A GRADIENT ALGORITHM FOR MINIMAX DESIGN

JAMES E. HELLER

UNIVERSITY OF ILLINOIS – URBANA, ILLINOIS
This work was supported in part by the Joint Services Electronics Program (U.S. Army, U.S. Navy, and U.S. Air Force) under Contract DAAB 07-67-C-0199; and in part by NSF GK-3893; also Air Force AFOSR 68-1579.

Reproduction in whole or in part is permitted for any purpose of the United States Government.

Distribution of this report is unlimited.
The minimax design of dynamic systems containing variable parameters is considered. The parameters fall into two categories: those which can be selected by the designer, \( x \), and the parameters which can neither be preset nor precisely measured, \( y \). For a given dynamic system the merits of the various parameter values are assumed to be summarized by a scalar performance index, \( J(x, y) \). Designs based on the minimax or worst case approach consist of selecting an \( x \) such that the maximum value of \( J(x, y) \) with respect to \( y \) is minimized. A gradient algorithm is presented for solving minimax problems. The term gradient is used here to denote the direction of steepest descent and is usually not a vector of partial derivatives evaluated at a point \((x, y)\). The algorithm is shown to converge to local solutions which satisfy the necessary conditions for a minimax solution. As the proofs of convergence do not restrict the solution to be a saddle point, the algorithm can be applied to both saddle point and non-saddle point solutions. In addition no convexity requirements on the function \( J(x, y) \) are necessary. Both the theoretical and the numerical aspects of the algorithm are discussed. Two examples are presented and the results of the examples illustrate the usefulness and simplicity of the procedure.

This report is a dissertation submitted in partial fulfillment of the requirements for the Ph.D. degree at the University of Illinois.
ACKNOWLEDGMENT

The author expresses sincere thanks to Dr. J. B. Cruz, Jr., his advisor, and Dr. J. Medanic' for their guidance and continued encouragement during the course of this research. He also appreciates the helpful discussions with Dr. P. Kokotović and Dr. W. R. Perkins. He further appreciates the discussions with Mrs. G. Metze, W. Imbriale, and S. Pieper. He expresses thanks to his wife Mary for her constant understanding, encouragement, and assistance during his graduate studies.

Finally, the author wishes to acknowledge the support of the Coordinated Science Laboratory, the National Science Foundation, and the Department of Electrical Engineering at the University of Illinois. Special thanks are due to Mrs. Sherry Kallembach for her expert typing of the manuscript.
# TABLE OF CONTENTS

1. INTRODUCTION ................................................. 1

2. A MINIMAX ALGORITHM ........................................... 6
   2.1 The Algorithm ........................................... 6
   2.2 Antenna Design Example .................................. 10

3. THEORY OF MINIMAX ............................................ 17
   3.1 Convergence of the Algorithm ............................. 17
   3.2 Termination of the Algorithm ............................. 21
   3.3 The Minimax Gradient ................................. 25

4. IMPLEMENTATION OF THE ALGORITHM ............................. 28
   4.1 Numerical Simplifications ............................... 28
   4.2 Orbital Transfer Example ................................. 30

5. CONCLUSION .................................................. 36

LIST OF REFERENCES ............................................. 38

APPENDIX (Figures 1 to 10) ..................................... 40

VITA ............................................................. 51
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Summary of the Minimax Algorithm</td>
<td>41</td>
</tr>
<tr>
<td>2</td>
<td>The Specified Symmetric Radiation Pattern</td>
<td>42</td>
</tr>
<tr>
<td>3</td>
<td>The Linear Array of Isotropic Radiators</td>
<td>43</td>
</tr>
<tr>
<td>4</td>
<td>Radiation Pattern of the Array at the Frequency ( f ) for Starting Value of ( d )</td>
<td>44</td>
</tr>
<tr>
<td>5</td>
<td>Envelope of Radiation Patterns over the Frequency Range for Starting Value of ( d )</td>
<td>45</td>
</tr>
<tr>
<td>6</td>
<td>Minimax Envelope of Radiation Patterns over the Frequency Range</td>
<td>46</td>
</tr>
<tr>
<td>7</td>
<td>Illustration of Satellite A</td>
<td>47</td>
</tr>
<tr>
<td>8</td>
<td>Final Trajectory of Satellite A</td>
<td>48</td>
</tr>
<tr>
<td>9</td>
<td>Portrayal of A Passing through the Admissible Sector of B</td>
<td>49</td>
</tr>
<tr>
<td>10</td>
<td>Trajectory of Satellite A through the Sector of B</td>
<td>50</td>
</tr>
</tbody>
</table>
1. INTRODUCTION

The selection of a mathematical model to represent a physical system is a major problem in computer aided system design. A precise mathematical model of the system is generally not available in most engineering applications. The form of the model may often be assumed, but the model parameters are known only a-priori to be contained in some admissible domain. This is true in a broad class of systems where over a short time period the parameters may be considered time invariant but change in the long run. Deterministic optimization theory can not be applied directly for the design of systems with uncertain time-invariant parameters.

The time invariant parameters of a given mathematical model can be divided into two classes, the controller parameters which can be selected by the designer and the plant parameters which can neither be preset nor precisely measured by the designer. The control parameter vector also includes the adjustable parameters of structures the designer can add to the original system. The added structure and model parameters adjustable by the designer will be denoted as the system controller while the remaining uncertain parameters are denoted as plant parameters.

Several approaches have been proposed for synthesizing a controller for systems with uncertain but time-invariant plant parameters which allow deterministic optimization theory to be applied. One approach is to determine at the time the system is to be used the equivalence class to which the plant parameter vector belongs. Two parameter vectors are equivalent, that is, belong to the same equivalence class, if they are associated with the same
optimal control. In the case of a fixed control structure with time invariant parameters the optimal control is the control parameter vector which is optimal for the specified performance index. After the plant parameter values are determined deterministic methods are used to obtain a control which is optimal with respect to the given cost or performance functional. A controller which combines both estimation and optimization together is designated an optimally adaptive controller. As parameter estimation may employ systems with high gains, it is usually susceptible to additive noise. A controller designed to be optimally adaptive in a noiseless case may not be capable of maintaining the performance functional near its minimum in a noisy situation.

The difficulties involved in parameter estimation and the problems of parameter estimation in the presence of noise have led to other approaches to the problem. A possible solution would be to design the controller such that for all plant parameter values the composite system would always give the same performance. This approach from the optimal standpoint is not too desirable as the structure does not allow the system performance to improve when the plant parameter values are favorable.

A desirable approach is to select a fixed control structure with adjustable parameters, \( x \). The performance function thus becomes a function of two vector variables, the controller parameter vector, \( x \), and the plant parameter vector, \( y \), and is denoted by \( J(x, y) \). A "super criterion", \( F(x) \), is then specified in which the dependence on the uncertain parameters, \( y \), is no longer explicitly contained. The control parameters which minimize the super criterion, \( F(x) \), are selected as the optimal control. The question as what to use as a super criterion thus arises. A super criterion of the form
with the proper choice of $p(y)$ could satisfy almost all cost criteria, where
$J(x,y)$ is the value of the original cost functional at the specified parameter values. The set $Y$ is the admissible domain of the uncertain but time invariant plant parameter vector. The term $p(y)$ is a weighting function and may be taken as a probability density function if optimality in a statistical sense is desired or selected to achieve other design objectives.

The difficulties in minimizing such an integral performance index are usually computationally prohibitive. In many cases the integral can be replaced by the upper bound

$$
\max_{y \in Y} [p(y) J(x,y)] \int \ldots \int dy_1 dy_2 \ldots dy_s
$$

or simply

$$
F(x) = \max_{y \in Y} [p(y) J(x,y)].
$$

The selection of an $x^*$ in the admissible domain, $X$, which minimizes such a super criterion is characterized by

$$
F(x^*) = \min_{x \in X} \max_{y \in Y} [p(y) J(x,y)].
$$

The term minimax or worst case design is used here to designate this upper bound approach. A minimax solution has been criticized as being too pessimistic.
but it has been shown\[^4,5\] that proper choice of a weighting function can overcome this criticism. By appropriatly modifying \(J(x, y)\) it can be assumed for notational simplicity that \(p(y) = 1.0\). This thesis will be restricted to the minimax design of systems with uncertain parameters.

The solution of a minimax problem is a saddle point if there exists a point \((x^*, y^*)\) such that

\[
J(x^*, y) \leq J(x^*, y^*) \leq J(x, y^*)
\]

and

\[
J(x^*, y^*) = \min_x \max_y J(x, y) = \max_y \min_x J(x, y)
\]

where \(x \in X\) and \(y \in Y\). Under the assumption that the solution to the problem is a saddle point the minimax problem may be considered as a particular differential game. Ho\[^6\] outlines variational methods to obtain the solution for classes of differential games usually considered in deterministic optimization theory. The relationship between differential games and certain automatic control problems is discussed in a number of articles in particular \[^7-9\] which also restrict themselves by assuming a saddle point solution.

A non-saddle point minimax solution, \((x^*, y^*)\), is characterized by

\[
\min_x \max_y J(x, y) \geq \max_y \min_x J(x, y)
\]

Cases for which equality holds in (1.7) are usually rare. The only relation involving an equality sign for non-saddle point which holds in all cases is
Very little has been done when considering the conditions under which a non-saddle point solution can be obtained. Neither the theoretical examinations \[10-13\] nor iteration schemes \[14-16\] previously presented gave a practical procedure for computer aided design of this type. In some articles \[15\] restrictions such as convexity of \(J(x,y)\) with respect to \(x\) were necessary in order to obtain a solution.

The main result of this thesis is to present a gradient method for the solution of minimax problems. The method is applicable for both saddle point and non-saddle point problems as the procedure converges to local solutions which satisfy the necessary conditions for a minimax point. Implementation of the algorithm is much easier than previously published methods. \[14,15\]
2. A MINIMAX ALGORITHM

2.1 The Algorithm

It is assumed that the merits of the various parameter values for a
given dynamic system can be summarized by a scalar performance index $J(x, y)$. The $r$ vector $x$ contains the controller parameters which can be selected by the designer and $y$ is a $s$ vector containing plant parameters whose values can neither be preset nor precisely measured. The admissible region for the $x$'s, $x \in \mathbb{R}^r$, is assumed to be a closed convex bounded set, and $Y$, the admissible space for the $y$'s, is a closed and bounded set in its respective Euclidean space $\mathbb{E}^s$. The usual metric topology generated by the Euclidean norm, square root of the sum of the squares, will be assumed. The set $Y$ for example could be a convex set of points such as a sphere or a finite set of points. Points in $Y$ for which no neighborhood of the point is contained in $Y$, for example the elements in a finite set, will be denoted as isolated points. The designer is thus faced with the problem of finding an admissible $x^*$ and $y^*$ such that

$$J(x^*, y^*) = \min_{x \in X} \max_{y \in Y} J(x, y).$$

The method to be presented is for obtaining such a minimax solution, $(x^*, y^*)$, when $J(x, y)$ satisfies the following conditions:

(a) $J(x, y)$ and the partial derivative of $J(x, y)$ with respect to $x$ are jointly continuous with respect to both $x$ and $y$ for all admissible parameter values.

(b) The function $J(x, y)$ for a fixed value of $x$ has only a finite number of local maxima in $Y$. 
Condition (a) is usually satisfied by most physical processes. Continuity with respect to $y$ is not necessary at isolated points. Condition (b) is sufficient for a solution to be realized computationally by the proposed algorithm. This condition as indicated by the examples is not so strong that it severely limits the use of the algorithm. A local maximum is here either a point $y_1 \in Y$ whose value of $J(x, y)$ is equal to or greater than the value of $J$ for all admissible points of $Y$ in the immediate neighborhood of $y_1$ or an isolated point. These conditions, (a) and (b), along with the assumptions about $X$ and $Y$ will be assumed throughout the remaining sections of the thesis.

The algorithm is as follows:

**Initialization**

The procedure is initialized with the choice of an admissible vector $x_0$ and the selection of four scalar parameters $\Delta > 0$, $0 < \beta < 1$, $h > 0$ and $\xi > 0$.

**The $i^{th}$ step**

The set $Y_\xi(x_i)$ is now determined, where

$$Y_\xi(x_i) \triangleq \{ y \in Y : [F(x_i) - J(x_i, y)] < \xi \text{ and } y \text{ is a local maximum of } J \}$$

and

$$F(x_i) \triangleq \max_{y \in Y} J(x_i, y).$$
The set \( Y_\varepsilon(x_i) \) consists of all points in \( Y \) globally maximizing \( J(x_i', y) \) and all points in \( Y \) which locally maximize \( J(x_i', y) \) such that the value of \( J \) at the local maximum is within \( \varepsilon \) distance of the value of \( J \) at the global maximum. The conditions placed on \( J \) as a function \( y \) for a given \( x \) guarantees the set \( Y_\varepsilon(x_i) \) is not empty.

The set of vectors \( S(x_i) \) is then obtained for the set \( Y_\varepsilon(x_i) \), \( Y_\varepsilon(x_i) = (y_1 \ldots y_m) \), where a vector is less than or equal to a scalar if every component is less than or equal to the scalar.

\[
S(x_i) \triangleq \{ y \in \Gamma : A_\varepsilon(x_i) \cdot y \leq -1 \} \tag{2.1.4}
\]

\[
A_\varepsilon(x_i) \triangleq \begin{bmatrix}
\frac{\partial J(x_i', y_1)}{\partial x} \\
\vdots \\
\frac{\partial J(x_i', y_m)}{\partial x}
\end{bmatrix}; \quad y \in Y_\varepsilon(x_i), i = 1, \ldots, m \tag{2.1.5}
\]

The set \( \Gamma \) is the cone of admissible directions and consists of all vectors \( y \) such that \( y = \alpha(x' - x) \), where \( x' \in X \) and \( \alpha \) is a positive scalar.

If the set \( S(x_i) \) is not empty it contains a smallest element with respect to the standard Euclidean norm. The smallest element is denoted by the vector \( y_i^* \) in \( S(x_i) \) and for all other \( y \in S(x_i) \)

\[
\|y_i^*\| < \|y\| \tag{2.1.6}
\]

The set \( S(x_i) \) is convex, thus the smallest element is unique since the norm function is strictly convex.
The direction $G_i$ defined by

$$G_i = \frac{1}{\|y_i^*\|} \cdot y_i^*$$  \hspace{1cm} 2.1.7

is a minimax generalization of the negative gradient direction usually defined in standard minimization. [17] A fixed step $h$ is now taken in the negative gradient direction and a point $x_{i+1}$ is obtained,

$$x_{i+1} = x_i + h \cdot G_i$$  \hspace{1cm} 2.1.8

If $F(x_{i+1})$ is greater than or equal to $F(x_i)$, $h$ is replaced by the product $\delta h$ and the step repeated. The parameter $h$ is also decreased if it causes a boundary condition to be violated, that is if $x_{i+1}$ is outside the admissible domain, $X$.

The control parameter vector $x_i$ is replaced by an admissible $x_{i+1}$ and the iteration begins again with the evaluation of $y_{x_{i+1}}$. The iteration procedure is illustrated in Figure 1.

**Termination**

The procedure terminates when either of two conditions are satisfied:

(a) The set $S(x_i)$ is empty.

(b) The norm of $y_i^*$ exceeds the value of $\Delta$.

Conditions (a) and (b) are based on the necessary conditions for a local minimax solution, as will be shown in Chapter 3. Condition (a) may only occur at a finite number of points in $X$ and numerically may be difficult to
satisfy. It will be shown that condition (b) is always satisfied as a local minimax solution is approached. A criterion for terminating in the proximity of a local minimax solution has not been proposed in the literature. Lack of such a termination criterion for non-saddle point problems presented the main obstacle for obtaining a minimax solution numerically.

A detailed justification and discussion concerning the implementation of the algorithm will be given in Chapters 3 and 4. First an example is presented to motivate the necessity of the various steps. Proofs concerning convergence of the algorithm will be given in Chapter 3.

2.2 Antenna Design Example*

The problem is to find the driving currents, $d_m$'s, for a linear array of seven discrete isotropic radiators so as to obtain the best approximation in terms of minimax design over a specified frequency range to a desired radiation pattern. The radiators have a given spacing, $x_o$, at the center frequency $f_o$. The radiation pattern for a linear array of discrete isotropic sources is given by

$$g(\theta) = \sum_{m=-M}^{M} d_m e^{j x_m \sin \theta} \tag{2.2.1}$$

where $d_m$ is the current distribution on the $m^{th}$ element, $2M+1$ is the number

---

*This example was proposed by Mr. W. Imbriale and Professor R. Mittra of the Department of Electrical Engineering and Antenna Laboratory at the University of Illinois.
of elements and \( x_m \) is the coordinate of the \( m^{th} \) element in terms of the \( \pi L/\lambda \) ratio.\(^{[18]} \) \( L \) is the length of the array and \( \lambda \) is the wavelength of the radiation.

The radiation pattern \( g(\theta) \) is generally specified in terms of both phase and magnitude but in the case studied only amplitude matching between a specified symmetric pattern, Figure 2, and the pattern produced by the symmetric array, Figure 3, were considered. The restriction of only matching the amplitude of the radiation patterns is not too severe as in a large class of design problems phase matching is not necessary. Synthesis of a desired pattern was carried out by comparing the Fourier series coefficients of the desired pattern to those of the pattern generated by the array at a given \( x \) and for a given current distribution. This method was selected as (2.2.1) leads to a Fourier series expansion of the magnitude squared of the symmetric array as follows:

\[
|g(\theta)|^2 = \sum_{n=0}^{\infty} \varepsilon_n 2^n \sum_{m=-M}^{M} \sum_{p=-M}^{M} d_m^* d_p J_{2n}(x_m-x_p) \cos[2n(\frac{\pi}{2} - \theta)]
\]

where \( \varepsilon_n = 1 \) for \( n = 0 \) and \( \varepsilon_n = 2 \) for \( n > 0 \). The \( J_{2n} \), \( n = 0, \infty \), are Bessel functions of the first kind. The * denotes complex conjugate. Equation 2.2.2 can be simply expressed as

\[
|g(\theta)|^2 = \sum_{n=0}^{\infty} \varepsilon_n 2^n \sum_{m=-M}^{M} \sum_{p=-M}^{M} d_m^* d_p J_{2n}(x_m-x_p) \cos[2n(\frac{\pi}{2} - \theta)]
\]

\*This derivation was presented by Mr. W. Imbriale in an unpublished note.
\[ |g(\theta)|^2 = \sum_{n=0}^{\infty} a_n \cos \left(2n\left(\frac{\pi}{2} - \theta\right)\right) \]

where
\[ a_n = \varepsilon_{2n} \sum_{m=-M}^{M} \sum_{p=-M}^{M} d_{m}^{*} d_{p} \times_{m,n} J_{2n}(x_{m,n} - x_{p}). \]

If the magnitude of the desired pattern, \(|p(\theta)|^2\), is expressed as a Fourier series
\[ |p(\theta)|^2 = \sum_{n=0}^{\infty} b_n \cos \left(2n(\pi/2 - \theta)\right) \]

the performance index or cost function, \(P_{I_n}\), is specified by
\[ P_{I_n}(x_0, d, f) = \sum_{i=0}^{n} (a_i - b_i)^2. \]

This \(P_{I_n}\) has been called \(J\) earlier but the notation was changed so as not to confuse it with Bessel functions. The functional \(P_{I_n}\) depends on the spacing at the center frequency \(f_0\), \(x_0\), current distribution, \(d\), and frequency, \(f\). It is desired to determine the set of currents such that for a fixed \(x_0\) at the center frequency
\[ \max_{f} P_{I_n}(x_0, d, f) \]

is minimized when \(f\) is allowed to vary over a specified range. This is a
worst case or minimax solution to the problem, where the frequency \( f \) is the
space of maximizing parameters and minimization is over the admissible set of
driving currents.

Reformulating the equation for \( a_n \), 2.2.4, in matrix notation, where
the \( t \) denotes transpose,

\[
a_n = e_{2n} j^{2n} \left\{ D_r^t J_{2n} - D_i^t J_{2n} \right\}.
\]

2.2.8

\( D_r \) and \( D_i \) are respectively the real and imaginary parts of the current
distribution vector

\[
d = \begin{bmatrix} d_{m-1} \\ \vdots \\ d_{-m} \end{bmatrix}
\]

2.2.9

and the matrix \( J_{2n} \) is defined by

\[
\begin{bmatrix} J_{2n} \end{bmatrix}_{ij} = J_{2n} (x_i - x_j)
\]

2.2.10

The gradient vector can be expressed by

\[
\frac{\partial P}{\partial D_i} = 2.0 \left( A - B \right)^t \frac{\partial A}{\partial D_i}
\]

2.2.11

\[
\frac{\partial P}{\partial D_r} = 2.0 \left( A - B \right)^t \frac{\partial A}{\partial D_r}
\]
where

\[ A = \begin{bmatrix} a_n & a_{n-1} & \vdots & a_0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & a_0 \end{bmatrix}, \quad B = \begin{bmatrix} b_n & b_{n-1} & \vdots & b_0 \end{bmatrix} \]

and

\[ \frac{\partial a_m}{\partial D_1} = 2.0 \epsilon_{2m} j^{2m} [D_1^t J_{2m}] \]

\[ \frac{\partial a_m}{\partial D_r} = 2.0 \epsilon_{2m} j^{2m} [D_r^t J_{2m}] \]

For the example considered, \( A \) and \( B \) contained sixteen components.

The initial value for \( d \) was selected by performing a minimization using the previously described performance index, \( P_{I_n} \), in order to match the given pattern at the center frequency \( f_0 \). The array contained seven elements. Thus there are only eight independent currents counting real and imaginary components because the array is symmetric. The pattern generated by the starting \( d \) at the center frequency is given in Figure 4 for the given \( x \) at this frequency of

\[ x_0^t = [-9.0, -7.0, -5.0, 0.0, 5.0, 7.0, 9.0]. \]

The envelope traced out by the patterns for the starting value of \( d \) as the frequency was varied over a two to one frequency range centered about \( f_0 \).
is given in Figure 5. This solution is not acceptable since the radiation pattern is no longer directional over the specified frequency range.

The value for $\xi$ was selected so that in the initial iterations only one element was contained in the set $Y_{\xi}(d)$. After a few iterations the current distribution vector, $d$, assumed values such that the set $Y_{\xi}(d)$ contained two elements. The function $P_{in}$ was convex with respect to frequency for the spacing considered. Thus the local maxima occurred at the extreme frequency limits, $y_L$ and $y_U$, placed on the allowable range. It is interesting to note that if $G_L(d)$ and $G_U(d)$ are the negative gradient vectors of the performance index with respect to $D_r$ evaluated at the extreme points

$$G_L(d) = -\frac{\partial P_{in}}{\partial D_r} (x_o, d, y_L)$$

$$G_U(d) = -\frac{\partial P_{in}}{\partial D_r} (x_o, d, y_U)$$

(2.2.15)

for all cases during the implementation when $Y_{\xi}(d)$ contained two elements,

$$G_L(d)^T G_U(d) < 0.$$  (2.2.16)

Similar equations are also valid for the imaginary parts of $d$. Any attempt to move in one of the directions in (2.2.15) would have resulted in a failure. This implies that another method such as the one presented in the algorithm is necessary for selecting the negative gradient direction.

After several iterations a current vector was obtained where
At such a point one of the previously reported procedures [16] would have terminated. This is because that method is only valid for saddle point solutions. The necessary conditions for a local minimax solution for this case as it will be shown in Chapter 3 are

\[ \Pi_n(x, d, y) = \Pi_n(x, d, y) \]

and

\[ G_u(d) = -aG_x(d) \]

for both real and imaginary parts of \( d \), where \( a \) is a positive scalar. To obtain a point \( d^* \) which approximately satisfies both these conditions a complex iteration procedure is necessary. The one presented in the algorithm is based on satisfying both necessary conditions.

The final solution obtained satisfied the necessary conditions for a local solution within reasonable numerical accuracy. The envelope of the patterns as the frequency is varied about \( f_0 \) is presented in Figure 6. Considerable improvement over Figure 5 is shown in the minimax envelope as the frequency is varied over the two to one frequency range. The final parameter values for \( d \) are,

\[ D_r = \begin{bmatrix} .086 \\ .106 \\ .087 \\ .155 \\ .087 \\ .106 \\ .086 \end{bmatrix} \quad \text{and} \quad D_1 = \begin{bmatrix} .064 \\ .079 \\ .066 \\ .116 \\ .066 \\ .079 \\ .064 \end{bmatrix} \]
3. THEORY OF MINIMAX

Various theoretical aspects of minimax will be investigated in this chapter and justification of all steps in the algorithm will be given. The function $J(x, y)$ and its first derivative with respect to $x$ are always assumed continuous with respect to admissible $x$ and $y$. The admissible domains for the $x$ and $y$ will be respectively a closed bounded convex set in $E_x$ and a closed bounded set in $E_y$.

3.1 Convergence of the Algorithm

The function $F(x)$ defined for admissible $x$ by

$$F(x) \triangleq \max_{y \in Y} J(x, y)$$

is continuous with respect to $x$ since $J(x, y)$ is continuous in both $x$ and $y$. The directional derivative in the $y$ direction, $||y|| = 1$, of the function $F(x)$ is defined as

$$D_y F(x) = \lim_{h \to 0} \frac{F(x + hy) - F(x)}{h}$$

if the limit exists. It was shown in Theorem I of Danskin [19] that in general $F(x)$ is not differentiable in $X$ but that the directional derivative $D_y F(x)$ of $F(x)$ in the direction $y$, $||y|| = 1$, at any point $x \in X$ exists and is continuous with respect to $y$. This result follows from the continuity of the partial derivative of $J$ with respect to $x$ in $x$ and $y$. The directional derivative as
given by \[19\] is
\[D_F(x) = \max_{\gamma \in \mathbb{Y}(x)} \left( \frac{\partial J(x, \gamma)}{\partial x} \right) \cdot \gamma\]  \hspace{1cm} 3.1.3

where \(\mathbb{Y}(x) = \mathbb{Y}_\xi(x)\) with \(\xi = 0\). The existence of directional derivatives allows a first order approximation of \(F\) for \(z \in \mathbb{X}\) about a point, \(x \in \mathbb{X}\), to be given by:
\[F(z) \approx \phi(z) = F(x) + D_F(x) \cdot h\]  \hspace{1cm} 3.1.4

where
\[h = ||z - x||\]

and
\[\gamma = \frac{1}{h} \cdot (z - x)\].  \hspace{1cm} 3.1.5

Given the matrix \(A(x)\) defined by
\[A(x) = \begin{bmatrix}
\frac{\partial J(x, \gamma_1)}{\partial x} \\
\vdots \\
\frac{\partial J(x, \gamma_m)}{\partial x}
\end{bmatrix}; \quad \gamma_i \in \mathbb{Y}(x)\]  \hspace{1cm} 3.1.6

then the first order approximation for \(F\), (3.1.4), may be given in matrix form.
\[\phi(z) = F(x) + \max \text{ entry of } \{A(x) \cdot (z - x)\}\]  \hspace{1cm} 3.1.7

The matrix \(A(x)\) therefore contains one row for each element in \(\mathbb{Y}(x)\). A cone of admissible vectors, \(C(x)\), for which \(\phi(z) < F(x)\) can now be defined, where \(\alpha > 0\) is a scalar and a relationship between a vector and a scalar means every
component of the vector satisfies the relationship with the scalar.

\[ C(x) \triangleq \{ \alpha (x'-x) : x' \epsilon x \text{ and } A(x)(x'-x) < 0 \} \quad 3.1.8 \]

If a point \( x^* \) is a local minimax solution then

\[
\text{maximum entry of } \{ A(x^*) \cdot (x-x^*) \} \geq 0 
\]

the vector

for all admissible \( x \) in the neighborhood of \( x^* \). This can be summarized in the form of a proposition involving \( C(x^*) \).

**Proposition 1.** A necessary condition for \( x^* \) to be a local minimax solution is that the cone of vectors \( C(x^*) \) having a vertex at \( x^* \) is equal to the empty set.

If a point does not satisfy proposition 1 other properties can be shown. The set \( C(x) \) contains all admissible directions in which the directional derivative is negative. If the directional derivative in the direction \( \begin{cases} \frac{1}{||\gamma||} \cdot \gamma \end{cases} \) is negative then a step \( h \) in this direction can be taken which results in a point \( z \) for which \( F(z) \) is less than \( F(x) \). If such an \( h \) did not exist then the directional derivative is equal to or greater than zero, a contradiction. Thus, the following proposition is in order:

**Proposition 2.** If at a point \( x \in X \) there exists a \( \gamma \in C(x) \) then there exists a scalar \( h > 0 \) such that

\[ F(x + h \gamma) < F(x). \quad 3.1.10 \]
Based on Proposition 2 a procedure can be formulated which converges to points satisfying Proposition 1. Assume \( x_0 \) is an arbitrary point in \( X \) and define a sequence in \( X \) \( \{x_n\} \) by

\[
x_{n+1} = x_n + h_n y_n
\]

where \( y_n \in C(x_n) \) if \( C(x_n) \) is not empty and \( h_n \) satisfies Proposition 2. If \( C(x_n) \) is empty for any \( n \) the corresponding \( x_n \) satisfies Proposition 1. The sequence \( \{F(x_n)\} \) is a strictly decreasing one and is bounded by \( \min_{x,y} \max_x \min_y J(x,y) \).

The sequence \( \{F(x_n)\} \) therefore converges to a limit. As \( X \) is closed and bounded, \( \{x_n\} \) must also converge to a limit point \( x^* \) which satisfies Proposition 1. If the point \( x^* \) did not satisfy Proposition 1 it would contradict the fact that \( F(x^*) \) is the limit of \( F(x_n) \) which follows from the continuity of \( F \). This can be summarized in the form of a proposition.

**Proposition 3.** The sequence of admissible points in \( X \), defined by equation (3.1.11), converges to a point \( x^* \) which satisfies the necessary conditions for a local minimax solution.

The algorithm presented in Figure 1 generates a sequence of admissible \( x_n \) which satisfies Proposition 3. Convergence to a point which satisfies the necessary conditions for a local minimax solution is guaranteed. To know a procedure eventually converges is comforting, but it is desirable to have a termination criterion which indicates when a point \( x_n \) is in the neighborhood of the solution \( x^* \). A neighborhood termination criterion will be presented in the next section.
3.2 Termination of the Algorithm

It is desirable to terminate the sequence when a point $x_n$ is in the neighborhood of a point which satisfies proposition 1. A set $Y(x_n)$ converges to a set $Y(x^*)$ if for every $y^* \in Y(x^*)$ there exists a distinct $y_n \in Y(x_n)$ such that the sequence $\{y_n\}$ converges to $y^*$ and there exists a one to one correspondence between elements of the sets $Y(x_n)$ and $Y(x^*)$ for each $n$. If $x_n$ is in the neighborhood of $x^*$ and $Y(x_n)$ converges to $Y(x^*)$ then from (3.1.6)

$$A(x_n) \approx A(x^*).$$

Termination is possible if a criterion is established to determine when the set $C(x_n)$ is approaching the empty set $C(x^*)$. The discussion of such a criterion will now be given.

Assume for all $n > N$ that $x_n$ is in the neighborhood of $x^*$ and $Y(x_n)$ converges to $Y(x^*)$. The matrices defined in equation 3.1.6, $A(x_n)$ and $A(x^*)$ are approximately the same as the partial of $J$ with respect to $x$ is continuous in both $x$ and $y$. The emptiness of the sets $C(x_n)$ is difficult to determine numerically so another set $C_k(x_n)$ will be defined. These sets as it will be shown are equivalent to $C(x_n)$ with respect to emptiness. The set $C_k(x)$ is defined by

$$C_k(x) \triangleq \{y \in C(x) : A(x) \cdot y \leq -k\}$$

It follows from the definition of $C_k(x)$ that it contains the same directions $(1/\|y\|) \cdot y$ as $C(x)$.
Proposition 4. The set $C_k(x)$ is empty if and only if the set $C(x)$ is empty.

Proof The emptiness of $C(x)$ implies $C_k(x)$ is empty as it is a subset of $C(x)$. Assume $C_k(x)$ is empty but not $C(x)$. Thus there exists a $y \in C(x)$ such that $A(x) \cdot y \leq -b < 0$. It then follows $A(x) \cdot (k/b \cdot y) \leq -k$ implying $k/b \cdot y \in C_k(x)$, a contradiction. Q.E.D.

Linear programming [20] or iterative methods [21] can be used to determine the emptiness of the sets $C_k(x_n)$. The next proposition illustrates how the minimal element of a non-empty set, $C_k(x_n)$, with respect to the standard Euclidean norm can be used as a stopping criterion.

Proposition 5. If there exists a sequence $\{x_n\}$ converging to a local minimax solution $x^*$ and if the sets $Y(x_n)$ converge to $Y(x^*)$ then $\|y_n^*\|$ of the non-empty $C_k(x_n)$ for a fixed $k$ approaches infinity, where $y_n^*$ is the minimal element with respect to norm of the set $C_k(x_n)$.

Proof Define the matrix $B_n$ by

$$B_n \triangleq A(x_n) - A(x^*)$$

As $n$ goes to infinity $B_n$ approaches a zero matrix. For any $n$ there exists a smallest $\zeta$ such that for each entry in $B_n$

$$|B_n|_{ij} < \zeta_n$$
For a given \( n \) and \( y^*_n \) it follows by definition that

\[
A(x_n) \cdot y^*_n \leq -k. \tag{3.2.5}
\]

The vector \( A(x^*) \cdot y^*_n \) contains at least one entry which is greater than or equal to zero as \( C_k(x^*) \) is empty. Assume the \( i^{th} \) row of \( A(x^*) \) causes such an entry.

\[
[A(x^*)]_i \cdot y^*_n \geq 0 \tag{3.2.6}
\]

Thus

\[
[B_n]_i \cdot y^*_n = [A(x_n)]_i y^*_n - [A(x^*)]_i y^*_n \leq [A(x_n)]_i y^*_n \leq -k \tag{3.2.7}
\]

Defining \( g_n \) as the largest component with respect to magnitude of \( y^*_n \) then

\[
r \xi_n g_n \geq k \tag{3.2.8}
\]

where \( \xi_n \) is defined by equation 3.2.4 and \( r \) is the dimensionality of \( x \). It then follows

\[
\| y^*_n \| \geq g_n \geq \frac{k}{r \xi_n} \tag{3.2.9}
\]

The scalar \( \xi_n \) goes to zero as \( n \) goes to infinity. The norm of the minimal element, \( \| y^*_n \| \), increases without bound as \( n \) goes to infinity. \( \text{Q.E.D.} \)
A difficulty encountered when using this criterion is that in general the limit of the set $Y(x_n)$ does not equal $Y(x^*)$. A simple example is given in Danskin [10] for which the limit is a strict subset of $Y(x^*)$. To avoid this difficulty the sets $Y_\xi(x_n)$ are used. The limit of $Y_\xi(x_n)$ as $n$ goes to infinity is $Y_\xi(x^*)$. The limit of $Y_\xi(x^*)$ as $\xi$ goes to zero is the set $Y(x^*)$. The set $C_k^\xi(x)$ is thus defined

$$C_k^\xi(x) \triangleq \{ y \in C_k(x) : A_\xi^k(x) \cdot y \leq -k \}$$

where $A_\xi(x)$ is given in (2.1.5). The norm of the smallest element of the cone $C_k^\xi(x_n)$ is now used as a terminating criterion under Proposition 5.

Several properties of $C_k(x)$ carry over to $C_k^\xi(x)$. The smallest element in $C_k^\xi(x)$ is larger than the smallest element of $C_k(x)$ as $C_k(x)$ contains $C_k^\xi(x)$. Emptiness of $C_k(x)$ implies emptiness of $C_k^\xi(x)$, a subset. The set $S(x)$ defined in (2.1.4) is equal to $C_1^\xi(x)$, the arbitrary $k$ was used here to generalize the proofs. This inclusion of the local maxima within $\xi$ distance of the global maxima in regard to the function value $F$ is a key issue in the success of a neighborhood termination criterion. The continuity of the sets $Y_\xi(x_n)$ follows directly from the next theorem.

**Theorem 1**

At every point $x$ in $X$ there exists a neighborhood $W \subset X$ containing the point $x$ such that for all $z \in W$ there is a $u \in Y_\xi(z)$ and a $y \in Y_\xi(x)$ where $\| u - y \| < \delta$ for any given scalar $\delta > 0$. 
Proof

If \( u^i \in Y_g(x) \) is an isolated point then trivially there exists a neighborhood \( W_i \) of \( x \) such that \( u^i \) is a local maximum for all \( z \in W_i \) and

\[
|J(z, u^i) - F(x)| < \xi
\]  
3.2.11

For a point \( u_j \in Y_g(x) \) where a neighborhood of \( u_j \) is contained in \( Y \), the admissible domain, there exists a real number \( r_j < \delta \) such that for all points \( u \in Y \) where \( \|u - u_j\| = r_j \)

\[
[F(x) - \xi] < J(x, u) < J(x, u_j)
\]  
3.2.12

Given \( J(x, y) \) is continuous with respect to both \( x \) and \( y \) there exists a \( W_j \in X \) such that for all \( z \in W_j \), at every \( u \), \( \|u - u_j\| = r_j \)

\[
[F(x) - \xi] < J(z, u) < J(z, u_j).
\]  
3.2.13

There exists a \( y_1 \in Y_g(z) \) which is within \( \delta \) of \( u_j \in Y_g(x) \) for each \( z \in W_j \) as a local maximum in \( y \) is contained inside the closed sphere. The neighborhood \( W \) of \( x \) is given by \( W = \cap_{i=1}^m W_i \). The set \( W \) is not empty as \( Y_g(x) \) contains only a finite number of elements. Q.E.D.

3.3 The Minimax Gradient

Proposition 2 indicates that any direction vector contained in \( C(x) \) determines an acceptable direction in which to move. As some directions may
be better to move in than others it is desirable to select the best, in some sense, possible direction. In accordance with the notation for classical minimization if \( \phi(x) \) denotes the linear approximation of \( F(x) \) about a point \( x \) given in (3.1.4) and (3.1.7) it is desirable to find a vector \( y* \in C(x) \) such that

\[
\phi(x + \frac{1}{\|y*\|} \cdot y*) = \min_{y \in C(x)} \phi(x + \frac{1}{\|y\|} \cdot y) \quad 3.3.1
\]

The vector \( \frac{1}{\|y*\|} \cdot y* \) which satisfies (3.3.1) will be defined as the negative gradient direction. The set \( C_k(x) \) contains the same directions as \( C(x) \) in terms of unit vectors. The following theorem indicates that the minimum element of non-empty \( C_k(x) \) specifies the negative gradient direction at \( x \in X \). If \( C_k(x) \) is empty a negative gradient direction does not exist and the point satisfies Proposition 1.

**Theorem 2**

If \( y* \) is the minimal element of \( C_k(x) \) with respect to the standard Euclidean norm then \( y* \) satisfies (3.3.1).

**Proof**

The maximum entry of the vector \( A(x) \cdot y* \) is \(-k\). Assume for an arbitrary \( y \in C_k(x) \) the maximum entry of \( A(x) \cdot y \) is \(-b\). It then follows that the maximum entry of the vector \( A(x) \cdot \frac{b}{k} y \) is \(-k \) and \( \frac{b}{k} y \) is also an element of \( C_k(x) \). Therefore \( \|y*\| \leq \frac{b}{k} \|y\| \). It follows directly that

\[
\text{maximum entry of } \left\{ A(x) \cdot \frac{1}{\|y*\|} \cdot y* \right\} \leq \text{maximum entry of } \left\{ A(x) \cdot \frac{b/k}{\|b/k \cdot y\|} \cdot y \right\} \quad 3.3.2
\]
and

\[
\text{maximum entry } \left\{ \begin{array}{c} A(x) \cdot \frac{1}{||y||} \cdot y \\ \end{array} \right\} = \text{maximum entry } \left\{ \begin{array}{c} A(x) \cdot \frac{b/k}{||b/k \cdot y||} \cdot y \\ \end{array} \right\}
\]

The statement of the theorem follows directly from (3.3.2), (3.3.3) and the definition of \( \phi(z) \) in (3.1.7). Q.E.D.

The set \( Y_x(x) \) takes into consideration regions in the \( Y \) set which may possibly be in the neighborhood of \( y \)'s which maximize \( J(z, y) \) for \( z \) very near to \( x \). Define the function \( \phi_x^*(z) \) based on \( Y_x(x) \) by

\[
\phi_x^*(z) = F(x) + \text{maximum entry } \left\{ A_x^*(x) \cdot (z-x) \right\}
\]

where \( A_x^*(x) \) is defined in (2.1.5). This function is also an upper bound for \( \phi(z) \) at the same value of \( x \). It is for these reasons that the \( Y_x^* \), the minimal element of \( C_k^x(x) \) specified by \( Y_x(x) \), is used instead of the minimal element of \( C_k(x) \) in defining the minimax negative gradient direction in the algorithm. It follows from Theorem 2 that a unit vector in the direction of the minimal element of \( C_k^x(x) \) is the negative gradient direction of \( \phi_x^*(z) \). Experience from the numerical examples indicates that this is a good direction to move in.

If the set \( Y_x(x) \) contains only one element then the minimax negative gradient direction coincides with the standard interpretation of the negative gradient direction, the negative of the partial derivative of \( J \) with respect to \( x \) evaluated at \( x \) and the \( y \in Y(x) \). The norm of the smallest element in this case is merely the reciprocal of the norm of the standard gradient.
4. IMPLEMENTATION OF THE ALGORITHM

4.1 Numerical Simplifications

In implementing the procedure presented in Figure 1 on a computer several simplifications are possible. It is best to start with the obtaining of the sets \( Y_{\xi}(x_{1}) \). For the case of \( Y \) containing only a finite number of points this is very simple. In many cases, such as in the next example or for convex functionals, it is possible to specify a finite set of points

\[
Y_{f} = \{y_{1}, y_{2}, \ldots, y_{f}\}
\]

such that for each \( x \in X \)

\[
\max_{y \in Y_{f}} J(x, y) \approx \max_{y \in Y} J(x, y).
\]

Application of the algorithm reduces to that of a finite set \( Y \) under this approximation and thus avoids the global search at each iteration. The calculation of the matrix \( A_{\xi}(x_{1}) \) completely specifies the set \( S(x_{1}) \). This requires the computation of the partial derivatives of \( J(x, y) \) with respect to \( x \) for the \( y \) vectors contained in the set \( Y_{\xi}(x_{1}) \).

The testing of the set \( S(x_{1}) \) to see if it is empty or to find the minimal element can be simplified in certain cases. Assume that \( y \in \Gamma \) in (2.1.4) can be expressed by inequality constraints,

\[
B y \leq c
\]
where $c$ is an $m$ vector and $B$ is a $r \times m$ matrix. (This form of constraint usually covers most problems studied.) The problem of finding the minimal element is equivalent to establishing that the set $D(x_i)$ is empty or find its minimal element where

$$D(x) = \{y: A y \leq e\}.$$  \hspace{2cm} 4.1.4

The augmented matrix $A$ and vector $e$ are given by

$$A = \begin{bmatrix} A_f(x) \\ B \end{bmatrix} \quad \text{and} \quad e = \begin{bmatrix} -1 \\ \vdots \\ -1 \\ -c \end{bmatrix}.$$  \hspace{2cm} 4.1.5

This same formulation also occurs when no constraints are put on acceptable directions. Either solving the quadratic programming problem of finding a $y$ which minimized the function

$$f(y) = y^t \cdot y$$  \hspace{2cm} 4.1.6

subject to the constraints

$$A y \leq e$$  \hspace{2cm} 4.1.7

or establishing such a solution does not exist is necessary.

The method of Hildreth and D'Esopo [22] for solving such a problem by mapping it into another quadratic programming problem using the Kuhn Tucker Theorem was applied. It is now necessary to find the vector $y^*$ which minimized the function
subject to the constraints $u \geq 0$. The solution of the original problem is given by

$$y^* = -\frac{1}{2} A^t u^* .$$

If the equations 4.1.7 are inconsistent it is easy to detect numerically that the limit of $H(u^*)$ decreases without bound when the Hildreth and D'Esopo iteration scheme is applied. The process is therefore terminated under these conditions. Other minimization techniques could be used to minimize $H(u)$ subject to the constraints. A minimization routine with boundary constraints was available and proved highly successful. Termination in this case occurred when $\|x_n\| > \Delta$ for all $n > N$ in the iteration scheme where

$$x_n = -\frac{1}{2} A^t u_n .$$

To numerically obtain the minimum element required considerable computer time in the case where the number of elements in $Y_\xi(x_1)$ is large.

The remaining steps in the implementation basically follow the block diagram in Figure 1. Implementation is relatively simple and no difficulties were encountered in solving the examples.

4.2 Orbital Transfer Example

This example considers two earth-orbiting satellites, one maneuverable and labelled A and the second passive and labelled B. Initially both satellite orbits are assumed to be circular and coplanar. At an initial time $t_0$ the
position of satellite A is known, however the initial position of satellite B is known only as contained in a particular sector. It is desired to transfer satellite A into another orbit such that A passes by satellite B as close as possible during its first revolution of the earth. In order to transfer satellite A two parameters are adjustable, the interval of time, $F_t$, starting from $t_0$ the satellite rocket engine is to be fired with a constant thrust and the angle, $\theta$, the engine makes with the radius, as shown in Figure 7. Both parameters are determined and set at time $t_0$ and remain fixed thereafter.

If $MD$ is the minimum distance between the two satellites during the first revolution for a given initial condition of B, in a worst case design $F_t$ and $\theta$ are selected such that the greatest $MD$ is minimized given satellite B was located in the admissible sector at time $t_0$.

A formal statement of the problem is as follows: The system parameters of satellite A are; the initial mass of the satellite, $M_o = 247$ slugs (8000 lbs.); the constant equivalent exit velocity of the rocket engine, $c = 1.0 \times 10^4$ ft/sec; the propellant mass flow, $\beta$, which is a constant for $t < F_t$, is .05 slugs/sec. The dynamic equations of satellite A are:

\[
\begin{align*}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= x_1 \cdot x_4^2 - \frac{GM}{x_1^2} - \frac{(C \cdot \beta \cdot \cos \theta)}{x_5} \\
\dot{x}_3 &= x_4 \\
\dot{x}_4 &= -2.0 \cdot x_2 \cdot x_4/x_1 + \frac{(C \cdot \beta \cdot \sin \theta)}{x_5} \\
\dot{x}_5 &= -\beta
\end{align*}
\]

where $\beta = .05$ for $t < F_t$ and 0 for $t \geq F_t$.  

The constant \( GM \) is \( 1.407 \times 10^{16} \text{ft}^3/\text{sec}^2 \) and \( \theta \) is the thrust direction angle. The initial conditions for satellite A are:

\[
x_1(t_0) = r_a = 4149 \text{ miles from the center of the earth}
\]
\[
x_2(t_0) = 0.0
\]
\[
x_3(t_0) = 0.0
\]
\[
x_4(t_0) = \omega_a
\]
\[
x_5(t_0) = M_a
\]

where \( \omega_a \) is the angular velocity of a satellite in a circular orbit at radius \( r_a \).

\[
\omega_a = \sqrt{GM/r_a^3}
\]

4.2.3

The quantity \((r_a, \phi_a)\) is the position of satellite A in earth centered polar coordinates and

\[
r_a = x_1(t)
\]
\[
\phi_a = x_3(t) \quad \text{(in radians)}.
\]

4.2.4

The position of satellite B is given by \((r_b, \phi_b)\), where

\[
\phi_b(t) = \omega_b(t-t_0) + \phi_b(t_0)
\]

4.1.5

\( \omega_b \) being the angular velocity of a satellite in a circular orbit of radius \( r_b \).

\[
\omega_b = \sqrt{GM/r_b^3}
\]

4.2.6
The initial conditions of satellite B are given only in the form of inequality constraints,

\[ 8359 \text{ miles} \leq r_b(t_o) \leq 8559 \text{ miles} \]

and

\[ 108.0^\circ \leq \phi_b(t_o) \leq 109.0^\circ \]

The distance between the satellites at any time is given by the law of cosines

\[ d(t) = \sqrt{r_a(t)^2 + r_b(t)^2 - 2r_a(t) \cdot r_b(t) \cdot \cos[\phi_a(t) - \phi_b(t)]} \]

The minimax problem to be solved is that of finding \( \theta \) and \( F_t \) such that

\[ J(\theta, F_t) = \max_{\phi_b(t_o)} \min_{t_o \leq t \leq T} d(t) \]

is minimized, where \( T \) is the time it takes \( A \) to go one revolution of the earth in the new orbit.

The gradient vector, \( \mathbf{G} \), for a given initial condition of satellite B

\[ \mathbf{G}^t = \left[ \frac{\partial \text{MD}}{\partial F_t}, \frac{\partial \text{MD}}{\partial \phi} \right] \]

was evaluated using finite differences and had the property
In the terminal stages of the algorithm the negative gradient direction as defined for the minimax case had the property

\[
\left| \frac{\partial MD}{\partial F_t} \right| \ll \left| \frac{\partial MD}{\partial \theta} \right| \tag{4.2.11}
\]

This change in magnitude aided in obtaining a solution. Large steps could be taken in the minimax negative gradient direction with considerable improvement in the cost function \( J(\theta, F_t) \). Any method using a standard gradient would have to chatter along a ridge to the solution. Because of the drastic change in magnitude of the components between the two gradients, the numerical accuracy of the method used to obtain the minimax negative gradient was limited.

A final solution of

\[
F_t = 1800.0 \text{ seconds} \tag{4.2.12}
\]

\[
\theta = 81.63^\circ \tag{4.2.13}
\]

was obtained and satellite A passed within 100 miles of satellite B for any admissible initial condition on B. This solution is impressive considering the sector of initial conditions is approximately 150 by 200 miles. The final trajectory of satellite A is given in Figure 8. A portrayal of the way A passes through the sector occupied by all possible locations of satellite B is given in Figure 9. The shape of the sector has changed because points on the lower boundary have a larger angular velocity than points on the upper boundary.
A plot of the way satellite A passes through the sector in a coordinate system which maintains the sector's original shape is given in Figure 10. This is a non-linear transformation from the real space, \((r_a(t), \phi_a(t))\), to the space \((r_a(t), \theta_a(t))\) where

\[
\theta_a(t) = \phi_a(t) - \left[ w_a^P \cdot (t-t_0) + \phi_a(t_0) \right]
\]

given

\[
w_a^P = \sqrt{\frac{GM}{r_a(t)^3}}
\]

This coordinate system preserves either horizontal or vertical distances but not combinations of both.
5. CONCLUSION

The minimax design of time invariant dynamic systems containing both parameters adjustable by the designer and uncertain parameters has been discussed. A gradient type algorithm was presented for use in obtaining minimax solutions numerically. The algorithm presented was shown to converge to local solutions which satisfy the necessary conditions for a minimax solution. Convergence of the algorithm was proven for both saddle point and non-saddle point solutions. Several of the numerical aspects which arise when implementing the algorithm were discussed. Two examples were presented which illustrate the usefulness and simplicity of the procedure.

The method offers several advantages over previous methods [14,15] when coded for a computer. The minimax gradient direction and the inclusion of local maxima in Y by the algorithm are new concepts which have not appeared in previous minimax studies. These concepts are shown to be helpful in the study of minimax problems. Their introduction also indicates new areas of study for minimax design.

Several extensions of the algorithm are possible. The first extension is to consider time varying parameters. Theoretically, similar results for the time varying case follow directly. Judging from the computer time necessary to carry out the various stages of the algorithm the solution time for the time varying case would be a prohibitive factor. In solving the two examples on a Control Data 1604, forty five minutes of machine time was necessary to solve the antenna design example and three hours were consumed in the orbital transfer problem.
It is also possible to take into consideration boundary constraints on the parameters adjustable by the designer other than the inequality constraints considered. Development of this area would follow the theory developed for gradient projection methods [22]. Finally, many of the ideas presented which are used to obtain the minimax negative gradient direction can be applied in developing a gradient projection or ridge tracking minimization procedure.
LIST OF REFERENCES


APPENDIX (Figures 1 to 10)
Figure 1. Summary of the Minimax Algorithm.
Figure 2. The Specified Symmetric Radiation Pattern.
Figure 3. The Linear Array of Isotropic Radiators.
Figure 4. Radiation Pattern of the Array at the Frequency $f_0$ for Starting Value of $d$. 
Figure 5. Envelope of Radiation Patterns over the Frequency Range for Starting Value of d.
Figure 6. Minimax Envelope of Radiation Patterns over the Frequency Range.
Figure 7. Illustration of Satellite A.
Figure 8. Final Trajectory of Satellite A.
Figure 9. Portrayal of A Passing through the Admissible Sector of B.
Figure 10. Trajectory of Satellite A through the Sector of B.
VITA

James Ernest Heller was born on September 28, 1943 in Elizabeth, New Jersey. He received the Bachelor of Science degree in Electrical Engineering with highest honors from Newark College of Engineering in June, 1965. He came to the University of Illinois in September, 1965 as a recipient of a National Science Foundation Traineeship in Electrical Engineering. From September, 1966 to February, 1967 he was a teaching assistant in the Department of Electrical Engineering. He received his Master of Science degree from the University of Illinois in February, 1967. From February, 1967 to February, 1969 he has been a research assistant in the Control Systems Group of the Coordinated Science Laboratory at the University of Illinois.

Mr. Heller is a member of Eta Kappa Nu and Tau Beta Pi. He is also a member of the Institute of Electrical and Electronic Engineers and a junior member of the National Society of Professional Engineers.
Distribution List as of November 1, 1968

1 Dr A.A. Deen
Asst Director (Research)
DG of Defense Res & Eng
Department of Defense
Washington, D.C. 20301

1 Office of Deputy Director
(Research and Technology)
ODD R&E-OSD
The Pentagon, Room 3-E-144
Washington, D.C. 20301

1 Director Advanced Research Projects Agency
Department of Defense
Washington, D.C. 20301

1 Director for Information Sciences
Advanced Research Projects Agency
Department of Defense
Washington, D.C. 20301

1 Director for Materials Sciences
Advanced Research Projects Agency
Department of Defense
Washington, D.C. 20301

1 Headquarters
Defense Communications Agency (333)
The Pentagon
Washington, D.C. 20305

20 Defense Documentation Center
Arts: DDDA
Cameron Station, Bldg. 5
Alexandria, Virginia 22314

1 Director
National Security Agency
Arts: Librarian C-338
Fort George G. Meade, Maryland 20755

1 Weapons Systems Evaluation Group
Arts: Col Daniel W. McElwee
Washington, D.C. 20305

1 National Security Agency
Arts: Asst Director (Research)
Office of Research
Fort George G. Meade, Maryland 20755

1 Central Intelligence Agency
Arts: OCH/BD Publications
Washington, D.C. 20305

1 Colonel
AFOSR
USAF, USAF
Room 10-191, The Pentagon
Washington, D.C. 20330

1 Aerospace Medical Division
Arts: Maj (Res) R.C. tenga
Brooks Air Force Base, Texas 78235

1 AILTL-9643
Nevayil AF, Alabama 36112

1 AFPTC (FETP-2)
Technical Library
Edwards AF, Calif. 93523

1 Lt SAMO (EHTA/SWS Albino)
AF Block Post Office
Los Angeles, California 90043

1 Lt Col Charles W. Waespy
USAF (AHFOS) Pentagon
Washington, D.C. 20330

1 NAR (FETP/Lt Stanback)
AFUPO
Los Angeles, California 90045

1 Det #8, OAR (LORC)
Air Force Post Office
Los Angeles, California 90045

1 AIC (H/3)
Wright-Patterson AF, Ohio 45433

1 Dr R.W. Noble
Wright-Patterson AF, Ohio 45433

1 Mr Peter Harvey
Wright-Patterson AF, Ohio 45433

1 AFAP (AFPS/MA. L. Larson)
Wright-Patterson AF, Ohio 45433

2 Commanding General
Arts: JUNVE-4K-X7
White Sands Missile Range
New Mexico, 88002

1 KANE (EHA), Griffin AF, New York 13442
Arts: Documents Library

1 Mr R.H. Webb (EWSA)
Rose Air Development Center
Griffiss AFB, New York 13442

1 Academy Library (AECF)
H. H. Air Force Academy
Colorado Springs, Colorado 80912

1 Mr Horton H. Payne, Chief
AFSC Scientific and Liaison Office
24 Federal Plaza
New York, N.Y. 10007

1 Lt Col Bernard S. Morgan
Frank J. Seiler Research Laboratory
U.S. Air Force Academy
Colorado Springs, Colorado 80912

1 Technical Library, AFET
EFP, MC-131
Patrick AFB, Florida 32925

1 AFETY (G51-1)
STING Office (For Library)
Patterson AFB, Florida 32925

1 Dr L. W. Helfingworth
AFRC (CNS)
L.D. Hancock Field
Bedford, Massachusetts 01731

1 AFRC (CNSH)
AFRC Research Library, Stop 29
L.D. Hancock Field
Bedford, Mass 01731

1 Colonel Robert E. Fontana
Head of Electrical Engineering
Air Force Institute of Technology
Wright-Patterson AF, Ohio 45433

1 Colonel A.H. Blue
RTS (KX51)
Bellingham Air Force Base, D.C. 30332

1 Dr L. Hiram
AFR (CNS)
Andrew AF, Maryland 20331

1 AFRC (CNS)
Andrew AF, Maryland 20331

1 Lt Col J.S. Bowers
AFRC (CNS)
Andrews AF, Maryland 20331

2 SRO (D211)
L.D. Hancock Field
Bedford, Mass 01731

1 AEC (ADO, DMC)
Arts: Library/Documents
Arnold AF, Texas 77389

2 European Office of Aerospace Research
Shell Building
47 Bus Centerstrasse
Brussels, Belgium

5 Lt Col Robert R. Latshaw
Chief, Electronics Division
Directorate of Engineering Sciences
Air Force Office of Scientific Research
Arlington, Virginia 22209

1 ANGS (NSC-11)
Elgin AF, Florida 33562

1 U.S. Army Research Office
Arts: Physical Sciences Division
Washington, D.C. 20315

1 Research Plans Office
U.S. Army Research Office
1005 Columbia Pike
Arlington, Virginia 22204

1 Commanding General
U.S. Army Materiel Command
Arts: AMRD-TF
Washington, D.C. 20315

1 Commanding General
U.S. Army Strategic Command
Fort Humphreys, Arizona 85413

1 Commanding Officer
Army Materiel & Mech. Res. Center
Waterstone Arsenal
Waterstone, Mass. 0172

1 Commanding Officer
U.S. Army Ballistics Research Laboratory
Arts: AMRD-942
Abraham Paving Ground
Abraham, Maryland 21005

1 Commandant
U.S. Army Air Defense School
Arts: Military Sciences Division
C & S Stages
Fort Bliss, Texas 79916

1 Commanding General
U.S. Army Missile Command
Arts: Technical Library
Redstone Arsenal, Alabama 35809

1 U.S. Army Munitions Command
Arts: Technical Information Command
Ficcicary Arsenal
Dover, New Jersey 07801

1 Commanding Officer
Harry Diamond Laboratories
Arts: Dr. Berthold Altman (AMRD-911)
Connecticut Ave & Van Ness St. N.W.
Washington, D.C. 20436

1 Commanding Officer
U.S. Army Security Agency
Arlington Hall
Arlington, Virginia 22212

1 Commanding Officer
U.S. Army Limited War Laboratory
Arts: Technical Director
Abraham Paving Ground
Abraham, Maryland 21005

1 Commanding Officer
Human Engineering Laboratories
Abraham Paving Grounds
Abraham, Maryland 21005

1 Commander
U.S. Army Command & General Staff
College
Army, Secretary
Fort Leavenworth, Kansas 6670

1 Commanding Officer
U.S. Army Research Office (Dwight)
Arts: COO-Mal (Richard S. Olah)
Box Chun, Duke Station
Durham, North Carolina 27706

1 Librarian
U.S. Army Military Academy
West Point, New York 10996
The minimax design of dynamic systems containing variable parameters is considered. The parameters fall into two categories: those which can be selected by the designer, \( x \), and the parameters which can neither be preset nor precisely measured, \( y \). For a given dynamic system the merits of the various parameter values are assumed to be summarized by a scalar performance index, \( J(x, y) \). Designs based on the minimax or worst case approach consist of selecting an \( x \) such that the maximum value of \( J(x, y) \) with respect to \( y \) is minimized. A gradient algorithm is presented for solving minimax problems. The term gradient is used here to denote the direction of steepest descent and is usually not a vector or partial derivatives evaluated at a point \( (x, y) \). The algorithm is shown to converge to local solutions which satisfy the necessary conditions for a minimax solution. As the proofs of convergence do not restrict the solution to be a saddle point, the algorithm can be applied to both saddle point and non-saddle point solutions. In addition no convexity requirements on the function \( J(x, y) \) are necessary. Both the theoretical and the numerical aspects of the algorithm are discussed. Two examples are presented and the results of the examples illustrate the usefulness and simplicity of the procedure.
<table>
<thead>
<tr>
<th>KEY WORDS</th>
<th>LINK A</th>
<th>LINK B</th>
<th>LINK C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimax</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worst Case Design</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optimal Control</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Automatic Control Systems</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>