Copyright © 1982, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.
SEMI-INFINITE OPTIMIZATION IN ENGINEERING DESIGN

by

E. Polak

Memorandum No. UCB/ERL M82/50

30 November 1981
SEMI-INFINITE OPTIMIZATION IN ENGINEERING DESIGN

by

E. POLAK

Memorandum No. UCB/ERL M82/50

30 November 1981

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720
SEMI-INFINITE OPTIMIZATION IN ENGINEERING DESIGN

by

E. Polak

Department of Electrical Engineering
and Computer Sciences
and the Electronics Research Laboratory
University of California
Berkeley, Ca. 94720

ABSTRACT.

It is shown that a large class of engineering design problems are expressible as nonlinear semi-infinite programming problems. Methods for the solution of these SIP problems are surveyed.

1. INTRODUCTION.

Many industries are facing rapid growth in the cost of raw materials and energy. More and more frequently, competitive product design is becoming an ever more time-consuming and difficult task because design complexity is increasing more rapidly than the number of design variables. All these factors contribute to increasing costs of design and manufacture. Computing power is the only important industrial resource which is declining in cost. Not surprisingly, the engineer is turning more and more frequently to the computer as a way of coping with his tasks.

The growing importance of computer-based tools for the analysis and design of devices, structures, machines and systems has resulted in the evolution of a relatively new interdisciplinary engineering specialty which is commonly referred to as computer-aided design (CAD). Evolution of the more sophisticated forms of CAD capability requires a team effort involving engineering design specialists, computer scientists, optimization experts, and numerical analysts. CAD techniques developed in one area of engineering are often of substantial relevance to other areas.

Both combinatorial and parametric optimization are used in CAD. The most commonly occurring parametric optimization problems in engineering design are usually expressed either as differentiable or as non-differentiable semi-infinite programming (SIP) problems. In this paper we shall illustrate by example how typical SIP problems arise in engineering design and survey some of the more promising methods for their solution.

2. FORMULATION OF ENGINEERING DESIGN PROBLEMS IN SIP FORM.

In this section we shall give four examples of engineering design problems which can be solved in SIP form: one in structural design, one in single-input single-output control system design, one in multi-input multi-output control system design, and one in electronic circuit design.


One of the conceptually simplest examples of SIP in engineering design is found in the design of braced frame buildings which are expected to withstand small earthquakes with no damage and large ones with repairable damage. A simple braced frame is shown in Fig. 1a, where the components of the vector design parameter \( x \) are the stiffnesses of the members indicated. The horizontal floors are assumed to be rigid and to concentrate the mass of the structure. The displacements of the three floors and "roof" form the components of the displacement vector \( y \). This lumped parameter model Fig. 1b obeys a vector second order differential equation of the form:
where $F(t)$ represents the seismic forces. For reasonable values of $x$, when $F$ is small, i.e., it represents a small earthquake, (2.1) is a linear differential equation, but when $F$ is large, the bending of steel introduces gross nonlinearities due to its hysteretic behavior. It is common to consider a whole family of excitations $\{F_k\}$, both large and small in carrying out a design. A simple cost to minimize is the cross section of the frame members, while the constraints are introduced to limit the relative floor displacements over the entire duration of a sample earthquake. Thus, we obtain a SIP of the form

$$\min \{f(x) \mid |y^j(t,x,F) - y^{j-1}(t,x,F)| \leq d^j(F), \forall t \in [0,T]\}$$

$$j = 1, 2, 3, \forall F \in \{F_k\}; x \geq L$$

In choosing a SIP algorithm for solving structural problems, one must bear in mind that the differential equations are highly nonlinear and quite difficult to integrate. For example, a simple four story frame may require several minutes of CPU time using a VAX 11/780 computer to simulate, while a quadratic programming search direction finding subprocedure requires only fractions of a second. Also, since simulation has to be carried out using standard computer codes, one finds that derivatives must be computed by finite differences, making it difficult to use second order algorithms. The algorithm in (33) was developed with these problems in mind. For further reading, see (3, 4, 21).

### 2.2 Design of SISO Control Systems.

Single-input single output (SISO) control systems are particularly easy to design by SIP techniques. Consider the block diagram of Fig. 2, which shows a plant, with transfer function $P(s)$ and a compensator, with transfer function $C(x,s)$, where $x$ is a three components vector of design parameters to be chosen. We assume that this system satisfies the following set of equations:

$$\ddot{y}(t,x) + 5\dot{y}(t,x) + 30y(t,x) = u(t,x)$$

$$\dot{u}(t,x) + 30u(t,x) = x^1\dot{e}(t,x) + x^2\ddot{e}(t,x) + x^3e(t,x)$$

$$e(t,x) = r(t) - y(t,x)$$

where $F(t)$ represents the seismic forces. For reasonable values of $x$, when $F$ is small, i.e., it represents a small earthquake, (2.1) is a linear differential equation, but when $F$ is large, the bending of steel introduces gross nonlinearities due to its hysteretic behavior. It is common to consider a whole family of excitations $\{F_k\}$, both large and small in carrying out a design. A simple cost to minimize is the cross section of the frame members, while the constraints are introduced to limit the relative floor displacements over the entire duration of a sample earthquake. Thus, we obtain a SIP of the form

$$\min \{f(x) \mid |y^j(t,x,F) - y^{j-1}(t,x,F)| \leq d^j(F), \forall t \in [0,T]\}$$

$$j = 1, 2, 3, \forall F \in \{F_k\}; x \geq L$$

In choosing a SIP algorithm for solving structural problems, one must bear in mind that the differential equations are highly nonlinear and quite difficult to integrate. For example, a simple four story frame may require several minutes of CPU time using a VAX 11/780 computer to simulate, while a quadratic programming search direction finding subprocedure requires only fractions of a second. Also, since simulation has to be carried out using standard computer codes, one finds that derivatives must be computed by finite differences, making it difficult to use second order algorithms. The algorithm in (33) was developed with these problems in mind. For further reading, see (3, 4, 21).
and hence that

\[ P(s) = \frac{1}{(s^3 + 5s^2 + 8s + 8)} \]  \hspace{1cm} (2.8)

\[ C(x,s) = \frac{x^1s^2 + x^2s + x^3}{s(s+30)} \]  \hspace{1cm} (2.7)

In the simplest case, there are three sets of constraints on the design. First, there are positivity and boundedness constraints on \( x \):

\[ 0 \leq x \leq b \]  \hspace{1cm} (2.8)

Next, there is a time domain constraint on the step response of the closed loop system (i.e., on the response \( y \) resulting from \( r(t) = 1 \) for \( t \geq 0 \) and \( r(t) = 0 \) for \( t < 0 \)), as shown in Fig. 3. It has the form

\[ rl(t) \leq y(t,x) \leq ru(t) \text{ for all } t \in [0,T] \]  \hspace{1cm} (2.9)

where \( rl(t) \) and \( ru(t) \) are piece-wise constant functions.

Finally, it may be required that the poles (eigenvalues) of the closed loop system lie in a parabolic region in the complex plane defined by

\[ z + aw^2 + c \leq 0, \]  \hspace{1cm} (2.10)

where \( s = z + iw \), and \( a, c \geq 0 \) (see Fig. 4a). This may be achieved by making use of a modified Nyquist stability criterion, as follows. We may write \( P(s) = 1/d(s) \) and \( C(x,s) = n(x,s)/m(s) \), where \( d, n \) and \( m \) are polynomials in \( s \). Then the characteristic polynomial of the closed loop system is

\[ X(x,s) = n(x,s) + d(s)m(s) \]  \hspace{1cm} (2.11)

and its zeros are the same as those of the normalized rational function

\[ Z(x,s) = X(x,s)/(s+50)^5 \]  \hspace{1cm} (2.12)
$z + a \omega^2 + c = 0$

Desired Region for Closed Loopholes

Im = $K(Re)^2 - d$

Fig. 4a

$Z(x, s)$ - plane

Im = $K(Re)^2 - d$

$\omega = \infty$

$\omega''$

$\omega'$

$Z(x_1 - a \omega^2 - c + i\omega)$

Fig. 4b
It follows from arguments as for the Nyquist stability criterion, that a sufficient condition that for the zeros \( s^* \) of \( Z(x,s) \) to satisfy \( z^* + awz^* + cz^* \leq 0 \) is that the following semi-infinite inequality hold (see Fig. 4b):

\[
d - KR_0 (Z(x,-aw^2 - c + iw)^2 + Im(Z(x,-aw^2 - c + iw)) \leq 0 \quad \text{for all } w \in [w', w'']
\]  

(2.13)

where \( d, K > 0 \), \( Re \) and \( Im \) denote the real and imaginary parts, respectively, of \( Z \), and \( w', w'' \) define a critical range of values in the complex plane. It is seen (see Fig. 4b) that (2.13) ensures that the image under \( Z(x, \cdot) \) of the modified Nyquist locus given by (2.10) does not encircle the origin.

Finally, as a performance criterion, one could use

\[
J(x) = \int_0^T u(t,x)^2 dt
\]

(2.14)

for \( r(t) \) a unit step.

Problems such as these led to the development of the algorithms (9, 5, 20). For other examples of SISO control design via optimization, see (12, 15, 18, 34, 35).

2.3. Design of MIMO Control Systems.

We shall now consider an elementary example from the design of multiple-input multiple-output (MIMO) control systems in the Laplace transform domain, i.e., via a technique which replaces all the differential equations by their Laplace transforms. Referring to the block diagram in Laplace transform relations given in Fig. 5, the closed loop system is made up of blocks whose inputs and outputs are all \( m \) dimensional vector functions. The vector \( r(s) \) denotes the Laplace transform of the closed loop system input \( r(t) \), the \( m \) vector \( y(s) \) denotes the Laplace transform of the closed loop system output \( y(t) \), and the \( m \) vector \( u(s) \) denotes the Laplace transform of the plant input \( u(t) \). Finally, the \( m \) vector \( d(s) \) denotes the Laplace transform of the output disturbance \( d(t) \). The plant transfer function \( P(s) \) is an \( m \times m \) matrix whose elements are proper rational functions in the complex variable \( s \) and one is required to design the \( m \times m \) compensator transfer function matrix \( C(x,s) \) whose elements are proper rational functions in the complex variable \( s \), with coefficients which depend on the \( n \) dimensional design vector \( x \). Among other things, the closed loop system is required to be stable and to remain stable under plant perturbations and to be insensitive to the disturbance \( d(s) \). Both of these requirements lead to similar, nondifferentiable, semi-infinite inequality constraints and hence we will describe only the latter.

For the purpose of expressing insensitivity to the disturbance \( d(s) \), we set \( r(s) = 0 \), which leads to the equation

\[
y(s) = [I + P(s)C(x,s)]d(s) = Q(x,s)d(s)
\]

(2.15)

\[
u(s) = -C(x,s)Q(x,s)d(s) = R(x,s)d(s).
\]

(2.16)
Let $q(x,w)$ denote the largest singular value of $Q(x,iw)$ and $s(x,w)$ denote the largest singular value of $R(x,iw)$, with $w$ real. The variable $w$ denotes frequency; the matrix $Q(x,iw)$ is complex valued. Since the largest singular values are matrix norms, to make the response $y$ of the system small for a large class of disturbances, without unduly saturating the system as a result of $u$ becoming too large, control system designers strive to keep both $q$ and $r$ small over appropriate frequency ranges. This leads to the following partial formulation of the MIMO control system design problem in SIP form:

$$\text{minimize } f(x) = \{ \max q(x,w) \mid w \in [w', w''] \}$$ (2.17)

subject to:

$$s(x,w) \leq b(w) \text{ for all } w \in [w', w'']$$ (2.18)

$$c \leq x \leq d$$ (2.19)

where $b(w)$ is a continuous function of the frequency $w$.

In addition, there are constraints expressing decoupling i.e., the requirement that when only a single component of the input vector is a nonzero function, only the corresponding component of the output vector is nonzero, as well as other constraints on time domain responses, all of which are semi-infinite in form. We note that the singular values $q$ and $s$ are non-differentiable and hence that the SIP problem corresponding to MIMO control system design is considerably more difficult than the SISO problem. For algorithms which solve problems of the for (2.17)–(2.19), see (F9). For other examples and a discussion of the use of singular values in control system design, see (9, 13, 28, 32).

2.4. Electronic Circuit Design.

A well known SIP problem in electronic circuit design is the so called design centering, tolerancing and tuning problem. Its formulation is developed in three stages, as follows. First, the responses (see Fig. 6) of the circuit to given inputs must satisfy semi-infinite constraints of the form

$$a(t) \leq y(t,x) \leq b(t) \text{ for all } t \in [t', t'']$$ (2.20)

where $x$ is an $n$ dimensional design vector. Next, due to production errors, when the circuit is built, the actual value of the design parameter will not be the nominal value $x$, but $x + g(e)$, where $e$ is a production error parameter, usually of much smaller dimension than $x$. To obtain 100% yield, the design engineer must therefore require that

$$a(t) \leq y(t,x + g(e)) \leq b(t) \text{ for all } t \in [t', t'']$$ (2.21)

for all $e \in E$
where $E$ is the production engineer's tolerance set. Now, it is not uncommon for the inequalities of the form of (2.21) to define an extremely small, or even empty feasible set and hence engineers have to resort to post manufacture tuning, i.e., post manufacture correction of the product. This leads to the introduction of a repair (tuning) parameter $r$ into the picture, so that (2.21) becomes replaced by

$$a(t) \leq y(t,x+g(e)+h(r)) \leq b(t) \text{for all } t \in [t',t'']$$

(2.22)

for all $e \in E$, for some $r \in R$

where $R$ defines the allowable range of the repair parameter. Clearly, (2.22) is an extremely intractable kind of semi-infinite inequality, particularly, since a realistic formulation must replace the single $y$ in (2.22) with the max over all such functions of interest. A certain amount of simplification can be obtained by replacing (2.22) by the following, slightly stricter inequality:

$$a(t) \leq \max_{e \in E} \min_{r \in R} \max_{k \in K} y^k(t,x+g(e)+h(r)) \leq b(t)$$

(2.23)

for all $t \in [t',t'']$

Of course, in addition, to the constraint of the form (2.23), the electronic circuit design problem will also involve a number of others that are much more benign. Thus, the electronic circuit designer sees his problem as that of choosing a design center $x$ so that the set $x + g(E)$, resulting from tolerances which must be allowed in production, is as much as possible in the feasible set specified for $x$, and of making sure that any part of $x + g(x)$, which is outside of the feasible set, can be brought back into it by post-manufacture tuning.

For further reading, see (6, 7, 16, 22, 29).

3. SIP ALGORITHMS FOR ENGINEERING DESIGN.

As we have seen in the previous section, SIP problems in engineering design are of the form

$$\min \{ f(x) | g^i(x) \leq 0 \ \forall i \in I; \ h^j(x,p) \leq 0 \ \forall j \in J, \ \forall p \in P;$$

$$\max \min \max \max_{e \in E} \max_{r \in R} \max_{t \in T} \max_{k \in K} y^k(x,e,r,t) \leq 0\}$$

(3.1)

where $f$, $g$, $h$, and $y$ are, usually, real valued, continuously differentiable functions; $P$, $E$, $R$ are compact multidimensional sets, $T$ is a compact interval, and $K$ is a finite set of indices. When the functions in (3.1) are not continuously differentiable, they are semi-smooth (19) or, at least, locally Lipschitz.

We shall now describe two algorithms which we have found to be very reliable in engineering design. The first is due to Gonzaga, Mayne, Polak and Trahan (P1,9) and it solves problems of the form (3.1) without the max min max term, when all functions are continuously differentiable, i.e., it solves problems for the form:

$$\min \{ f(x) | g^i(x) \leq 0 \ \forall i \in I; \ h^j(x,p) \leq 0 \ \forall p \in P; \ \forall j \in J\}$$

(3.2)

The algorithm in (9) is a refinement of the algorithm in (20). In turn, it underwent some refinement, as described in (5), the most significant of which is the introduction of weights $w$ which can be used to improve the conditioning of the problem. For simplicity of exposition, we may as well consider the case of (3.2), where the "simple" inequalities specified by the $g^i$ are not present and there is only one semi-infinite inequality, i.e., $j = \{ 1 \}$, so that the index $j$ can be dropped. For this case, we define, with $e \geq 0$, $t \geq 0$, $w \geq 0$,

$$M(x) \max \{ h(x,p) | p \in P \}$$

(3.3)

$$M(x)_* = \max \{ M(x), 0 \}$$

(3.4)

$$P_e(x) = \{ p \in P | h(x,p) \geq M(x)_* - e, \ \text{and} \ p \ \text{is a local maximizer of}$$

$$h(x, \cdot) \ \text{in} \ P \}$$

(3.5)
where \( u^f, u^p \) solve (3.8). Finally, with \( 0 < c < 1 \), and \( q > 0 \), we define
\[
E = \{0, 1, c, c^2, c^3, \ldots\}
\]
and
\[
e(x) = \max\{e \in E | W_e(x) \geq sq\}
\]

We note that the parameter \( e \) in (3.5) controls the number of local maximizers that will be used in the computation of the search direction \( d_e(x) \). The parameters \( u^f, u^p \) in (3.6) are used to control the angle between the search direction \( d_e(x) \) and the corresponding gradients: when \( u^p \) is increased, the angle between \(-d_e(x)\) and \( \nabla h(x,p) \) is decreased. The parameter \( t \) controls the phase I - phase II aspect of the algorithm: when \( t \) is large, the algorithm places greater emphasis on achieving feasibility than when \( t \) is small. Finally, the parameter \( q \) is used to control the value of the parameter \( e \) that will be used in the computation of the search direction \( d_e(x) \): the smaller \( q \), the larger will \( e(x) \) be. In practice, this rather large number of parameters turns out to be rather easy to adjust, provided the computations are carried out in an interactive environment.

Finally, introducing two parameters \( a, b \in (0, 1) \) for the Armijo step size rule, we are ready to state the algorithm which will be recognized as a phase I - phase II method of feasible directions, see (23).

Algorithm 1:

Parameters: \( a, b, c \in (0, 1), q > 0, K > 0 \).

Data: \( x_0 \).

Step 0: Set \( i = 0 \).

Step 1: Compute \( d_e(x_i)(x_i) \).

Step 2: If \( M(x_i) > 0 \), compute the largest \( s \in \{K, bK, b^2K, \ldots\} \) such that
\[
M(x_i + sd_e(x_i)(x_i)) - M(x_i) \leq -saW_e(x_i)(x_i).
\]
Else, compute the largest \( s \in \{K, bK, b^2K, \ldots\} \) such that
\[
M(x_i + sd_e(x_i)(x_i)) \leq 0
\]
and
\[
f(x_i + sd_e(x_i)(x_i)) - f(x_i) \leq -saW_e(x_i)(x_i).
\]

Step 3: Set
\[
x_{i+1} = x_i + sd_e(x_i)(x_i)
\]
set \( i = i + 1 \), and go to Step 1.
It can be shown (see (9)) that every accumulation point of Algorithm 1 is feasible and satisfies a standard F. John type condition of optimality.

Algorithm 1 has been used as is and also as a vehicle for stabilizing local superlinearly converging algorithms such as sequential quadratic programming (27). The stabilization is accomplished as follows. A simple comparison test is used to determine whether the local version is in its region of convergence. If the local algorithm is in its region of convergence, it is allowed to proceed, otherwise Algorithm 1 is used to drive the iterates into the region of convergence of the local method. The result is an algorithm with mathematically demonstrable global convergence properties (see (27)).

For non-differentiable problems of the form (17) -(19), a modification of Algorithm 1 has been proposed in (28). The algorithm in (28) substitutes e-approximations to the generalized gradients for the gradients used in Algorithm 1, and uses the outer decompositions algorithm (10), to be described below, to replace the intervals \([w',w'']\) with finite sets of points.

Finally, for differentiable problems for which derivative computations have not been incorporated into simulation codes, we find a derivative free version of Algorithm 1 in (33).

An important tool for dealing with problems of the form (3.1), either totally or partially, is the family of outer approximations algorithms, see, e.g (10,14,16). Probably the most elaborate use of this tool can be found in (29), where it is used to "unwind" (3.1) into sequences of differentiable, finitely constrained optimization problems. We shall describe an example of these methods in terms of the simplified problem

\[
\min \{f(x) \mid h(x,p) \leq 0, \forall p \in P\} 
\]

where \(f\) and \(h\) are real, at least, locally Lipschitz continuous functions defined on \(R^n\). The idea is to decompose (3.14) into a sequence of problems in which \(P\) is replaced by finite sets \(P\), which, hopefully, remain of small cardinality.

**Algorithm 2:**

**Parameters:** \(t \in (0,1)\) and a double subscripted sequence \(\{s_{jk}\}\) satisfying

(i) \(s_{kk} = 0\) for all \(k\) and \(s_{jk} > 0\) for all \(k > j\);

(ii) \(s_{jk} \to s_j^*\) as \(k \to \infty\).

(iii) \(s_j^* \to 0\) as \(j \to \infty\).

**Data:** A finite set \(P_0 \in P\).

**Step 0:** Set \(k = 0\).

**Step 1:** Compute by means of Algorithm 1 an \(x_k\) such that

\[
h(x_k,p) \leq t^k, \forall p \in P, e(x_k) \leq t^k.
\]

**Step 2:** Computer a

\[
p_k \in \arg\max \{h(x,p) \mid p \in P\}.
\]

**Step 3:** Set

\[
P_{k+1} = \{p\} \cup \{p_j \in P_k \mid h(x_j,p_j) \geq s_{jk}\},
\]

set \(k = k + 1\) and go to Step 1.

We note that the parameter \(t\) is used to control the precision with which the \(k\)th problem is solved. It is necessary to solve the successive problems with progressively greater precision. In (3.14) an exponential increase of precision is specified, however, one may increase the precision at a slower rate. The double subscripted sequence \(s_{jk}\) is used for determining which of the \(p_j\) that were included in \(P_k\) need not be carried over into \(P_{k+1}\), i.e., they form part of the constraint dropping scheme. When one wishes to drop lots of constraints, the \(s\) should decrease to zero very slowly as \(k\) increases. There are many possibilities for such a sequence, for example, one may use \(s_{jk} = 100/\left[1/(1+j)^{1/10}-1/(1+k)^{1/10}\right]\). When computing
in environment, it is possible to exercise highly intelligent control over the constraint dropping mechanism.

In view of our comments on the effect of parameters, it should be clear that algorithms of the type discussed above are very difficult to use in batch mode in an engineering design where function evaluations consume many minutes of computer time. It has been our experience that the computations can be made much more efficient by parameter adjustment in an interactive computing environment, such as the ones described in (2,19).

4. CONCLUSION.

We hope that this brief survey of SIP problems arising in engineering and of some of the algorithms that have been used for their solution will stimulate the interest of SIP researchers in the very challenging problems that occur in engineering design.

5. ACKNOWLEDGEMENT: This research was supported by the National Science Foundation under grants ECS-79-13148 and CEE-81-05790 and the Joint Services Electronics Program under grant F49620-79-C-0178.

6. REFERENCES.


Bandler, J. W., P. C. Liu and H. Tromp:


Davison, E. J., and I. J. Ferguson:


Gonzaga, C., E. Polak, and R. Trahan:


Gonzaga, C., and E. Polak:


Hettich, R.,


Karmarkar J. S., and D. D. Siljak:


Laub, A. J.:


Mayne, D. Q., E. Polak, and R. Trahan:


Mayne, D. Q., and E. Polak:


Mayne, D. Q., E. Polak, and A. Voreadis:


Mayne, D. Q., E. Polak, and A. J. Heunis:


Mifflin, R.,


