INTERNAL REPORTS IN

SIMULATION, OPTIMIZATION AND CONTROL

No. SOC-229

PROOF OF GLOBAL CONVERGENCE AND RATE OF CONVERGENCE FOR A ONE-DIMENSIONAL MINIMAX ALGORITHM

H.L. Abdel-Malek, J.W. Bandler and R.M. Biernacki

February 1979

FACULTY OF ENGINEERING
McMASTER UNIVERSITY
HAMILTON, ONTARIO, CANADA

PROOF OF GLOBAL CONVERGENCE AND RATE OF CONVERGENCE FOR A ONE-DIMENSIONAL MINIMAX ALGORITHM

H.L. Abdel-Malek, J.W. Bandler and R.M. Biernacki

Abstract This report studies the convergence properties of the one-dimensional minimax algorithm developed by Abdel-Malek and Bandler to handle biquadratic functions. It is shown, as expected, that the algorithm converges from any set of starting conditions. Furthermore, under mild conditions we show that the rate of convergence is at least of second order. The rare case when the minimax solution is defined by functions whose derivatives vanish at that solution is considered in some detail. Minor modifications to improve the algorithm are suggested.

This work was supported by the Natural Sciences and Engineering Research Council of Canada under Grant A7239.

H.L. Abdel-Malek is with the Department of Engineering Physics and Mathematics, Faculty of Engineering, Cairo University, Giza, Egypt.

J.W. Bandler is with the Group on Simulation, Optimization and Control and Department of Electrical Engineering, McMaster University, Hamilton, Canada, L8S 4L7.

R.M. Biernacki is on leave from the Institute of Electronics Fundamentals, Warsaw Technical University, Warsaw, Poland.

I. INTRODUCTION

Abdel-Malek and Bandler [1] have presented some new results involving the biquadratic function obtained from the modulus squared of the bilinear network function. They presented a globally convergent and extremely efficient minimax algorithm for obtaining minimax solutions to sets of biquadratic error functions.

This report studies the convergence properties of the algorithm in some detail. The global convergence is verified, and the rate of convergence under different conditions is examined.

II. ALGORITHM

We begin by briefly describing the one-dimensional algorithm given by Abdel-Malek and Bandler [1].

Consider the minimax problem

$$\underset{\phi}{\text{minimize}} \quad \max_{1 \leq i \leq m} \quad e_i(\phi), \tag{1}$$

where the $e_i(\phi)$ are biquadratic functions of the form

$$e_i(\phi) = \frac{A_i + 2B_i\phi + C_i\phi^2}{1 + 2D_i\phi + E_i\phi^2} \, . \tag{2}$$

Fig. 1 illustrates the algorithm for solving this problem. The following steps set it out in sufficient detail, with appropriate definitions to be used subsequently in the convergence proofs.

<u>Step 1</u> Initialize ϕ.

<u>Step 2</u> Find

$$\delta = \max_i e_i(\phi) \tag{3}$$

<u>Step 3</u> Find intervals $I_i \triangleq [\check{\phi}_\ell, \hat{\phi}_\ell]$ and $\check{i}_\ell$, $\hat{i}_\ell$, $\ell = 1, 2, \ldots, k$, such that, for all i, $e_i(\phi) \leq \delta$ for $\phi \in \cup I_\ell$, and that there exists i such that $e_i(\phi) > \delta$ for $\phi \notin \cup I_\ell$, $e_{\check{i}}(\check{\phi}_\ell) = e_{\hat{i}}(\hat{\phi}_\ell) = \delta$.

<u>Comment</u> This is carried out using the algorithm presented by Abdel-Malek and Bandler [1].

<u>Step 4</u> Find $\check{g}_\ell$ and $\hat{g}_\ell$, $\ell = 1, 2, \ldots, k$, given by

$$\check{g}_\ell = \left. \frac{de_{\check{i}_\ell}}{d\phi} \right|_{\check{\phi}_\ell}, \tag{4}$$

$$\hat{g}_\ell = \left. \frac{de_{\hat{i}_\ell}}{d\phi} \right|_{\hat{\phi}_\ell}, \tag{5}$$

<u>Comment</u> These are simply the sensitivities at the extreme points of each valid interval. It is to be noted that $\check{g}_\ell \leq 0$ and $\hat{g}_\ell \geq 0$.

<u>Step 5</u> If $k = 1$, set $j \leftarrow 1$ and go to Step 7.

<u>Step 6</u> Find j such that

$$\Delta_j \geq \Delta_\ell, \quad \ell = 1, 2, \ldots, k, \tag{6}$$

where

$$\Delta_\ell = \begin{cases} \hat{g}_\ell \check{g}_\ell (\hat{\phi}_\ell - \check{\phi}_\ell)/(\check{g}_\ell - \hat{g}_\ell). \\ 0 \text{ if } \check{g}_\ell = \hat{g}_\ell = 0. \end{cases} \tag{7}$$

<u>Comment</u> In this step we select the jth interval which appears to be the

most promising one in terms of the expected improvement in the minimax optimum based on linearization. Δ_ℓ will always be positive unless either $\check{g}_\ell = 0$, $\hat{g}_\ell = 0$ or $\check{\phi}_\ell = \hat{\phi}_\ell$.

<u>Step 7</u> Set

$$\phi^* \leftarrow (\check{g}_j\check{\phi}_j - \hat{g}_j\hat{\phi}_j)/(\check{g}_j - \hat{g}_j) \text{ if } \check{i}_j \neq \hat{i}_j \text{ and } \Delta_j \neq 0. \qquad (8)$$

<u>Comment</u> If the extremes of the jth interval are defined by two different functions, the new value ϕ is taken as the intersection of the linear approximation to the two functions.

<u>Step 8</u> Set ϕ^* to the minimizing point of the function $e_{\check{i}_j}$ if $\check{i}_j = \hat{i}_j$.

<u>Step 9</u> Set

$$\phi^* \leftarrow (\check{\phi}_j + \hat{\phi}_j)/2 \text{ if } \phi^* \notin (\check{\phi}_j, \hat{\phi}_j) \text{ or } \Delta_j = 0.$$

<u>Comment</u> This is a default value to obviate any numerical problem which may arise in Step 6 or Step 7, for example, $\hat{g}_j = 0$.

<u>Step 10</u> Set $\phi \leftarrow \phi^*$.

<u>Step 11</u> Stop if k = 1 and if $(\hat{\phi}_1 - \check{\phi}_1)$ is sufficiently small.

<u>Step 12</u> Go to Step 2.

In the following sections, superscript n will denote the index of iteration of the algorithm.

III. THE MAIN RESULT

Theorem

If $I^n \triangleq [\check{\phi}^n, \hat{\phi}^n]$ is a unique interval such that $e_i(\phi) \leq \delta^n$ for $i = 1, 2, \ldots, m$, then $|\hat{\phi}^n - \check{\phi}^n| \to 0$ as $n \to \infty$. The rate of convergence is at least of second order.

Proof of Convergence

Let us consider two different functions $e_{\check{i}}(\phi)$ and $e_{\hat{i}}(\phi)$ which define the extreme points $\check{\phi}^n$, $\hat{\phi}^n$ of I^n. The proof is obvious from Step 8 of the algorithm if only one function is considered. Without loss of generality, we can assume that $e_i(\phi) \leq e_{\check{i}}(\phi)$, for $\phi \in \check{I}^n$ and $e_i(\phi) \leq e_{\hat{i}}(\phi)$ for $\phi \in \hat{I}^n$, $i = 1, 2, \ldots, m$, where $\check{I}^n \triangleq [\check{\phi}^n, \phi_{min}]$, $\hat{I}^n = [\phi_{min}, \hat{\phi}^n]$ and ϕ_{min} is the unique intersection point of $e_{\check{i}}(\phi)$ and $e_{\hat{i}}(\phi)$ in the interval I^n. There is also no loss of generality if we assume $\check{\phi}^{n+1} = \phi^{n*}$ for all n and that $\check{g}^n < 0$, $\hat{g}^n > 0$, since there is only a finite set of ϕ for which the derivative is zero. We will show that there exists a value $\gamma < 1$ such that

$$|\hat{\phi}^{n+1} - \check{\phi}^{n+1}| \leq \gamma \, |\hat{\phi}^n - \check{\phi}^n| \text{ for any n.} \tag{9}$$

Since $\hat{g}^n > 0$ the interval I^{n+1} can be estimated as follows. We have

$$\hat{\phi}^{n+1} - \check{\phi}^{n+1} = \hat{\phi}^{n+1} - \phi^{n*} < \hat{\phi}^n - \phi^{n*} = \frac{\check{g}^n}{\check{g}^n - \hat{g}^n} (\hat{\phi}^n - \check{\phi}^n). \tag{10}$$

If $e_{\hat{i}}(\phi)$ is such that $e'_{\hat{i}}(\phi) \geq \xi$ for any $\phi \in \hat{I}^n$, where ξ is a sufficiently small positive number, we will find γ as

$$\gamma = \frac{\eta}{\eta+\xi} < 1, \tag{11}$$

where

$$\eta = \max_{\phi \in \check{I}^n} -e'_{\check{i}}(\phi).$$

The above estimate is not possible only if $e'_{\hat{i}}(\phi) \to 0$ and $e'_{\check{i}}(\phi) \to c \neq 0$ for $\phi \to \phi_{min}$[1]. In this case the function $e_{\hat{i}}(\phi)$ becomes convex in the interval $[\phi_{min}, \hat{\phi}^n]$ for any $n \geq N$ when N is sufficiently large. Then the interval I^{n+1} can be estimated as follows

$$\hat{\phi}^{n+1} - \check{\phi}^{n+1} < \hat{\phi}^{nL} - \phi^{n*}, \tag{12}$$

where $\hat{\phi}^{nL}$ is the intersection point of the linearization of $e_{\hat{i}}(\phi)$ at the point $\hat{\phi}^n$ and the line $\delta^{n+1} = e_{\check{i}}(\phi^{n*})$.

From the appropriate geometrical relations (see Fig. 2) we obtain

$$\frac{\hat{\phi}^{nL} - \phi^{n*}}{\hat{\phi}^n - \phi^{n*}} = \frac{\Delta^n + e_{\check{i}}(\phi^{n*}) - e_{\check{i}}(\check{\phi}^n)}{\Delta^n}. \tag{13}$$

After some manipulations, we find that

$$\hat{\phi}^{nL} - \phi^{n*} = \frac{R^n}{\Delta^n}(\hat{\phi}^n - \phi^{n*}) = \frac{R^n}{\hat{g}^n}, \tag{14}$$

where R^n is the second-order remainder of Taylor's formula for the function $e_{\check{i}}(\phi^{n*})$ at the point $\check{\phi}^n$. It can be written in the form

$$\hat{\phi}^{nL} - \phi^{n*} = \left(\frac{1}{2}\check{q}^n \hat{g}^n \frac{\hat{\phi}^n - \check{\phi}^n}{(\check{g}^n - \hat{g}^n)^2}\right)(\hat{\phi}^n - \check{\phi}^n), \tag{15}$$

where $\check{q}^n$ is the second derivative of $e_{\check{i}}(\phi)$ at some $\phi \in \check{I}^n$. Since $\check{q}^n$ is limited by a number ζ and $\hat{g}^n \to 0$, $\check{g}^n \to c \neq 0$, we can find a sufficiently

---

[1] If both $e'_{\check{i}}(\phi) \to 0$ and $e'_{\hat{i}}(\phi) \to 0$ the rates of convergence of $\check{g}^n$ and $\hat{g}^n$ are of the same order and it is possible to find an estimation of $\hat{g}^n/\check{g}^n$ such that (9) is satisfied. See Appendix for details.

large number $N_1 \geq N$ such that the number

$$\gamma = \frac{1}{2} \zeta \hat{g}^{N_1} \frac{\hat{\phi}^{N_1} - \check{\phi}^{N_1}}{c^2} < 1 \qquad (16)$$

satisfies the condition (9) for all $n \geq N_1$. But according to (10) the interval I^{N_1} can be reached after a finite number of steps since $\hat{g}^n$ and $-\check{g}^n$ are greater than sufficiently small positive numbers for any $n < N_1$.

QED

Rate of Convergence I

Because of the estimate (12) and equality (15) we have already proved that the convergence is at least of the second order in the foregoing case. This result can be generalized for any case when the function $e_{\hat{i}}(\phi)$ becomes convex in a neighbourhood of ϕ_{min}. This is because the neighbourhood in question can be reached after a finite number of steps, following which the estimate (12) is valid. The only exception is for the case when ϕ_{min} is the minimizing point of both functions $e_{\check{i}}(\phi)$ and $e_{\hat{i}}(\phi)$ since the denominator of (15) approaches zero if $n \to \infty$. We consider this case in the Appendix.

Now, let us consider the remaining case of concavity of the function $e_{\hat{i}}(\phi)$ on the whole interval $\hat{I}^n$. It is easily seen that the sequence $\{\hat{g}^n\}$ is strictly increasing and $\lim_{n\to\infty} \hat{g}^n = e'_{\hat{i}}(\phi_{min}) \triangleq \hat{g}$. The interval I_{n+1} can then be estimated as follows (see Fig. 3).

$$\hat{\phi}^{n+1} - \check{\phi}^{n+1} < \hat{\phi}^{nLg} - \check{\phi}^{n*}, \qquad (17)$$

where $\hat{\phi}^{nLg}$ is the intersection point of line $\delta^{n+1} = e_{\check{i}}(\phi^{n*})$ and the straight line through the point $(\hat{\phi}^n, e_{\hat{i}}(\hat{\phi}^n))$ with slope $\hat{g}$. From geometrical relations we have

$$\frac{\hat{\phi}^{nLg} - \phi^{n*}}{\hat{\phi}^{n} - \phi^{n*}} = \frac{\Delta'^{n} + e_{\check{i}}(\phi^{n*}) - e_{\check{i}}(\check{\phi}^{n})}{\Delta'^{n}}, \tag{18}$$

where

$$\Delta'^{n} = \frac{\hat{g}}{\hat{g}^{n}} \Delta^{n}.$$

After some manipulations we obtain

$$\hat{\phi}^{nLg} - \phi^{n*} = \frac{\hat{g}-\hat{g}^{n}}{\hat{g}} (\hat{\phi}^{n} - \phi^{n*}) + \frac{R^{n}}{\hat{g}}, \tag{19}$$

where R^{n} is the second order remainder of Taylor's formula of the function $e_{\check{i}}(\phi^{n*})$ at the point $\check{\phi}^{n}$.

Using Taylor's expansion for the function $e_{\hat{i}}'(\phi)$ we find

$$\hat{g}^{n} = \hat{g} + \hat{q}^{n}(\hat{\phi}^{n} - \phi_{min}), \tag{20}$$

where $\hat{q}^{n}$ is the second order derivative of $e_{\hat{i}}(\phi)$ at some point $\phi \in \hat{I}^{n}$, so (19) can be estimated by

$$\hat{\phi}^{nLg} - \phi^{n*} \leq \frac{-\hat{q}^{n}}{\hat{g}} (\hat{\phi}^{n} - \check{\phi}^{n*})^2 + \frac{\check{q}^{n}}{2\hat{g}} (\phi^{n*} - \check{\phi}^{n})^2$$

$$\leq \frac{1}{\hat{g}} \left(|\hat{q}^{n}| + \frac{1}{2} |\check{q}^{n}| \right) (\hat{\phi}^{n} - \check{\phi}^{n})^2 . \tag{21}$$

From (17) and (21) it is seen that the convergence of the algorithm is at least of the second order.

QED

Rate of Convergence II

The quadratic convergence of the algorithm can also be proved in a somewhat different, but equivalent way. In the following we assume that both $\check{g}^{n}$ and $\hat{g}^{n}$ do not approach zero but we do not need distinguish the case of convexity from the case of concavity of $e_{\hat{i}}(\phi)$.

Using the second order Taylor formula we can write

$$e_{\check{i}}(\check{\phi}^{n+1}) = e_{\check{i}}(\check{\phi}^{n}) + \check{g}^{n} (\check{\phi}^{n+1} - \check{\phi}^{n}) + \frac{1}{2} \check{q}^{n} (\check{\phi}^{n+1} - \check{\phi}^{n})^2, \tag{22}$$

$$e_{\hat{i}}(\hat{\phi}^{n+1}) = e_{\hat{i}}(\hat{\phi}^{n}) + \hat{g}^{n} (\hat{\phi}^{n+1} - \hat{\phi}^{n}) + \frac{1}{2} \hat{q}^{n} (\hat{\phi}^{n+1} - \hat{\phi}^{n})^2. \tag{23}$$

(It should be noted that $\hat{q}^n$ in (23) can be different than $\hat{q}^n$ in (20).) Knowing that $e_{\check{i}}(\check{\phi}^n) = e_{\hat{i}}(\hat{\phi}^n)$ for any n, (22) and (23) give the relation

$$\hat{g}^n (\hat{\phi}^{n+1} - \hat{\phi}^n) - \check{g}^n (\check{\phi}^{n+1} - \check{\phi}^n) = \frac{1}{2} \left[\check{q}^n (\check{\phi}^{n+1} - \check{\phi}^n)^2 - \hat{q}^n (\hat{\phi}^{n+1} - \hat{\phi}^n)^2\right]. \tag{24}$$

Using (6), the left hand side of (24) can be written as

$$\phi^{n*} (\check{g}^n - \hat{g}^n) + \hat{g}^n \hat{\phi}^{n+1} - \check{g}^n \check{\phi}^{n+1}.$$

Now, if we assume as before, for example, that $\phi^{n*} = \check{\phi}^{n+1}$, (24) can be rewritten as

$$\hat{g}^n (\hat{\phi}^{n+1} - \check{\phi}^{n+1}) = \frac{1}{2} \left[\check{q}^n (\check{\phi}^{n+1} - \check{\phi}^n)^2 - \hat{q}^n (\hat{\phi}^{n+1} - \hat{\phi}^n)^2\right]. \tag{25}$$

From the above we have the estimate

$$|\hat{\phi}^{n+1} - \check{\phi}^{n+1}| \le \frac{1}{2\hat{g}^n} \left(|\check{q}^n| (\check{\phi}^{n+1} - \check{\phi}^n)^2 + |\hat{q}^n| (\hat{\phi}^{n+1} - \hat{\phi}^n)^2\right)$$

$$\le \frac{1}{2\hat{g}^n} (|\check{q}^n| + |\hat{q}^n|) (\hat{\phi}^n - \check{\phi}^n)^2. \tag{26}$$

The final estimate is based on the fact that both $\check{\phi}^{n+1}$ and $\hat{\phi}^{n+1}$ are interior points of I^n so $(\check{\phi}^{n+1} - \check{\phi}^n)^2 < (\hat{\phi}^n - \check{\phi}^n)^2$ and $(\hat{\phi}^{n+1} - \hat{\phi}^n)^2 < (\hat{\phi}^n - \check{\phi}^n)^2$. Since second derivatives $\check{q}^n$ and $\hat{q}^n$ are bounded and $\hat{g}^n$ does not approach zero the factor on the right side of (26) has a finite limit, so the convergence of the algorithm is at least of the second order.

QED

IV. GLOBAL CONVERGENCE

According to the comment after Step 6, Δ_ℓ^n is always positive if $|\hat{\phi}_\ell^n - \check{\phi}_\ell^n| > 0$. We can omit the cases when $\hat{g}_\ell^n = 0$ and/or $\check{g}_\ell^n = 0$ since $\check{g}_\ell^n < 0$ and $\hat{g}_\ell^n > 0$ almost everywhere and Step 9 secures us against these situations. Moreover, it is easy to notice that $\Delta_\ell^n \to 0$ if $|\hat{\phi}_\ell^n - \check{\phi}_\ell^n| \to 0$.

Let us consider two intervals I_1^n and I_2^n which are found by the algorithm in the nth iteration. Let us assume that $\bar{\phi} \in I_2^n$ is a unique global minimax optimum. According to (7) and using the following notation

$$a_i^n = \min(-\check{g}_i^n, \hat{g}_i^n); \quad b_i^n = \max(-\check{g}_i^n, \hat{g}_i^n)$$

where i = 1, 2 is the index of the interval, we have

$$\Delta_1^n = \frac{a_1^n b_1^n}{a_1^n + b_1^n} (\hat{\phi}_1^n - \check{\phi}_1^n) \le \frac{b_1^n}{2} (\hat{\phi}_1^n - \check{\phi}_1^n) \tag{27}$$

and

$$\Delta_2^n = \frac{a_2^n b_2^n}{a_2^n + b_2^n} (\hat{\phi}_2^n - \check{\phi}_2^n) \ge \frac{a_2^n}{2} (\hat{\phi}_2^n - \check{\phi}_2^n). \tag{28}$$

Thus,

$$\frac{\Delta_1^n}{\Delta_2^n} \le \frac{b_1^n}{a_2^n} \cdot \frac{\hat{\phi}_1^n - \check{\phi}_1^n}{\hat{\phi}_2^n - \check{\phi}_2^n} \tag{29}$$

Since $\bar{\phi} \in I_2^n$ is a unique global minimax optimum $|\hat{\phi}_2^n - \check{\phi}_2^n| \to \text{const} \neq 0$ if $|\hat{\phi}_1^n - \check{\phi}_1^n| \to 0$ so that $(\hat{\phi}_1^n - \check{\phi}_1^n)/(\hat{\phi}_2^n - \check{\phi}_2^n) \to 0$. The left hand side of (29) can converge to a value different than zero only if $a_2^n \to 0$ if

$|\hat{\phi}_1^n - \check{\phi}_1^n| \to 0.$[2] But this means that there is a local minimum of at least one of the functions $e_{\check{i}_2}(\phi)$ or $e_{\hat{i}_2}(\phi)$ of value equal to the local minimum value at $\phi_{1min} \in I_1^n$ so that $\bar{\phi} \in I_2^n$ is not the unique global minimax optimum. Otherwise, since $\Delta_1^n/\Delta_2^n \to 0$ the algorithm will select the second interval according to (6).

V. CONCLUSIONS

We have studied the convergence properties of the minimax algorithm of Abdel-Malek and Bandler in some detail. The global convergence was verified. Furthermore, under mild conditions it was shown that the rate of convergence is at least of second order. The rare case when the minimax solution is defined by functions whose derivatives vanish at that solution has been considered. Minor modifications to improve the minimax algorithm have also been suggested.

---

[2] The left hand side of (29) can converge to a value different than zero also when $b_1^n \to \infty$. But this means that the pole of every function $e_i(\phi)$ $i = 1, 2, \ldots, m$ exists at the point ϕ_{1min}, so $\bar{\phi}$ cannot be the unique global optimum point. Moreover, since we consider error functions as magnitudes of network functions this case is of no interest.

APPENDIX

In this Appendix we investigate the particular case when the intersection point ϕ_{min} is the minimizing point of both the functions $e_{\check{i}}(\phi)$ and $e_{\hat{i}}(\phi)$ so that $\check{g}^n \to 0$ and $\hat{g}^n \to 0$ if $n \to \infty$. We show that $\hat{g}^n/\check{g}^n \to$ const $\neq 0$ and examine the rate of convergence of the algorithm.

Let us consider a sufficiently small neighbourhood of ϕ_{min} in which both $e_{\check{i}}(\phi)$ and $e_{\hat{i}}(\phi)$ are convex $(n \geq N)$. Using Taylor's expansion at the point ϕ_{min} we have

$$\delta^n = e_{\check{i}}(\check{\phi}^n) = e_{\check{i}}(\phi_{min}) + \frac{1}{2} e''_{\check{i}}(\tilde{\phi}) (\check{\phi}^n - \phi_{min})^2$$

$$= e_{\hat{i}}(\hat{\phi}^n) = e_{\hat{i}}(\phi_{min}) + \frac{1}{2} e''_{\hat{i}}(\tilde{\tilde{\phi}}) (\hat{\phi}^n - \phi_{min})^2,$$

where $\tilde{\phi} \in \check{I}^n$ and $\tilde{\tilde{\phi}} \in \hat{I}^n$. Because $e_{\check{i}}(\phi_{min}) = e_{\hat{i}}(\phi_{min})$ we get

$$e''_{\check{i}}(\tilde{\phi}) (\check{\phi}^n - \phi_{min})^2 = e''_{\hat{i}}(\tilde{\tilde{\phi}}) (\hat{\phi}^n - \phi_{min})^2. \tag{30}$$

Since the second derivatives $e''_{\check{i}}(\tilde{\phi})$ and $e''_{\hat{i}}(\tilde{\tilde{\phi}})$ are positive and bounded we have

$$0 < \alpha^{\ell} \leq \frac{|\hat{\phi}^n - \phi_{min}|}{|\check{\phi}^n - \phi_{min}|} \triangleq \alpha^n \leq \alpha^u < \infty, \tag{31}$$

where

$$(\alpha^{\ell})^2 = \frac{\min\limits_{\phi \in \check{I}^N} e''_{\check{i}}(\phi)}{\max\limits_{\phi \in \hat{I}^N} e''_{\hat{i}}(\phi)},$$

and

$$(\alpha^u)^2 = \frac{\max\limits_{\phi \in \check{I}^N} e''_{\check{i}}(\phi)}{\min\limits_{\phi \in \hat{I}^N} e''_{\hat{i}}(\phi)} .$$

In order to simplify the remaining notation let us consider subsidiary functions

$$f_1(\phi) \triangleq e_{\hat{i}} (\phi + \phi_{min}), \qquad f_2(\phi) \triangleq e_{\check{i}} (\phi_{min} - \phi) \tag{32}$$

so the minimizing point of $f_1(\phi)$ and $f_2(\phi)$ is at the origin. The corresponding boundary points are

$$\begin{aligned} \phi_1^n &= \hat{\phi}^n - \phi_{min} > 0, \\ \phi_2^n &= \phi_{min} - \check{\phi}^n > 0. \end{aligned} \tag{33}$$

According to (31) we have

$$\alpha^\ell \ \phi_2^n \leq \phi_1^n = \alpha^n \ \phi_2^n \leq \alpha^u \ \phi_2^n. \tag{34}$$

Since $f_1(\phi)$ and $f_2(\phi)$ are biquadratic functions of the form

$$f_i(\phi) = \frac{A_i + 2B_i \ \phi + C_i \ \phi^2}{1 + 2D_i \ \phi + E_i \ \phi^2} , \ i = 1, 2, \tag{35}$$

the derivatives $f'_i(\phi)$ can be expressed as

$$f'_i(\phi) = 2 \frac{(B_i - A_i D_i) + (C_i - A_i E_i) \ \phi + (C_i D_i - B_i E_i) \ \phi^2}{(1 + 2D_i \ \phi + E_i \ \phi^2)^2} . \tag{36}$$

Because $f'_i(0) = 0$, $i = 1, 2$ and under the assumption of irreducibility of $f_i(\phi)$ we find

$$B_i - A_i D_i = 0,$$

$$G_i \triangleq 2 (C_i - A_i E_i) \neq 0, \tag{37}$$

$$H_i \triangleq 2 (C_i D_i - B_i E_i) = G_i D_i \neq 0,$$

for i = 1,2.

According to definitions (32) we have

$$\hat{g}^n = \frac{G_1 \phi_1^n + H_1 (\phi_1^n)^2}{(1+2D_1\phi_1^n+E_1(\phi_1^n)^2)^2},$$

$$\check{g}^n = \frac{G_2 \phi_2^n + H_2 (\phi_2^n)^2}{(1+2D_2\phi_2^n+E_2(\phi_2^n)^2)^2}. \tag{38}$$

The first derivative $f_1'(\phi)$ is strictly decreasing if $\phi \to 0^+$. Thus, (34) gives the estimate

$$f_1' (\alpha^\ell \phi_2^n) \leq \hat{g}^n \leq f_1' (\alpha^u \phi_2^n). \tag{39}$$

Using (36) and (39) we find

$$\alpha^\ell \left|\frac{G_1}{G_2}\right| \leq \lim_{n\to\infty} \left|\frac{\hat{g}^n}{\check{g}^n}\right| \leq \alpha^u \left|\frac{G_1}{G_2}\right|. \tag{40}$$

From (37) and (40) we come to the conclusion that it is always possible to find γ such that

$$\frac{\check{g}^n}{\check{g}^n - \hat{g}^n} \leq \gamma < 1, \tag{41}$$

for any n. Thus, according to (10) we have proved that $|\hat{\phi}^n - \check{\phi}^n| \to 0$ when $\hat{g}^n \to 0$ and $\check{g}^n \to 0$.

Now, we note that using the notation of (31) we can write

$$\hat{\phi}^n - \check{\phi}^n = (\hat{\phi}^n - \phi_{min}) + (\phi_{min} - \check{\phi}^n)$$

$$= (1 + \alpha^n)\,(\phi_{min} - \check{\phi}^n). \qquad (42)$$

Similarly, for the (n+1)th iteration we have

$$\hat{\phi}^{n+1} - \check{\phi}^{n+1} = (1 + \alpha^{n+1})\,(\phi_{min} - \phi^{n*}), \qquad (43)$$

which can be expressed as

$$\hat{\phi}^{n+1} - \check{\phi}^{n+1} = (1 + \alpha^{n+1})\,[(\phi_{min} - \check{\phi}^n) - (\phi^{n*} - \check{\phi}^n)]$$

$$= \left[\frac{-\check{g}^n}{\hat{g}^n} - \alpha^n\right] \frac{1 + \alpha^{n+1}}{(1+\alpha^n)\left[1 + \frac{-\check{g}^n}{\hat{g}^n}\right]}\,(\hat{\phi}^n - \check{\phi}^n). \qquad (44)$$

Using (30) and checking the second derivatives of the functions of (35) we can find the limit

$$\lim_{n\to\infty} \alpha^n = \sqrt{\frac{G_2}{G_1}}\,. \qquad (45)$$

Since

$$\lim_{n\to\infty} \frac{-\check{g}^n}{\hat{g}^n} = \frac{G_2/G_1}{\lim_{n\to\infty} \alpha^n} = \sqrt{\frac{G_2}{G_1}} \qquad (46)$$

it is obvious that the second factor of (44) converges to the finite limit

$$\frac{\sqrt{G_1}}{\sqrt{G_1} + \sqrt{G_2}}\,.$$

But the sequence in the square brackets of (44) approaches zero. Intuitively, this sequence converges at least linearly with $(\hat{\phi}^n - \check{\phi}^n)$. To prove this let us consider the expression $(-\check{g}^n/\hat{g}^n - \alpha^n)$ in terms of

(38), (33) and (34)

$$\frac{-\check{g}^n}{\hat{g}^n} - \alpha^n = \frac{G_2\ \phi_2^n\ (1+D_2\phi_2^n)\ (1+2D_1\phi_1^n+E_1(\phi_1^n)^2)^2}{G_1\ \phi_1^n\ (1+D_1\phi_1^n)\ (1+2D_2\phi_2^n+E_2(\phi_2^n)^2)^2} - \frac{\phi_1^n}{\phi_2^n}$$

$$= \frac{\left(\frac{G_2\ \phi_2^n}{G_1\ \phi_1^n} - \frac{\phi_1^n}{\phi_2^n}\right) + \phi_2^n\ P(\phi_2^n)}{M(\phi_2^n)}, \tag{47}$$

where $M(\phi_2^n) = (1+D_1\alpha^n\phi_2^n)(1+2D_2\phi_2^n+E_2(\phi_2^n)^2)^2$ and $P(\phi_2^n)$ is a polynomial in ϕ_2^n.

In order to find the required relation between G_1, G_2, ϕ_1^n and ϕ_2^n we use the equality $f_1(\phi_1^n) = f_2(\phi_2^n)$ and the conditions (37), i.e.,

$$\frac{A+2AD_1\phi_1^n+C_1(\phi_1^n)^2}{1+2D_1\phi_1^n+E_1(\phi_1^n)^2} = \frac{A+2AD_2\phi_2^n+C_2(\phi_2^n)^2}{1+2D_2\phi_2^n+E_2(\phi_2^n)^2}, \tag{48}$$

where $A = A_1 = A_2$ since $f_1(0) = f_2(0)$.

From (48), and using (34), we find

$$\frac{G_2\ \phi_2^n}{G_1\ \phi_1^n} - \frac{\phi_1^n}{\phi_2^n} = 2\ D_2\ \phi_1^n - 2\ D_1\ \frac{G_2}{G_1}\ \phi_2^n + 2\ \frac{C_1E_2-C_2E_1}{G_1}\ \phi_1^n\ \phi_2^n$$

$$= \phi_2^n\ Q(\phi_2^n), \tag{49}$$

where $Q(\phi_2^n)$ is a polynomial in ϕ_2^n. Finally, we have

$$\frac{-\check{g}^n}{\hat{g}^n} - \alpha^n = \phi_2^n\ \frac{P(\phi_2^n) + Q(\phi_2^n)}{M(\phi_2^n)}, \tag{50}$$

where M(0) = 1. Using (50), (42) and (44) we obtain

$$\hat{\phi}^{n+1} - \check{\phi}^{n+1} = \left\{ \frac{P(\phi_{min}-\check{\phi}^n) + Q(\phi_{min}-\check{\phi}^n)}{M(\phi_{min}-\check{\phi}^n)} \cdot \frac{1 + \alpha^{n+1}}{(1+\alpha^n)^2 \left(1 + \frac{-\check{g}^n}{\hat{g}^n}\right)} \right\} (\hat{\phi}^n - \check{\phi}^n)^2, \tag{51}$$

where the sequence in square brackets has a finite limit. Hence, the convergence of the algorithm is at least of second order when ϕ_{min} is the minimizing point of the both functions $e_{\check{i}}(\phi)$ and $e_{\hat{i}}(\phi)$.

The convergence of the algorithm can be improved, albeit for rare cases, by adding the following Step 6a.

Step 6a

Set ϕ^* to the minimizing point of $e_{\check{i}_j}(\phi)$. If $e_{\check{i}_j}(\phi^*) \geq e_{\hat{i}_j}(\phi^*)$ and $\phi^* \in I_j$ go to Step 10.

The above Step 6a excludes the case when $\check{g}^n \to 0$ and $\hat{g}^n \to 0$ and the case when only $\check{g}^n \to 0$. If we also want to exclude the case when only $\hat{g}^n \to 0$ we should add the step

Step 6b

Set ϕ^* to the minimizing point of $e_{\hat{i}_j}(\phi)$. If $e_{\hat{i}_j}(\phi^*) \geq e_{\check{i}_j}(\phi^*)$ and $\phi^* \in I_j$ go to Step 10.

REFERENCE

[1] H.L. Abdel-Malek and J.W. Bandler, "Centering, tolerancing, tuning and minimax design employing biquadratic models," Faculty of Engineering, McMaster University, Hamilton, Canada, Report SOC-211, 1978.

FIGURE CAPTIONS

Fig. 1 Illustration of the behaviour of the one-dimensional minimax algorithm [1]. Note that the algorithm switches from interval 1 to interval 2, based on predictions of the decrease in the maximum.

Fig. 2 Two functions which define the minimax optimum. The point $\hat{\phi}^{nL}$ at which the linear approximation at $\hat{\phi}^{n}$ takes the value of $e_{\check{i}}(\phi^{n*})$ is indicated.

Fig. 3 The case when $e_{\hat{i}}(\phi)$ is concave on the interval $\hat{I}^{n}$. The point $\hat{\phi}^{nLg}$ at which a linearization based on the gradient $\hat{g}$ of $e_{\hat{i}}(\phi)$ at the optimum takes the value $e_{\check{i}}(\phi^{n*})$ is indicated.

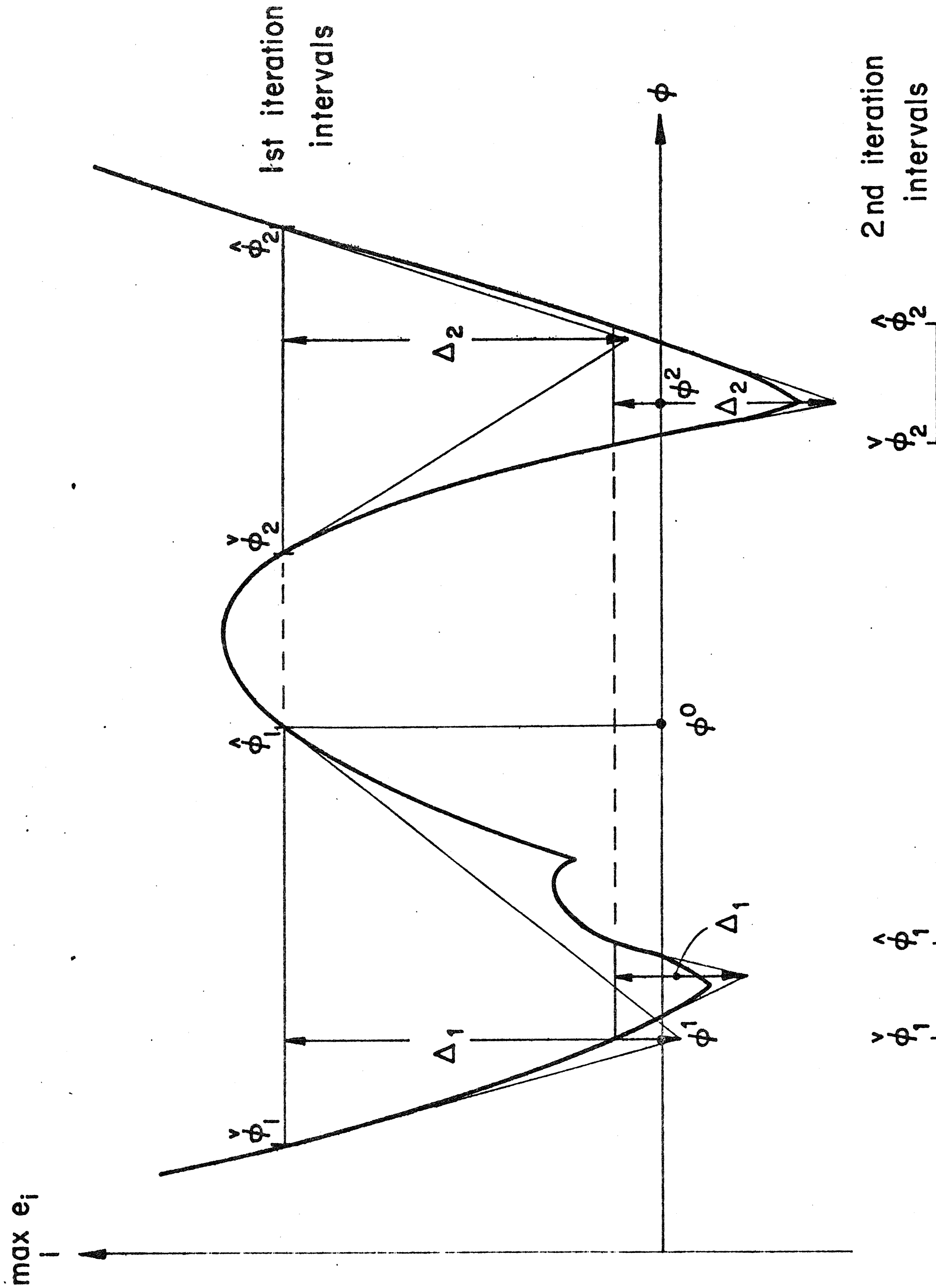
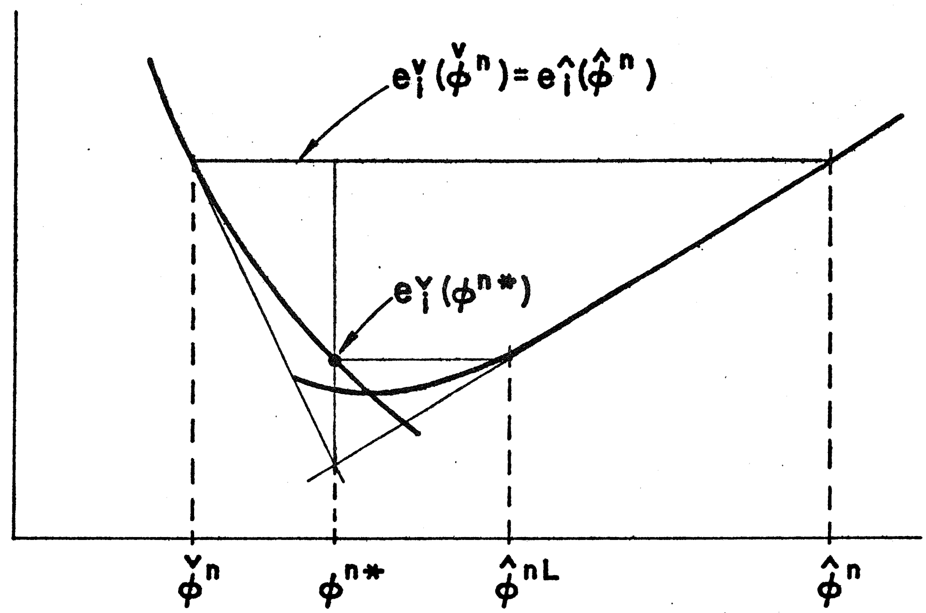
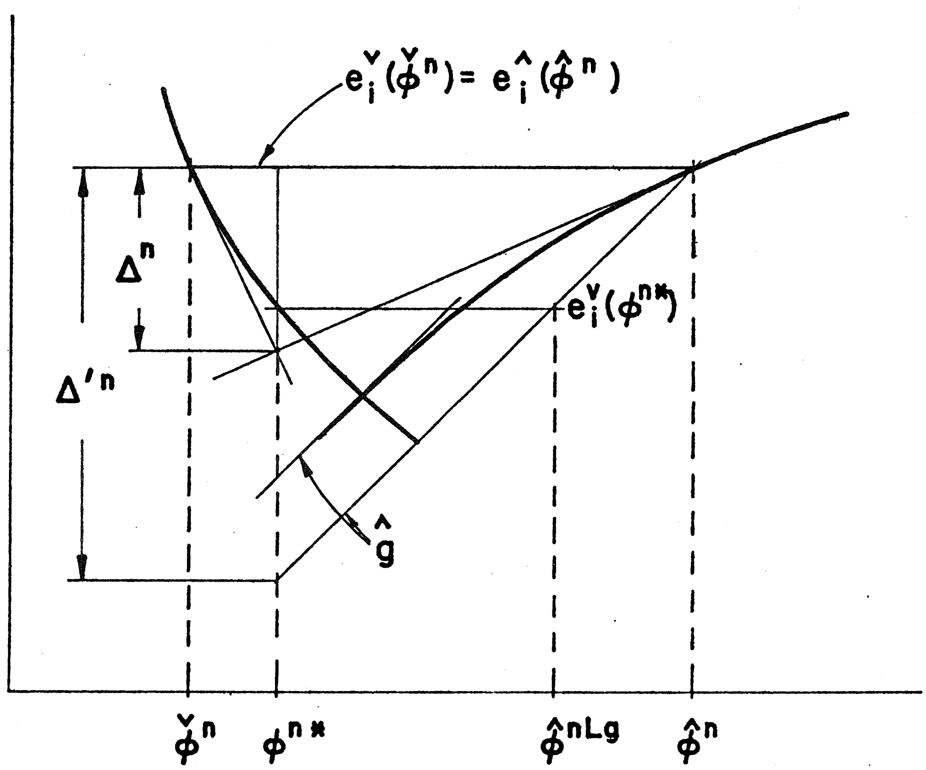

1st iteration
intervals
2nd iteration
intervals
φ
max e_i
Δ₂
Δ₁
φ²
φ⁰
φ¹

Fig. 1



$e_i^{\vee}(\check{\phi}^n) = e_i^{\wedge}(\hat{\phi}^n)$
$e_i^{\vee}(\phi^{n*})$
$\check{\phi}^n$
ϕ^{n*}
$\hat{\phi}^{nL}$
$\hat{\phi}^n$

Fig. 2



$e_i^{\vee}(\check{\phi}^n) = e_i^{\wedge}(\hat{\phi}^n)$
Δ^n
Δ'^n
$e_i^{\vee}(\phi^{n*})$
$\hat{g}$
$\check{\phi}^n$
ϕ^{n*}
$\hat{\phi}^{nLg}$
$\hat{\phi}^n$

Fig. 3

SOC-229

PROOF OF GLOBAL CONVERGENCE AND RATE OF CONVERGENCE FOR A ONE-DIMENSIONAL MINIMAX ALGORITHM

H.L. Abdel-Malek, J.W. Bandler and R.M. Biernacki

February 1979, No. of Pages: 20

Revised:

Key Words: Design centering, tolerancing, tuning, minimax design, biquadratic functions, one-dimensional optimization

Abstract: This report studies the convergence properties of the one-dimensional minimax algorithm developed by Abdel-Malek and Bandler to handle biquadratic functions. It is shown, as expected, that the algorithm converges from any set of starting conditions. Furthermore, under mild conditions we show that the rate of convergence is at least of second order. The rare case when the minimax solution is defined by functions whose derivatives vanish at that solution is considered in some detail. Minor modifications to improve the algorithm are suggested.

Description:

Related Work: SOC-211.

Price: $ 6.00.