LOCKHEED MARTIN ENERGY RESEARCH LIBRARIES

3 4456 0452455 7

ORNL/CSD/TM-161

UCC-ND

NUCLEAR DIVISION

UNION CARBIDE

Steepest Descent for Systems of Nonlinear Partial Differential Equations

J. W. Neuberger

OAK RIDGE NATIONAL LABORATORY
CENTRAL RESEARCH LIBRARY
CIRCULATION SECTION
4500N ROOM 175
LIBRARY LOAN COPY
DO NOT TRANSFER TO ANOTHER PERSON
If you wish someone else to see this report, send in name with report and the library will arrange a loan.
UCN-7969 (3 9-77)

OPERATED BY
UNION CARBIDE CORPORATION
FOR THE UNITED STATES
DEPARTMENT OF ENERGY

Printed in the United States of America. Available from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road, Springfield, Virginia 22161
NTIS price codes—Printed Copy: A03; Microfiche A01

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

ORNL/CSD/TM-161
Dist. Category UC-32

STEEPEST DESCENT FOR SYSTEMS OF NONLINEAR PARTIAL DIFFERENTIAL EQUATIONS

J. W. Neuberger

Sponsor: G. W. Westley

Date Published - October 1981

NOTICE This document contains information of a preliminary nature. It is subject to revision or correction and therefore does not represent a final report.

COMPUTER SCIENCES DIVISION
at
Oak Ridge National Laboratory
Post Office Box X
Oak Ridge, Tennessee 37830

Union Carbide Corporation, Nuclear Division
operating the
Oak Ridge Gaseous Diffusion Plant Oak Ridge National Laboratory
Oak Ridge Y-12 Plant Paducah Gaseous Diffusion Plant
under Contract No. W-7405-eng-26
for the
Department of Energy

LOCKHEED MARTIN ENERGY RESEARCH LIBRARIES

3 4456 0452455 7

CONTENTS

ACKNOWLEDGMENTS iv
ABSTRACT 1
I. INTRODUCTION. 1
II. STEEPEST DESCENT METHOD - SINGLE EQUATION 1
III. A STEEPEST DESCENT METHOD - GENERAL CASE. 7
IV. STUDY OF THE PROJECTION P 11
V EXAMPLES. 13
VI. CONVERGENCE RESULTS 18
VII. DISCUSSION. 21
REFERENCES 23

ACKNOWLEDGMENTS

Funding for the support of this project is provided by Technical Applications.

STEEPEST DESCENT FOR SYSTEMS OF NONLINEAR PARTIAL DIFFERENTIAL EQUATIONS

J. W. Neuberger

ABSTRACT

This report describes a steepest descent method for numerical solution of nonlinear partial differential equations. The method is independent of type and is applicable to a wide variety of flow problems.

I. INTRODUCTION

This report describes a general iterative numerical method for systems of nonlinear partial differential equations. It has been used for a variety of problems including transonic flow, Navier-Stokes, nonlinear wave motion, minimal surfaces. It is a type-independent method which has been especially effective in cases where type is determined by nonlinearities and may change from one subregion to another.

II. STEEPEST DESCENT METHOD - SINGLE EQUATION

In order to illustrate the general method, it is first applied to a single nonlinear partial differential equation on the square region $\Omega = [0,1] \times [0,1]$.

$$\begin{aligned} F[\partial u/\partial x,\ \partial u/\partial y,\ u(x,y),\ x,y] &= 0,\ (x,y) \in \Omega \\ u(x,y) &= \omega(x,y),\ x,\ y \in \Gamma\ , \end{aligned} \tag{1.1}$$

where Γ is a designated curve in Ω, ω is a given continuously differentiable function on Ω, and F is a continuously differentiable function of its arguments.

The numerical approximation to the problem is as follows: If n is a positive integer, denote by G the rectangular grid obtained by dividing each side of Ω into n equal pieces. Denote by K the collection of all functions z on G. Define

$$||z|| = \left[\sum_{(x,y)\varepsilon G} z(x,y)^2 \right]^{1/2}, \quad z\varepsilon K .$$

Define difference operators D_1 and D_2 approximating $\partial/\partial x$, $\partial/\partial y$, respectively: for $z\varepsilon K$, $(x,y)\varepsilon G$,

$$(D_1 z)\ (x,y) = \begin{cases} [z(x+1/n,y) - z(x-1/n,y)]n/2 \text{ if } x = i/n,\ i=1,\ldots,n-1 \\ [z(x+1/n,y) - z(x,y)]n \text{ if } x=0 \\ [z(x-1/n,y) - z(x,y)]n \text{ if } x=1 . \end{cases}$$

Define D_2 similarly.

To obtain a finite difference approximation to Eq. (1.1), consider the problem of finding $z\varepsilon K$ so that

$$\begin{aligned} &F[(D_1 z)(x,y),\ (D_2 z)(x,y),\ z(x,y),\ x,y] = 0,\ (x,y)\varepsilon G \\ &z(x,y) = r(x,y),\ (x,y)\varepsilon \Gamma' , \end{aligned} \tag{2.1}$$

where Γ' is a subset of G approximating Γ and $r\varepsilon K$ such that $r(x,y) = \omega(x,y)$, $(x,y)\varepsilon\Gamma'$.

As an initial step in developing a method for solving Eq. (2.1), define a real-valued ϕ on $K_0 = \{v\varepsilon K | v(x,y) = 0,\ (x,y)\varepsilon\Gamma'\}$:

$$\phi(v) = (1/2) \sum_{(x,y)\varepsilon G} F[(D_1(v+r))(x,y),\ (D_2(v+r))(x,y), (v+r)(x,y),\ x,y]^2 \tag{2.2}$$

for $v \varepsilon K_0$. Note that if $v \varepsilon K_0$ is found so that $\phi(v) = 0$, then $z = v+r$ satisfies Eq. (2.1). The problem then becomes one of minimizing ϕ. Although a minimum of ϕ is not necessarily a zero of ϕ, we hope that it yields a sufficiently close approximation to Eq. (2.1).

Since the grid G has $(n+1)^2$ points, K has dimension $(n+1)^2$. The problem of minimizing ϕ becomes an unconstrained optimization problem on the space K_0, which has dimension less than $(n+1)^2$.

The manner in which a minimization of ϕ is attempted is of <u>critical</u> importance. A method is described in two steps. The first step yields a plausible method which does not, however, work very well in practice. The second step is based upon the first step and yields a modification which produces much better results. The second step makes use of the calculations in the first step. The first step has also the purpose of motivating the second step. Most of the new material of this report centers around this second step. Some numerical results are given in Refs. [4] and [6].

The first step proceeds as follows: For $v \varepsilon K_0$ define T_v so that

$$T_v(x,y) = [D_1(v+r)(x,y),\ D_2(v+r)(x,y),\ (v+r)(x,y),(x,y)], x,y \varepsilon G. \quad (2.3)$$

Then Eq. (2.2) may be expressed as

$$\phi(v) = ||F(T_v)||^2/2 \ . \quad (2.4)$$

Now take a Fréchet derivative of ϕ, i.e., calculate a Jacobian matrix for ϕ:

$$\phi'(v)h = \langle F_1(T_v)D_1h + F_2(T_v)D_2h + F_3(T_v)h,\ F(T_v)\rangle, h \varepsilon K_0 \ . \quad (2.5)$$

where F_1, F_2, F_3 denote partial derivatives of F in the first, second and third arguments of F, respectively, and where for q, $s\varepsilon K$, $<q,s>$ is the inner product

$$\sum_{(x,y)\varepsilon G} q(x,y)s(x,y) \ .$$

In order to obtain a first gradient for ϕ, rewrite Eq. (2.5):

$$\begin{aligned}\phi'(v)h &= <D_1h,F_1(T_v)F(T_v)> + <D_2h,F_2(T_v)F(T_v)> \\ &\qquad + <h,F_3(T_v)F(T_v)> \\ &= <h,\pi(D_1^*(F_1(T_v)F(T_v)) + D_2^*(F_2(T_v)F(T_v)) \\ &\qquad + F_3(T_v)F(T_v))>, \ h\varepsilon K_0 \ ,\end{aligned}$$

where π is the orthogonal projection of K onto K_0;

$$(\pi z)(x,y) = \begin{cases} 0 & (x,y)\varepsilon\Gamma' \\ z(x,y) & (x,y)\varepsilon G/\Gamma', \ (x,y)\varepsilon G \ . \end{cases}$$

Hence, the gradient $\nabla\phi$ is given by

$$(\nabla\phi)(v) = \pi(D_1^*(F_1(T_v)F(T_v)) + D_2^*(F_2(T_v)F(T_v)) + F_3(T_v)F(T_v)), v\varepsilon K_0. \quad (2.6)$$

Here D_1 is regarded as a matrix on the vector space K with the natural basis; D_1^* is then the transpose matrix of D_1.

A steepest descent scheme based upon Eq. (2.6) is the problem of finding g on $[0,\infty)$ to K such that

$$g(0) = r, \ g'(t) = -(\nabla\phi)(g(t)), \ t \geq 0, \qquad (2.7)$$

This steepest descent method, however, has difficulties since the gradient, Eq. (2.6), approximates a differential operator and is likely to be unstable numerically. We now proceed with step 2 which produces a second gradient which is a modification of Eq. (2.6).

The key element of the construction of a second gradient is the introduction of a second norm for K. For $v \varepsilon K$, define

$$||v||_S = (||v||^2 + ||D_1 v||^2 + ||D_2 v||^2)^{1/2} \quad . \tag{2.8}$$

To this norm corresponds the inner product

$$\langle v,z\rangle_S = \langle v,z\rangle + \langle D_1 v, D_1 z\rangle + \langle D_2 v_1 D_2\rangle, \; v,z \varepsilon K.$$

In order to derive an expression for the gradient of Q relative to the norm $|| \; ||_S$, some additional notation will be useful. Denote the space KxKxK by H and denote by H_0' the subspace of H consisting of all elements of the form $Dz \equiv \begin{pmatrix} z \\ D_1 z \\ D_2 z \end{pmatrix}$ such that $z \varepsilon K_0$. Denote by P the orthogonal projection of H onto H_0'. Now P has the explicit form

$$P = D(\pi D^* D|_{R(\pi)})^{-1} \pi D^* \; , \tag{2.9}$$

where π is the orthogonal projection of K onto K_0 (this fact is a special case of a proposition in the next section). In Eq. (2.9), D is from the $(n+1)^2$ dimensional space K into the $3(n+1)^2$ dimensional space H. It may be regarded as a $(n+1)^2 \times 3(n+1)^2$ matrix; D^* then represents the transpose of D.

Denote by π_0 the transformation from H to K so that

$$\pi_0 \begin{pmatrix} f \\ g \\ h \end{pmatrix} = f \text{ for } \begin{pmatrix} f \\ g \\ h \end{pmatrix} \varepsilon H \; .$$

Details on the numerical calculation of P are given in Section IV. We are now in a position to calculate a gradient for ϕ regarded as a function from K (<u>under the norm</u> $||\cdot||_S$) to R. We start from Eq. (2.5):

$$\phi'(v)h = \langle F_1(T_v)D_1h + F_2(T_v)D_2h, + F_3(T_v)h, F(T_v)\rangle$$

$$= \left\langle \begin{pmatrix} h \\ D_1h \\ D_2h \end{pmatrix}, \begin{pmatrix} F_3(T_v)F(T_v) \\ F_1(T_v)F(T_v) \\ F_2(T_v)F(T_v) \end{pmatrix} \right\rangle_H$$

$$= \left\langle \begin{pmatrix} h \\ D_1h \\ D_2h \end{pmatrix}, P\begin{pmatrix} F_3(T_v)F(T_v) \\ F_1(T_v)F(T_v) \\ F_2(T_v)F(T_v) \end{pmatrix} \right\rangle_H$$

$$= \langle h,q\rangle_S, \quad h,v\varepsilon K_0, \text{ where}$$

$$r \equiv \pi_0 \, P\begin{pmatrix} F_3(T_v)F(T_v) \\ F_1(T_v)F(T_v) \\ F_2(T_v)F(T_v) \end{pmatrix}.$$

Hence the gradient $\nabla_S\phi$, <u>calculated according to the inner product</u> $\langle\ ,\ \rangle_S$ <u>on</u> K, is given by

$$(\nabla_S\phi)(v) = P\begin{pmatrix} F_3(T_v)F(T_v) \\ F_1(T_v)F(T_v) \\ F_2(T_v)F(T_v) \end{pmatrix}.$$

Using Eq. (2.9),

$$(\nabla_S\phi)(v) = (\pi D^*D|_{R(\pi)})^{-1}\pi D^*\begin{pmatrix} F_3(T_v)F(T_v) \\ F_1(T_v)F(T_v) \\ F_2(T_v)F(T_v) \end{pmatrix}.$$

Observe that Eq. (2.6) may be rewritten

$$(\nabla\phi)(v) = \pi D^* \begin{pmatrix} F_3(T_v)F(T_v) \\ F_1(T_v)F(T_v) \\ F_2(T_v)F(T_v) \end{pmatrix}$$

so that

$$(\nabla_s\phi)(v) = \left(\pi D^* D|_{R(\pi)}\right)^{-1}(\nabla\phi)(v),\ v\varepsilon K_0 \ . \tag{2.10}$$

Relative to the gradient function $\nabla_s\phi$ we have in place of Eq. (2.7) the steepest descent process g:

$$g(0) = r, g'(t) = (\nabla_s\phi)(g(t)),\ t \geq 0 \ . \tag{2.7'}$$

The effect of the term $\left(\pi D^* D|_{R(\pi)}\right)^{-1}$ connecting the two gradients is to smooth $\nabla\phi$ in a particularly natural way. Typically in the literature[1] smoothing is introduced in steepest descent methods. Sometimes, in this writer's opinion, such smoothing alters the basic problem unacceptably. The smoothing introduced in going from the gradient, Eq. (2.6) to the gradient, Eq. (2.10) is natural to the problem since it arises from a widely accepted measure (i.e., $||\ ||_s$) of a function and its derivatives. It has produced good results numerically in all problems on which it has been used.

III. A STEEPEST DESCENT METHOD - GENERAL CASE

We now broaden our development to a much more general case. Suppose that each of m and n is a positive integer and Ω is a bounded region in R^m. Denote by F a $C^{(2)}$ function with domain $(R^n)^m \times R^n \times \Omega$. We

seek solutions $u:\Omega \to R^n$ to

$$F(\partial u/\partial x_1,\ldots,\partial u/\partial x_n,u(x),x) = 0,\ x\varepsilon\Omega \tag{3.1}$$

where, in addition, u satisfies linear inhomogeneous boundary conditions.

As in the scalar-valued case, pick a C^1 function ω on $\overline{\Omega}$. We want to require our solution u to agree with ω in certain respects. To this end, pick a space C of functions on Ω and require as boundary conditions

$$u - \omega \ \varepsilon\ C\ . \tag{3.2}$$

Examples are given in Section V concerning how to choose, in the numerical analogue of Eq (3.1), such subspaces in order to meet concrete boundary conditions.

The numerical approximation to this general problem is obtained in much the same way as the special one in Section II. Pick a rectangular grid G' (with even mesh spacing $\delta>0$) which covers $\overline{\Omega}$. Denote by G the intersection of G' and $\overline{\Omega}$. Denote by K the space of all R^n – valued functions u on G with

$$||u|| \equiv \left(\sum_{p\varepsilon G} ||u(p)||^2\right)^{1/2}$$

where $u(p)\ \varepsilon R^n$ and $||u(p)||$ denotes the ordinary Euclidean norm of u(p), $p\varepsilon G$. For $u\varepsilon K$, define $D_i u$ such that if $p\varepsilon G$,

$$(D_i u)(p) = \begin{cases} (u(p+\delta e_i) - u(p-\delta e_i))/(2\delta) & \text{if } p \pm \delta e_i \varepsilon G \\ (u(p+\delta e_i) - u(p))/\delta & \text{if } p-\delta e_i \notin G \\ (u(p) - u(p-\delta e_i))/\delta & \text{if } p+\delta e_i \notin G,\ i=1,\ldots,m \end{cases}$$

For $u \varepsilon K$ define

$$||u||_s = (||u||^2 + ||D_1 u||^2 + ..+ ||D_m u||^2)^{1/2}$$

and define Du to be $(u, D_1 u, \ldots, D_m u)$. Take K_0 to be a finite dimensional approximation to the space C of functions used in Eq. (3.2). Define $r \varepsilon k$ to approximate on G the function ω.

Define

$$T_v(p) = [D_1(v+r)(p), \ldots, D_m(v+r)(p), (v+r)(p), p], \; p \varepsilon G, \; v \varepsilon K_0 \; .$$

Our finite difference approximation to the system defined by Eqs. (3.1) and (3.2) becomes the problem of finding $v \varepsilon K_0$ such that

$$F[D_1(v+r))(p), \ldots, (D_m(v+r))(p), (v+r)(p), p] = 0 \qquad (3.3)$$

$$\text{i.e., } F(T_v(p)) = 0, \; p \varepsilon G \; ,$$

or at least that of finding $v \varepsilon K_0$ such that

$$\phi(v) \equiv (1/2)||F(T_v)||^2$$

is minimum.

Proceeding as in the case of a single equation, for $v, h \varepsilon K_0$,

$$\phi_1(v)h = \langle F_1(T_v)D_1 h, \ldots, F_m(T_v)D_m h, F_{m+1}(T_v)h, F(T_v)\rangle \; , \qquad (3.4)$$

where F_i is the partial derivative of F in its ith argument. Note in this connection that the ith argument of F operates on vectors in R^n, so $F_i(T_v(p))$ is a linear transformation from R^n to R^n. Denote by $F_i(T_v)^*$ the function on G so that

$$F_i(T_v)^*(p) = (F_i(T_v(p)))^*, \; p \varepsilon G, \; i=1, \ldots, m+1.$$

Using this notation,

$$\phi'(v)h = \langle D_1 h,\ F_i(T_v)*F(T_v)\rangle + \ldots + \langle D_m h,\ F_m(T_v)*F(T_v)\rangle$$

$$+ \langle h, F_{m+1}(T_v)*F(T_v)\rangle\ .$$

As before, define $H = K^{m+1}$, H_0' to be the image of K_0 under D and P to be the orthogonal projection of H onto H_0'. Then

$$\phi'(v)h = \left\langle \begin{pmatrix} h \\ D_1 h \\ \vdots \\ D_m h \end{pmatrix}, \begin{pmatrix} F_{m+1}(T_v)*F(T_v) \\ F_1(T_v)*F(T_v) \\ \vdots \\ F_m(T_v)*F(T_v) \end{pmatrix} \right\rangle$$

$$= \left\langle \begin{pmatrix} h \\ D_1 h \\ \vdots \\ D_m h \end{pmatrix} = P \begin{pmatrix} F_{m+1}(T_v)*F(T_v) \\ F_1(T_v)*F(T_v) \\ \vdots \\ F_1(T_v)*F(T_v) \end{pmatrix} \right\rangle$$

$$= \langle h, z\rangle_S$$

where

$$z = (\nabla_S \phi)(v) \equiv \pi_0 P \begin{pmatrix} F_m{+}1(T_v)*F(T_v) \\ F_1(T_v)*F(T_v) \\ \vdots \\ F_m(T_v)*F(T_v) \end{pmatrix},$$

$$\pi_0 \begin{pmatrix} g \\ D_1 g \\ \vdots \\ D_m g \end{pmatrix} = g,\ g \varepsilon K\ .$$

Using the gradient function $\nabla_s\phi$, we have the steepest descent process:

$$g(0) = r,\ g'(t) = -(\nabla_s\phi)(g(t)),\ t \geq 0\ . \tag{3.5}$$

We seek a minimum of ϕ as $\lim_{t\to\infty} g(t)$. Of the terms making up $(\nabla_s\Phi)(v)$, $v\varepsilon K_0$, the expressions $F_{m+1}(T_v)^*F(T_v)$ and $F_i(T_v)^*F(T_v)$, $i=1,\ldots,m$ are readily computed. This involves simple manipulations of v, r, F together with various partial derivatives F_i, $i=1,\ldots,m$. The main computational effort is with the projection P. It will be seen in the next section that P can be computed in an organized way once the projection π of K onto K_0 is determined from concrete boundary conditions. This projection is simply given in reasonable examples as may be seen in Section V.

IV. STUDY OF THE PROJECTION P

Suppose that H, K, H_0' and D are as in the preceding section.

Theorem

$$P = D(\pi D^*D|_{R(\pi)})^{-1}\ \pi D^*\ .$$

Proof

We show first that $\pi D^*D|_{R(\pi)}$ has an inverse. For $z\varepsilon K$, $D^*Dz = D^*(z,D_1z,\ldots,D_mz) = (I+D_1^*D_1 + \ldots + D_m^*D_m)Z$. Thus D^*D is a symmetric positive transformation which is $\geq I$, the identity transformation on K. Denote $\pi D^*D|_{R(\pi)}$ by E. If $z\varepsilon R(\pi)$, then since $\pi z=z$, $\langle Ez,z\rangle = \langle \pi D^*Dz,z\rangle = \langle D^*Dz,z\rangle = ||Dz||^2 \geq ||z||^2$. Therefore $E \geq I_0$, the identity on K_0 and so E^{-1} exists and $E^{-1} \leq I_0$. Hence $P = DE^{-1}\pi D^*$ exists. Clearly R(P), the range of P is a subset of H_0' since if $z\varepsilon H$, $Pz = DE^{-1}\pi D^*z\ \varepsilon H_0'$ inasmuch as $R(E^{-1})\varepsilon K_0$. Suppose $z\ \varepsilon\ H_0'$. Then $z = Dr$ for some $r\varepsilon K_0$ and so

$Pz = DE^{-1}\pi D^*Dr = DE^{-1}(\pi D^*D|_{R\pi})r = Dr = z$. Therefore P is fixed on the image of K_0 under D. By calculation, $P^* = P$, $P^2 = P$ and so P <u>is</u> an orthogonal projection. In summary, since P is an orthogonal projection on H which is fixed on H_0' and has range in H_0', it must be the orthogonal projection of H onto H_0'.

<u>Theorem</u>

If $f \equiv (f_0, f_1, \ldots, f_m) \in H$, then $Pf = (z, D_1 z, \ldots, D_m z)$ such that $z \in K_0$ and

$$||Dz - f|| = \left|\left| \begin{pmatrix} z \\ D_1 z \\ \vdots \\ D_m z \end{pmatrix} - \begin{pmatrix} f_0 \\ f_1 \\ \vdots \\ f_m \end{pmatrix} \right|\right|$$

is minimized.

This is a direct consequence of the fact P is the orthogonal projection of H onto H_0'.

Theorem. If $f = (f_0, f_1, \ldots, f_m) \in H$ then $Pf = (z, D_1 z, \ldots, D_m z)$ where z is the unique solution in K_0 to

$$\pi(z + D_1^* D_1 z + \ldots + D_m^* D_m z) = \pi(f_0 + D_1^* f_1 + \ldots + D_m^* f_m) \ . \qquad (4.1)$$

This follows since $Dz = Pf = D(\pi D^*D|_{R(\pi)})^{-1}\pi D^* f$ and so $z = (\pi D^*D|_{R(\pi)})^{-1}\pi D^* G$ and hence Eq. (4.1) holds. This theorem forms the computational basis for determination of z. For given $(f_1, f_1, \ldots, f_m)$, Eq. (4.1) is essentially a large linear system for the unknown z. In applications this is usually a sparse system. It may be solved by Gauss-Siedel iteration, partial Cholesky decomposition or any of a

variety of other schemes. A FORTRAN listing[5] of such a scheme has been provided (in subroutine PT of the code listed in this reference).

V. EXAMPLES

The consideration of the following rather general problem on a bounded region $\Omega \subset R^2$ will be useful for the examples given in this report.

$$f(z,\partial z/\partial x,\partial z/\partial y)\partial^2 z/\partial x^2 + 2g(z,\partial z/\partial x,\partial z/\partial y)\partial^2 z/\partial x\partial y + h(z,\partial z/\partial x,\partial z/\partial y)\partial^2 z/\partial y^2 = 0 \qquad (5.1)$$

$$z = \omega \text{ on } \Gamma_1, \quad \partial z/\partial n = \partial\omega/\partial n \text{ on } \Gamma_2$$

where Γ_1 and Γ_2 are two piecewise differentiable curves in $\overline{\Omega}$, ω is a given C^1 function on $\overline{\Omega}$ and $\partial z/\partial n$ is the derivative of z normal to Γ^2. Rewrite Eq. (5.1) as a system

$$\begin{aligned} &f(z,u,v)\partial u/\partial x + g(z,u,v)(\partial u/\partial y+\partial u/\partial x) + h(z,u,v)\partial v/\partial y = 0 \\ &\partial z/\partial x - u = 0 \\ &\partial z/\partial y - v = 0 \qquad (5.2) \\ &z = \omega \text{ on } \Gamma_1 \\ &\begin{pmatrix} u \\ v \end{pmatrix} \cdot \overline{n} = \begin{pmatrix} \partial\omega/\partial x \\ \partial\omega/\partial y \end{pmatrix} \cdot \overline{n} \quad \text{on } \Gamma_2 \end{aligned}$$

where $\overline{n}$ is a normal vector function on Γ_2.

The numerical approximation to Eq. (5.2) is obtained by reinterpreting z,u,v as functions on a grid G approximating the region Ω. Take

K to be the space of triples of such grid functions. Take D_1 and D_2 as in Section II for this grid. Take s to be a member of K which agrees with ω on G. Take Γ_1' and Γ_2' to be subsets of G approximating the curves Γ_1 and Γ_2, respectively. Our numerical problem then becomes

$$\begin{aligned} &f(z,u,v)D_1u + g(z,u,v)(D_2u+D_1v) + h(z,u,v)D_2v=0 \\ &D_1z - u = 0 \\ &D_2z - v = 0 \\ &z(p) = s(p),\ p\varepsilon\ \Gamma_1',\ \begin{pmatrix} u(p) \\ v(p) \end{pmatrix} \cdot \bar{n}(p) = \begin{pmatrix} (D_1s)(p) \\ (D_2s)(p) \end{pmatrix} \cdot \bar{n}(p),\ p\varepsilon\Gamma_2' . \end{aligned} \tag{5.3}$$

Define F: $R^3 \times R^3 \times R^3 \to R^3$ by

$$F\left(\begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix}, \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix}, \begin{pmatrix} a \\ b \\ c \end{pmatrix}\right) = \begin{pmatrix} f(a,b,c)b_1 + g(a,b,c)(b_2+c_1) \\ \qquad + h(a,b,c)c_2 \\ a_1 - b \\ a_2 - c \end{pmatrix}$$

for $\begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix}, \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix}, \begin{pmatrix} a \\ b \\ c \end{pmatrix} \varepsilon R^3$.

The relevant subspace K_0 of K for this problem is defined as all $\begin{pmatrix} z \\ u \\ v \end{pmatrix} \varepsilon K$ such that $z(p) = 0$, $p\varepsilon\Gamma_1'$, $\begin{pmatrix} u(p) \\ v(p) \end{pmatrix} \cdot \bar{n}(p) = 0$, $p\varepsilon\Gamma_2'$. The orthogonal projection π from K onto K_0 is easy to calculate:

$$\pi \begin{pmatrix} z \\ u \\ v \end{pmatrix} = \begin{pmatrix} \bar{z} \\ \bar{u} \\ \bar{v} \end{pmatrix} \quad \text{where}$$

$$\bar{z}(p) = \begin{cases} z(p), & p \in \Gamma_1' \\ 0, & p \in \Gamma_1 \end{cases},$$

$$\begin{pmatrix} \bar{u}(p) \\ \bar{z}(p) \end{pmatrix} = \begin{cases} \begin{pmatrix} u(p) \\ v(p) \end{pmatrix}, & p \notin \Gamma_2' \\ \begin{pmatrix} a \\ b \end{pmatrix} & p \in \Gamma_2' \end{cases}$$

where $\begin{pmatrix} a \\ b \end{pmatrix}$ is the nearest element of R^2 to $\begin{pmatrix} u(p) \\ v(p) \end{pmatrix}$ such that $\begin{pmatrix} a \\ b \end{pmatrix} \cdot \bar{n}(p) = 0$. Now let us consider some specific examples.

Example 1. Minimal surface equation:

$$(1+(\partial z/\partial y)^2)\partial^2 z/\partial x^2 - 2\partial z/\partial x \; \partial z/\partial y \; \partial^2 z/\partial x \partial y$$

$$+ (1+(\partial z/\partial x)^2)\partial^2 z/\partial x^2 = 0 \text{ on } \Omega ,$$

$$z = \omega \text{ on } \Omega .$$

Example 2:

$$(a^2-(\partial z/\partial x)^2)\partial^2 z/\partial x^2 - 2\partial z/\partial x \; \partial z/\partial y \; \partial^2 z/\partial x \partial y$$

$$+ (a^2-(\partial z/\partial y)^2)\partial^2 z/\partial y^2 = 0 .$$

Where $\underline{a}$ stands for $a_\infty^2 + ((\gamma-1)(2)(z_\infty^2-(\partial z/\partial x)^2 - (\partial z/\partial y)^2)$, a_∞ and z_∞ are given positive numbers.

A solution z is elliptic when

$$\Delta \equiv (\partial z/\partial x)^2(\partial z/\partial y)^2 - (a^2-(\partial z/\partial x)^2)(a^2-(\partial z/\partial y)^2)$$

$$= a^2((\partial z/\partial x)^2 + (\partial z/\partial y)^2 - a^2) < 0 ,$$

and hyperbolic when $\Delta > 0$. The hyperbolic region indicates supersonic flow and the elliptic region indicates subsonic flow. The quantity a is the "local speed of sound", $\gamma = 1.4$ for air, z_∞ is the speed of flow at infinity, and a_∞ is the speed of sound at infinity.

Boundary conditions for the problem on all of R^2 are:

$$z(x,y) \sim x ,$$

$$\text{and } \partial z/\partial x \sim u_\infty ,$$

$$\partial z/\partial y \sim 0$$

$$\text{as } x^2 + y^2 \to \infty$$

together with

$$\frac{\partial z/\partial y}{\partial z/\partial x} = \begin{cases} f'_+(x,y) & (x,y) \in \Gamma_3 \\ f'_-(x,y) & (x,y) \in \Gamma_4 \end{cases}$$

where Γ_3, Γ_4 describe upper and lower contours, respectively, of an air foil. For numerical calculations, the problem is put in a box (say [0,1] x [0,1]) with the first boundary conditions replaced by

$$z(x,y) = x, \quad y = 0 \text{ or } 1$$

$$\partial y/\partial x = u_\infty, \quad x = 0 \text{ or } x = 1 \text{ or } y = 0 \text{ or } y = 1$$

$$\partial z/\partial y = 0, \quad x= 0 \text{ or } x = 1 \text{ or } y = 0 \text{ or } y = 1 .$$

References [2], [5].

Example 3. Navier-Stokes equation.

$$
\begin{aligned}
-\nu\,\nabla\cdot\omega + u\cdot\omega + \nabla p &= 0 \\
\nabla u &= \omega \qquad\qquad (5.4) \\
\nabla\cdot u &= 0\ ,
\end{aligned}
$$

where p,u,ω are unknown functions on $\Omega\subset R^2$, p being real valued, $u:\Omega\to R^2$ and $\omega:\Omega\to L(R^2,R^2)$ the space of 2x2 matrices. The number ν is a known viscosity. Renaming the unknowns u and ω,

$$u = \begin{pmatrix} r \\ s \end{pmatrix},\quad \omega = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}.$$

Equation (5.4) is rewritten

$$
\begin{aligned}
&-\begin{pmatrix} \alpha_x+\beta_y \\ \gamma_x+\delta_y \end{pmatrix} + \begin{pmatrix} \alpha r+\beta s \\ \gamma r+\delta s \end{pmatrix} + \begin{pmatrix} p_x \\ p_y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}\ , \\
&\begin{pmatrix} r_x & r_y \\ s_x & s_y \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \\
&r_x + s_y = 0
\end{aligned}
$$

where $\alpha_x = \partial\alpha/\partial x$ etc. For appropriate F this may be written in the form of Eq. (3.1). Various boundary conditions may be incorporated (and a corresponding space K_0 defined) as in Examples 1 and 2.

VI. CONVERGENCE RESULTS

This section contains some comments on the steepest descent process Eq. (3.5). A first fact to note is the following:

Theorem.

There is a unique function y from $[0,\infty)$ to K such that $g(0) = r$, $g'(t) = -(\nabla_S\phi)(g(t))$, $t \geq 0$.

Proof.

Since F is a $C^{(2)}$ function, $\nabla_S\phi$ is $C^{(1)}$. Hence for some $\delta > 0$ there is a unique solution g to the problem $g(0) = r$, $g'(t) = -(\nabla_S\phi)(g(t))$, $0 \leq t \leq \delta$. Consider the union of all such intervals $[0,\delta]$ and suppose that this union is not $[0,\infty)$. Denote by τ the least upper bound of this union. Define $J(z) = F(T_z)$, $z\varepsilon K$, using the notation introduced before Eq. (3.3). Define $\alpha:[0,\tau) \to R$ by $\alpha(t) = (1/2)\ ||J(g(t))||^2$, $t\varepsilon[0,\tau)$. Then

$$\alpha'(t) = \langle J'(g(t))\ g'(t),\ J(g(t))\rangle$$

$$= -\langle J'(g(t))\ J'(g(t))^*\ J(g(t)),\ J(g(t))\rangle$$

$$= -\ ||J'(g(t))^*\ J(g(t))||^2 = -\ ||g'(t)||^2,\ t\varepsilon[0,\tau).$$

Thus $\alpha'(t) \leq 0$, $t\varepsilon[0,\tau)$ and $\alpha(a) - \alpha(b) = \int_a^b ||g'||^2$, $0 \leq a < b < \tau$. By Swartz's inequality,

$$(\int_a^b ||g'||)^2 \leq \int_a^b ||g'||^2 \int_a^b 1 = (b-a) \int_a^b ||g'||^2$$

and so

$$\int_a^b ||g'|| \leq (b-a)^{1/2} \left(\int_a^b ||g'||^2\right)^{1/2}$$

$$= (b-a)^{1/2} \ (\alpha(a)-\alpha(b))^{1/2} \leq (b-a)^{1/2} \ \alpha(a)^{1/2}$$

$$\leq \tau^{1/2} \ ||J(r)||/\sqrt{2}.$$

Therefore, $\lim_{t\to\tau} \int_0^t ||g'||$ exists (and does not exceed $\sqrt{\tau}\ ||J(r)||/\sqrt{2}$). Hence $S \equiv \lim_{t\to\tau} g(t)$ exists. But there is $\delta_1 > 0$ so that there is a unique solution h to $h(\tau) = s$, $h'(t) = -(\nabla_S\phi)(h(t))$, $t\varepsilon[\tau,\delta_1+\tau]$. Define $k: [0,\delta_1+\tau] \to K$ such that

$$k(t) = \begin{cases} g(t) & 0 \leq t < \tau \\ h(t) & \tau \leq t \leq \tau + \delta_1 \ . \end{cases}$$

Clearly k satisfies uniquely

$$k(0) = r, \ k'(t) = -(\nabla_S\phi)\ (k(t)), \ 0 \leq t \leq \delta_1 + \tau \ ,$$

a contradiction in view of the definition of τ.

We now consider directly the question of the existence of

$$\lim_{t\to\tau} g(t) \ .$$

Where g is the unique solution of Eq. (3.4).

The following theorem gives a convergence criterion for this process. There are two purposes for giving the condition:

i) it provides a practical way for discovering rate of convergence from a sequence of calculated approximations,

ii) it is related to a condition which can be verified in some cases.

Theorem.

Suppose $\{t_i\}_{i=0}^{\infty}$ is a sequence of numbers so that $0 \leq t_{i+1} - t_i = t_i - t_{i-1} \leq 1$, $i=0,1,2,\ldots$. Suppose moreover that the

$$\sum_{i=0}^{\infty} (\beta(t_{i+1}) - \beta(t_i))^{1/2}$$

converges where $\beta(t) \equiv (1/2)\ ||J(g(t))||^2$, $t \geq 0$. Then $u = \lim_{t\to\infty} g(t)$ exists and $(\nabla_s\phi)\ (u) = 0$.

Proof.

If $0 \leq a < b$, then as seen in the proof of the previous theorem

$$\int_a^b ||g'|| \leq (b-a)^{1/2}\ \left(\int_a^b ||g'||^2\right)^{1/2} \leq (b-a)(\beta(a)-\beta(b))^{1/2}\ .$$

Hence $\int_{t_i}^{t_{i+1}} ||g'|| \leq (\beta(t_{i+1}) - \beta(t_i))^{1/2}$, $i=0,1,2,\ldots$ and so the convergence of $\sum_{i=0}^{\infty} (\beta(t_{i+1}) - \beta(t_i))^{1/2}$ implies the existence $\int_0^{\infty}||g'||$ and consequently of $u = \lim_{t\to\infty} g(t)$. If $(\nabla_s\phi)(u) \neq 0$, then $\lim_{t\to\infty} g'(t) = -(\nabla_s\phi)(u) \neq 0$. But this is in contradiction to the existence of $\int_0^{\infty}||g'||$ and so the theorem is established.

As a practical matter, since $\beta(t) \equiv (1/2)||J(g(t))||^2 = \frac{1}{2}||F(T_{g(t)-r})||^2$ and $F(T_{g(t)-r}(p)) = F(D_1(g(t)\ (p),\ldots,D_m(g(t))(p), g(t)(p),p)$, $p\varepsilon G$, it follows that $\beta(t)$ is a measure of the amount $g(t)$ misses being a solution u to Eq. (3.1). In actual computation $\beta(t)$ may be readily computed and hence also the ratios $(\beta(t_{i+1}) - \beta(t_{i+2}))^{1/2}$ / $(\beta(t_{i+1})) - \beta(t_i))^{1/2}$ for a succession of integers i. Evidence of convergence is present if these quantities are bounded below some number less than 1.

VII. DISCUSSION

Methods described in this report are probably not competitive with many existing schemes developed for particular problems - for example Laplace's equation. The strength of these methods rests in part on the following three items:

i) The method is widely applicable. Existing codes already cover a wide class of problems; e.g., systems which cover a general second-order quasilinear equation as in Examples (1) and (2). Alteration of these codes to cover much broader classes of problems is a routine matter.

ii) The method is not dependent upon type. In Example (2) for some choices of parameters a_∞, u_∞, the solution will be hyberbolic in some parts of the region and elliptic in others. The nonlinearities in a sense determine where these regions lie. Any purely hyperbolic method or purely elliptic method would be inappropriate. A wide collection of flow problems have similar characteristics. Some of these problems cannot at present, it seems, be treated by other methods.

iii) Full boundary data (i.e., boundary conditions which determine a unique solution) are not required by the method. If "insufficient" data are given, the method tends to converge to the solution "nearest" to the starting estimate. This can be very useful when sufficient boundary conditions are not known.

The numerical development given here has been in terms of finite differences. Corresponding finite element procedures have been coded by graduate students Craig Beasley and Morris Liaw of North Texas State University. Indications are that the given steepest descent method will work very well in a finite element setting.

REFERENCES

1. K. W. Brodlie, *Unconstrained Minimization*, The State of the Art in Numerical Analysis, Academic Press, 1977.
2. G. F. Carey, "Variational Principles for the Transonic Airfoil Problem", Computer Methods *App. Mech. Eng.*, 13:129-40, 1978.
3. J. W. Neuberger, *An Iterative Method for Approximating Solutions to Nonlinear Partial Differential Equations*, Applied Nonlinear Analysis, Academic Press, 1979.
4. J. W. Neuberger, "Finite Dimensional Potential Theory Applied to Numerical Analysis of Linear Systems, *Lin. Algebra Appl.* 35:193-202, 1981.
5. J. W. Neuberger, "A Type-Independent Iterative Method for Systems of Nonlinear Partial Differential Equations: Application to the Problem of Transonic Flow, *Comp. Math. Appls.*, 6: 67-78, 1980.
6. J. W. Neuberger, "A Type-Independent Numerical Method Applied to Navier-Stokes Equations", *Proceedings, Ninth Conference on Numerical* Simulations of Plasmas, Northwestern University, 1980.
7. J. M. Ortega and W. C. Rheinboldt, "Iterative Solution of Non-Linear Equations in Several Variables", Academic Press, 1970.

ORNL/CSD/TM-161
Dist. Category UC-32

INTERNAL DISTRIBUTION

1. B. R. Becker
2. L. Berry
3. J. A. Carpenter
4. H. P. Carter/
   G. E. Whitesides/CS Library
5. J. E. Drake
6. R. E. Funderlic
7. M. T. Heath
8. J. A. Holmes
9. J. P. Holmes
10. W. F. Lawkins
11. J. B. Mankin

12-16. J. W. Neuberger
17. D. W. Noid
18. R. V. O'Neill
19. J. E. Park
20. R. J. Renka
21. A. D. Soloman
22. J. N. Tunstall
23. C. C. Travis
24. G. W. Westley
25. H. G. Wood, III
26. Central Research Library
27. Document Reference Section Y-12
28-29. Laboratory Records
30. Laboratory Records - RC
31. ORNL Patent Office

EXTERNAL DISTRIBUTION

32. J. N. Rogers, Division 8324 Sandia Laboratories, Livermore, California 94550.

33. Office of Assistant Manager for Energy Research and Development, Department of Energy, Oak Ridge Operations, Oak Ridge, Tennessee 37830.

34. Division of Engineering, Mathematics and Geosciences, Department of Energy, Washington, D.C. 20545.

35. Professor Luther White, Department of Mathematics, University of Oklahoma, Norman, Oklahoma 73019

36. P. D. Lax, New York University, Courant-251 Mercer Street, New York, New York 10012.

37. G. Carey, Engineering Mechanics Department, Science Building 45, Austin, Texas 78712.

38. Phillip Walker, Department of Mathematics, University of Houston, Houston, Texas 77004.

39. J. Goldstein, Department of Mathematics, Tulane University, New Orleans, Louisiana 70118.

40. J. R. Dorroh, Department of Mathematics, Louisana State University, Baton Rouge, Louisiana 70803.

41. S. Busenberg, Department of Mathematics, Harvey Mudd College, Claremont, California 91711.

42. J. L. Lions, College du France, Place Marcelin, Berthelat, Paris (5) France.

43. R. Glowinski, 30 Allee Blaise Pascal, Chevreuse (76460) France.

44-225. Given distribution as shown in TID-4500 under UC-32, Mathematics and Computers