1]$$

$$AB = [\min(a_1 b_1, a_1 b_2, a_2 b_1, a_2 b_2), \max(a_1 b_1, a_1 b_2, a_2 b_1, a_2 b_2)]$$

$$A/B = [a_1, a_2][1/b_2, 1/b_1], \quad 0 \notin B.$$

From these interval operations there follows the so-called *inclusion principle* of interval arithmetic: $C \subseteq A \wedge D \subseteq B \Rightarrow C \circ D \subseteq A \circ B$, where \circ denotes the arithmetic operations $\{+, -, *, /\}$ from above. Note, that there are no general inverse elements according to these operations. Moreover, there is only a so-called *subdistributive law* $A(B + C) \subseteq AB + AC$ in \mathbb{I} .

2.3 Interval Analysis

The main tool for the treatment of optimization problems using interval arithmetic is the concept of *inclusion functions*. Let f be a real valued function over an interval X and $F^*(X) := \left[\min_{x \in X} (f(x)), \max_{x \in X} (f(x)) \right]$ be the *range* of f over X . An interval function F then is called an inclusion function for f if $F^*(A) \subseteq F(A)$ holds for any $A \in X$. An example where the range of a function can be given exactly is the polynomial $f(x) = x - x^2$ with $F^*(X) = \frac{1}{4} - (X - \frac{1}{2})^2$. Other examples are elementary functions like \sin , \log , etc., whose monotonicity intervals are known exactly. Their range over an interval X consequently would be denoted by $\text{SIN}(X)$, $\text{LOG}(X)$, etc. Implementations of such elementary interval functions can be found in a few programming languages like FORTRAN-SC [11] or Triplex-Algol-60 [3]. Nevertheless, it is generally impossible to calculate F^* analytically or numerically for a given function f . Therefore, one is very interested in good approximations F of F^* . The inclusion principle of interval arithmetic allows the straightforward construction of such inclusion functions and is called the *fundamental principle of interval mathematics*:

The *natural interval extension* of f to X is defined as that expression $F(X)$ which is obtained from the expression $f(x)$ by replacing each occurrence of the variable x by the interval X , the arithmetic operators $\{+, -, *, /\}$ of \mathbb{R} by their counterparts of \mathbb{I} , respectively, and each elementary function by its range over X . Due to the inclusion principle of the interval operators, F is an inclusion function for f .

Natural interval extensions are not unique, but depend on the way how the expression for f is written down. The reason is that intervals X occurring more than once in F are treated independently by the interval operators $\{+, -, *, /\}$. For the previous example $f(x) = x - x^2$, $F_1(X) := X(1 - X)$ and $F_2(X) := X - XX$ are valid natural interval extensions with $F^*(X) \subseteq F_1(X) \subseteq F_2(X)$ for any $X \in \mathbb{I}$.

Functions of several variables can be treated analogously, leading to interval functions over rectangular regions X .

3. Global Optimization by Interval Mathematics

In the following sections the global optimization problem is stated in a manner that makes it suitable for a treatment with interval analytical methods. One of these

methods is Hansen's algorithm, which tends to be very simple and, nevertheless, very powerful for the safeguarded inclusion of all global minimizers of the objective function.

3.1 The Bound-Constrained Global Minimization Problem

Let $X \in \mathbb{I}^n$ be an interval vector and f be a real valued function over X which is a rectangular box in \mathbb{R}^n and is supposed to contain all feasible points $x = (\xi_1, \dots, \xi_n)^T$ of the minimization problem.

The point x_l is called a *local minimizer* of f if $f(x_l) \leq f(x)$ holds for any positive number ε with $\|x_l - x\| < \varepsilon$ and $x \in X$.

On the other hand, x_g is called a *global minimizer* of f if $f(x_g) \leq f(x)$ holds for any $x \in X$.

One-sided constraints have to be mapped into a bounded form by a inverse coordinate transformation like, for example, $\xi_i > 0 \rightarrow 1/[\underline{X}_i, \infty] = [0, 1/\underline{X}_i]$ taking account of possible resulting singularities $f(0)$ in the objective function, of course.

Non-linear constraints of the form $c(x) = 0$, $c \in \mathbb{R}^m$, however, have to be treated by the mean of so-called *penalty functions* which are to be incorporated into the objective function. For example, $f(x) \rightarrow f(x) + p(x) = f(x) + \sum_{i=1}^m c_i^2(x)$.

3.2 Hansen's Algorithm

The following global minimization method was originally described by Eldon R. Hansen in [9] and [8]. It represents an interval-based bisection algorithm and is predestinated to treat minimization problems as those of the previous section. It is presumed that an interval inclusion function F for the objective function f and an initial box B are provided. This box is splitted repeatedly into so-called subboxes X_ℓ where the inclusion function is evaluated: $F_\ell := F(X_\ell)$. The range of the objective function f over each subbox is thus determined with 100% reliability, due to the fundamental principle of interval mathematics: $f(x) \in F_\ell$ for any $x \in X_\ell$. Moreover, f is calculated in the center of each subbox: $f_\ell := f(\text{mid}(X_\ell))$. These data are stored chronologically in the 'list of boxes': $\mathcal{L} := \{(X_\ell, F_\ell, f_\ell) | \ell = 1, \dots, L\}$. This list originally is of length $L = 1$ and contains merely the initial box $X_1 = B$. Note that the index ℓ denotes the ordering of the boxes X_ℓ in the list \mathcal{L} while the index i is used to indicate its i -th interval component X_i .

The bisection takes place at the longest side of the first element X_1 and produces two subboxes which are added as new elements $(X_{L+1}, F_{L+1}, f_{L+1})$ and $(X_{L+2}, F_{L+2}, f_{L+2})$ at the end of the list \mathcal{L} . The value $f_g := \min_\ell(f_\ell)$ is considered an approximation for the global minimum of f . The list is shortened and renumbered by application of following rules:

1. Cancelling the splitted box:

The intersected box X_1 is cancelled because its subboxes X_{L+1} and X_{L+2} with $X_1 = X_{L+1} \cup X_{L+2}$ are used in the further process.

2. Midpoint test:

Each box X_ℓ with $F_\ell > f_g$ is cancelled because then $f(x) > f_g$ holds for any $x \in X_\ell$ due to the inclusion property of F .

For piecewise differentiable functions, a remarkable acceleration of convergence can be achieved by applying the so-called ‘monotonicity test’: Let G be a inclusion function for the gradient $g := \nabla f$. Let further X be a box from \mathcal{L} with $G_i(X) > 0$ for some coordinate index i . Since G_i includes the range of g_i over X , f is strictly increasing along the i -th coordinate direction. Now, the following distinctions have to be made:

1. $\underline{B}_i < \underline{X}_i$ Elimination:

The i -th lower bound of X lies inside the initial box B . For this reason, in X there cannot be any minimizer of f . The subbox X is discarded from the list \mathcal{L} .

2. $\underline{B}_i = \underline{X}_i$ Degeneration:

The subbox X has a common surface together with the initial box B . This surface possibly contains a global minimizer of f . Therefore, X is being ‘degenerated’ with respect to the i -th coordinate direction: $X = (X_1, \dots, X_i, \dots, X_n) \rightarrow X = (X_1, \dots, \underline{B}_i, \dots, X_n)$. After recalculating the values of F and f , the degenerated subbox replaces the original element in \mathcal{L} .

The case $G_i < 0$ is treated analogously, using the upper bound \bar{B}_i .

The algorithm can be characterized by the two kinds of steps bisection and elimination. The bisection steps cause a continuous expansion of the list but the elimination steps possibly would shorten it, provided that the midpoint test or the monotonicity test are successful. Nevertheless, in the case of a constant function $f \equiv c$, the list is always expanding since f contains an infinite number of global minimizers and no subboxes would ever be discarded. On the other hand, a pathological function like $f(x) := \sin(1/x)$ contains a countable but infinite number of global minimizers in every open interval $(0, \varepsilon]$ with $\varepsilon > 0$. The algorithm tries to include all these minimizers. Therefore, the list grows continuously, although many subboxes would be discarded, too.

The procedure is stopped if at least one of the following conditions holds:

1. The length L of the list \mathcal{L} exceeds a maximum number.
2. The number of function evaluations f_ℓ exceeds an upper bound.
3. A function value f_g was reached, which is smaller than a lower bound f_{\min} .
4. A minimum span is reached by every box in \mathcal{L} .

The advantages of this global minimization method can be summarized as follows:

1. After the procedure has stopped, the list of boxes contains every global minimizer with absolute reliability.

2. The required computer memory is adjusted automatically to the topology of the objective function. There is no need to cover the initial box with a fine grid or numerous random points since in \mathbb{I} it represents already a single point.
3. Discontinuous functions can be treated, too. For functions which are at least piecewise differentiable, the convergence speed of the method is increasing remarkably due to the monotonicity test.
4. For functions which are at least twice piecewise differentiable, second order information could provide further acceleration of convergence by investigation of the interval Hessian matrix. However, this has not been carried out by the author, yet.
5. A box to be eliminated after the midpoint test but failing the monotonicity test tends to contain local minimizers. If the upper bound of the objective function in this box is sufficiently small and the corresponding systems are well suited for a technical realization, the box can be stored and investigated by a local optimization method, afterwards.

3.3 Numerical Example

The so-called ‘three-hump-camel-back function’

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad x = (\xi_1, \xi_2)^T \mapsto 2\xi_1^2 - 1.05\xi_1^4 + \frac{1}{6}\xi_1^6 - \xi_1\xi_2 + \xi_2^2$$

has two saddle points $x_s \approx (\pm 1.07, \pm 0.54)^T$ and three local minimizers $x_{1,2} \approx (\pm 1.78, \pm 0.87)^T, x_3 = (0, 0)^T$. In the box $B = ([-2, +2], [-2, +2])^T$ the difference between the local minima $f(x_{1,2}) = 0.30$ and the global minimum $f(x_3) = 0$ is merely 3%. Natural interval extensions of this function and its gradient are:

$$\begin{aligned} F(X) &= 2X_1^2 - 1.05X_2^4 + \frac{1}{6}X_1^6 - X_1X_2 + X_2^2 \\ G(X) &= (4X_1 - 4.2X_1^3 + X_1^5 - X_2, 2X_2 - X_1)^T. \end{aligned}$$

Figure 1 shows the application of Hansen’s method to this function. After 2000 bisection steps the list only contains the 4 subboxes

$$\begin{aligned} &([0, 2.38 \times 10^{-7}], [0, 2.38 \times 10^{-7}])^T, \\ &([-1.19 \times 10^{-7}, 0], [0, 2.38 \times 10^{-7}])^T, \\ &([-2.38 \times 10^{-7}, 0], [-2.38 \times 10^{-7}, 0])^T, \\ &([0, 2.38 \times 10^{-7}], [-2.38 \times 10^{-7}, 0])^T. \end{aligned}$$

The lowest function value $f_g = 5.86 \times 10^{-14}$ was calculated in the center of the second box. Note that of Hansen’s techniques, only the midpoint test and the monotonicity test have been applied and that convergence could be accelerated significantly by using second order information, e.g. by calculating the interval Hessian matrix, too.

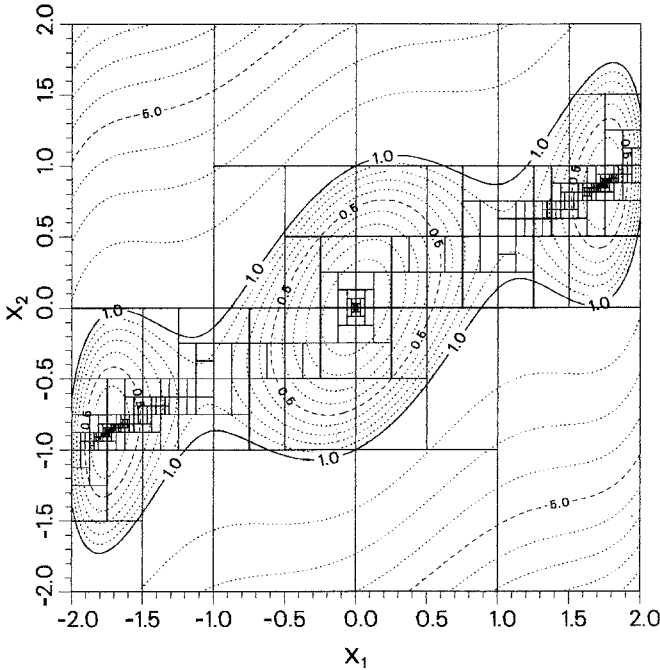


Figure 1. Hansen's method for global minimization applied to the so-called 'three-hump camel-back function'. The global minimizer $x = (0, 0)^T$ was included after 2000 bisections in the box $X_g = [-1.19 \times 10^{-7}, 0], [0, +2.38 \times 10^{-7}]^T$

4. Construction of Interval Inclusion Functions by Multinomial Fitting

The main disadvantage of Hansen's method is the need of an interval inclusion function F for the objective function f . There are two ways of obtaining such an inclusion function. If a computer language like, for example, FORTRAN-SC [11] is available, the data type INTERVAL, the interval operators $\{+, -, *, /\}$ and most of the elementary interval functions like SIN, LOG, etc. are provided. The computer program for the objective function f then can be written down using interval arithmetic and the natural interval extension of f is done automatically by the compiler. This process can be compared with the arithmetic of complex numbers in standard FORTRAN-77, where the data type COMPLEX is already supported and the programmer does not have to deal with the different meanings of the operators $\{+, -, \times, /\}$ in \mathbb{R} and \mathbb{C} .

Unfortunately, computer languages supporting interval mathematics are not wide-spread, yet. Therefore, pre-compilers have been developed which replace each occurrence of the real arithmetic operators and elementary functions in every expression containing pre-declared interval data by SUBROUTINE and FUNCTION calls to their interval arithmetic counterparts, respectively. The output is generated in a standard computer language, such as FORTRAN-77 by

the AUGMENT pre-compiler [22]. Nevertheless, the programs produced in this manner, become much longer and therefore much slower than the original code. Besides, like the programming languages including interval extensions, such pre-compilers merely are available on a few computers.

For this reason, an interval-based modelling technique will be developed in the following sections, which supplies an approximation of an interval inclusion function for f and which can be applied to conventionally, i.e. in real arithmetic, programmed objective functions, too.

4.1 The Choice of the Modelling Function

Choosing an appropriate modelling function, the following points have to be taken in account:

1. Of the modelling function the ability to approximate a wide variety of objective functions is requested. This implies a sufficient complexity of the modelling function which ensures its capability to handle singularities or steep areas of the objective function. Although the application of problem-oriented modelling functions seems to be more efficient, this cannot be done automatically, however. The same holds for the corresponding interval extensions which had to be carried out every time a new modelling function is to be chosen.
2. The modelling function has to be piecewise differentiable if Hansen's monotonicity test is to be applied.
3. The coefficients representing the degrees of freedom of the modelling function must not become too numerous in higher dimensions. Note that this demand limits the complexity of the modelling function requested in 1.
4. The interval extension of the modelling function is required to be a close approximation to its range of values.

A heuristic approach to such a modelling function which satisfies closely the requests from above has been established by fitting affine quadratic forms $q(x) := x^T(Ax + b) + c$ (being discussed in the following section) to a sample of transformed objective function data $\text{arsinh}(f(x))$.

The following considerations led to this special choice:

1. The complete multiquadratic form $q(x)$ uses merely $\binom{n+2}{n} = (n+1)(n+2)/2$ degrees of freedom in n linearly independent dimensions and it is able to approximate either objective functions with regions of the saddle type or the basin type. The interval extension $q(x) \rightarrow Q(X)$ can be carried out very efficiently, as will be shown later.
2. Objective functions where a steepness more than linear occurs cannot be approximated by quadrics in a satisfactory way, however. Therefore, the objective function is transformed by the bijective nonlinear transformation $f(x) \rightarrow \text{arsinh}(f(x))$ which takes account of exponentially growing functions, too. The interval exten-

sion of the modelling function $\sinh(q(x)) \rightarrow \text{SINH}(Q(X))$ also can be done straight forward.

It must be emphasized that using this method the absolute reability of isolating *all* global minimizers is lost, in contrast to Hansen’s original method. Since the interval extension of the objective function is not carried out directly but versus the modelling function, the approximation error is interval extended, too. However, this lack may be overcome partly by the following method for estimating the quality of the approximation.

Due to the definition of an interval inclusion function the modelling function had to satisfy the following conditions:

1. Inclusion principle:
 $f(x) \in \text{SINH}(Q(X))$ and
 $\forall f(x) \in \forall Q(X) \text{COSH}(Q(X))$ for any $x \in X$.
2. Inclusion isotonicity:
 $\text{SINH}(Q(A)) \subseteq \text{SINH}(Q(X))$ and
 $\forall Q(A) \text{COSH}(Q(A)) \subseteq \forall Q(X) \text{COSH}(Q(X))$ for any $A \subseteq X$.

According to the construction of the modelling function as natural interval extension, the first condition is sufficient but the second one is only a necessary condition. Since the error $E(X) := [\min_{x \in X}(f(x) - q(x)), \max_{x \in X}(f(x) - q(x))]$ is generally non known, the inclusion property cannot be guaranteed. The inclusion isotonicity can, however, be checked by the following method: In every bisection step the ranges F_ℓ and G_ℓ over the subboxes $X_\ell, \ell = L + 1, L + 2$ are compared with those of the original splitted box X_1 . If $F_\ell \not\subseteq F_1$ or $G_\ell \not\subseteq G_1$ it follows that the approximation of the objective function is not sufficiently accurate in this region. Such subboxes are suspended from the midpoint test and the monotonicity test, respectively, and therefore are preserved from a premature elimination or degeneration. The same holds for subboxes including the lowest function value f_g , so far.

In the next sections we will discuss the details of the approximation such as the choice of the sample points, their unisolvence property, the transformation of sample data points and a suitable interval extension of the affine multiquadratic form.

4.2 Interpolation with Affine Multiquadratic Forms

If $c \in \mathbb{R}, b \in \mathbb{R}^n$ and $\mathcal{A} \in \mathcal{M}(n \times n)$ is an upper triangular matrix, then $q: \mathbb{R}^n \rightarrow \mathbb{R}, x = (\xi_1, \dots, \xi_n)^T \mapsto q(x) := x^T(\mathcal{A}x + b) + c$ is called an *affine* or *complete multiquadratic* form. The $m = \binom{n+2}{n} = (n+1)(n+2)/2$ multinomials $\varphi_k(x) := \xi_{i(k)}^{\lambda_1(k)} \xi_{j(k)}^{\lambda_2(k)}$ with $\lambda_1(k) + \lambda_2(k) \leq 2$ for any $k = 1, \dots, m$ are linearly independent.

The interpolation problem $q(x_k) := y_k, k = 1, \dots, m$ is solved by the coefficient vector $d = (\delta_1, \dots, \delta_m)^T = (c, (\beta_1, \dots, \beta_n)^T, \delta_{n+2}, \dots, \delta_m)^T \in \mathbb{R}^m$ of the linear equation system

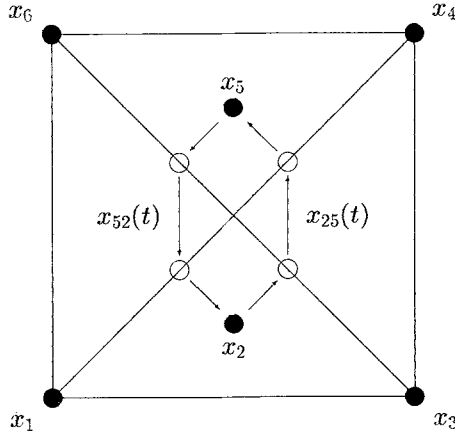


Figure 2. Example for the unisolvence property of the quadratic form in \mathbb{R}^2 for a set of sample points $x = (\xi_1, \xi_2)^T$. The points x_2 and x_5 of the unisolvent set $\{\bullet\}$ are exchanged via the curves x_{25} and x_{52} . There are configurations $\{\bullet, \circ\}$ where the solution of the interpolation problem $y_i = x_i^T \mathcal{A} x_i + b^T x_i + c$ is not unique. Here, all sample points lie on the diagonals $\xi_1 = \xi_2$ and $\xi_1 = 1 - \xi_2$

$$\begin{pmatrix} \varphi_1(x_1) & \dots & \varphi_m(x_1) \\ \vdots & \ddots & \vdots \\ \varphi_1(x_m) & \dots & \varphi_m(x_m) \end{pmatrix} \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_m \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

This system of equations is called unisolvent if and only if the generalized Gram determinant does not vanish:

$$\det(\varphi_i(x_k)) = \begin{vmatrix} \varphi_1(x_1) & \dots & \varphi_1(x_m) \\ \vdots & \ddots & \vdots \\ \varphi_m(x_1) & \dots & \varphi_m(x_m) \end{vmatrix} \neq 0.$$

In the case of the affine multiquadratic form such multisolvant sets are the real valued solutions of $q(x) = 0$ which define hyper-surfaces in \mathbb{R}^n . Figure 2 shows an unisolvent set of the quadratic form in \mathbb{R}^2 . The points 2 and 5 can be exchanged producing the intermediate set $\{(\xi_1, \xi_2)^T | \xi_1 = \xi_2 \vee \xi_1 = 1 - \xi_2\}$ which is multisolvant. For example, together with the quadratic form $q_1(x) = x^T \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ 0 & \alpha_{22} \end{pmatrix} x + (\beta_1, \beta_2)^T x + c$ there is an infinite number of solutions $q_2(x) = x^T \begin{pmatrix} r & \alpha_{12} \\ 0 & \alpha_{22} + \alpha_{11} - r \end{pmatrix} x + (\alpha_{11} + \beta_1 - r, r - \alpha_{11} + \beta_2)^T x + c$ for any $r \in \mathbb{R}$.

4.3 Transformation of the Objective Function

For objective functions of order $\mathcal{O}(x) > 2$ approximations by affine multiquadratic forms would produce unsatisfactory approximations. Figure 4 shows this behaviour for the power function $f(x) = x^7$. The interval extension of q would predict a

misleading minimizer in the neighbourhood of $x = 1$. In this region the monotonicity test of Hansen's method would always fail and the midpoint test would discard the area not till many bisection steps have been carried out.

Without introducing additional degrees of freedom in the modelling function or a non-linear least-squares-fit, this lack can be overcome by the application of the bijective transformation $\sinh \circ q: \mathbb{R}^n \rightarrow \mathbb{R}, x \mapsto \sinh(q(x))$.

The coefficients of the multiquadratic form q allow a wide variety of functions to be fitted. Three typical examples are shown in Fig. 3. For this sake, the inverse transformation $\tilde{y} = \operatorname{arsinh}(y) = \ln(y + \sqrt{1 + y^2})$ is applied to the original data points $y = f(x)$ of the objective function. The multiquadratic form q then is fitted to these transformed data using Shepard's method described in one of the next sections. The power function $f(x) = x^7$ of Fig. 4 is obviously better approximated and the monotonicity is correctly reproduced within a wide region.

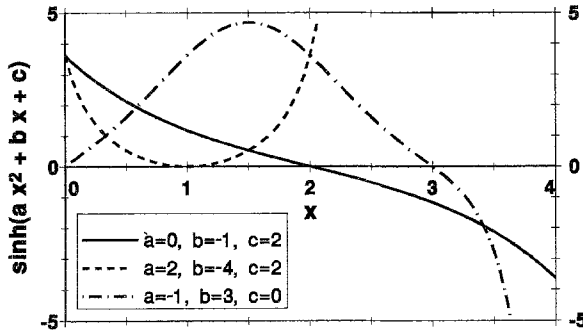


Figure 3. Plots of the function $\sinh(ax^2 + bx + c)$ with three different coefficient sets (a, b, c) . Point symmetric functions (solid line) as well as axial symmetric functions (dashed lines) can be fitted

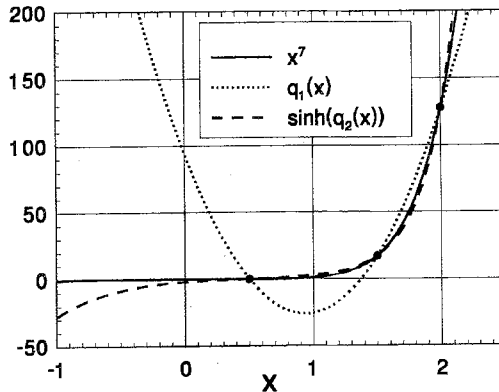


Figure 4. Approximation with quadratic form and hyperbolic sine. The power function x^7 is interpolated between the sample points $x_1 = 0.5, x_2 = 1.5$ and $x_3 = 2$. The monotonicity is not reproduced correctly by the quadratic form $q_1(x) = a_1x^2 + b_1x + c_1$. An interpolation of the data $\operatorname{arsinh}(x_i), i = 1, 2, 3$, with $q_2(x) = a_2x^2 + b_2x + c_2$, however, leads to a better result

4.4 The Choice of the Sample Points

The set of abscissas $\{x_k\}$ for the interpolation problem has to be unisolvent and equally spaced in the interval vector X . In this section, a method for regular distributions of sample points in several dimensions is presented. For simplification, the interval vector X is normalized to the unit cube $K := [0, 1] \times \dots \times [0, 1]$ which can always be done by an affine coordinate transformation and vice versa. A placement of the sample points only on the corners of K is not the best solution because the number of corners 2^n exceeds the number of sample points $(n + 1)(n + 2)/2$ needed for interpolation by affine multiquadratic forms in dimensions $n > 3$. Besides, the total volume of K is not covered regularly by this strategy. The task of covering a multidimensional volume ‘regularly’ has been investigated by many authors with mostly different intentions such as numerical integration by Monte-Carlo techniques or global optimization problems. A measure for such a regular distribution of m points $\{x_1, \dots, x_m\}$ in the unit hypercube is the so-called *discrepancy* function

$$\mathcal{D}(x_1, \dots, x_m) := \max_{x \in K} |m^{-1}u(x_1, \dots, x_m; x) - v(x)|$$

where $u(x_1, \dots, x_m; x)$ denotes the number of $k \in \{1, \dots, m\}$ with $0 \leq x_k < x = (\xi_1, \dots, \xi_n)^T$ (componentwise), and $v(x) = \xi_1 \xi_2 \dots \xi_n$ denotes the volume of the box $[0, x]$. There is no finite distribution $\{x_k\}$ with $\mathcal{D}(x_1, \dots, x_m) < \mathcal{O}(m^{-1} \ln^{n-1} m)$ [16]. For example, the hypercubic grid in $K \in \mathbb{R}^n$ which is defined by the coordinate sections $\xi_i = i/(k + 1), i = 1, \dots, k$ with $k \in \mathbb{N}$ and which contains $m = k^n$ mesh points has the discrepancy $\mathcal{D}_{grid}(x_1, \dots, x_m) = 1/(k + 1) = (m^{1/n} + 1)^{-1} \approx m^{-1/n}$ if $m \gg 1$. A random point distribution of these m points, on the other hand, has a discrepancy $\mathcal{D}_{rand}(x_1, \dots, x_m)$ of order $m^{-1/2}$ which is smaller than \mathcal{D}_{grid} in dimensions $n \geq 3$ [20].

Distributions which have even less discrepancy are the so-called LP_τ -sequences of I. M. Sobol [18], [19]. They can be defined as follows:

1. $V \in K$ is called a binary interval vector if $V_i = [j_i, j_i + 1]2^{-\lambda_i}$ with $\lambda_i \in \mathbb{N}$ and $0 \leq j_i < 2^{\lambda_i} - 1$ holds for any of its components.
2. A countable set \mathcal{P} , consisting of 2^v points in K is called a P_τ -mesh for $v, \tau \in \mathbb{N}$ with $v \geq \tau + 1$ if every binary interval vector in K of volume $\prod_{i=1}^n 2^{-\lambda_i} = 2^{\tau-v}$ contains exactly 2^τ points of \mathcal{P} .
3. A mesh (x_k) with $k \in \mathbb{N}$ is called a LP_τ -sequence if every partial mesh (x_κ) with $j2^\sigma \leq \kappa < (j + 1)2^\sigma, j, \sigma \in \mathbb{N}$ and at least $2^{\tau+1}$ elements defines a P_τ -mesh.

The LP_τ -sequences have the following remarkable properties:

$$\begin{aligned} \mathcal{D}_{LP_\tau}(x_1, \dots, x_m) &= \mathcal{O}(m^{-1} \ln^n m) && \text{for any } m \in \mathbb{N}, \\ \mathcal{D}_{LP_\tau}(x_1, \dots, x_m) &= \mathcal{O}(m^{-1} \ln^{n-1} m) && \text{if } m = 2^v, v \in \mathbb{N}. \end{aligned}$$

The projection of a LP_τ -sequence of a hypercube $K \in \mathbb{R}^n$ onto one of its lateral faces $K \in \mathbb{R}^{n-1}$ is a LP_τ -sequence, too.

A LP_τ -sequence represents a series of P_τ -meshes due to its definition. This is shown in Fig. 5 for a LP_0 -sequence in \mathbb{R}^2 . The obvious mirror symmetry according to

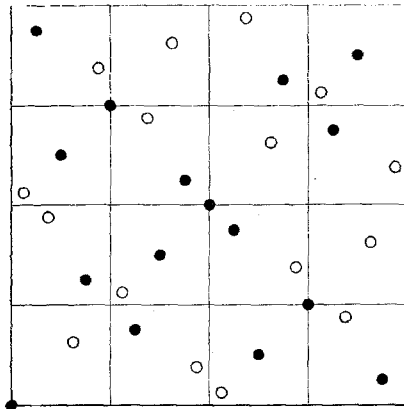


Figure 5. LP_0 -sequence by Sobol in \mathbb{R}^2 . Only the first 32 elements are presented. Points on the nodes belong to the right upper mesh. The sequences $\{\bullet\}$ and $\{\circ\}$ consisting of the points 1–16 and 17–32 build a P_0 -mesh each of its own, i.e. in every mesh of volume $1/16$ there is exactly one point of the two meshes, respectively. Both P_0 -meshes build a P_1 -mesh together

the diagonal $x_1 = x_2$ makes clear the difference between these so-called ‘quasi random numbers’ and pseudo random numbers.

The construction of LP_τ -sequences is very a extensive task and therefore is not presented in this publication. The interested reader is referred to the original papers of Sobol [18] and [19]. An algorithm suitable for digital computers has been described by Antonov and Saleev [2] and its implementation in FORTRAN-77 can be found in [4].

4.5 Approximation of Irregularly Spaced Data with Shepard’s Method

Although the LP_τ -sequences of Sobol provide unisolvent sets for the described interpolation method, this unisolvence property may not be maintained during the bisection steps of Hansen’s method where new sample points are generated continuously. Restoring the unisolvence property for a multisolvent set requires complex strategies, i.e. the replacement of one or more sample points by more distant ones. In higher dimensions this procedure can be very tedious and time consuming. Therefore the strict interpolation is carried out only in the initial box K using the unisolvent Sobol set. Afterwards, the objective function is approximated by a linear least-squares-fit: $\sum_{k=1}^M w_k \|f_k - q_k\|_2^2 = \min$, where f_k and q_k denote the values of the objective function and the multiquadratic form at the sample point x_k , respectively. After Shepard [17] a suitable weight function in the box X is defined by $x = \min(X)$ and $w_k = \left(1 - \left(\frac{\|x - x_k\|_p}{d}\right)^{e_1}\right)^{e_2}$, where $\|x\|_p := \sqrt[p]{\sum_{i=1}^n |\xi_i|^p}$ is an arbitrary p -norm with $p \geq 1$, and $d = \max_{k=1, \dots, M} \|x - x_k\|_p$ is the radius of the smallest sphere containing all M sample points. $e_1, e_2 \in \mathbb{N} \setminus \{0\}$ are arbitrary exponents for the present. Since at least one sample point lies on the surface of the sphere,

where $w \equiv 0$, the number M has to be chosen $M > m = (n + 1)(n + 2)/2$. The other parameters have been set as follows:

1. $M = m + 1 = (n^2 + 3n + 4)/2$

If the according set of sample points is multisolvent, this number is increased step-by-step.

2. $p = \infty$

$\|x\|_\infty = \max_{i=1, \dots, n} |x_i|$ (maximum-norm). This norm can be calculated considerably faster than the Euclidian norm $\|\cdot\|_2$. Between the two norms the relation $\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty$ holds. A disadvantage of the maximum-norm, however, is the lack of its invariance according to rotations of the coordinate system.

3. $e_1 = 2, e_2 = 1$

Considering the small number of sample points needed for the multiquadratic fitting, by choosing these values for e_1 and e_2 the weights do not decrease too much for increasing distances of sample points.

The sample points themselves are chosen by calculating and storing the distance $d_i = \|x_i - x\|_\infty$ from the current boxes midpoint x for every point produced by Hansen's optimization method, so far. Afterwards, these distances are sorted by increasing order and only the first M elements are used for the approximation technique described above. Therefore, not only the midpoint x of the current box, but also midpoints of adjacent boxes are used. In the beginning of the algorithm even far-away boxes have to be considered, since the initialization procedure only provides the minimum number $m = (n + 1)(n + 2)/2$ of sample points necessary for the Lagrange interpolation method. Note that due to this strategy, the procedure starts using very few information for forming the modelling function. This fact bears an additional uncertainty of loosing some minimizers, however. Nevertheless, during the optimization progress, the amount of sample points and therefore the reliability of the modelling function is increasing dynamically.

4.6 Interval Extension of the Modelling Function

For the approximation of an interval inclusion function of F a suitable interval extension Q of the multiquadratic form q has to be chosen. Q is required to be a close representation of the range Q^* over the box X : $\text{dist}(Q(X), Q^*(X)) = \min$ for any $X \in K$. For this reason, a multiple occurrence of the coordinate intervals X_i has to be avoided. Therefore, the natural interval extension $X^T \mathcal{A} X + b^T X + c$ is not the best solution. Taking advantage of the relation $X \in K \Rightarrow X_i^2 = X_i X_i$ for any $i = 1, \dots, n$ and the subdistributive law in \mathbb{I} , Horner's method is a suitable representation of Q :

$$Q(X) = c + \sum_{i=1}^n X_i \left(\beta_i + \sum_{k=i}^n \alpha_{ik} X_k \right).$$

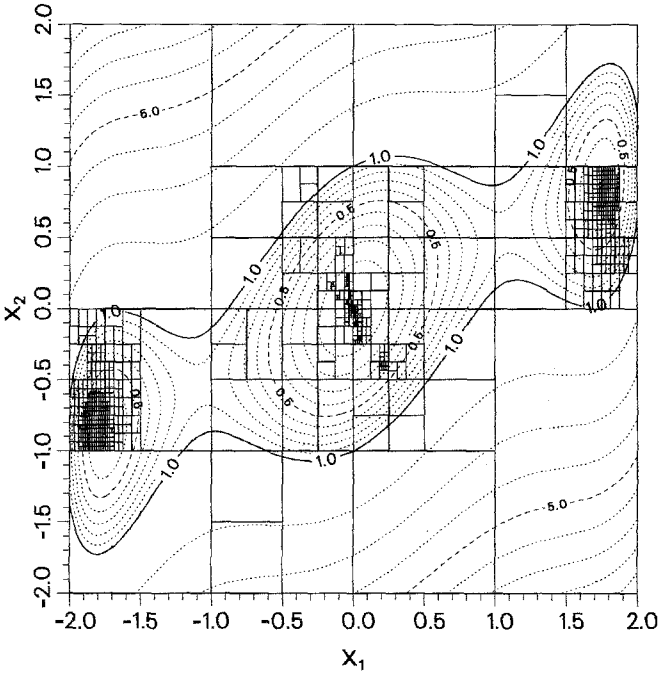


Figure 6. Inclusion of the three-hump-camel-back function by quadratic interval functions. The global minimizer $x = (0, 0)^T$ is included in the box $X_q = ([-1.22, +1.22], [-1.22, +1.22])^T \times 10^{-4}$ after 2000 bisection steps

The hyperbolic sine is strictly increasing in \mathbb{R} and therefore $\text{SINH}(Q) = [\sinh(q_1), \sinh(q_2)]$. The hyperbolic cosine, on the other hand, has to be distinguished in \mathbb{R}_+ and \mathbb{R}_- :

$$\text{COSH}(Q) = \begin{cases} [\cosh(q_1), \cosh(q_2)] & \text{if } Q > 0 \\ [\cosh(q_2), \cosh(q_1)] & \text{if } Q < 0 \\ [\min(\cosh(q_1), \cosh(q_2)), \max(\cosh(q_1), \cosh(q_2))] & \text{if } Q \ni 0 \end{cases}$$

The i -th component of the interval gradient is calculated as

$$(\nabla \text{SINH}(Q(X)))_i = \left(b + \sum_{k=1}^n (a_{ik} + a_{ki})X_k \right) \text{COSH}(Q(X)).$$

The method has been applied to the three-hump-camel-back function again and the bisection process is shown in Fig. 6.

5. Results

Our optimization method has been applied successfully to a practical example which has already been treated by local methods [6], [10]. The task was to reduce

the diameter of a electron beam spot in an lithography device consisting of four magnetic lenses. The system was modelled by 19 parameters defining the geometries, positions and currents of the lenses together with the mask and wafer position. Further details of the modelling technique can be found in [14] and [15].

The optimization algorithm has been programmed in FORTRAN-77 and runs on a vector computer Convex-220 of our local computer centre. It took about 150 CPU-hours and 10000 function calls to isolate the global minimizer with sufficient accuracy. The according minimum leads to remarkably better system properties compared with the results of the local methods of [6] and [10]. In [14] and [15] a more detailed discussion of the results is given.

6. Conclusion

In contrast to Monte-Carlo methods or deterministic algorithms like dynamic programming, in the present state of the art Hansen's method seems to be the only reliable tool for the safeguarded inclusion of all global minimizers of a real valued objective function in a rectangular region of the parameter space. In every stage of the minimization process the solution is included in the union of subboxes calculated, so far. The algorithm is well suited for the implementation on a digital computer since memory requirements are very small due to the representation of large parameter regions as single points in the interval vector space \mathbb{I}^n .

The modelling of the objective function by complete multiquadratic interval functions allows the application of the optimization method to conventionally (that means in real arithmetic) programmed system simulations. The method, however, is still very time consuming owing to both, the numerous system calculations and the subsequent modelling process which requires an extensive sorting procedure and the solution of a linear system of equations in many dimensions. The main disadvantage is the loss of the safeguarded inclusion property of the modelling interval function since the approximation error cannot be estimated with certainty. Nevertheless, this drawback seems to be of less importance if *non-singular*, i.e. smooth objective functions are investigated, and the inclusion isotonicity test represents a sufficient mean to avoid the elimination of regions possibly containing global minimizers. High-level programming languages supporting interval arithmetic, on the other hand, allow the construction of genuine inclusion functions and enable the application of Hansen's method with none of the restrictions mentioned above.

Since interval mathematics is a powerful tool not only for the treatment of global optimization problems but also for the automatical control of round-off errors in daily numerical calculations, the propagation of interval arithmetic as well as its software based implementations and the development of hardware based realizations, similar to the modern vector processors, has to be forced along. On the other hand, our investigation has shown that even without such advantageous tools the global optimization of realistic and highly complex systems has become feasible.

References

- [1] Alefeld, G., Herzberger, J.: Einführung in die Intervallrechnung. Mannheim Wien Zürich: B.I.-Verlag 1974.
- [2] Antonov, I. A., Saleev, V. M.: An economic method of computing LP_r -sequences. U.S.S.R Computational Mathematics and Mathematical Physics 19(1), 252–256 (1979).
- [3] Apostolatos, N., Kulisch, U., Krawczyk, R., Lortz, B., Nickel, K., Wippermann, H.-W.: The algorithmic language Triplex-Algol 60. Numerische Mathematik 11, 175–180 (1968).
- [4] Bratley, P., Fox, B. L.: Implementing Sobol's quasirandom sequence generator. Algorithm 659. ACM Transactions on Mathematical Software 14(1), 88–100 (1988).
- [5] Gill, P. E., Murray, W., Wright, M. H.: Practical optimization. Orlando (Florida): Academic Press 1981.
- [6] Glatzel, U.: Optimierung elektronenoptischer Systeme, Mai 1987. Diplomarbeit.
- [7] Glatzel, U., Lenz, F.: Optimization of electrostatic immersion lenses. Optik 79(1), 15–18 (1988).
- [8] Hansen, E. R.: Global optimization using interval analysis—the multi-dimensional case. Numerische Mathematik 34, 247–270 (1979).
- [9] Hansen, E. R.: Global optimization using interval analysis—the one-dimensional case. J. Optim. Theory and Appl. 29, 331–344 (1979).
- [10] Lenz, F.: Optimization of the imaging properties of an electron optical demagnifying system consisting of four magnetic lenses. Optik 78(4), 135–140 (1988).
- [11] Metzger, M., Walter, W.: FORTRAN-SC Eine FORTRAN-Erweiterung für wissenschaftliches Rechnen. In: Kulisch, U. (ed.) Wissenschaftliches Rechnen mit Ergebnisverifikation Braunschweig, Wiesbaden: Vieweg 1989.
- [12] Moore, R. E.: Interval analysis. Englewood Cliffs, N.J.: Prentice-Hall 1966.
- [13] Moore, R. E.: On computing the range of a rational function of n variables over a bounded region. Computing 16, 1–15 (1976).
- [14] Munack, H.: Global optimization of an electron beam lithography system using interval arithmetic. Optik (to be printed).
- [15] Munack, H.: Globale Optimierung elektronenoptischer Systeme mittels Intervallarithmetik. PhD thesis, Tübingen, 1992.
- [16] Roth, K. F.: On irregularities of distributions. Mathematika 1, 73–79 (1954).
- [17] Shepard, D.: A two-dimensional interpolation function for irregularly spaced data. In 23rd Nat. Conf. ACM pages 517–524. ACM (1968).
- [18] Sobol, I. M.: On the distribution of points in a cube and the approximate evaluation of integrals. U.S.S.R. Computational Mathematics and Mathematical Physics 7(4), 790–793 (1967).
- [19] Sobol, I. M.: Uniformly distributed sequences with an additional uniform property. U.S.S.R Computational Mathematics and Mathematical Phys 16(5), 236–242 (1976).
- [20] Sobol, I. M.: On the systematic search in a hypercube. SIAM Journal of Numerical Analysis 19(5), 790–793 (1979).
- [21] Szilágyi, M.: A new approach to electron optimization. Optik 48(2), 215–224 (1977).
- [22] Yohe, J. M.: Portable software for interval arithmetic. Computing [Suppl.] 2, 211–229 (1980).

Harry Munack
 Lehrstuhl für Theoretische Elektronenphysik
 Institut für Angewandte Physik
 Universität Tübingen
 Auf der Morgenstelle 12
 D-W-7400 Tübingen 1
 Federal Republic of Germany