Computer Sciences Department
The University of Wisconsin
1210 West Dayton Street
Madison, Wisconsin 53706

APPROXIMATE SOLUTIONS AND ERROR BOUNDS FOR QUASILINEAR ELLIPTIC BOUNDARY VALUE PROBLEMS*

by

To-yat Cheung

Technical Report #114

March 1971

*This research was supported in part by NSF Research Grant GJ0362.

APPROXIMATE SOLUTIONS AND ERROR BOUNDS FOR QUASILINEAR ELLIPTIC BOUNDARY VALUE PROBLEMS*

To-yat Cheung

ABSTRACT

An error bound for a quasilinear elliptic boundary value problem (including the case of nonlinear differential boundary conditions) is obtained as a positively weighted sum of the absolute defects of the operator equations. Once an approximate solution is computed, using linear programming, by minimizing this error bound over a discrete grid, a corresponding realistic error bound over the whole domain of definition can also be obtained by solving an associated linear program.

*This research was supported in part by NSF Research Grant GJ0362.

1. INTRODUCTION

This paper is concerned with the numerical determination of approximate solutions and error bounds by linear programming for the following quasilinear elliptic boundary value problem:

$$Q(g_j): \begin{cases} L_2[u] + g_2(x,y,u) = r_2(x,y) \text{ in } R_2 & (1.1) \\ L_1[u] + g_1(x,y,u) = r_1(x,y) \text{ in } R_1 & (1.2) \\ u = r_0(x,y) \text{ in } R_0 & (1.3) \end{cases}$$

where $L_2 \equiv -(a(x,y)\frac{\partial^2}{\partial x^2} + b(x,y)\frac{\partial^2}{\partial x \partial y} + c(x,y)\frac{\partial^2}{\partial y^2})$ and $a\frac{\partial^2}{\partial x^2} + b\frac{\partial^2}{\partial x \partial y} + c\frac{\partial^2}{\partial y^2}$ is a linear uniform elliptic differential operator such that $L_2[c] \equiv 0$ for any constant c. $L_1 \equiv \frac{\partial}{\partial \nu}$ is an outward nontangential directional derivative, and g_j, $j = 1,2$, may be nonlinear in x, y and u. R_2 is a bounded, simply-connected and open domain in E^2. Its boundary is composed of two mutually disjoint parts R_0 and R_1, each (possibly empty) consisting of a finite number of smooth arcs. Each point of R_1 lies on the boundary of an open ball which lies entirely in R_2. (This is the inside sphere property. See Friedman [1964, p. 55]). $r_j \in C^0(R_j)$, $j = 0,1,2$.

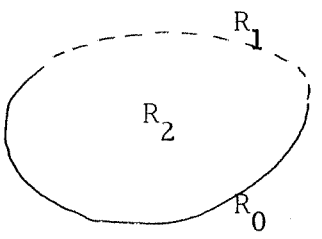

Figure 1

In Section 2 some preliminary material is given. In Section 3, an error bound for a given approximate solution to the problem $Q(g_j)$ is derived, making use of the conditionally inverse-positive property of the operators L_1 and L_2 and the local Lipschitz conditions of the functions g_1 and g_2. This error bound is a positively weighted sum of the absolute defects of the operator equations. In Section 4, by means of this error bound, a theoretical constrained minimization problem is formulated by which we can determine an approximate solution and a corresponding error bound. Section 5 discusses the various advantages of such approach. In Section 6, a computational scheme, making use of linear programming, is suggested to solve this constrained minimization problem. Section 7 presents some numerical results and extensions.

Rosen [1970] considered a particular case of the problem $Q(g_j)$, in which the quasilinear boundary condition (1.2) is not included. Though our approach is slightly different, his results motivated this research work.

2. PRELIMINARY

Def. 1 Let f be a function defined on a set $X \subset E^2$. Set

$$\|f\|_X \equiv \inf_{(x,y)\in X} |f(x,y)|.$$

If there is no ambiguity, the subscript X may be omitted.

Def. 2 $\bar{R} \equiv R_0 \cup R_1 \cup R_2$.

Def. 3 $V(\bar{R}) \equiv C^0(R_0) \cap C^1(R_1) \cap C^2(R_2)$.

Def. 4 $g_j(u) \equiv g_j(x,y,u)$; $g_j'(u) \equiv \dfrac{\partial g_j(x,y,u)}{\partial u}$

Def. 5 For $j = 1,2$, suppose g_j' exists and is bounded. For a fixed $v \in V(\bar{R})$ and a constant $\xi > 0$, define

$$p_j \equiv p_j(\xi, v, x, y) = \min_{\eta} \{g_j'(\eta) \mid |\eta - v| \le \xi\}$$

for each $(x,y) \in R_j$,

and

$$\hat{g}_j(u) \equiv \hat{g}_j(x,y,u) = \begin{cases} g_j(u) & v+\xi \ge u \ge v-\xi \text{ in } R_j \\ g_j(v+\xi) + (u-v-\xi)g_j'(v+\xi) & u > v+\xi \text{ in } R_j \\ g_j(v-\xi) - (v-\xi-u)g_j'(v-\xi) & v-\xi > u \text{ in } R_j \end{cases} \tag{2.1}$$

The following lemma is obvious:

Lemma 1

Let p_j and $\hat{g}_j$ be defined as in Def. 5. Then $\hat{g}_j$ is differentiable with respect to u, and, for a fixed $(x,y) \in R_j$ and arbitrary $u \in V(\bar{R})$, we have

$$\hat{g}_j'(u) \equiv \frac{\partial \hat{g}_j(x,y,u)}{\partial u} \geq p_j(\xi,v,x,y) \quad (x,y) \in R_j, \ j = 1,2 \qquad (2.2)$$

or equivalently

$$\hat{g}_j(x,y,v_1) - \hat{g}_j(x,y,v_2) \geq p_j(\xi,v,x,y)(v_1-v_2)$$
$$(x,y) \in R_j, \quad j = 1,2 \qquad (2.3)$$

where $v_1 \geq v_2$, $v_1, v_2 \in V(\bar{R})$.

In particular,

$$g_j(x,y,v_1) - g_j(x,y,v_2) \geq p_j(\xi,v,x,y)(v_1-v_2)$$
$$(x,y) \in R_j, \quad j = 1,2 \qquad (2.4)$$

where $v_1 \geq v_2$, $|v_i - v| \leq \xi$, $v_i \in V(\bar{R})$, $i = 1,2$.

Def. 6 (2.4) is called a one-sided local Lipschitz condition.

Def. 7 The problem $Q(g_j)$ is said to be inverse-positive if, for every $v_1, v_2 \in V(\bar{R})$, we have

$$\left.\begin{array}{l} L_j[v_1] + g_j(x,y,v_1) \geq L_j[v_2] + g_j(x,y,v_2) \text{ in } R_j, \ j = 1,2 \\ v_1 \geq v_2 \qquad\qquad\qquad\qquad \text{in } R_0 \end{array}\right\} \Longrightarrow v_1 \geq v_2 \text{ on } \bar{R}.$$

Notation: We say $[(L_j + g_j, R_j), j = 1,2; (I, R_0)]$ is inverse-positive, where I is an identity operator.

An important feature of an inverse-positive problem is that it can have at most one solution.

The elliptic problem $Q(g_j)$ is not always inverse-positive. For the linear case, where $g_j(x,y,u) \equiv uk_j(x,y)$, $j = 1,2$, the following conditionally inverse-positive property, as a consequence of the maximum principle of Hopf [1952] and a theorem by Oleinik [1952], is stated in theorem 13 of Protter and Weinberger [1967, p. 78].

Lemma 2

Consider the linear problem $Q(uk_j)$, where k_j are bounded in R_j, $j = 1,2$. Let there exist a positive function $\mu \in V(\bar{R})$ such that

$$(L_j + k_j)[\mu] \geq 0 \quad \text{in} \quad R_j, \; j = 1,2$$

and that the three conditions (i) $(L_1 + k_1)[\mu] \equiv 0$ in R_1; (ii) $(L_2 + k_2)[\mu] \equiv 0$ in R_2; and (iii) R_0 is empty, do not hold simultaneously. Then $Q(uk_j)$ is inverse-positive.

3. DERIVATION OF ERROR BOUND

By means of the local Lipschitz condition (2.4) and the above (Lemma 2) conditionally inverse-positive property, we can derive an error bound for any approximate solution to $Q(g_j)$. In fact, we have

Theorem 3

For the problem $Q(g_j)$, the following assumptions are made:

1) For any $u \in V(\bar{R})$, $g_j'(u)$ exists and is bounded on $\bar{R}_j$ (the closure of R_j), $j = 1, 2$.

2) Let $\hat{g}_j$ be defined in Def. 5 and $Q(\hat{g}_j)$ have a solution in $V(\bar{R})$.

3) $v \in V(\bar{R})$ is an approximate solution to $Q(g_j)$. λ_j, $j = 0, 1, 2$, are scalars such that

$$\begin{cases} \lambda_j \geq |L_j[v] + g_j(v) - r_j| & \text{in } R_j,\ j = 1, 2 \\ \lambda_0 \geq |v - r_0| & \text{in } R_0 \end{cases} \tag{3.1}$$

4) For $p_j(\xi, v, x, y)$ as defined in Def. 5 and

$$\bar{p}_j \equiv \bar{p}_j(\xi, v) = \begin{cases} \min\left\{ \inf_{(x,y) \in R_j} p_j(\xi, v, x, y),\ \lambda_j/\lambda_0 \right\} & \text{if } \lambda_0 \neq 0 \\ \min\left\{ \inf_{(x,y) \in R_j} p_j(\xi, v, x, y),\ 0 \right\} & \text{if } \lambda_0 = 0, \end{cases} \tag{3.2}$$

there exists a solution to the following system of differential inequalities.

$$\left\{ \begin{array}{l} (L_j + p_j(\xi, v, x, y))[\mu] \geq 1 \text{ in } R_j,\ j = 1,2 \\ \qquad\qquad \mu \geq 0 \text{ on } \bar{R} \end{array} \right\} \tag{3.3}$$

$$\xi \geq \lambda_0 + \mu \sum_{i=1}^{2} (\lambda_i - \lambda_0 \bar{p}_i(\xi, v)) \text{ on } \bar{R}\,. \tag{3.4}$$

Then, there exists exactly one solution $u \in V(\bar{R})$ of $Q(g_j)$ such that

$$|u(x,y) - v(x,y)| \leq \rho(x,y) \equiv \lambda_0 + \mu(x,y) \sum_{i=1}^{2} (\lambda_i - \lambda_0 \bar{p}_i) \text{ on } \bar{R} \tag{3.5}$$

Proof. We first prove that

$$[(L_j + \hat{g}_j', R_j),\ j = 1,2;\ (I, R_0)] \text{ is inverse-positive.} \tag{3.6}$$

In fact, by (2.2) and (3.3), there exists a positive $\mu \in V(\bar{R})$ such that

$$(L_j + \hat{g}_j')[\mu] \geq (L_j + p_j)[\mu] \geq 1 \quad \text{in } R_j,\ j = 1,2\,. \tag{3.7}$$

Hence, (3.6) follows from Lemma 2, with k_j replaced by $\hat{g}_j'$.

Now, let u be a solution of $Q(\hat{g}_j)$. For $j = 1,2$, let $\hat{g}_j(v) - \hat{g}_j(u) = \hat{g}_j'(\bar{v}_j)(v-u)$ where $\bar{v}_j = \theta_j v + (1-\theta_j)u$, $0 < \theta_j < 1$. Since $\hat{g}_j(v) = g_j(v)$, we have

$$\begin{aligned} (L_j + \hat{g}_j'(\bar{v}_j))[v + \rho - u] &= (L_j + \hat{g}_j)[v] - (L_j + \hat{g}_j)[u] + (L_j + \hat{g}_j'(\bar{v}_j))[\rho] \\ &= (L_j + g_j)[v] - r_j + \{(L_j + \hat{g}_j'(\bar{v}_j))[\lambda_0 + \mu \sum_{i=1}^{2} (\lambda_i - \lambda_0 \bar{p}_i)]\} \end{aligned}$$

By (3.1), (3.2), (3.7) and (2.2),

$$\begin{aligned} (L_j + \hat{g}_j'(\bar{v}_j))[v + \rho - u] &\geq -\lambda_j + \{\lambda_0 \hat{g}_j'(\bar{v}_j) + \sum_{i=1}^{2} (\lambda_i - \lambda_0 \bar{p}_i)\} \\ &\geq -\lambda_j + \{\lambda_j + (\lambda_i - \lambda_0 \bar{p}_i)\} \quad i \neq j \\ &\geq 0 \text{ in } R_j\,. \end{aligned}$$

On the boundary segment R_0, (3.1) implies

$$\begin{aligned} v + \rho - u &= v + [\lambda_0 + \mu \Sigma (\lambda_i - \lambda_0 \bar{p}_i)] - u \\ &\geq v + \lambda_0 + 0 - r_0 \geq 0 \end{aligned} \tag{3.9}$$

Hence, by (3.6), (3.8) and (3.9) we get

$$v + \rho \geq u \qquad \text{on } \bar{R}.$$

Similarly, we have

$$v - \rho \leq u \qquad \text{on } \bar{R}.$$

Next, by (3.4)

$$|v - u| \leq \rho \leq \xi \qquad \text{on } \bar{R}. \tag{3.10}$$

But then (2.1) implies that u is also a solution of $Q(g_j)$.

Lastly, suppose that $Q(g_j)$ has two solutions u_1 and $u_2 \in V(\bar{R})$ satisfying (3.10). For $j = 1,2$, set $g_j'(\bar{v}_j)(u_1 - u_2) = g_j(u_1) - g_j(u_2)$, where $\bar{v}_j = \theta_j u_1 + (1 - \theta_j)u_2$, $0 < \theta_j < 1$. Then, $|\bar{v}_j - v| \leq \xi$,

$$\hat{g}_j'(\bar{v}_j) = g_j'(\bar{v}_j) \qquad \text{in } R_j,\ j = 1,2,$$

and

$$\begin{cases} (L_j + \hat{g}_j')[u_1 - u_2] = (L_j + g_j)[u_1] - (L_j + g_j)[u_2] = 0 & \text{in } R_j,\ j = 1,2 \\ u_1 - u_2 = 0 & \text{in } R_0 \end{cases} \tag{3.11}$$

Since $[(L_j + \hat{g}_j', R_j),\ j = 1,2;\ (I, R_0)]$ is inverse-positive, (3.11) has the only solution $u_1 - u_2 \equiv 0$ on $\bar{R}$. ∎

The basic idea of theorem 3 is as follows: For $g_j' \equiv g_j'(v_j)$, if $[(L_j + g_j', R_j),\ j = 1,2;\ (I, R_0)]$ is not inverse-positive in the whole domain of u, we want to find a set $Z = \{u \mid v - \rho \leq u \leq v + \rho, (x,y) \in \bar{R}\}$ in which it has this property (i.e. locally inverse-positive) and hence $Q(g_j)$ has at most one solution in Z. This is done by first taking an approximate solution v as the 'center' of Z. If ρ is determined in theorem 3, then it can be a possible 'radius' of Z. The constrained minimization problem formulated in the next section is devised for finding the 'smallest possible' ρ such that Z may 'trap' a solution of $Q(g_j)$.

4. CONSTRAINED MINIMIZATION PROBLEM

In this section, we derive from the error bound formula (3.5) of the last section a constrained minimization problem. In solving this problem by some numerical techniques and linear programming method, we can obtain an approximate solution and the corresponding error bound for the problem $Q(g_j)$.

Notations

For a given function v^k, define

$$g_j(v^k) \equiv g_j(x,y,v^k);$$

$$q_j^k \equiv g_j'(v^k) \equiv \frac{\partial g_j(x,y,v^k)}{\partial u} \text{ ; and}$$

$$G_j^k \equiv v^k q_j^k - g_j(v^k) + r_j .$$

Constrained minimization problem

Suppose $Q(g_j)$ satisfies the Assumptions 1) and 2) of Theorem 3.

Given suitable initial approximation v^0 and constants ξ^0, $\hat{\mu}^0$ and $\bar{p}^0$, $j = 1,2$, the k^{th} cycle of the following iterative process starts with known v^{k-1}, ξ^{k-1}, $\hat{\mu}^{k-1}$ and $\bar{p}_j^{k-1}$:

Step 1 Let $v^k(x,y) \in V(\bar{R})$ and δ_j^k, $j = 0,1,2$, solve

$$\left.\begin{aligned} &\min_{v,\delta_j} \quad \{(1 - \hat{\mu}^{k-1} \sum_{j=1}^{2} \bar{p}_j^{k-1})\, \delta_0 + \hat{\mu}^{k-1} \sum_{j=1}^{2} \delta_j\} \\ &\text{subject to:} \\ &\delta_j \geq (L_j + q_j^{k-1})[v] - G_j^{k-1} \geq -\delta_j, \quad \text{in } R_j,\ j = 1,2 \\ &\delta_0 \geq v - r_0 \geq -\delta_0 \qquad\qquad \text{in } R_0 \end{aligned}\right\} \tag{4.1}$$

Step 2 For $j = 1,2$, evaluate

$$\lambda_0^k = \| v^k - r_0 \|_{R_0}, \quad \lambda_j^k = \| L_j[v^k] + g_j(v^k) - r_j \|_{R_j} \tag{4.2}$$

$$p_j^k(x,y) \equiv p_j(\xi^{k-1}, v^k, x, y) = \min_{\eta} \{ g_j'(\eta) \,\big|\, |\eta - v^k| \leq \xi^{k-1} \} \quad \text{for fixed } (x,y) \tag{4.3}$$

and

$$\bar{p}_j^k = \begin{cases} \min \{ \inf_{(x,y)\in R_j} p_j^k(x,y),\ \lambda_j^k/\lambda_0^k \} & \text{if } \lambda_0^k \neq 0 \\ \inf_{(x,y)\in R_j} p_j^k(x,y) & \text{if } \lambda_0^k = 0 \end{cases} \tag{4.4}$$

Step 3 Let the scalar $\hat{\mu}^k$ and the function $\mu^k \in V(\bar{R})$ solve

$$\min_{\mu,\hat{\mu}} \left\{ \hat{\mu} \;\middle|\; \begin{aligned} &(L_j + p_j^k(x,y))[\mu] \geq 1 \qquad \text{in } R_j,\ j = 1,2 \\ &\hat{\mu} \geq \mu \geq 0 \qquad \text{on } \bar{R} \end{aligned} \right\} \tag{4.5}$$

Set $\xi^k = \lambda_0^k + \hat{\mu}^k \sum_{j=1}^{2} (\lambda_j^k - \lambda_0^k \bar{p}_j^k)$.

Theorem 4

In the above iterative process, if at the k^{th} cycle we have $\xi^k \le \xi^{k-1}$, then the error bound (3.5) holds with

$$\rho(x,y) \equiv \lambda_0^k + \mu^k(x,y) \sum (\lambda_j^k - \lambda_0^k \bar{p}_j(\xi^k, v^k)) \tag{4.6}$$

where $\bar{p}_j(\xi, v) = \inf_{(x,y)\in R_j} p_j(\xi, v, x, y)$.

Proof. Clearly we have only to show that (3.3) and (3.4) hold for $\xi = \xi^k$, $v = v^k$ and $\mu = \mu^k$.

Since $\hat{\mu}^k \ge \mu^k \ge 0$, and, for fixed v^k and (x,y), p_j and $\bar{p}_j$ are both monotone non-increasing functions of ξ, it follows from Step 3 that

$$(L_j + p_j(\xi^k, v^k, x, y))[\mu^k] \ge (L_j + p_j(\xi^{k-1}, v^k, x, y))[\mu^k] \ge 1$$

in R_j, $j = 1,2$

$$\xi^k = \lambda_0^k + \hat{\mu}^k \sum (\lambda_j^k - \lambda_0^k \bar{p}_j(\xi^{k-1}, v^k)) \ge \lambda_0^k + \mu^k(x,y)$$

$$\sum (\lambda_j^k - \lambda_0^k \bar{p}_j(\xi^k, v^k)) \text{ on } \bar{R}. \quad \blacksquare$$

5. DISCUSSION

(1) By (4.1) we see that the approximate solutions v^k are obtained by minimizing a positively weighted sum of the absolute defects of the linearized differential equations and the identity boundary equation. Also, by its structure, the error bound ρ is a monotonic decreasing function of each of the quantities λ_j and $\mu(x,y)$. λ_j (see (4.2)) are absolute defects of the operator equations. $\mu(x,y)$ (see (4.5)) is obtained by minimizing an upper bound. Hence, we may say that the approximate solution is obtained by minimizing its error bound in certain sense.

A similar error bound, which depends on the maximum of the absolute defects instead of a sum of them, can also be derived. But it can be shown that the error bound discussed in this paper is more realistic. See Cheung [1970].

(2) Suppose, instead of the local Lipschitz condition (2.4), g_j satisfies a global Lipschitz condition, i.e. there exists a bounded function k_j, independent of u, such that

$$g_j(v_1) - g_j(v_2) \geq (v_1 - v_2)k_j \quad \text{in } R_j, \text{ for all } v_1 \geq v_2,\ v_i \in V(\bar{R}),$$
$$i = 1,2 \tag{5.1}$$

Theorem 3 and the constrained minimization problem of Section 4 still hold, with $p_j = k_j$. However, a local Lipschitz condition has at least two advantages:

i) The global Lipschitz condition (5.1) may not exist; whereas, we can always construct a local Lipschitz condition provided $g_j'(u)$ is bounded for every u. For example, consider $g(x,y,u) \equiv e^{-u}$. Then no global Lipschitz condition exists, but

$$g(x,y,v_1) - g(x,y,v_2) \geq -e^{-(v-\xi)}(v_1 - v_2)$$

where $v_i \in V(\bar{R})$, $|v_i - v| \leq \xi$, $i = 1,2$ and $v_1 \geq v_2$.

ii) A local Lipschitz condition gives a 'better' error bound than a global Lipschitz condition. In order to show this, we first prove

Lemma 5

Let $\hat{\mu}^i$ be the minimum value of the minimization problem:

$$Q^i: \min_{\mu,\hat{\mu}} \left\{ \hat{\mu} \;\middle|\; \begin{array}{l} (L_j + k_j^i)[\mu] \geq 1 \quad \text{in } R_j,\ j = 1,2 \\ \hat{\mu} \geq \mu \geq 0 \quad \text{on } \bar{R} \end{array} \right\}, \quad i = 1,2$$

(Notice that the superscript i does not imply iteration here).

If $k_j^1 \geq k_j^2$, $j = 1,2$, then $\hat{\mu}^1 \leq \hat{\mu}^2$.

Proof. Let μ^2 be a minimizing function of Q^2. Then

$$\begin{cases} (L_j + k_j^1)[\mu^2] \geq (L_j + k_j^2)[\mu^2] \geq 1 \quad \text{in } R_j,\ j = 1,2 \\ \hat{\mu}^2 \geq \mu^2 \geq 0 \quad \text{on } \bar{R} \end{cases}$$

i.e. $\hat{\mu}^2$ and μ^2 also satisfy the constraints of Q^1. It follows that $\hat{\mu}^1 \leq \hat{\mu}^2$. ∎

Now, for $j = 1,2$, let $k_j^1 = p_j(\xi,v,x,y)$ be the local Lipschitz function as defined in (2.4) and $k_j^2 = k_j$ be the global Lipschitz function as defined in (5.1). Obviously, $p_j \geq k_j$. It follows from Lemma 5 that $\hat{\mu}^1 \leq \hat{\mu}^2$. Generally, we may hope that $\mu^1 \leq \mu^2$ on $\bar{R}$ also. This implies a better error bound for local Lipschitz condition. (see (4.6)).

6. COMPUTATIONAL METHOD

In this section, we consider a computational method of solving the constrained minimization problem of Section 4. Let $\{\phi_i(x,y)\}_{i=1}^{m} \subset V(\bar{R})$ be a set of suitably chosen functions. Assume the function μ and the approximate solution v to be of the form

$$\mu(\beta;x,y) = \sum_{i=1}^{m} \beta_i \phi_i(x,y) \quad \text{and} \quad v(\alpha;x,y) = \sum_{i=1}^{m} \alpha_i \phi_i(x,y) .$$

Let D_j and D_j^* be two discretizations of the region R_j, where D_j^* has finer grid sizes than D_j (Figure 2).

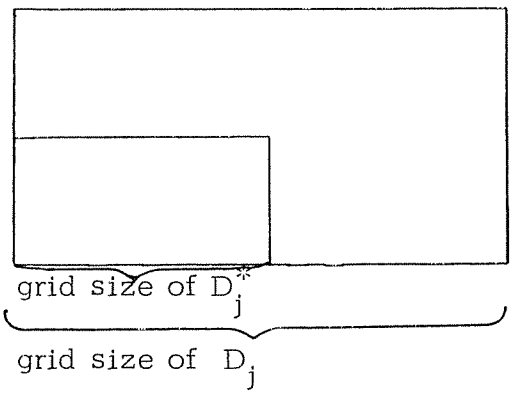

Figure 2

Computational Method

Given suitable initial approximation v^0 and constants ξ^0, $\hat{\mu}^0$ and $\bar{p}_j^0$, the k^{th} cycle of the following iterative process starts with known ξ^{k-1}, $\hat{\mu}^{k-1}$, $\bar{p}_j^{k-1}$ and $v^{k-1}(x,y) \equiv \sum \alpha_i^{k-1} \phi_i(x,y)$:

Step 1 Let α_i^k, $i = 1,\ldots,m$, and δ_j^k, $j = 0,1,2$, solve

$$\left.\begin{aligned} &\min_{\alpha_i,\,\delta_j} \left\{ (1 - \hat{\mu}^{k-1} \sum_{j=1}^{2} \bar{p}_j^{k-1}) \delta_0 + \hat{\mu}^{k-1} \sum_{j=1}^{2} \delta_j \right\} \\ &\text{subject to:} \\ &\delta_j \geq \sum \alpha_i (L_j + q_j^{k-1})[\phi_i] - G_j^{k-1} \geq -\delta_j \quad \text{in } D_j,\ j = 1,2 \\ &\delta_0 \geq \sum \alpha_i \phi_i - r_0 \geq -\delta_0 \quad \text{in } D_0 \end{aligned}\right\} \tag{6.1}$$

Step 2 For $j = 1,2$, evaluate

$$\lambda_0^k = \| v^k - r_0 \|_{D_0^*}, \quad \lambda_j^k = \| L_j[v^k] + g_j(v^k) - r_j \|_{D_j^*}$$

$$p_j^k(x,y) \equiv p_j(\xi^{k-1}, v^k, x, y) \approx \min_{\eta} \{ g_j'(\eta) \mid |\eta - v^k| \leq \xi^{k-1} \}$$

for fixed (x,y).

and

$$\bar{p}_j^k = \begin{cases} \min \left\{ \min_{(x,y)\in D_j^*} p_j^k(x,y),\ \lambda_j^k/\lambda_0^k \right\} & \text{if } \lambda_0^k \neq 0 \\ \min_{(x,y)\in D_j^*} p_j^k(x,y) & \text{if } \lambda_0^k = 0 . \end{cases}$$

Step 3 Choose suitable small positive constants ε_j, $j = 0,1,2$.

Let β_i^k, $i = 0,1,\ldots,m$, solve

$$\min_{\beta_i} \left\{ \beta_0 \;\middle|\; \begin{array}{l} \sum_{i=1}^{m} \beta_i (L_j + p_j^k(x,y))[\phi_i] \geq 1 + \varepsilon_j \text{ in } D_j,\ j = 1,2 \\ \\ \beta_0 \geq \sum_{i=1}^{m} \beta_i \phi_i \geq \varepsilon_0 \text{ on } \bar{D} \equiv D_0 \cup D_1 \cup D_2 \end{array} \right\} \quad (6.2)$$

Let $\hat{\mu}^k = \max_{D^*} \sum_{i=1}^{m} \beta_i^k \phi_i$ and $\xi^k = \lambda_0^k + \hat{\mu}^k \sum_{j=1}^{2} (\lambda_j^k - \lambda_0^k \bar{p}_j^k)$, where $D^* = D_0^* \cup D_1^* \cup D_2^*$.

Choice of initial values for the parameters ξ^0, $\hat{\mu}^0$, $\bar{p}_j^0$ and approximation v^0

By Theorem 4, we see that the initial ξ^0 should overestimate the error of the initial approximation v^0. Usually if ξ^0 is large enough we should have $\xi^0 \geq \xi^1$ and hence an error bound is obtained at a single cycle. However difficulty may arise that if ξ^0 is too large, (6.2) may have no feasible solution.

Without better values, we may set $\hat{\mu}^0 = 1$, $\bar{p}_j^0 = 0$, $j = 1,2$ and $v^0(x,y) \equiv 0$ (or 1) and $\xi^0 = 1$.

Choice of the parameters ε_j, $j = 0,1,2$

In (6.2), we add the positive quantities ε_j to the right sides. If the density of discretization is fine enough and the differential operators satisfy some Lipschitz conditions, it can be shown (for detail,

see Cheung [1970]) that a solution $\mu = \Sigma \beta_i \phi_i$ of the discretized problem (6.2) also satisfies the inequalities (4.5) over the whole region. Therefore, the error bound is valid over $\bar{R}$ instead of over D^* only.

Criterion for terminating the iterative process

The iterative process may stop whenever $\xi^k \leq \xi^{k-1}$. However, in practice, this is usually satisfied at the first cycle. To obtain better accuracy, we may use the following criterion:

Let $\delta_2^k = \| L_2[v^k] + q_2^{k-1} v^k - G_2^{k-1} \|_{D_2}$ and $\lambda_2^k = \| L_2[v^k] + g_2(v^k) - r_2 \|_{D_2^*}$.

For a preassigned quantity ε_N (convergence tolerance of Newton's method), the iterative process is stopped at the k^{th} cycle when

$$\frac{|\lambda_2^k - \delta_2^k|}{\lambda_2^k} \leq \varepsilon_N .$$

Since λ_2^k and δ_2^k are quantities obtained during the iterative process, only little additional computation is required.

Linear programming formulation

It is easy to show that both (6.1) and (6.2) are linear programs of the form (for detail, see Rosen [1970]):

$$\min_{\pi} \quad \{d'\pi \mid A'\pi \geq -c\} \tag{6.3}$$

where d, π and c are vectors and A is a matrix. $'$ denotes the transpose. With $w = -\pi$, (6.3) is equivalent to the dual problem of a standard linear program (see Dantzig [1963]):

$$\max_{w} \quad \{d'w \mid A'w \leq c\} \tag{6.4}$$

Instead of solving (6.4) directly, most available computer linear programming code (e.g. SIMPLX [1969]) are derived so as to solve its primal problem

$$\min_{z} \quad \{c'z \mid Az = d, \quad z \geq 0\} \tag{6.5}$$

(6.3) and (6.5) have the following relations which are well known in the duality theory of linear programming:

(i) If (6.5) has an optimal solution z^* with optimal base B^*, then $\pi^* = -w^* = -(B^*)^{-1}z^*$ is an optimal solution to (6.3). In some computer linear programming codes, π^* is one of the output data. Hence, we can directly obtain an optimal solution to (6.3) by solving (6.5).

(ii) If (6.5) has an infinite (negative) solution, (6.3) has no feasible solution. This fact can be used to test the inverse-positive property of the given problem $Q(g_j)$.

Sizes of the linear programs

For the linear program (6.1), the dimension of the matrix A is $(m + 3) \times 2n$, where m is the number of base functions $\{\phi_i\}_{i=1}^{m}$ and n is the total number of grid points over the three meshes D_j, $j = 0,1,2$.

For the linear program (6.2), the dimension of A is $(m + 1) \times (3n - n_0)$, where n_0 is the number of grid points on D_0.

7. EXTENSION AND NUMERICAL RESULTS

Extension

Throughout our previous discussions, it was assumed that there is only one boundary differential operator. In fact we can consider the more general case where there are $(J-1)$ of them. Let the interior elliptic operator be defined over a bounded simply-connected open domain R_J. Then, by similar argument, we have the error bound

$$\rho(x,y) = \lambda_0 + \mu(x,y) \sum_{j=1}^{J} (\lambda_j - \lambda_0 \bar{p}_j)$$

Obviously, all previous results can be generalized.

Extensions to parabolic and hyperbolic problems had been considered by Cheung [1970].

Example (mixed BVP on square domain)

$$\begin{cases} -\Delta u + g_3(u) = r_3 & 0 < x < .6, \quad 0 < y < .6, \quad R_3 \\ -\dfrac{\partial u}{\partial x} + g_2(u) = r_2 & x = 0, \quad 0 < y < .6, \quad R_2 \\ \dfrac{\partial u}{\partial y} + g_1(u) = r_1 & 0 < x < .6, \quad y = .6, \quad R_1 \\ u = e^{.1x} \cos x & 0 \le x \le .6, \quad y = 0 \\ u = e^{.06} \cos(.6+y) & x = .6, \quad 0 \le y \le .6 \end{cases} \Big\} R_0$$

where $g_1 = u$, $g_2 = .002 \cos u$, $g_3 = -.002yu^2$;

$r_1 = e^{.1x}\{\cos(x+.6) - \sin(x+.6)\}$;

$r_2 = \sin y - .1 \cos y + .002 \cos(\cos y)$; and

$r_3 = \{1.99 - .002\, ye^{.1x} \cos(x+y)\}\, e^{.1x} \cos(x+y) + .2e^{.1x} \sin(x+y)$.

Exact solution $u = e^{.1x} \cos(x+y)$

Algorithm It is not known whether this problem is inverse-positive. The computational method discussed in Section 6 is applied, starting with $v \equiv 1$, $\xi^0 = .4$, $\hat{u}^0 = 1$ and $\bar{p}_1^0 = \bar{p}_2^0 = \bar{p}_3^0 = 0$. $\|\cdot\|_{R_j}$ is approximated by $\|\cdot\|_{D_j^*}$, where D_j^*, $j = 0,1,2$, are uniform meshes on the boundary all with the same grid size .0125, and D_3^* is an interior uniform square mesh with grid size .025. The termination criterion is applied with $\varepsilon_N = .0016$.

Discretization method The iterative linear programs (6.1) and (6.2) are solved (see Section 6) with $\varepsilon_j = .0001$, $j = 0,1,2,3$. D_j^n is the same as D_j^*, $j = 0,1,2$, but has grid size .025. D_3^n is the same as D_3^* but has grid size .075.

Function space $S^{(3+4,3+4)}$ (bicubic elementary splines with 5 knots in both the x- and y-directions. See Appendix).

Computer and LP code CDC 3600, RS MSUB (Clasen [1961]).

Time 3 minutes 50 seconds.

Numerical results After a single iteration, the termination criterion is satisfied and the following results are obtained (omitting superscripts):

$$\lambda_0 = 9.00\ E\text{-}5,\ \lambda_1 = 8.29\ E\text{-}4,\ \lambda_2 = 1.03\ E\text{-}5,\ \lambda_3 = 3.98\ E\text{-}3,$$

$$\bar{p}_1 = 1.14\ E\text{-}1,\ \bar{p}_2 = -2.00\ E\text{-}3,\ \bar{p}_3 = -3.37\ E\text{-}3,\ \xi = 4.89\ E\text{-}3,$$

Each entry in Table 1 is the coefficient of a basis function which is the product of the corresponding functions at the top row and on the leftmost column.

Table 1

Coefficients of Approximate Solution

	1	x	x^2	$(x)_+^3$	$(x-.15)_+^3$	$(x-.30)_+^3$	$(x-.45)_+^3$
1	0.99992	.09999	-0.49572	-0.03230	0.00940	0.03836	0.00298
y	0.00002	-1.00046	-0.05499	-0.06568	0.47911	-0.46823	0.53879
y^2	-0.50062	-0.04464	-0.19463	1.99857	-3.90312	3.84170	-3.03943
$(y)_+^3$	0.01101	0.15202	1.15686	-4.95939	9.61352	-9.26311	4.88510
$(y-.15)_+^3$	0.02729	0.01980	-1.33839	5.54766	-0.10710	9.73219	-0.41881
$(y-.30)_+^3$	0.02284	-0.01531	0.22011	-0.79819	1.27428	0.03618	-9.98217
$(y-.45)_+^3$	0.02248	0.00901	-0.23179	0.63576	-0.67888	-0.95072	0.11270

Table 2 shows the values of the approximate solution, the actual error and error bound at some points distributed fairly uniformly over the whole region. Figure 3 shows the errors along 4 horizontal lines at equal distance. The error curve for $y = 0$ oscillates as expected for uniform approximation to boundary data. This is not true for the other 3 curves since we only minimize the absolute defects of the differential equations there.

Table 2

Approximate Solution v, Error (u-v) and Error Bound ρ

	x	y	v	u - v	ρ
*	.0	.0	.999924	-7.62 E-5	2.81 E-4
†*	.1	.0	1.004934	-6.99 E-5	8.05 E-5
o*	.475	.0	.932643	8.99 E-5	9.41 E-5
	.1	.075	.994559	-6.43 E-5	8.76 E-4
*	.225	.075	.977045	-3.01 E-5	6.30 E-4
	.525	.075	.869820	-3.68 E-6	2.19 E-4
	.475	.1	.880016	3.50 E-7	3.49 E-4
	.0	.175	.984655	-7.17 E-5	2.11 E-3
	.375	.175	.885076	-2.53 E-5	8.23 E-4
	.6	.175	.758518	-8.01 E-5	9.38 E-5
*	.0	.3	.955263	-7.39 E-5	2.94 E-3
	.125	.3	.922430	-6.87 E-5	2.37 E-3
*	.3	.3	.850423	-4.78 E-5	1.61 E-3
	.475	.3	.749145	-2.99 E-5	7.86 E-4
*	.0	.375	.930432	-7.59 E-5	3.37 E-3
*	.3	.375	.804433	-4.99 E-5	1.93 E-3
	.575	.375	.616117	6.81 E-6	2.79 E-4
	.1	.475	.847555	-7.09 E-5	3.39 E-3
	.3	.475	.736133	-4.53 E-5	2.35 E-3
	.475	.475	.609980	-3.10 E-7	1.21 E-3
	.6	.475	.505219	6.94 E-5	8.33 E-5
ǂ*	.0	.6	.825268	-6.76 E-5	4.45 E-3
*	.15	.6	.742688	-5.94 E-5	3.71 E-3
*	.3	.6	.640502	-3.88 E-5	2.89 E-3
*	.45	.6	.520475	1.85 E-6	1.88 E-3
*	.6	.6	.384839	7.40 E-5	2.80 E-4

* grid points; o max. error; ǂ max. error bound; † min. error bound

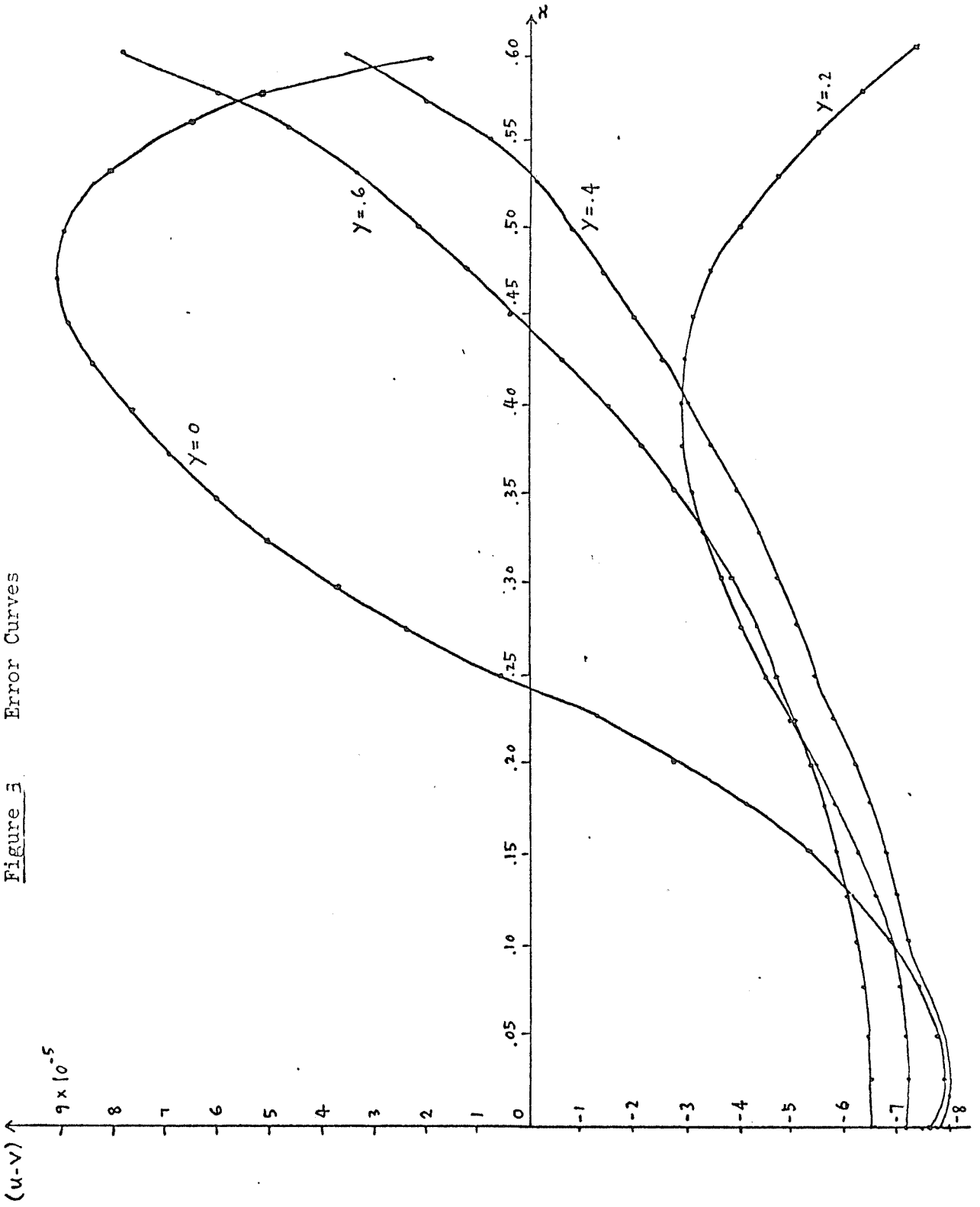

Figure 3 Error Curves

APPENDIX

Monovariate splines

In the following we consider the space of spline functions of degree m with $n + 1$ knots in terms of the basic splines defined by

$$(x)_+^m = \begin{cases} x^m & \text{if} \quad x > 0 \\ 0 & \text{if} \quad x \leq 0 . \end{cases}$$

Elementary splines

Let $x_0, x_1, \ldots, x_n$ be a set of knots over $[x_0, x_n]$. An arbitary spline of degree m is given by

$$\sum \alpha_i \phi_i(x) = \sum_{i=0}^{m-1} \alpha_i x^i + \sum_{i=0}^{n-1} \alpha_{m+i} (x - x_i)_+^m .$$

This space is of dimension $m + n$.

Given a function $\theta(x) \in C^m([0,a])$, it was proved in Theorem 1.2 of Cheung [1968, pp. 8-10] that there exists a sequence of splines of fixed degree m and with uniform knots which converges (as the number of knots tends to infinity) to $\theta(x)$ together with its derivatives up to the order m. This is the denseness property of splines. On the other hand, the evaluation, differentiation and integration of splines are simple. Also, when m is small, evaluation of the values of the

functions and their derivatives usually will not lead to large relative round-off errors. (For references, see Ahlberg, Nilson and Walsh [1967] and Schoenberg [1969].)

Bivariate splines

A convenient way of obtaining splines of two variables is to form products of one-dimensional splines. Let $\{\phi_i(x)\}_{i=1}^{m+n}$ be a basis for a space of dimension $m + n$ and $\{\varphi_j(y)\}_{j=1}^{M+N}$ be a basis for another space of dimension $M + N$, then $\{\phi_i(x)\varphi_j(y)\}_{i=1,m+n}^{j=1,\ldots,M+N}$ would form a basis for the product space of dimension $(m+n)\cdot(M+N)$.

Notations: If ϕ_i and φ_j are elementary splines the product space is denoted by $S^{(m+n,M+N)}(\bar{R})$.

Bivariate splines had been used to approximate the solutions of partial differential equations by Birkhoff, Schultz and Varga [1967] and Schultz [1970].

REFERENCES

1. J. H. Ahlberg, E. N. Nilson and J. L. Walsh, The Theory of Splines and Their Application, Academic Press, New York, 1967.

2. G. Birkhoff, M. H. Schultz and R. S. Varga, "Piecewise Hermite interpolation in one and two variables with applications to partial differential equations," Numer. Math., 11 (1968), pp. 232-256.

3. T. Y. Cheung, "Spline approximation of the Cauchy problem $\frac{\partial^{p+q}u}{\partial x^p \partial y^q} = f(x,y,u,\ldots)$", Computer Sci. Tech. Report #29, Univ. of Wisconsin, Madison, 1968.

4. T. Y. Cheung, "Quasilinear partial differential equations with inverse-positive property," Ph.D. Thesis, Dept. of Computer Sciences, Univ. of Wisconsin, Madison, 1970.

5. R. J. Clasen, "RS MSUB, linear programming subroutine FORTRAN coded," RAND, Sept., 1961.

6. G. B. Dantzig, Linear Programming and Extensions, Princeton Univ. Press, Princeton, 1963.

7. A. Friedman, Partial Differential Equations of Parabolic Type, Prentice-Hall, Inc., Inglewood Cliffs, 1964.

8. E. Hopf, "A remark on linear elliptic differential equations of the second order," Proc. of Amer. Math. Soc., 3 (1952) pp. 791-793.

9. O. A. Oleinik, "On properties of some boundary problems for equations of elliptic type," Math. Sbornik, N.S. 30 (72) (1952), pp. 695-702.

10. M. H. Protter and H. R. Weinberger, Maximum Principles in Differential Equations, Prentice-Hall, Inc., Englewood Cliffs, 1967.

11. J. B. Rosen, "Approximate solutions and error bounds for quasi-linear elliptic boundary value problems," SIAM J. Numer. Anal. 7 (1970), pp. 81-104.

12. I. J. Schoenberg, Approximations with Special Emphasis on Spline Functions, Proc. of a Symposium by Math. Research Center, Univ. of Wisconsin, Academic Press, New York, 1969.

13. SIMPLX, SIMPLX - Linear programming subroutine reference manual, Univ. of Wisconsin Computing Center, Madison, 1969.