Sobolev Gradients and Differential Equations

J. W. Neuberger

Author address:

Department of Mathematics
University of North Texas
Denton, TX 76203
USA

E-mail address: jwn@unt.edu

Contents

Preface

What is it that we might expect from a theory of differential equations? Let us look first at ordinary differential equations.

THEOREM 1. *Suppose that n is a positive integer and G is an open subset of $R \times R^n$ containing a point $(c, x) \in R \times R^n$. Suppose also that f is a continuous function on G for which there is $M > 0$ such that*

$$\|f(t,x) - f(t,y)\| \leq M\|x - y\|, \ (t,x), \ (t,y) \in G. \tag{0.1}$$

Then there is an open interval (a, b) containing c for which there is a unique function y on (a, b) so that

$$y(c) = x, \ y'(t) = f(t, y(t)), \ t \in (a, b).$$

This result can be proved in several constructive ways which yield, along the way, error estimates which give a basis for numerical computation of solutions. Now this existence and uniqueness result certainly does not solve all problems (for example, two point boundary value problems) in ordinary differential equations. Nevertheless, it provides a position of strength from which to study a wide variety of differential equations. First of all the fact of existence of a solution gives us something to study in a qualitative, numerical or algebraic setting. The constructive nature of arguments for the above result gives one a good start toward discerning properties of solutions.

It would generally be agreed that (1) one would want a similar position of strength for partial differential equations and (2) there is not such a theory. It is sometimes argued that there cannot be such a theory since partial differential equations show such great variety. To such an argument one might reply that the same opinion about ordinary differential equations was probably held not so much longer than a century ago.

This monograph is devoted to a description of Sobolev gradients for a wide variety of problems in differential equations. These gradients are used in descent processes to find zeros or critical points of real-valued functions; these zeros or critical points provide solutions to underlying differential equations. Our gradients are generally given constructively. We do not require that full boundary conditions (*i.e.*, conditions which are necessary and sufficient for existence and uniqueness) be known beforehand. The processes tend to converge in some (perhaps noneuclidean) sense to the nearest solution. The methods apply in cases which are mixed hyperbolic and elliptic — even cases in which regions of hyperbolicity and ellipticity are determined by nonlinearities. Applications to the problem of transonic flow will illustrate this. We emphasize numerical emulation of function space processes as well as theoretical convergence results.

So, do we arrive at a position of strength for fairly general partial differential equations? We claim to have only a shadow of such a theory. In particular, there is a great deal of work to be done in finding convergence results for descent processes using Sobolev gradients and showing how even the abstract convergence results given here actually apply to concrete systems of partial differential equations. My main hope for this work is that some who are really good at hard estimates for partial differential equations will find some interesting problems here and consequently will contribute to this program. I will try to learn enough to be able to join you in such activity when this book is finished. What we offer here is mainly constructions of gradients which have shown themselves to be viable numerically. Our computational practice is considerably ahead of our theory at present.

Several people have read portions of this monograph, have found a number of errors and have suggested improvements. Simeon Reich has been particularly helpful as have been a number of this writer's colleagues and former students. The influence of Robert Renka on this work is more than what is indicated in our joint work represented in several chapters.

CHAPTER 1

Several Gradients

These notes contain an introduction to the idea of Sobolev gradients and how they are of use in differential equations. Numerical considerations are at once a motivation, an investigative tool and an application for this work.

First recall some facts about ordinary gradients. Suppose that for some positive integer n, ϕ is a real-valued $C^{(1)}$ function on R^n. It is customary to define $\nabla\phi$ as the function on R^n so that if $x = (x_1, x_2, ..., x_n) \in R^n$, then

$$(\nabla\phi)(x) = (\phi_1(x_1, \ldots, x_n), \ldots, \phi_n(x_1, \ldots, x_n)) \tag{1.1}$$

where we write $\phi_i(x_1, ..., x_n)$ in place of $\partial\phi/\partial x_i$, $i = 1, 2, ..., n$. A property of $\nabla\phi$ (or an equivalent definition) is that if $x \in R^n$, then $(\nabla\phi)(x)$ is the element of R^n so that

$$\phi'(x)h = \langle h, (\nabla\phi)(x)\rangle, \ h \in R^n, \tag{1.2}$$

where $\phi'(x)h = \lim_{t\to 0}(\phi(x + th) - \phi(x))/t, \ x, h \in R^n$ and $\langle \ , \ \rangle$ denotes the standard inner product on R^n. Another property (and again a possible equivalent definition) is that for $x \in R^n$, $(\nabla\phi)(x)$ is the element $h \in R^n$ for which

$$\phi'(x)h \text{ is maximum subject to } \|h\| = |\phi'(x)|, \tag{1.3}$$

where $|\phi'(x)| = \sup_{k\in R^n, \|k\|=1} \phi'(x)k$. In Hilbert spaces other than R^n, (1.2),(1.3) provide equivalent definitions for a gradient. For some Banach spaces X without an inner product, (1.3) is available even though (1.2) is not (see comments on duality in [**86**]). Moreover, (1.3) generalizes to some cases in which $(\nabla\phi)(x)$ may be defined as the element $h \in X$ which provides a critical point to $\phi'(x)h$ subject to the constraint $\beta(h, x) - c = 0$, where β is some specified function. We will see that if β is defined in terms of ϕ itself, (more specifically $\beta(h, x) = \phi(x+h), \ x, h \in X$) then this process leads to Newton's method.

A central theme in these notes is that a given function ϕ has a variety of gradients depending on choice of metric. More to the point, these various gradients have vastly different numerical and analytical properties even when arising from the same function. Related ideas have appeared in several places. In [**79**] there is the idea of variable metrics in which in a descent process, different metrics are chosen as the process develops. More recently, Karmarkar [**40**] has used the idea with great success in a linear programming algorithm. In [**42**] and others, Karmarkar's ideas are developed further. This writer has developed this idea (with differential equations in mind) in a series of papers starting in [**55**] (or maybe in [**51**]) and leading to [**69**],[**71**],[**72**]. Variable metrics are closely related to the conjugate gradient method [**37**]. Some other classical references to steepest descent are [**17**],[**20**],[**94**]. The present work contains an exposition of some of the earlier work of this writer. It also contains some results which ordinarily would have been published separately.

For some quick insight into our point of view, recall how various inner product norms on R^n are related. Recall that Q is a positive definite symmetric bilinear function on $R^n \times R^n$ if and only if there is a positive definite symmetric matrix $A \in L(R^n, R^n)$ so that

$$Q(x,y) = \langle Ax, y\rangle,\ x, y \in R^n. \tag{1.4}$$

Such bilinear functions constitute the totality of all possible inner products which may be associated with the linear space R^n. If Q satisfies (1.4), we denote $Q(x,y)$ by $\langle x,y\rangle_A,\ x,y \in R^n$.

The following question gives our first look at a Sobolev gradient: Suppose for a given function we substitute $\langle\ ,\ \rangle_A$ for $\langle\ ,\ \rangle$ in (1.2). To what sort of gradient are we led? Given $x \in R^n$, what member $(\nabla_A\phi)(x)$ gives us the identity:

$$\phi'(x)h = \langle h, (\nabla_A\phi)(x)\rangle_A\ ,\ h \in R^n? \tag{1.5}$$

We think of R^n made into two different normed linear spaces, one with the standard Euclidean norm $\|\ \ \|$, and the other with $\|x\|_A = \langle x,x\rangle_A^{1/2}, x \in R^n$. We calculate, for $x \in R^n$, $\phi'(x)h = \langle h, (\nabla_A\phi)(x)\rangle_A = \langle Ah, (\nabla_A\phi)(x)\rangle = \langle h, A(\nabla_A\phi)(x)\rangle$, for all h in R^n. But this implies that $A(\nabla_A\phi)(x) = (\nabla\phi)(x)$ and hence

$$(\nabla_A\phi)(x) = A^{-1}(\nabla\phi)(x), x \in R^n, \tag{1.6}$$

since

$$\phi'(x)h = \langle h, (\nabla\phi)(x)\rangle, x, h \in R^n.$$

The inverse in (1.6) exists since A is positive definite.

The relationship (1.6) is typical of our development although in some instances A is nonlinear; in others there is a separate 'A' for each x; in still others, these two are combined. These will all be investigated in what follows.

To continue our introduction, we point out two related versions of steepest descent. The earliest reference we know to steepest descent is Cauchy [**17**]. The first version is discrete steepest descent, the second is continuous steepest descent. By discrete steepest descent we mean an iterative process

$$x_n = x_{n-1} - \delta_{n-1}(\nabla_A\phi)(x_{n-1}), n = 1, 2, 3, ..., \tag{1.7}$$

where x_0 is given and δ_{n-1} is chosen optimally to minimize, if possible,

$$\phi(x_{n-1} - \delta(\nabla_A\phi)(x_{n-1})),\ \delta \in R.$$

On the other hand, continuous steepest descent consists of finding a function $z : [0,\infty) \to R$ so that

$$z(0) = x \in R^n, z'(t) = -(\nabla_A\phi)(z(t)),\ t \geq 0. \tag{1.8}$$

Continuous steepest descent may be interpreted as a limiting case of (1.8) in which, roughly speaking, various δ_n tend to zero (rather than being chosen optimally). Conversely, (1.7) might be considered (without the optimality condition on δ) as a numerical method (Euler's method) for approximating solutions to (1.8).

Using (1.7) we seek $u = \lim_{n\to\infty} x_n$ so that

$$\phi(u) = 0 \tag{1.9}$$

or

$$(\nabla_A \phi)(u) = 0. \tag{1.10}$$

Using (1.8) we seek $u = \lim_{t\to\infty} z(t)$ so that (1.9) or (1.10) holds. Before we consider more general forms of gradients (for example where A in (1.6) is nonlinear), we give in Chapter 2 an example which we hope convinces the reader that there are substantial issues concerning Sobolev gradients. We hope that Chapter 2 provides motivation for further reading even though later development's do not depend logically upon Chapter 2. We close this introduction by recalling two applications of steepest descent.

(1) Many differential equations have a variational principal, *i.e.* there is a function ϕ such that x satisfies the differential equation if and only if x is a critical point of ϕ. In such cases we try to use steepest descent to find a zero of a gradient of ϕ.

(2) In other problems we write a system of nonlinear differential equations in the form

$$F(x) = 0 \tag{1.11}$$

where F maps a Banach space H of functions into another such space K. In some cases one might define for some $p > 1$, $\phi : H \to R$ by

$$\phi(x) = \|F(x)\|^p/p, x \in H.$$

Then one might seek x satisfying this equation by means of steepest descent.

Problems of both kinds are considered. The following chapter contains an example of the second kind.

CHAPTER 2

Comparison of Two Gradients

This chapter gives a comparison between conventional and Sobolev gradients for a finite dimensional problem associated with a simple differential equation. On first reading one might examine just enough to understand the statements of the two theorems. Nothing in the following chapters depends on the techniques of the proofs of these results.

Suppose that ϕ is a $C^{(2)}$ real-valued function on R^n and $\nabla_A \phi$ is the gradient associated with ϕ by means of the positive definite symmetric matrix A. A measure of the worth of $\nabla_A \phi$ in regard to a descent process (similar comments apply to ascent problems associated with a maximization process) is

$$\sup_{x \in R^n, \phi(x) \neq 0} \phi(x - \delta_x (\nabla_A \phi)(x))/\phi(x) \tag{2.1}$$

where, for each $x \in R^n$, $\delta_x \in R$ is chosen optimally, i.e. to minimize

$$\phi(x - \delta(\nabla_A \phi)(x)),\ \delta > 0. \tag{2.2}$$

Generally, the smaller the value in (2.1), the greater the worst case improvement in each discrete steepest descent step. We remark that $(\nabla_A \phi)(x)$ is a descent direction at x (unless $(\nabla_A \phi)(x) = 0$), since if $f(\delta) = \phi(x - \delta(\nabla_A \phi)(x))$, $\delta \geq 0$, then $f'(0) = -\|(\nabla_A \phi)(x)\|_A^2 < 0$. We will use (2.1) to compare two gradients arising from the same function ϕ. We choose ϕ which arises from a discretization of a differential equation into a finite difference problem. We pick the simple problem

$$y' = y \text{ on } [0, 1]. \tag{2.3}$$

For each positive integer n and $\gamma_n = 1/n$, define $\phi_n : R^{n+1} \to R$ so that if

$$x = (x_0, x_1, ..., x_n) \in R^{n+1},$$

then

$$\phi_n(x) = \sum_{i=1}^{n} ((x_i - x_{i-1})/\gamma_n - (x_i + x_{i-1})/2)^2/2. \tag{2.4}$$

Consider first the conventional gradient $\nabla \phi_n$ of ϕ_n. Pick $y \in C^{(3)}$ so that at least one of the following hold:

$$y'(0) - y(0) \neq 0,\ y'(1) - y(1) \neq 0. \tag{2.5}$$

Condition (2.5) amounts to the requirement that $y' - y$ **not** be in the domain of the formal adjoint of

$$L : Lz = z' - z, \ z \text{ absolutely continuous on } [0,1] \tag{2.6}$$

(cf. [**28**]). We define a sequence of points $\{w^n\}_{n=1}^{\infty}$, $w^n \in R^{n+1}$, $n = 1, 2, \dots$, which are taken from y in the sense that for each positive integer n, w^n is the member of R^{n+1} so that

$$w_i^n = y(i/n), \ i = 0, 1, \dots, n. \tag{2.7}$$

What we will show is that our measure of worth (2.1) deteriorates badly if we in succession choose $\phi = \phi_n$ and $x = w^n$. More specifically we have:

Theorem 2.

$$\lim_{n \to \infty} (\phi_n(w^n - \delta_n(\nabla\phi_n)(w^n))/\phi_n(w^n) = 1,$$

where for each positive integer n, δ_n is chosen optimally in the sense of (2.2).

This theorem expresses what many must have sensed in using straightforward steepest descent on differential equations. If one makes a definite choice for y with, say $y'(0) - y(0) \neq 0$, then one finds that the gradients $(\nabla\phi_n)(w^n)$, even for n quite small, have very large first component relative to all the others (except possibly the last one if $y'(1) - y(1) \neq 0$). This in itself renders $(\nabla\phi_n)(w^n)$ an unpromising object with which to perturb w^n in order that

$$w^n - \delta(\nabla\phi_n)(w^n)$$

should be a substantially better approximation to a zero of ϕ_n than is w^n.

Proof. (Theorem 2) Denote by n a positive integer. Denote $1/n$ by γ_n, $(1/\gamma_n + 1/2)$ by c, $(1/\gamma_n - 1/2)$ by d. Denote by Q_n the transformation from R^{n+1} to R^n so that if $x = (x_0, x_1, \dots, x_n)$, then

$$Q_n x = (r_1, \dots, r_n), \text{ where } \quad r_i = -cx_{i-1} + dx_i, i = 1, \dots, n.$$

Observe that in terms of Q_n,

$$\phi_n(x) = \|Q_n x\|^2/2, \ x \in R^{n+1}, \tag{2.8}$$

so that if $x, h \in R^{n+1}$ then

$$\phi_n'(x)h = \langle Q_n h, Q_n x\rangle = \langle h, Q_n^t Q_n x\rangle$$

and hence

$$\nabla\phi_n = Q_n^t Q_n. \tag{2.9}$$

Thus if $\delta > 0$, $x \in R^{n+1}$ and $Q_n x \neq 0$,

$$(\phi_n(x - \delta(\nabla\phi_n)(x)))/\phi_n(x) = \|Q_n(x - \delta Q_n^t Q_n x)\|^2/2.$$

This expression is a quadratic in δ and has its minimum at

$$\delta_n = \langle Q_n x, Q_n Q_n^t Q_n x\rangle/\|Q_n Q_n^t Q_n x\|^2.$$

In particular

$$\phi_n(w^n - \delta_n(\nabla\phi_n)(w^n))/\phi_n(w^n) = 1 - \|Q_n^t g_n\|^4/(\|Q_n Q_n^t g_n\|^2\| \cdot \|g_n\|^2)$$

where $g_n = Q_n w^n$. Inspection yields that the following limits all exist and are positive: $\lim_{n\to\infty} \|Q_n^t g_n\|^4/n^4$, $\lim_{n\to\infty} \|Q_n Q_n^t g_n\|^2/n^4$, $\lim_{n\to\infty} \|g_n\|^2/n$ It follows that $\phi_n(w^n - \delta_n(\nabla\phi_n)(w^n))/\phi_n(w^n) = 1$, and the argument is finished. □

Suppose that ϕ is defined so that

$$\phi(z) = \|Lz\|^2/2, z \in L_2([0,1]),\ z \text{ absolutely continuous.}$$

Now suppose in addition that z is $C^{(2)}$ but z is not in the domain of the formal adjoint of L (i.e., either $z'(0) - z(0) \neq 0$ or $z'(1) - z(1) \neq 0$). Then there is no $v \in L_2$ so that

$$\phi'(z)h = \langle h, v\rangle_{L_2([0,1])} \text{ for all } \quad h \in C^2([0,1]).$$

Hence, in a sense, there is nothing that the gradients

$$\{(\nabla\phi_n)(w^n)\}_{n=1}^{\infty}$$

are approximating. One should expect really bad numerical performance of these gradients (and one is not at all disappointed).

This example represents something common. Sobolev gradients give us an organized way to modify these poorly performing gradients in order to obtain gradients with good numerical and analytical properties. We give now a construction of Sobolev gradients corresponding to the ones above. We then give a theorem which indicates how they are judged by (2.1). Choose a positive integer n. We indicate a second gradient for ϕ_n which will be our first Sobolev gradient for a differential equation. For R^{n+1} we choose the following norm:

$$\|x\|_{A_n} = [\sum_{i=1}^{n}(((x_i - x_{i-1})/\gamma_n)^2 + ((x_i + x_{i-1})/2)^2)]^{1/2},\ x \in R^{n+1} \tag{2.10}$$

It is easy to see that this norm carries with it an inner product

$$\langle\ ,\ \rangle_{A_n}$$

on R^{n+1}. This inner product is related to the standard inner product $\langle\ ,\ \rangle$ on R^{n+1} by

$$\langle x, y\rangle_{A_n} = \langle A_n x, y\rangle$$

where if $x = (x_o, x_1, ..., x_n)$ then

$$A_n x = z$$

so that

$$z_0 = ((c^2 + d^2)/2)x_0 - cdx_1,$$

$$z_i = -cdx_{i-1} + (c^2 + d^2)x_i - cdx_{i+1},\ i = 1, ..., n-1,$$

$$z_n = -cdx_{n-1} + ((c^2 + d^2)/2)x_n,$$

in which $c = (1/\gamma_n + 1/2)$, $d = (1/\gamma_n - 1/2)$. Accordingly, using (2.8) we have that the gradient $\nabla_{A_n}\phi_n$ of ϕ_n is given by

$$(\nabla_{A_n}\phi_n)(x) = A_n^{-1}(\nabla\phi_n)(x). \tag{2.11}$$

In contrast to Theorem 2 we have

THEOREM 3. *If* $x \in R^{n+1}$ *then*

$$\phi_n(x - \delta_n(\nabla_{A_n}\phi_n)(x)) \le \phi_n(x)/9, \ n = 1, 2, ...$$

This indicates that $\nabla_{A_n}\phi_n$ performs much better numerically than $\nabla\phi_n$. Before a proof is given we have

LEMMA 1. *Suppose* n *is a positive integer,* $\lambda > 0$*,* $v \in R^n$*,* $v \neq 0$*. Suppose also that* $M \in L(R^n, R^n)$ *is so that* $Mv = \lambda v$ *and* $Mx = x$ *if* $\langle x, v\rangle = 0$*. Then* $\langle My, y\rangle^2/(\|My\|^2\|y\|^2) \ge 4\lambda/(1+\lambda)^2 y \in R^n$*,* $y \neq 0$*.*

PROOF. Pick $y \in R^n$, $y \neq 0$ and write $y = x + r$ where r is a multiple of v *and* $\langle x, r\rangle = 0$. Then

$$\langle My, y\rangle^2/(\|My\|^2\|y\|^2) = (\langle x + \lambda r, x + r\rangle/(\|x + \lambda r\|)^2\|x + r\|^2) =$$

$$(\|x\|^2 + \lambda\|r\|^2)^2/((\|x\|^2 + \lambda^2\|r\|^2)(\|x\|^2 + \|r\|^2)) =$$
$$(\sin^2\theta + \lambda cos^2\theta)^2/(\sin^2\theta + \lambda^2\cos^2\theta), \tag{2.12}$$

where

$$\sin^2\theta = \|x\|^2/(\|x\|^2 + \|r\|^2), \cos^2\theta = \|r\|^2/(\|x\|^2 + \|r\|^2).$$

Expression (2.12) is seen to have its minimum for $\cos^2\theta = 1/(1+\lambda)$. The lemma readily follows. □

PROOF. (Theorem 3). Denote $Q_nA_n^{-1}Q_n^t$ by M_n. Using (2.8), (2.9), (2.11), for $x \in R^{n+1}$, $\delta \ge 0$,

$$\phi_n(x - \delta(\nabla_{A_n}\phi_n)(x)) = \|g\|^2 - 2\delta\langle g, M_ng\rangle + \delta^2\|M_ng\|^2,$$

where $g = Q_nx$. This expression is minimized by choosing

$$\delta = \delta_n = \langle g, M_ng\rangle/\|M_ng\|^2,$$

so that with this choice of δ,

$$\phi_n(x - \delta_n(\nabla_{A_n}\phi_n)(x))/\phi_n(x) = 1 - \langle g, M_ng\rangle^2/(\|M_ng\|^2\|g\|^2).$$

To get an expression for M_ng, we first calculate $u = A_n^{-1}Q_n^tg$. To accomplish this we intend to solve

$$A_nu = Q_n^tg \tag{2.13}$$

for $u = (u_0, u_1, ..., u_n)$. Writing $(g_1, ..., g_n)$ for g, (2.13) becomes the system

$$((c^2 + d^2)/2)u_0 - cdu_1 = -cg_1,$$
$$-cdu_{i-1} + (c^2 + d^2)u_i - cdu_{i+1} = dg_i - cg_{i+1}, i = 1, ..., n-1,$$
$$-cdu_{n-1} + ((c^2 + d^2)/2)u_n = dg_n. \tag{2.14}$$

From equations $1, ..., n-1$ it follows, using standard difference equation methods, that there must be α and β so that

$$u_0 = \alpha r^0 + \beta s^0$$

$$u_i = \alpha r^i + \beta s^i + (1/d) \sum_{k=1}^{i} r^{i-k} g_k, i = 1, ..., n, \quad (2.15)$$

where $r = c/d$, $s = d/c$. The first equation of (2.14) implies that $\alpha = \beta$. The last equation in (2.14) may then be solved for α, leaving us with

$$u_0 = -\eta(r^0 + s^0)$$

$$u_i = -\eta(r^i + s^i) + (1/d)(\sum_{k=1}^{i} r^{i-k} g_k), i = 1, ..., n, \quad (2.16)$$

where $\eta = (\sum_{k=1}^{n} r^{n-k} g_k)/(d(r^n - s^n))$. Note that $M_n g = Q_n u$. After some simplification of (2.16), we are led to the fact that

$$M_n g = \langle g, z^{(n)} \rangle z^{(n)} + g$$

where $z^{(n)} = (z_1, ..., z_n)$ is defined by

$$z_i = (\sum_{k=0}^{n-1} s^{2k})^{-1/2} s^{i-1}, \ i = 1, ..., n.$$

Note that $\|z^{(n)}\| = 1$ and that M_n satisfies the conditions of the Lemma with $\lambda = 2$. Accordingly,

$$1 - \langle g, M_n g\rangle^2/(\|M_n g\|^2 \|g\|^2) \leq 1 - 4\lambda/(1+\lambda)^2$$

$$= (\lambda - 1)^2/(\lambda + 1)^2 = 1/9.$$

Hence

$$\phi_n(x - \delta_n(\nabla_{A_n}\phi_n)(x))/\phi_n(x) \leq 1/9,$$

so our argument is complete. □

The above follows an argument in [**67**] for a slightly different case.

The inequality in Theorem 3 implies a good rate of convergence for discrete steepest descent since it indicates that the ratio of the norm of a new residual and the norm of the old residual is no more than 1/3. We will see that Theorem 3 can be extended by continuity to a function space setting. No such extension of Theorem 2 seems possible. Actual programming of the process in Theorem 2 leads to a very slowly converging iteration. The number of steps required to reach a fixed accuracy increases very fast as n increases with perhaps 500,000 iterations required for $n = 100$. By contrast, the process in Theorem 3 requires about 7 steps, independently of the choice of n. A computer code connected with Theorem 3 is given in Chapter 18. The results of the present chapter illustrates the observation that for a given problem, analytical and numerical difficulties always come in pairs.

CHAPTER 3

Continuous Steepest Sescent in Hilbert Space: Linear Case

In this chapter we consider continuous steepest descent for linear operator equations in Hilbert space.

Suppose $G \in L(H, K)$, $g \in K$ and $Fx = Gx - g$, $x \in H$. The following shows that if there is a solution v to $Fv = 0$, then a solution may be found by means of continuous steepest descent.

THEOREM 4. *Suppose there is $v \in H$ so that $Gv = g$ and*

$$\phi(y) = \|Gy - g\|_K^2/2, \ y \in H.$$

Suppose also that $x \in H$ and z is the function on $[0, \infty)$ so that

$$z(0) = x, \ z'(t) = -(\nabla\phi)(z(t)), \ t \geq 0. \tag{3.1}$$

Then $u = \lim_{t\to\infty} z(t)$ exists and $Gu = g$.

PROOF. First note that

$$\phi'(y)h = \langle Gh, Gy - g\rangle_K = \langle h, G^*(Gy - g)\rangle_H$$

so that

$$(\nabla\phi)(y) = G^*(Gy - g), \ y \in H. \tag{3.2}$$

Restating (3.1) we have that

$$z(0) = x, \ z'(t) = -G^*Gz(t) + G^*g, \ t \geq 0.$$

From the theory of ordinary differential equations in a Banach space we have that

$$z(t) = \exp(-tG^*G)x + \int_0^t \exp(-(t-s)G^*G)G^*g \ ds, \ t \geq 0. \tag{3.3}$$

Now by hypothesis, v is such that $Gv = g$. Replacing g in the preceding by Gv and using the fact that

$$\int_0^t \exp(-(t-s)G^*G)G^*Gv \ ds = \exp(-(t-s)G^*G)v \,\big|_{s=0}^{s=t} = v - \exp(-tG^*G)v$$

we have that $z(t) = \exp(-tG^*G)x + v - \exp(-tG^*G)v$. Now since G^*G is symmetric and nonnegative, $\exp^{-tG^*G}$ converges strongly, as $t \to \infty$, to the orthogonal projection P onto the null space of G (see the discussion of the spectral theorem in [**91**], Chap. VII). Accordingly, $u \equiv \lim_{t\to\infty} z(t) = Px + v - Pv$. But then

$$Gu = GPx + Gv - GPv = Gv = g$$

since $Px, Pv \in N(G)$. □

Notice that abstract existence of a solution v to $Gv = g$ leads to a concrete function z whose asymptotic limit is a solution u to

$$Gu = g.$$

The reader might see [**49**] in connection with these problems. In case g is not in the range of G there is the following:

THEOREM 5. *Suppose $G \in L(H, K)$, $g \in K$, $x \in H$ and z satisfies* (3.1). *Then*

$$\lim_{t\to\infty} Gz(t) = g - Qg$$

where Q is the orthogonal projection of K onto $R(G)^\perp$.

PROOF. Using (3.3)

$$Gz(t) = G(\exp(-tG^*G)x) + G(\int_0^t \exp(-(t-s)G^*G)G^*g\ ds),\ t \geq 0$$

$$= \exp(-tGG^*)Gx + \int_0^t \exp(-(t-s)GG^*)GG^*g\ ds$$

$$= \exp(-tGG^*)Gx + \exp(-(t-s)GG^*)g \,\Big|_{s=0}^{s=t}$$

$$= \exp(-tGG^*)Gx + g - \exp(-tGG^*)g \to g - Qg$$

since $\exp(-tGG^*)$ converges strongly to Q, the orthogonal projection of K onto $N(G^*) = R(G)^\perp$, as $t \to \infty$. □

Next we note the following characterization of the solution u obtained in Theorem 4

THEOREM 6. *Under the hypothesis of Theorem 4, if $x \in H$ and z is the function from $[0, \infty)$ to H so that* (3.1) *holds, then $u \equiv \lim_{t\to\infty} z(t)$ has the property that*

$$\|u - z(t)\|_H < \|y - z(t)\|_H,\ t \geq 0,\ y \in H \text{ such that } \quad Gy = g,\ y \neq u.$$

PROOF. Suppose $w \in H$ and $Gw = g$. For u as in the argument for Theorem 4, notice that $u - w \in N(G)$ and $x - u = (I - P)(x - v)$. Hence $\langle x - u, u - w\rangle_H = 0$ since $I - P$ is the orthogonal projection onto $N(G)^\perp$. Consequently $\|x - w\|_H^2 = \|x - u\|_H^2 + \|u - w\|^2$ and so u is the nearest element to x which has the property that $Gu = g$. Now if $t \geq 0$, then $x - z(t) = (I - \exp(-t^*GG))(x - v)$. Since

$$Pe^{-tG^*G} = P,$$

([**91**], Chap VII) it follows that $P(x - z(t)) = 0$ and hence u is the also the nearest element to x which has the property that $Gu = g$. □

Another way to express Theorem 5 is that P, the orthogonal projection onto $N(G)$ provides an invariant for steepest descent generated by $-\nabla\phi$ in the sense that $Px = Pz(t)$, $t \geq 0$. An invariant (or a set of invariants) for steepest descent in nonlinear cases would be very interesting. More about this problem will be indicated in chapter 14.

CHAPTER 4

Continuous Steepest Descent in Hilbert Spaces: Nonlinear Case

Denote by H a real Hilbert space and suppose that ϕ is a $C^{(2)}$ function on all of H. For this chapter denote by $\nabla\phi$ the function on H so that if $x \in H$, then

$$\phi'(x)h = \langle h, (\nabla\phi)(x)\rangle_H, \ h \in H,$$

where ϕ' denotes the Fréchet derivative of ϕ. Here we have adopted (1.2) as our point of view on gradients; (1.3) is equivalent but (1.1) is not adequate. Here we will seek zeros of ϕ by means of continuous steepest descent, *i.e.*, we seek

$$u = \lim_{t\to\infty} z(t) \ \text{ exists and } \ \phi(u) = 0. \tag{4.1}$$

1. Global Existence

We first want to establish global existence for the steepest descent in this setting.

THEOREM 7. *Suppose that ϕ is a nonnegative $C^{(2)}$ function on H. If $x \in H$, there is a unique function z from $[0,\infty)$ to H such that*

$$z(0) = x, \ z'(t) = -(\nabla\phi)(z(t)), \ t \geq 0. \tag{4.2}$$

PROOF. Since ϕ is a $C^{(2)}$ function, it follows that $\nabla\phi$ is a $C^{(1)}$ function. From basic ordinary differential equation theory there is $d_0 > 0$ so that the equation in (4.2) has a solution on $[0, d_0)$. Suppose that the set of all such numbers d_0 is bounded and denote by d its least upper bound.

Denote by z the solution to the equation in (4.2) on $[0, d)$. We intend to show that $\lim_{t\to d} z(t)$ exists. To this end note that if $0 \leq a < b < d$ then

$$\|z(b) - z(a)\|_H^2 = \|\int_a^b z'\|_H^2 \leq (\int_a^b \|z'\|_H)^2 \leq (b-a)\int_a^b \|z'\|_H^2 \tag{4.3}$$

Note also that

$$\phi(z)'(t) = \phi'(z(t))z'(t) = \langle z'(t), (\nabla\phi)(z(t)\rangle_H = -\|(\nabla\phi)(z(t))\|_H^2 \tag{4.4}$$

so

$$\phi(z(b)) - \phi(z(a)) = \int_a^b \phi(z)' = -\int_a^b \|z'\|_H^2$$

Hence

$$\int_a^b \|z'\|_H^2 \leq \phi(z(a)),\ b \geq a. \tag{4.5}$$

Using (4.3) and (4.5)

$$\|z(b) - z(a)\| \leq \int_a^b \|z'\|_H \leq (d\phi(a))^{1/2},\ a \leq b < d.$$

But this implies that $\int_a^{d-} \|z'\|_H$ exists and so $q \equiv \lim_{t \to d} z(t)$ exists. But again by the basic existence for ordinary differential equations, there is $c > d$ for which there is a function y on $[d, c)$ such that

$$y(d) = q,\ y'(t) = -(\nabla\phi)(y(t)),\ t \in [d, c).$$

But the function w on $[0, c)$ so that $w(t) = z(t),\ t \in [0, d), w(d) = q,\ w(t) = y(t),\ t \in (d, c)$, satisfies

$$w(0) = x, w'(t) = -(\nabla\phi)(w(t)),\ t \in [0, c),$$

contradicting the nature of d since $d < c$. Hence there is a solution to (4.2) on $[0, \infty)$. Uniqueness follows from the basic existence and uniqueness theorem for ordinary differential equations. □

Note that in the above result and in some others to follow in this chapter, the hypothesis that $\phi \in C^{(2)}$ can be replaced with the weaker assumption that $\nabla\phi$ is locally lipschitian on H. See [**5**] for another discussion of steepest descent in Hilbert spaces.

A function ϕ as in the hypothesis of Theorem 7 generates a one parameter semigroup of transformations on H. Specifically, define T_ϕ so that if $s \geq 0$, then $T_\phi(s)$ is the transformation from H to H such that

$$T_\phi(s)x = z(s),\ \text{where}\ \ z \text{ satisfies (4.2)}.$$

THEOREM 8. *$T_\phi(t)T_\phi(s) = T_\phi(t+s)$, $t, s \geq 0$ where $T_\phi(t)T_\phi(s)$ indicates composition of the transformations $T_\phi(t)$ and $T_\phi(s)$.*

PROOF. Suppose $x \in H$ and $s > 0$. Define z satisfying (4.2) and define y so that $y(s) = T_\phi(s)x$, $y'(t) = -(\nabla\phi)(y(t))$, $t \geq s$. Since $y(s) = z(s)$ and y, z satisfy the same differential equation on $[s, \infty)$, we have by uniqueness that $y(t) = z(t)$, $t \geq s$ and so have the truth of the theorem. □

2. Gradient Inequality

Note that our proof of Theorem 7 yields that

$$\int_0^\infty \|(\nabla\phi)(z)\|_H^2 < \infty \tag{4.6}$$

under the hypothesis of that theorem. Now the conclusion

$$\int_0^\infty \|(\nabla\phi)(z)\|_H < \infty \tag{4.7}$$

implies, as in the argument for Theorem 7, that $\lim_{t\to\infty} z(t)$ exists. Clearly (4.6) does not imply (4.7). For an example, take $\phi(x) = e^{-x}$, $x \in R$. Note also that under the hypothesis of Theorem 7

$$\{\phi(z(t)) \ : \ t \geq 0\} \text{ is bounded} \tag{4.8}$$

since $\phi(z(t)) \leq \phi(x), t \geq 0$. Following are some propositions which lead to the conclusion (4.7).

Definition 1. ϕ is said to satisfy a gradient inequality on $\Omega \subset H$ provided there is $c > 0$ such that

$$\|(\nabla\phi)(x)\|_H \geq c\phi(x)^{1/2}, \ x \in \Omega. \tag{4.9}$$

We will see that this condition gives compactness in some instances and leads to (4.1) holding (Theorems 4.2,4.4,4.5,4.8).

Theorem 9. *Suppose ϕ is a nonnegative $C^{(2)}$ function on $\Omega \subset H$ and ϕ satisfies* (4.9). *If $x \in \Omega$ and z satisfies* (4.2) *then* (4.1) *holds provided that $R(z) \subset \Omega$.*

Proof. Suppose z satisfies (4.2), (4.9) holds on Ω and $R(z) \subset \Omega$. Note that if $(\nabla\phi)(x) = 0$, then the conclusion holds. Suppose that $(\nabla\phi)(x) \neq 0$ and note that then $(\nabla\phi)(z(t)) \neq 0$ for all $t \geq 0$ (for if $(\nabla\phi)(z(t_0)) = 0$ for some $t_0 \geq 0$, then the function w on $[0,\infty)$ defined by $w(t) = z(t_0)$, $t \geq 0$, would satisfy $w' = (\nabla\phi)(w)$ and $w(t_0) = z(t_0)$, the same conditions as z; but $z \neq w$ and so uniqueness would be violated). Now by (4.4),

$$\phi(z)' = -\|(\nabla\phi)(z)\|_H^2$$

and so, using (4.9)

$$\phi(z)'(t) \leq -c^2\phi(z(t))$$

and

$$\phi(z)'(t)/\phi(z(t)) \leq -c^2, \ t \geq 0.$$

Accordingly,

$$\ln(\phi(z(t))/\phi(x)) \leq -c^2 t$$

and

$$\phi(z(t)) \leq \phi(x)\exp(-c^2 t), \ t \geq 0.$$

Therefore

$$\lim_{t\to\infty} \phi(z(t)) = 0.$$

Moreover if n is a positive integer,

$$\left(\int_n^{n+1} \|z'\|_H\right)^2 \leq \int_n^{n+1} \|z'\|_H^2 = \phi(z(n)) - \phi(z(n+1)) \leq \phi(x)\exp(-c^2 n)$$

and so

$$\int_0^\infty \|z'\|_H \leq \phi(x)^{1/2} \sum_{n=0}^\infty \exp(-c^2 n/2) = \phi(x)^{1/2}/(1-\exp(-c^2/2)).$$

Consequently, $\|z'\|_H \in L_1([0,\infty)]$ and so $u = \lim_{t\to\infty} z(t)$ exists. Moreover $\phi(u) = 0$ since $0 = \lim_{t\to\infty} \phi(z(t))$. □

The following lemma leads to short arguments for Theorems 10 and 11.

LEMMA 2. *Suppose ϕ is a nonnegative $C^{(2)}$ function on H, $c > 0$ and Ω is an open subset of H so that (4.9) holds. Suppose that $x \in H$ and z satisfies (4.2). Then there does not exist $\epsilon > 0$ and sequences $\{s_i\}_{i=1}^{\infty}, \{r_i\}_{i=1}^{\infty}$ so that $\{[s_i, r_i]\}_{i=1}^{\infty}$ is a sequence of pairwise disjoint intervals with the property that*

$$(1)\ \ s_i < r_i < s_{i+1},$$

$$(2)\ \ s_i, r_i \to \infty \text{ as } i \to \infty,$$

$$(3)\ \ \|z(r_i) - z(s_i)\|_H \geq \epsilon$$

$$(4)\ \ z(t) \in \Omega \text{ if } t \in [s_i, t_i],\ \text{for } i = 1, 2, \ldots.$$

PROOF. Suppose that the hypothesis holds but that the conclusion does not. Denote by ϵ a positive number and by

$$\{s_i\}_{i=1}^{\infty}, \{r_i\}_{i=1}^{\infty}$$

sequences so that $\{[r_i, s_i]\}_{i=1}^{\infty}$ is a sequence of pairwise disjoint intervals so that (1)-(4) hold. Note that $(\nabla\phi)(z(t)) \neq 0$ for $t \geq 0$ since if not then $(\nabla\phi)(z(t)) = 0$ for all $t \geq 0$ and hence z is constant and consequently, (3) is violated. Now for each positive integer i,

$$\epsilon^2 \leq \|z(r_i) - z(s_i)\|_H^2 = (\int_{s_i}^{r_i} \|z'\|_H)^2$$

$$\leq (\int_{s_i}^{r_i} \|z'\|_H)^2 \leq (r_i - s_i) \int_{s_i}^{r_i} \|z'\|_H^2.$$

As in the proof for Theorem 7, if $0 \leq a < b$,

$$\phi(z(a)) = \phi(z(b)) + \int_a^b \|z'\|_H^2.$$

Hence $\int_0^{\infty} \|z'\|_H^2$ exists and therefore

$$\lim_{i\to\infty} \int_{s_i}^{r_i} \|z'\|_H^2 = 0.$$

Consequently, $\lim_{i\to\infty}(r_i - s_i) = \infty$ since

$$\epsilon^2 \leq (r_i - s_i) \int_{s_i}^{r_i} \|z'\|_H^2,\ i = 1, 2, \ldots.$$

Since $\phi(z)'(t) = -\|(\nabla\phi)(z(t)\|_H^2 \leq -c^2\phi(z(t)), t \geq 0$, it follows that

$$\phi(z)'(t)/\phi(z(t)) \leq -c^2,\ t \geq 0$$

and so for each positive integer i,

$$\phi(z(t)) \leq \phi(z(s_i)) \exp(-c^2(t - s_i)),\ t \in [s_i, r_i] \text{ and in particular,}$$

$$\phi(z(r_i)) \le \phi(z(s_i))\exp(-c^2(r_i - s_i)).$$

Therefore, since $\phi(z(r_i)) \ge \phi(z(s_{i+1}))$, $i = 1, 2, \ldots$, it follows that

$$\lim_{i\to\infty} \phi(z(s_i)) = 0.$$

Denote by i a positive integer so that $r_i - s_i > 1$, denote $[r_i - s_i]$ by k and denote $s_i, s_i + 1, ..., s_i + k, r_i$ by $q_0, q_1, ..., q_{k+1}$. Then

$$\epsilon \le \|z(r_i) - z(s_i)\|_H \le \sum_{j=0}^{k} \|z(q_{j+1}) - z(q_j)\|_H \le \sum_{j=0}^{k} \int_{s_i+j}^{s_i+j+1} \|z'\|_H$$

$$\le \sum_{j=0}^{k} (\int_{s_i+j}^{s_i+j+1} \|z'\|_H^2)^{1/2} = \sum_{j=0}^{k} (\phi(z(s_i + j)) - \phi(z(s_i + j + 1)))^{1/2}$$

$$\le \sum_{j=0}^{k} \phi(z(s_i + j))^{1/2} \le \sum_{j=0}^{k} (\phi(z(s_i)) \exp(-c^2 j)^{1/2}$$

$$= (\phi(z(s_i)) \sum_{j=0}^{k} \exp(-c^2 j))^{1/2}$$

$$\le \phi(z(s_i))^{1/2} \sum_{j=0}^{k} \exp(-jc^2/2)$$

$$\le \phi(z(s_i))^{1/2} (1 - \exp(-c^2/2))^{-1}$$

since $\phi(z(a + 1)) \le \phi(z(a)) \exp(-c^2), a = s_i, s_i + 1, \ldots s_i + k - 1$ under our hypothesis. But since $\phi(z(s_i)) \to 0$ as $i \to \infty$ and one arrives at a contradiction. $\square$

The following rules out some conceivable alternatives:

THEOREM 10. *Suppose that ϕ is a nonnegative $C^{(2)}$ function on all of H which satisfies* (4.9) *for every bounded subset Ω of H. If z satisfies* (4.2) *then either*

$$(i) \ \ (4.1) \ holds$$

or else

$$(ii) \ \ \lim_{t\to\infty} \|z(t)\|_H = \infty.$$

PROOF. Suppose that $x \in H$, z satisfies (4.2) and (4.9) holds for every bounded subset Ω of H. Suppose furthermore that $R(z)$ is not bounded but nevertheless does not satisfy (ii) of the theorem. Then there are $r, s > 0$ so that $0 < s < r$, and two unbounded increasing sequences $\{r_i\}_{i=1}^{\infty}, \{s_i\}_{i=1}^{\infty}$ so that (i) $s_i < r_i < s_{i+1}$, (ii) $\|z(s_i)\|_H = s$, $\|z(r_i)\|_H = r$, (iii) $s \le \|z(t)\|_H \le r$, $t \in$

$[s_i, r_i]$, $i = 1, 2, \ldots$ But by the Lemma 2, this is impossible and the theorem is established. □

A similar phenomenon has been indicated in [**2**] for semigroups related to monotone operators.

THEOREM 11. *Under the hypothesis of Theorem 10, suppose that* $x \in H$ *and* z *satisfies* (4.2). *Suppose also that* u *is an* ω*-limit point of* z*, i.e.,*

$$u = \lim_{i\to\infty} z(t_i)$$

for some increasing unbounded sequence of positive numbers $\{t_i\}_{i=1}^{\infty}$*. Then* (4.1) *holds.*

PROOF. If z has an ω-limit point then (ii) of Theorem 10 can not hold and hence (i) must hold. □

We remind the reader of the Palais-Smale condition (see [**80**]) on ϕ : ϕ satisfies the Palais-Smale condition provided it is true that if $\{x_i\}_{i=1}^{\infty}$ is a sequence for which $\lim_{i\to\infty}(\nabla\phi)(x_i) = 0$ and $\{\phi(x_i)\}_{i=1}^{\infty}$ is bounded, then $\{x_i\}_{i=1}^{\infty}$ has a convergent subsequence. This leads us to the following:

THEOREM 12. *Suppose that* ϕ *is a nonnegative* $C^{(2)}$ *function on* H *which satisfies the Palais-Smale condition and also satisfies* (4.9) *for every bounded subset* Ω *of* H*. Then* (4.1) *holds.*

PROOF. Since by the proof of Theorem 7, (4.5) holds, it follows that there is an unbounded increasing sequence $\{t_i\}_{i=1}^{\infty}$ of positive numbers such that

$$\lim_{i\to\infty} (\nabla\phi)(z(t_i)) = 0.$$

Note also that

$$\phi(x) \geq \phi(z(t)) \geq 0, \ t \geq 0$$

holds. It follows from the (PS) condition that there is an increasing sequence $\{n_i\}_{i=1}^{\infty}$ of positive integers such that $\{z(t_{n_i})\}_{i=1}^{\infty}$ converges. But this rules out (ii) of Theorem 10 so that (i) of that theorem must hold. □

We note that in the above, we do not use the full strength of the (PS) condition; the conclusion '$\{x_i\}_{i=1}^{\infty}$ has a bounded subsequence' is sufficient for the purpose of Theorem 4.6.

Suppose now that K is a second real Hilbert space and that F is a $C^{(2)}$ function from H to K. Define ϕ by

$$\phi(x) = \|F(x)\|_K^2/2, \ x \in H. \tag{4.10}$$

In this case we have

$$\phi'(x)h = \langle F'(x)h, F(x)\rangle_K = \langle h, F'(x)^* F(x)\rangle_H, \ x, h, \in H,$$

so that

$$(\nabla\phi)(x) = F'(x)^* F(x), x \in H, \tag{4.11}$$

where $F'(x)^*$ denotes the Hilbert space adjoint of $F'(x), x \in H$. We see that if $\Omega \subset H$ so that $F'(x)^*$ is uniformly bounded below for $x \in \Omega$, i.e., there is $d > 0$ so that $\|F'(x)^* g\|_H \geq d\|g\|_K, x \in \Omega,\ g \in K$, then ϕ satisfies (4.6) with $c = 2^{-1/2}d$. One can do a little better with

THEOREM 13. *Suppose there exist $M, b > 0$ so that if $g \in K$ and $x \in \Omega$, then for some $h \in H, \|h\|_H \leq M$,*

$$\langle F'(x)h, g\rangle_K \geq b\|g\|_K.$$

Then (4.6) holds with $c = 2^{-1/2}b/M$.

PROOF. Suppose $x \in \Omega$. Then

$$\|(\nabla\phi)(x)\|_H = \sup_{h\in H, \|h\|_H = M} \langle h, F'(x)^* F(x)\rangle_H / M =$$

$$\sup_{h\in H, \|h\|_H = M} \langle F'(x)h, F(x)\rangle_K / M \geq (b/M)\|F(x)\|_K$$

since by hypothesis there is $h \in H$ such that $\|h\|_H \leq M$ and $\langle F'(x)h, F(x)\rangle_K \geq b\|F(x)\|_K$. □

Applications of Theorem 13 may be as follows: Many systems of nonlinear differential equations may be written (for appropriate H and K) as the problem of finding $u \in H$ such that $F(u) = 0$. The problem of finding h given $g \in K$, $u \in H$ such that

$$F'(u)h = g$$

then becomes a systems of linear differential equations. An abundant literature exists concerning existence of (and estimates for) solutions of such equations (cf **[38]**,**[98]**). Thus linear theory holds the hope of providing gradient inequalities in specific cases.

Another general result on continuous steepest descent is

THEOREM 14. *Suppose that F is a $C^{(2)}$ function from H to K, $x \in K$ and $r, c > 0$ exist so that*

$$\|(\nabla\phi)(y)\|_H \geq c\|F(y)\|_K, \|y - x\|_H \leq r,$$

where ϕ is defined by (4.10). Suppose also that (4.2) holds. Then (4.1) holds if $\|F(x)\|_K \leq cr$.

PROOF. Suppose $x \in H$ and denote by z the function so that (4.2) holds. If $(\nabla\phi)(x) = 0$, then $z(t) = x$ for all $t \geq 0$ and so (4.1) holds. Suppose now that $(\nabla\phi)(x) \neq 0$. Note that

$$(\nabla\phi)(z(t)) \neq 0,\ t \geq 0,$$

since if there were $t_0 > 0$ such that $(\nabla\phi)(z(t_0)) = 0$, then the function w so that $w(t) = z(t_0),\ t \geq 0$ would satisfy

$$w(t_0) = z(t_0), w'(t) = -(\nabla\phi)(w(t)),\ t \geq 0,$$

a contradiction to uniqueness since z also satisfies the above. Define α as the function which satisfies

$$\alpha(0) = 0,\ \alpha'(t) = 1/\|(\nabla\phi)(z(\alpha(t)))\|_H,\ t \in [0, d), \tag{4.12}$$

where d is as large as possible, possibly $d = \infty$.

Define v so that

$$v(t) = z(\alpha(t)),\ t \in D(\alpha). \tag{4.13}$$

Then

$$v'(t) = \alpha'(t)z'(\alpha(t)) = -\alpha'(t)(\nabla\phi)(z(\alpha(t)))$$

$$= -(1/\|(\nabla\phi)(z(\alpha(t)))\|_H)(\nabla\phi)(z(\alpha(t)))$$

and so

$$v(0) = x, v'(t) = -(1/\|(\nabla\phi)(v(t))\|_H)(\nabla\phi)(v(t)),\ t \in D(v) = D(\alpha) \tag{4.14}$$

Note that

$$\|v(t) - x\|_H = \|v(t) - v(0)\|_H = \|\int_0^t v'\|_H \le \int_0^t \|v'\|_H = t,\ t \in D(v).$$

Hence

$$\phi(v)' = \phi'(v)v' = -\langle(\nabla\phi)(v), (\nabla\phi)(v)\rangle_H/\|(\nabla\phi)(v)\|_H = -\|(\nabla\phi)(v)\|_H, \tag{4.15}$$

and so if $t \in D(v)$ and $t \le r$, then

$$\phi(v)'(t) = -\|(\nabla\phi)(v(t))\|_H \le -c\|F(v(t))\|_K = -c(2\phi(v(t)))^{1/2},$$

and thus

$$\phi(v)'(t)/\phi(v(t))^{1/2} \le -c2^{1/2}.$$

This differential inequality is solved to yield

$$2\phi(v(t))^{1/2} - 2\phi(v(0))^{1/2} \le -2^{1/2}ct,\ t \in D(v)\, t \le r.$$

But this is equivalent to

$$\|F(v(t))\|_K \le \|F(x)\|_K - ct \le c(r - t),\ t \in D(v),\ t \le r. \tag{4.16}$$

It follows from (4.16) that $d \le r$ since if not, there would be $t \in [0, r]$ such that $F(v(t_0)) = 0$ and hence $(\nabla\phi)(v(t_0)) = 0$ and consequently, $(\nabla\phi)(z(\alpha^{-1}(t_0))) = 0$, a contradiction. From this it follows that $R(v) \subset B_r(x)$. Since $D(\alpha)$ is a bounded set, and α is increasing, it must be that $\lim_{t\to d} \alpha(t) = \infty$ since $R(\alpha)$ is not bounded. Therefore

$$R(z) = R(v(\alpha^{-1})) \subset B_r(x).$$

It follows then from Theorem 9 that (4.2) holds. □

As an application of Theorem 14 there is the following implicit function theorem due to A. Castro and this writer [**16**].

THEOREM 15. *Suppose that each of H and K is a Hilbert space, $r, Q > 0, G$ is a $C^{(1)}$ function from H to K which has a locally lipschitzian derivative and $G(0) = 0$. Suppose also that there is $c_0 > 0$ so that if $u \in H, \|u\|_H \leq r$, and $g \in K, \|g\|_K = 1$, then*

$$\langle G'(u)v, g\rangle_K \geq c_0 \text{ for some } v \in H \text{ with } \|v\|_H \leq Q. \tag{4.17}$$

If $y \in K$ and $\|y\|_K < rc_0/Q$ then

$$u = \lim_{t\to\infty} z(t) \text{ exists and satisfies } G(u) = y \text{ and } \|u\|_H \leq r$$

where z is the unique function from $[0,\infty)$ to H so that

$$z(0) = 0, z'(t) = -(G'(z(t)))^*(G(z(t)) - y), t \geq 0. \tag{4.18}$$

.

PROOF. Define $c = c_0/Q$. Pick $y \in K$ such that $\|y\|_K < rc$ and define $F : H \to K$ by $F(x) = G(x) - y, x \in H$. Then $\|F(0)\|_K = \|y\|_K < rc$. Noting that $F' = G'$ we have by Theorem 14, for z satisfying (4.18),

$$u = \lim_{t\to\infty} z(t) \text{ exist and } F(u) = 0, i.e., G(u) = y.$$

By the argument for Theorem 14 it is clear that $\|u\|_H \leq r$. □

3. Convexity

It has long been recognized that convexity of ϕ is an important consideration in the study of steepest descent. For the next theorem we take ϕ to be, at each point of H, convex in the gradient direction at that point. More specifically there is

THEOREM 16. *Suppose ϕ is a nonnegative $C^{(2)}$ function on H so that there is $\epsilon > 0$ such that*

$$\phi''(x)((\nabla\phi)(x), (\nabla\phi)(x)) \geq \epsilon\|(\nabla\phi)(x)\|_H^2,\ x \in H. \tag{4.19}$$

Suppose also that $x \in H$, $(\nabla\phi)(x) \neq 0$ and (4.2) holds. Then

$$u = \lim_{t\to\infty} z(t) \text{ exists and } (\nabla\phi)(u) = 0.$$

PROOF. Define $g = \phi(z)$ where for $x \in H$, z satisfies (4.2). Note that $g' = \phi'(z)z' = -\|(\nabla\phi)(z)\|_H^2$ and

$$g'' = \langle(\nabla\phi)'(z)z', z'\rangle_H.$$

Note that if each of $h, k, y \in H$ then

$$\phi''(y)(h,k) = \langle(\nabla\phi)'(y)h, k\rangle_H.$$

Using (4.19),

$$g''(t) = 2\langle(\nabla\phi)'(z(t))z'(t), z'(t)\rangle_H \geq \epsilon\|z'(t)\|_H^2 = -\epsilon g'(t),\ t \geq 0,$$

we have that

$$-g''(t)/g'(t) \geq 2\epsilon,\ t \geq 0$$

and hence

$$-\ln(-g'(t)/(-g'(0)) \geq 2\epsilon t,\ t \geq 0.$$

and consequently,

$$0 \leq -g'(t) \leq -g'(0)\exp(-2\epsilon t), t \geq 0 \tag{4.20}$$

and so

$$\lim_{t\to\infty} g'(t) = 0.$$

From (4.20) it follows that if $0 \leq a < b$, then

$$g(a) - g(b) \leq (-g'(0))(\exp(-2\epsilon a) - \exp(-2\epsilon b))/(2\epsilon),$$

i.e.

$$g(a) \leq -g'(0)\exp(-2\epsilon a)/(2\epsilon).$$

But $g'(t) = -\|z'(t)\|_H^2$, $t \geq 0$, and so $-\int_a^b g' = \int_a^b \|z'\|_H^2$, $0 < a < b$. Therefore

$$g(a) - g(b) = \int_a^b \|z'\|_H^2 \text{ and hence}$$

$$\int_a^b \|z'\|_H^2 \leq (-g'(0))(\exp(-2\epsilon a) - \exp(-2\epsilon b))/(2\epsilon).$$

Therefore,

$$-g'(0)(\exp -2\epsilon a)/(2\epsilon) \geq \int_a^{a+1} \|z'\|_H^2 \geq (\int_a^{a+1} \|z'\|_H)^2.$$

Hence, $\int_0^\infty \|z'\|_H$ exists and consequently

$$u = \lim_{t\to\infty} z(t)$$

exists. Since $\lim_{t\to\infty} g'(t) = 0$ and $-\|(\nabla\phi)(z(t))\|_H^2 = g'(t)$, it follows that

$$(\nabla\phi)(u) = \lim_{t\to\infty}(\nabla\phi)(z(t)) = 0.$$

□

The above is close to arguments found in [**94**]. Relationships between convexity and gradient following are extensive. Even without differentiability assumptions on ϕ, a subgradient of ϕ may be defined. This subgradient becomes a monotone operator. The theory of one parameter semigroups of nonlinear contraction mappings on Hilbert space then applies. The interested reader might see [**10**] for such developments. We emphasize here that for many problems of interest in the present context, ϕ is not convex. The preceding theorem is included mainly for comparison.

We note, however, the following connection between convexity conditions and gradient inequalities.

THEOREM 17. *Suppose F is a $C^{(2)}$ function from H to K, $u \in H$, $F(u) = 0$ and $s, d > 0$ so that*

$$\|F'(x)^* F(x)\|_H \geq d\|F(x)\|_K, \ \text{if } \|x-u\|_H \leq s.$$

Then there exist $r, \epsilon > 0$ so that

$$\phi''(x)(\nabla\phi)(x), (\nabla\phi)(x)) \geq \epsilon\|(\nabla\phi)(x)\|_H^2 \ \text{if } \|x-u\|_H < r.$$

LEMMA 3. *Suppose that $T \in L(H,K)$, $y \in K$, $y \neq 0$, $d > 0$ and $\|T^*y\|_H \geq d\|y\|_K$. Then $\|TT^*y\|_K \geq (d^2/|T^*|)\|T^*y\|_H$.*

PROOF. (Lemma 3) First note that

$$\|TT^*y\|_K = \sup_{\|k\|_K=1} \langle TT^*y, k\rangle_K$$
$$\geq \langle TT^*y, (1/\|y\|_K)y\rangle_K = \|T^*y\|_H^2/\|y\|_K \geq d^2\|y\|_K.$$

Hence

$$\|TT^*y\|_K/\|T^*y\|_H = (\|TT^*y\|_K/\|y\|_K)/(\|T^*y\|_H/\|y\|_K) \geq d^2/|T^*|$$

since $\|T^*y\|_H/\|y\|_K \leq |T^*|$. □

PROOF. (Of Theorem 17) Note that if $x, h \in H$ then

$$\phi''(x)(h,h) = \|F'(x)h\|_K^2 + \langle F''(x)(h,h), F(x)\rangle_K.$$

Choose $r_1, M_1, M_2 > 0$ so that if $\|x-u\|_H \leq r_1$, then

$$|F'(x)| \leq M_1 \text{ and } |F''(x)| \leq M_2.$$

Pick $r > 0$ so that $r \leq \min(s, r_1)$ and

$$\|F(x)\|_K \leq \alpha \equiv d^2/(2M_1M_2) \text{ if } \|x-u\|_H \leq r.$$

Then using the lemma and taking x, h so that

$$\|x-u\|_H \leq r \text{ and } \quad h = (\nabla\phi)(x),$$

we have

$$\|F'(x)h\|_K \geq (d^2/M_1)\|(\nabla\phi)(x)\|_H^2$$

and

$$|\langle F''(x)(h,h), F(x)\rangle_K\| \leq M_2\alpha\|(\nabla\phi)(x)\|_H^2$$

so that

$$\phi''(h,h) \geq ((d^2/M_1) - M_2\alpha)\|(\nabla\phi)(x)\|_H^2 = \epsilon\|(\nabla\phi)(x)\|_H^2,$$

where $\epsilon \equiv (d^2/M_1) - M_2\alpha = d^2/(2M_1) > 0$. □

In this next section we give some examples of instances in which a gradient inequality (4.9) is satisfied for some ordinary differential equations. First note that if $F : H \to K$ is a $C^{(1)}$ function, Ω is a bounded subset of H, $c > 0$,

$$\|F'(x)^*k\|_H \geq c\|k\|_K, k \in K, x \in \Omega \tag{4.21}$$

and

$$\phi(x) = \|F(x)\|_K^2/2, x \in K,$$

then it follows that

$$\|(\nabla\phi)(x)\|_H \geq 2^{1/2}c\phi(x)^{1/2}, x \in \Omega,$$

i.e., (4.9) holds.

Another result for which existence of an ω-limit point implies convergence is the following. The chapter on convexity in [**25**] influenced the formulation of this result.

THEOREM 18. *Under the hypothesis of Theorem 7, suppose that $x \in H$ and z satisfies* (4.2). *Suppose also that u is an ω-limit point of z at which ϕ is locally convex. Then*

$$u = lim_{n\to\infty} z(t) \text{ exists}$$

and $(\nabla\phi)(u) = 0$.

PROOF. Suppose that $\{t_i\}_{i=1}^{\infty}$ is an increasing sequence of positive numbers so that

$$u = \lim_{i\to\infty} z(t_i) \tag{4.22}$$

but that it is not true that

$$u = \lim_{t\to\infty} z(t). \tag{4.23}$$

Then z is not constant and it must be that $\phi(z)$ is decreasing. Define α so that

$$\alpha(t) = \|z(t) - u\|_H^2/2, \ t \geq 0.$$

Note that

$$\alpha'(t) = \langle z'(t), z(t) - u\rangle_H = -\langle(\nabla\phi)(z(t)), z(t) - u\rangle_H$$
$$= \langle(\nabla\phi)(z(t)), u - z(t)\rangle_H, \ t \geq 0.$$

If for some $t_0 \geq 0, \alpha'(t) \leq 0$ for all $t \geq t_0$, then (4.23) would follow in light of (4.22). So suppose that for each $t_0 \geq 0$ there is $t > t_0$ so that $\alpha'(t) > 0$. For each positive integer n, denote by s_n the least number so that $s_n \geq t_n$ and so that if $\epsilon > 0$ there is $t \in [s_n, s_n + \epsilon]$ such that $\alpha'(t) > 0$. Note that

$$\phi(z(s_n)) \leq \phi(z(t_n)), \ n = 1, 2, \ldots.$$

. Denote by $\{q_n\}_{n=1}^{\infty}$ a sequence so that $q_n > s_n, \alpha'(q_n) > 0$ and $\|z(s_n) - z(q_n)\| < 1/n, \ n = 1, 2, \ldots$ and note that $\lim_{n\to\infty} z(q_n) = u$ since

$$\|z(s_n) - u\|_H \leq \|z(t_n) - u\|_H, \ \|z(s_n) - z(q_n)\|_H < 1/n, n = 1, 2, \ldots$$

and $lim_{n\to\infty} z(t_n) = u$.

Now if n is a positive integer,

$$0 < \alpha'(q_n) = \langle(\nabla\phi)(z(q_n)), u - z(q_n)\rangle_H = \phi'(z(q_n))(u - z(q_n))$$

and so there is $p_n \in [z(q_n), u]$ so that $\phi(p_n) > \phi(z(q_n))$. But $\phi(z(q_n)) > \phi(u)$ since $\phi(z)$ is decreasing and (4.22) holds. Thus $\phi(z(q_n)) < \phi(p_n) > \phi(u)$. Since p_n is between $z(q_n)$ and u, it follows that ϕ is not convex in the ball with center u and radius $\|z(q_n) - u\|_H$. But $\lim_{n\to\infty} \|z(q_n) - u\|_H = 0$ so ϕ is not locally convex at u, a contradiction. Thus the assumption that (4.23) does not hold is false and so (4.23) holds. Since

$$\int_0^\infty ||(\nabla\phi)(z)||^2 < \infty$$

it follows that $(\nabla\phi)(u) = 0$. □

4. Examples

The next two theorems show that (4.21) holds for a family of ordinary differential operators.

THEOREM 19. *Suppose $H = H^{1,2}([0,1])$ and g is a continuous real-valued function on R. Suppose also that Q is a bounded subset of H. There is $c > 0$ so that*

$$||\pi P\binom{g(y)k}{k}||_H \geq c||k||_K, k \in K = L_2([0,1]),\ y \in Q,$$

where P is the orthogonal projection of $L_2([0,1])^2$ onto

$$\{\binom{u}{u'} : u \in H\}$$

and

$$\pi\binom{f}{g} = f,\ f, g \in K.$$

PROOF. We give an argument by contradiction. Suppose there is $y_1, y_2, ...$ $\in H$, $k_1, k_2, ... \in K$ such that $||k_n||_K = 1$ and

$$||\pi P\binom{g(y_n)k_n}{k_n}||_H < 1/n,\ n = 1, 2, \ldots \tag{4.24}$$

Denote

$$\pi P\binom{g(y_n)k_n}{k_n} \text{ by } \quad u_n,\ n = 1, 2, \ldots$$

For each positive integer n, denote by v_n that element of H so that

$$u_n + v_n' = g(y_n)k_n, u_n' + v_n = k_n, v_n(0) = 0 = v_n(1) \tag{4.25}$$

(such a decomposition is possible since

$$\{\binom{u}{u'} : u \in H\}^\perp = \{\binom{v'}{v} : v \in H, v(0) = 0 = v(1)\}).$$

Then (4.24) may be restated as

$$||u_n||_H < 1/n,\ n = 1, 2, ...$$

which implies that

$$\int_0^1 u_n^2, \int_0^1 {u_n'}^2 \to 0 \text{ as } n \to \infty. \tag{4.26}$$

From (4.25) it follows that if n is a positive integer then

$$\int_0^1 u_n + v_n(t) = \int_0^t g(y_n)k_n$$

and so

$$v_n(t) = (k_n - u_n')(t) = -\int_0^t u_n + \int_0^t g(y_n)u_n' + \int_0^t g(y_n)v_n,\ t \geq 0,$$

i.e.,

$$v_n(t) = \int_0^t h_n + \int_0^t g(y_n)v_n,\ t \geq 0 \tag{4.27}$$

where $h_n = -u_n + g(y_n)u_n'$.

From (4.27) we have, using Gronwall's inequality,

$$|v_n(t)| \leq |\int_0^t h_n| + \int_0^t \exp^{\int_s^t g(y_n)|h_n(s)|ds}, t \in [0,1],\ n = 1, 2, \ldots. \tag{4.28}$$

Using (4.26) and (4.28) together we see that if $t \in [0,1]$

$$\lim_{n\to\infty} v_n(t) = 0 \text{ since } \quad \lim_{n\to\infty} \|h_n\|_K = 0$$

inasmuch as $\{y_n\}_{n=1}^{\infty}$ is uniformly bounded in H. We thus arrive at a contradiction

$$1 = \|k_n\|_K \leq \|u_n'\|_K + \|v_n\|_K \to 0 \text{ as } \quad n \to \infty.$$

□

THEOREM 20. *Suppose G is a $C^{(1)}$ real-valued function on R and Q is a bounded subset of $H = H^{1,2}([0,1])$. Suppose furthermore that $F : H \to K = L_2([0,1])$ is defined by*

$$F(y) = y' + G(y),\ y \in H.$$

There is $c > 0$ so that

$$\|F'(y)^* k\|_H \geq c\|k\|_K,\ y \in Q,\ k \in K.$$

PROOF. For $y \in Q$, $h \in H$ we have that

$$F'(y)h = h' + G'(y)h$$

and hence if $k \in K$,

$$\langle F'(y)h, k\rangle_K = \langle h' + G'(y)h, k\rangle_K = \langle \binom{h}{h'}, \binom{G'(y)k}{k}\rangle_{K\times K}$$

and so

$$F'(y)^* k = \pi P\binom{G'(y)k}{k}.$$

□

The conclusion then follows from the preceding theorem. One can envision higher order generalizations of this theorem.

5. Higher Order Sobolev Spaces for Lower Order Problems

Sometimes it may be interesting to carry out steepest descent in a Sobolev space of higher order than absolutely required for formulation of a problem. What follows is an indication of how that might come about. This work is taken from [**73**]

Suppose m is a positive integer, Ω is a bounded open subset of R^m and ϕ is a $C^{(1)}$ function from $H^{1,2}(\Omega)$ to $[0,\infty)$ which has a locally lipschitzian derivative. For each positive integer k we denote the Sobolev space $H^{k,2}(\Omega)$ by H_k. We assume that Ω satisfies the cone condition (see [**1**] for this term as well as other matters concerning Sobolev spaces) in order to have that H_k is compactly embedded in $C_B^{(1)}(\Omega)$ for $2k > m+2$.

If k is a positive integer then denote by $\nabla_k\phi$ the function on H_1 so that

$$\phi'(y)h = \langle h, (\nabla_k\phi)(y)\rangle_{H_k}, y \in H_1, h \in H_k. \tag{4.29}$$

We may do this since for each $y \in H_1, \phi'(y)$ is a continuous linear functional on H_1 and hence its restriction to H_k is a continuous linear functional on H_k. Each of the functions, $\nabla_k\phi, k = 1, 2, \ldots$ is called a Sobolev gradient of ϕ.

LEMMA 4. *If k is a positive integer then $\nabla_k\phi$ is locally lipschitzian on H_k as a function from H_1 to H_k.*

PROOF. Suppose $w \in H_k$. Denote by each of r and L a positive number so that if $x, y \in H_1$ and $\|x-w\|_{H_1}, \|y-w\|_{H_1} \le r$, then

$$\|(\nabla_1\phi)(x) - (\nabla_1\phi)(y)\|_{H_1} \le L\|x-y\|_{H_1}. \tag{4.30}$$

Now suppose that $x, y \in H_1$ and

$$\|x-w\|_{H_1}, \|y-w\|_{H_1} \le r.$$

Then

$$\begin{aligned}
&\|(\nabla_k\phi)(x) - (\nabla_k\phi)(y)\|_{H_k} \\
&= sup_{h\in H_k,\|h\|_{H_k}=1}\langle(\nabla_k\phi)(x) - (\nabla_k\phi)(y), h\rangle_{H_k} \\
&= sup_{h\in H_k,\|h\|_{H_k}=1}(\phi'(x)h - \phi'(y)h) \\
&= sup_{h\in H_k,\|h\|_{H_k}=1}\langle(\nabla_1\phi)(x) - (\nabla_1\phi)(y), h\rangle_{H_1} \\
&\le sup_{h\in H_k,\|h\|_{H_k}=1}\|h\|_{H_1}\|\nabla_1\phi)(x) - (\nabla_1\phi)(y)\|_{H_1} \\
&\le sup_{h\in H_k,\|h\|_{H_k}=1}\|h\|_{H_k}\|\nabla_1\phi)(x) - (\nabla_1\phi)(y)\|_{H_1} \le L\|x-y\|_{H_1}.
\end{aligned}$$

□

In particular, $\nabla_k\phi$ is continuous as a function from H_1 to H_k.

THEOREM 21. *In addition to the above assumptions about ϕ, suppose that*

$$\phi'(y)y \ge 0,\ y \in H_1. \tag{4.31}$$

If k is a positive integer, $x \in H_k$ and

$$z(0) = x, z'(t) = -(\nabla_k\phi)(z(t)),\ t \ge 0, \tag{4.32}$$

then $R(z)$, the range of z, is a subset H_k and is bounded in H_k.

PROOF. First note that one has existence and uniqueness for (4.32) since the restriction of $\nabla_k\phi$ is locally lipschitzian as a function from H_k to H_k and ϕ is bounded below (see [**71**]). Since $x \in H_k$, for z as in (4.32), it must be that $R(z) \subset H_k$ and

$$(\|z\|^2_{H_k}/2)'(t) = \langle z'(t), z(t)\rangle_{H_k}$$
$$= -\langle(\nabla_k\phi)(z(t)), z(t)\rangle_{H_k} = -\phi'(z(t))z(t) \le 0, t \ge 0.$$

Thus $\|z\|^2_{H_k}$ is nonincreasing and so $R(z)$ is bounded in H_k. ☐

Assume for the remainder of this section that $2k > m+2$. Observe that for z as in Theorem 1, $R(z)$ is precompact in $C^{(1)}_B$ and hence also in H_1. For $x \in H_k$ and z satisfying (4.32) denote by Q_x the H_1 $\omega-$ limit set of z, i.e.,

$$Q_x = \{y \in H_1 : y = H_1 - \lim_{n\to\infty} z(t_n), \{t_n\}_{n=1}^{\infty}\text{increasing, unbounded}\}.$$

THEOREM 22. *If $x \in H_k$ and $y \in Q_x$, then $(\nabla_k\phi)(y) = 0$.*

PROOF. Note that

$$(\phi(z))'(t) = \phi'(z(t))z'(t) = -\|(\nabla_k\phi)(z(t)\|^2_{H_k}$$

and so

$$\phi(z(0)) - \phi(z(t)) = \int_0^t \|(\nabla_k\phi)(z(t)\|^2_{H_k}$$

and hence

$$\int_0^\infty \|(\nabla_k\phi)(z(t))\|^2_{H_k} < \infty. \tag{4.33}$$

Thus if

$$u = H_1 - \lim_{t\to\infty} z(t)$$

exists, then by (4.33), $(\nabla_k\phi)(u) = 0$ since $\nabla_k\phi$ is continuous as a function from H_1 to H_k and z is continuous as a function from $[0,\infty)$ to H_1. Thus the conclusion holds in this case.

Suppose now that

$$H_1 - \lim_{t\to\infty} z(t) \text{ does not exist}$$

and that $y \in Q_x$ but $(\nabla_k\phi)(y) \ne 0$. Then $(\nabla_1\phi)(y) \ne 0$ also. Denote by each of α, M a positive number so that $\|(\nabla_k\phi)(x)\|_{H_1} \ge M$ if $\|x - w\|_{H_1} \le \alpha$. Then there are

$$0 < t_1 < s_1 < t_2 < s_2 < \cdots$$

such that $lim_{i\to\infty}t_i = \infty$ and

$$\|z(t_i) - y\|_{H_1} < \alpha,\ \|z(s_i) - y\|_{H_1} > \alpha,\ i = 1, 2, \ldots.$$

Thus there are positive unbounded increasing sequences $\|a_i\|_{i=1}^{\infty}, \|b_i\|_{i=1}^{\infty}$ so that $a_i < b_i < a_{i+1}, i = 1, 2, \ldots$ and so that if $\epsilon > 0$ then there are $t \in [a_i - \epsilon, a_i), s \in$

$(b_i, b_i + \epsilon]$ so that $\|z(t) - y\|_{H_1} > \alpha, \|z(s) - y\|_{H_1} > \alpha$. Since (4.33) holds it follows that $\int_0^\infty \|(\nabla_k \phi)(z(t)\|_{H_1}^2 < \infty$. Hence

$$\alpha \leq \|z(b_i) - z(a_i)\|_{H_1}^2 = \|\int_{a_i}^{b_i} z'\|_{H_1}^2$$
$$\leq (\int_{a_i}^{b_i} \|z'\|_{H_1})^2 \leq (b_i - a_i)(\int_{a_i}^{b_i} \|z'\|_{H_1}^2) \leq (b_i - a_i)(\int_{a_i}^{b_i} \|z'\|_{H_k}^2).$$

Hence $b_i - a_i \geq \alpha/(\int_{a_i}^{b_i} \|z'\|_{H_k}^2)$ and so $lim_{i \to \infty}(b_i - a_i) = \infty$ since $\int_0^\infty \|z'\|_{H_k}^2 < \infty$. But

$$\|(\nabla_k \phi)(z(t))\|_{H_1} \geq M, \; t \in [a_i, b_i], \; i = 1, 2, \ldots$$

and thus

$$\infty > \int_0^\infty \|z'\|_{H_1}^2 \geq \sum_{i=1}^\infty \int_{a_i}^{b_i} \|z'\|_{H_1}^2 = \infty,$$

a contradiction. Thus $(\nabla_k \phi)(y) = 0$. □

Possible applications are to problems in which either a zero or a critical point of ϕ is sought. In **[71]** there is a family of problems of the form

$$\phi(x) = \|F(x)\|_{H_1}^2, x \in H_1,$$

where F is a function from H_1 to $L_2(\Omega)$ which has a locally lipschitzian derivative. In these problems a system of partial differential equations is represented by the problem of finding $u \in H_1$ so that

$$F(u) = 0.$$

In **[71]** conditions are given under which critical points of ϕ are also zeros of ϕ.

Condition (4.31) is too strong to apply to most systems of partial differential equations. We note however, that this condition does not imply convexity. We hope that Theorem 21 will lead to results in which (4.31) is weakened while still allowing our conclusions.

CHAPTER 5

Orthogonal Projections, Adjoints and Laplacians

Adjoints as in (4.11) have a major place in construction of Sobolev gradients, both in function space and in finite dimensional settings. In this chapter we develop some background useful in understanding these adjonts.

Suppose H, K are Hilbert spaces and $T \in L(H, K)$, the space of all bounded linear transformations from H to K. It is customary to denote by T^* the member of $L(K, H)$ so that

$$\langle Tx, y\rangle_K = \langle x, T^*y\rangle_H, \ x \in H, \ y \in K. \tag{5.1}$$

For applications to differential equations, we will often take H to be a Sobolev space (which is also a Hilbert space) and K to be an L_2 space. In order to illustrate how gradient calculations depend upon adjoint calculations, we deal first with the simplest setting for a Sobolev space. Our general reference for linear transformations on Hilbert spaces is [**91**].

In [**102**], Weyl deals with the problem of determining when a vector field is the gradient of some function. He introduces a method of orthogonal projections to solve this problem for all square integrable (but not necessarily differentiable) vector fields. Our construction of Sobolev gradients is related to Weyl's work.

1. A Construction of a Sobolev Space.

We rely heavily on [**1**] for references to Sobolev spaces. In this section we give, however, a construction of the simplest Sobolev space. Take $K = L_2([0,1])$ and define $H = H^{1,2}([0,1])$ to be the set of all first terms of members of Q where

$$Q = \{\binom{u}{u'} : \ u \in C^1([0,1])\} \tag{5.2}$$

and the closure $cl(Q)$ is taken in $K \times K$. The following is a crucial fact:

Lemma 5. *$cl(Q)$ is a function in the sense that no two members of $cl(Q)$ have the same first term.*

Proof. Suppose that $\binom{f}{g}, \binom{f}{h} \in cl(Q)$ and $k = g - h$. Then $\binom{0}{k} \in cl(Q)$. Denote by $\{\binom{f_n}{f_n'}\}_{n=1}^{\infty}$ a sequence in Q which converges to $\binom{0}{k}$. If $m, n \in Z^+$, denote by $c_{m,n}$ a member of $[0,1]$ so that $|(f_m - f_n)(c_{m,n})| \le |(f_m - f_n)(t)|$, $t \in [0,1]$. Then if $t \in [0,1]$,

$$f_m(t) - f_n(t) = f_m(c_{m,n}) - f_n(c_{m,n}) + \int_{c_{m,n}}^{t} (f_m' - f_n')$$

and so

$$|(f_m(t) - f_n(t)| \le \|f_m - f_n\|_K + |\int_{c_{m,n}}^t (f'_m - f'_n)|$$

$$\le \|f_m - f_n\|_K + (\int_{c_{m,n}}^t (f'_m - f'_n)^2)^{1/2}$$

$$\le \|f_m - f_n\|_K + (\int_0^1 (f'_m - f'_n)^2)^{1/2} \to 0 \text{ as } m, n \to 0.$$

Hence $\{f_n\}_{n=1}^\infty$ converges uniformly to 0 on [0,1] since it already converges to 0 in K. Note that if $t, c \in [0, 1]$,

$$(\int_c^t f'_n - \int_c^t k)^2 \le (\int_c^t |f'_n - k|)^2 \le \int_c^t (f'_n - k)^2 \le \|f'_n - k\|_K^2 \to 0 \text{ as } n \to \infty,$$

and so

$$\lim_{n\to\infty} \int_c^t f'_n = \int_c^t k.$$

Therefore since

$$f_n(t) - f_n(c) = \int_c^t f'_n,$$

it follows that

$$0 = \int_c^t k, \ c, t \ \in [0, 1].$$

But this implies that $k = 0$ and hence $g = h$. □

If $\binom{f}{g} \in Q$, then we take by definition $f' = g$ and also define

$$\|f\|_H = (\|f\|_K^2 + \|f'\|_K^2)^{1/2}. \tag{5.3}$$

If $f \in C^1$ then this definition is consistent with the usual definition. In fact, if $g \in K$, $c \in R$ and $f(t) = c + \int_0^t g$, $t \in [0, 1]$, then $f \in H$ and $\binom{f}{g} \in Q$ and so in the above sense, $f' = g$. Moreover, every member of H arises in this way.

To illustrate our point of view on adjoints of linear differential operators, we consider the member T of $L(H, K)$ defined simply by

$$Tf = f', \ f \in H. \tag{5.4}$$

In regard to T, we have the following:

Problem 5.2. Find a construction for T^* as a member of $L(H, K)$.

Solution. We identify a subset of $Q^\perp$ as

$$M = \{\binom{v'}{v} \mid v \in C^1([0, 1]), \ v(0) = 0 = v(1)\}. \tag{5.5}$$

It is an elementary problem in ordinary differential equations to deduce that if $f, g \in C([0, 1])$, then there are uniquely $\binom{u}{u'} \in Q$, $\binom{v'}{v} \in M$ so that

$$\binom{u}{u'} + \binom{v'}{v} = \binom{f}{g}. \tag{5.6}$$

Explicitly with $C(t) = \cosh(t), S(t) = \sinh(t),\ t \in R$,

$$u(t) = [C(1-t)\int_0^t (C(s)f(s) + S(s)g(s))\ ds \tag{5.7}$$

$$+C(t)\int_t^1 (C(1-s)f(s) - S(1-s)g(s))\ ds]/S(1),\ t \in [0,1],$$

$$v(t) = [S(1-t)\int_0^t (C(s)f(s) + S(s)g(s))\ ds$$

$$-S(t)\int_t^1 (C(1-s)f(s) - S(1-s)g(s))\ ds]/S(1), t \in [0,1].$$

Hence we see that M and Q are mutually orthogonal and their direct sum is dense in $K \times K$. Therefore $K \times K$ is the direct sum of Q and M. Since (5.7) may be extended by continuity to any $\binom{f}{g} \in K \times K$, it follows that

$$P\tbinom{f}{g} = \tbinom{u}{u'} \text{ where } u \text{ is given by (5.7).} \tag{5.8}$$

Hence we have an explicit expression for P.

We now finish our solution of Problem 5.2. Suppose that $f \in H$ and $g \in K$. Then

$$\langle Tf, g\rangle_K = \langle f', g\rangle_K = \langle \tbinom{f}{f'}, \tbinom{0}{g}\rangle_{K\times K} = \langle \tbinom{f}{f'}, P\tbinom{0}{g}\rangle_{K\times K} = \langle f, \pi P\tbinom{0}{g}\rangle_H$$

where $\pi\binom{r}{s} = r$ if $r, s \in K$. Hence $T^*g = \pi P\binom{0}{g}$. Since P is explicitly given, it follows that we have found an explicit expression for T^*.

2. A Formula of Von Neumann

Adjoints as just calculated have a close relation with the adjoints of unbounded closed linear transformations. Recall that from **[91]**,**[100]**, for example, if W is a closed linear transformation on H to K (i.e., $\{\binom{x}{Wx} : x \in D(W)\}$ is a closed subset of $H \times K$), then an adjoint W^* of W is defined by:

(*i*) $D(W^t) = \{y \in K \mid \exists\ z \in H \text{ such that } \langle Wx, y\rangle_K = \langle x, z\rangle_H,\ x \in D(W)\}$,

and

$$(ii)\ \ W^t y = z,\ y, z \text{ as in } (i).$$

From this definition it follows that if $x \in D(W)$, $y \in D(W^t)$, then

$$\langle \tbinom{x}{Wx}, \tbinom{-W^t y}{y}\rangle_{H\times K} = \langle Wx, y\rangle_K - \langle x, W^t y\rangle_H = 0. \tag{5.9}$$

Furthermore, it is an easy consequence of the definition that

$$\{\tbinom{x}{Wx} \mid x \in D(W)\}^\perp = \{\tbinom{-W^t y}{y} \mid y \in D(W^t)\}$$

and consequently that if $\binom{r}{s} \in H \times K$, then there exists uniquely $x \in D(W)$, $y \in D(W^t)$ such that

$$\tbinom{x}{Wx} + \tbinom{-W^t y}{y} = \tbinom{r}{s}. \tag{5.10}$$

From [**100**] we have the following:

THEOREM 23. *Suppose W is a closed, densely defined linear transformation on the Hilbert space H to the Hilbert space K. Then the orthogonal projection of $H \times K$ onto*

$$\{\binom{x}{Wx} \ : \ x \in D(W)\} \tag{5.11}$$

is given by the 2×2 maxtix v so that:

$$v_{11} = (I + W^tW)^{-1}, \ v_{12} = W^t(I + WW^t)^{-1} \tag{5.12}$$

$$v_{21} = W(I + W^tW)^{-1}, \ v_{22} = I - (I + WW^t)^{-1}$$

PROOF. Note first that if r, s are as in (5.10) and $s = 0$, then $x - W^ty = r$, $Wx+y = 0$ and consequently $(I+W^tW)y = r$. Hence the range of $(I+W^tW) = H$. Since if $x \in D(W^tW)$,

$$\langle (I + W^tW)x, x\rangle_H \geq \langle x, x\rangle_H$$

it follows that $(I + W^tW)^{-1} \in L(H, H)$, and $|(I + W^tW)^{-1}| \leq 1$. Similar properties hold for $(I + WW^t)^{-1}$. It is easily checked that the matrix indicated (5.12) is idempotent, symmetric, fixed on the set (5.11) and has range that set (using that $W(I + W^tW)^{-1}x = (I + WW^t)^{-1}Wx$, $x \in D(W)$; $W^t(I + WW^t)^{-1}y = (I + W^tW)^{-1}y$, $y \in D(W^t)$). Hence the matrix in (5.12) is the orthogonal projection onto the set in (5.11). □

3. Relationship between Adjoints

To see a relationship between the adjoints W^t, W^* of the above two sections, take W to be the closed densely defined linear transformation on K defined by $Wf = f'$ for exactly those members of K which are also members of $H^{1,2}([0, 1])$. Then the projection P in Section 1 is just the orthogonal projection onto the set in (5.11) which in this case is the same as clQ. See [**28**] for an additional description of adjoints of linear differential operators when they are considered as densely defined closed operators.

This relationship may be summarized by the following:

THEOREM 24. *Suppose that each of H and K is a Hilbert space, W is a closed densely defined linear transformation of H to K. Suppose in addition that J is the Hilbert space whose points consist of $D(W)$ with*

$$\|x\|_J = (\|x\|_H^2 + \|Wx\|_K^2)^{1/2}, \ x \in J. \tag{5.13}$$

Then the adjoint W^* of W **(with W regarded as a member of** *$L(J, K)$) is given by*

$$W^*y = \pi P\binom{0}{y}, \ y \in J,$$

where P is the orthogonal projection of $H \times K$ onto $\{\binom{x}{Wx} \ : \ x \in J\}$ and $\pi\binom{r}{s} = r$, $\binom{r}{s} \in H \times K$.

PROOF. If $x \in D(W)$,

$$\langle Wx, y\rangle_K = \langle \binom{x}{Wx}, \binom{0}{y}\rangle_{H\times K} = \langle \binom{x}{Wx}, P\binom{0}{y}\rangle_{H\times K} =$$

$$\langle x, \pi P\binom{0}{y}\rangle_J, \text{ so that } W^* y = \pi P\binom{0}{y}.$$

□

We emphasize that W has two separate adjoints: one regarding W as a closed densely defined linear transformation on H to H and the other regarding W as a bounded linear transformation on $D(W)$ where the norm on $D(W)$ is the graph norm (5.13). One sometimes might want to use the separate symbols W^t, W^* for these two distinct (but related objects). At any rate, it is an obligation of the writer to make clear in a given context which adjoint is being discussed.

4. General Laplacians

Suppose that H is a Hilbert space and H' is a dense linear subspace of H which is also a Hilbert space under the norm $\|\ \|_{H'}$ in such a way that

$$\|x\|_H \le \|x\|_{H'},\ x \in H'.$$

Following Beurling-Deny [**8**],[**9**], for the pair (H, H') there is an associated transformation called the laplacian for (H, H'). It is described as follows. Pick $y \in H$ and denote by f the functional on H corresponding to y :

$$f(x) = \langle x, y\rangle_H,\ x \in H.$$

Denote by k the restriction of f to H'. Then

$$|k(x)| = |\langle x, y\rangle_H| \le \|x\|_H\|y\|_H \le \|x\|_{H'}\|y\|_H,\ x \in H'.$$

Hence k is a continuous linear functional on H' and so there is a unique member z of H' such that

$$k(x) = \langle x, y\rangle_{H'},\ x \in H'.$$

Define $M : H \to H'$ by $My = z$ where y, z are as above. We will see that M^{-1} exists; it will be called the laplacian for the pair (H, H'). If H'_0 is a closed subspace of H' whose points are also dense in H, we denote the corresponding transformation by M_0.

THEOREM 25. *The following hold for M as defined above:*
(a) $R(M)$, the range of M is dense in H,
(b) $|M|_{L(H,H')} \le 1$,
(c) M^{-1} exists,
(d) M, considered as a transformation from $H \to H$, is symmetric.

PROOF. Suppose first that there is $z \in H'$ so that $\langle z, Mx\rangle_{H'} = 0, x \in H$. Then $0 = \langle z, Mz\rangle_{H'} = \langle z, z\rangle_H$ and so $z = 0$, a contradiction. Thus $cl_{H'}R(M) = H'$. But then $H' = cl_{H'}(R(M)) \subset cl_H(R(M))$. Hence $H = cl_H(H') = cl_H(R(M))$ and (a) is demonstrated.

To show that (b) holds, suppose that $x \in H$. Then

$$\|Mx\|_{H'} = sup_{z\in H', z\neq 0}\langle z, Mx\rangle_{H'}/\|z\|_{H'} =$$

$$sup_{z\in H', z\neq 0}\langle z, x\rangle_H/\|z\|_{H'} \leq sup_{z\in H, z\neq 0}\langle z, x\rangle_H/\|z\|_H,$$

and so $|M|_{L(H,H')} \leq 1$.

To show that (c) holds, suppose that $x \in H$ and $Mx = 0$. Then

$$0 = \langle z, Mx\rangle_{H'} = \langle z, x\rangle_H, z \in H'.$$

But this implies that $z = 0$ since the points of H' are dense in H.

To see that (d) holds observe that if $x, z \in H$, then

$$\langle z, Mx\rangle_H = \langle Mz, Mx\rangle_{H'} = \langle Mz, x\rangle_H.$$

□

5. A Generalized Lax-Milgram Theorem

In this section we present a slight extension of the Lax-Milgram [**43**] Theorem. The work of this section is taken from [**74**]. Denote by each of H, H_0', H' a Hilbert space with

$$H_0' \subset H' \subset H$$

so that

$$\|x\|_{H_0'} = \|x\|_{H'}, x \in H_0'$$

and

$$\|x\|_H \leq \|x\|_{H'}, x \in H'.$$

Suppose also that the points of H_0', H' are dense in H. Define P_0 to be the orthogonal projection of H' onto H_0' and denote the complementary projection $I - P_0$ by Q_0. For example take

$$H = L_2([0,1]), H' = H^{1,2}([0,1])$$

and

$$H_0' = \{f \in H' : f(0) = 0 = f(1)\}.$$

Returning to the general case of H, H', H_0', define

$$\beta(u) = (1/2)\|u\|_{H'}^2 - \langle u, g\rangle_H,\ u \in H'. \tag{5.14}$$

THEOREM 26. *Suppose $g \in H, w \in H'$ and $\beta : H' \to R$ is defined by (5.14). Then the minimum of $\beta(u)$ subject to the condition $Q_0u = Q_0w$ is achieved by*

$$u = Q_0w + M_0g. \tag{5.15}$$

The condition $Q_0u = Q_0w$ may be regarded as a generalized boundary condition; it is equivalent to asking the $u - w \in H_0'$.

PROOF. Define $q = Q_0w$ and define $\gamma : H_0' \to R$ by

$$\gamma(y) = \beta(y + q),\ y \in H_0'.$$

Note that

$$\gamma'(y)k = \beta'(y+q)k = \langle y + q, k\rangle_{H'} - \langle k, g\rangle_H,\ k \in H_0'.$$

Note also that since

$$\gamma''(y)(k,k) = \|k\|_H^2,\ k, y \in H_0',$$

it follows that γ is (strictly) convex. Now

$$\beta(u) = (1/2)\|u\|_{H'}^2 - \langle u, g\rangle_H \geq (1/2)\|u\|_{H'}^2 - \|u\|_H\|g\|_H \geq$$
$$(1/2)\|u\|_{H'}^2 - \|u\|_{H'}\|g\|_H = \|u\|_{H'}(\|u\|_{H'}/2 - \|g\|_H),\ u \in H'$$

so β and hence γ is bounded from below.

Since γ is convex and bounded from below it has an absolute minimum if and only if it has a critical point. Moreover, such a critical point would be the unique point at which β attains its minimum. Observe that

$$\gamma'(y)k = \langle y + q, k\rangle_{H'} - \langle k, g\rangle_H = \langle y, k\rangle_{H'} - \langle k, g\rangle_H$$

since $\langle q, k\rangle_{H'} = 0,\ k \in H_0'$. Now

$$\langle k, g\rangle_H = \langle y, k\rangle_{H'},\ k \in H_0'$$

if and only if $y = M_0 g$. Choosing y in this way thus yields a critical point of γ. Consequently $u = y + q$ is the point of H' at which β attains its minimum. Therefore

$$u = Q_0 w + M_0 g$$

is the point at which β attains its minimum and the theorem is proved. □

6. Laplacians and Closed Linear Transformations

We now turn to a somewhat more concrete case of the above - a case which is closer to the example.

Suppose that each of H and K is a Hilbert space and T is a closed and densely defined linear transformation on H to K. Let H' be the Hilbert space whose points are those of $D(T)$ where

$$\|x\|_{H'} = \|\binom{x}{Tx}\|_{H\times K},\ x \in D(T). \tag{5.16}$$

Suppose that the linear transformation T_0 is a closed, densely defined restriction of T (see [**91**] for a discussion of closed unbounded linear operators from one Hilbert space to another). Denote by H_0' the Hilbert space whose points are those of $D(T_0)$ where

$$\|x\|_{H_0'} = \|\binom{x}{T_0x}\|_{H\times K},\ x \in D(T_0). \tag{5.17}$$

Then H, H', H_0' fit the hypothesis of Theorem 26.

We remind the reader of an equivalent definition of T^t. The domain of T^t is

$$\{y \in H : x \to \langle Tx, y\rangle_K \text{ is continuous }\}.$$

For $y \in D(T^t), T^t y$ is the element of H such that

$$\langle Tx, y\rangle_K = \langle x, T^t y\rangle_H,\ x \in D(T).$$

The definition of adjoint applies just as well when T is replaced by T_0.

One can choose T to be a differential operator in such a way that the resulting space H' is one of the classical Sobolev spaces which is also a Hilbert space. Then the restriction T_0 of T can be chosen so that H_0' is a subspace of H' consisting of those members of H' which satisfy zero boundary conditions in some sense (much more variety than this can be accommodated). In the example, T is the derivative operator whose domain consists of the elements of $H^{1,2}([0,1])$. In other cases T might be a gradient operator.

THEOREM 27. *Suppose $g \in H, w \in H'$ and β satisfies* (5.14).

$$\beta(u) = (1/2)\|u\|_{H'}^2 - \langle g, u\rangle_H, \ u \in H'.$$

Then the element of H' which renders β minimum is the unique solution u to

$$(I + T_0{}^tT)u = g, \ Q_0u = Q_0w$$

where Q_0 is as in Theorem 26 in its relationship with H', H_0'.

In the example, $(I + T_0{}^tT)$ is the differential operator so that

$$(I + T_0{}^tT)u = u - u''$$

for all u in its domain (without any boundary conditions on its domain - that is it is the maximal operator associated with its expression).

PROOF. From Theorem 26, the minimum u of β, subject to $Q_0u = Q_0w$, may be written

$$u = Q_0w + M_0g.$$

It is clear that for u defined in this way, $Q_0u = Q_0w$ since $R(M_0) \subset R(P_0)$ and $Q_0 = I - P_0$. It remains to show that

$$(I + T_0^tT)u = g.$$

We first show that

$$(I + T_0^tT)Q_0w = 0.$$

To this end, first note that

$$\langle Q_0w, x\rangle_{H_0'} = 0, \ x \in H_0'$$

since $x = P_0x, x \in H_0'$. This may be rewritten

$$\langle \binom{Q_0w}{TQ_0w}, \binom{x}{T_0x}\rangle_{H\times K} = 0, \ x \in D(T_0).$$

But this is equivalent to

$$\langle T_0x, TQ_0w\rangle_K = \langle x, -Q_0w\rangle_H, \ x \in D(T_0)$$

and hence

$$TQ_0w \in D(T_0^t)$$

and

$$T_0^tTQ_0w = -Q_0w$$

that is,

$$(I + T_0^tT)Q_0w = 0.$$

Next we show that

$$(I + T_0^tT)M_0g = g.$$

To do this first note that $M_0 g \in D(T_0)$ since $M_0 g \in H_0'$ and so $TM_0 g = T_0 M_0 g$. Using the definition of M_0,

$$\langle x, g\rangle_H = \langle x, M_0 g\rangle_{H_0'},$$

and so

$$\langle x, g\rangle_H = \langle x, M_0 g\rangle_H + \langle T_0 x, T_0 M_0 g\rangle_K,$$

that is

$$\langle T_0 x, T_0 M_0 g\rangle_K = \langle x, g - M_0 g\rangle_H, \ x \in D(T_0).$$

But this implies that

$$T_0 M_0 g \in D(T_0^t)$$

and

$$T_0^t T_0 M_0 g = g - M_0 g,$$

that is

$$(I + T_0^t T_0) M_0 g = g$$

and the argument is complete. □

The expression

$$(I + T_0^t T_0) \tag{5.18}$$

is the inverse of M_0 and is called the laplacian associated with the pair (H, H_0'). Similarly the expression

$$(I + T^t T) \tag{5.19}$$

is the laplacian associated with the pair (H, H'). The expression

$$(I + T_0^t T) \tag{5.20}$$

plays the role of maximal operator associated with the triple H, H', H_0'. Theorem 27 gives that $R(I+T_0^t T) = H$. One may observe that $N(I+T_0^t T)$ is the orthogonal complement of H_0' in H'.

7. Remarks on Higher Order Sobolev Spaces

Work of this section is largely taken from [**73**]. We may use results in this chapter to specify adjoints related to more general Sobolev spaces $H^{m,2}(\Omega)$ where Ω is an open subset of R^n, $n, m \in Z^+$. For $j = 1, 2, ..., m$ denote by $S(j, n)$ the vector space of all j-linear symmetric functions on R^n. Take $H = L_2(\Omega)$ and

$$K = L_2(\Omega, S(1, n)) \times \cdots \times L_2(\Omega, S(m, n)) \tag{5.21}$$

where $L_2(\Omega, S(j, n))$ denotes the space of square integrable functions from Ω to $S(j, n)$, $j = 1, ..., m$. More precisely, $e_1, ..., e_n$ denotes any orthonormal basic for R^n and $v \in S(j, n)$, then

$$\|v\|_{S(j,n)} = (\sum_{p_1=1}^{n} \cdots \sum_{p_j=1}^{n} v(e_{p_1}, ..., v_{p_j})^2)^{1/2}. \tag{5.22}$$

As noted in (Weyl,[**103**], p 139),[**52**], this norm is independent of particular orthonormal basis. For $z \in L_2(\Omega, S(j,n))$, define

$$\|z\|_{L_2(\Omega,S(j,n))} = (\int_\Omega \|z\|^2_{S(j,n)})^{1/2}. \quad (5.23)$$

Thus the norm on K is the Cartesian product norm on $L_2(\Omega, S(j,n))$, $j = 1, ..., m$. See also [**78**] in regard to the above construction.

For $u \in C^m(\Omega)$, denote by Du the m-tuple $(u', u^{(2)}, ..., u^{(m)})$ consisting of the first m Fréchet derivatives of u and take

$$Q = \{(\begin{smallmatrix} u \\ Du \end{smallmatrix}) | u \in C^m(\Omega)\}.$$

From [**1**], the closure of Q in $H \times K$ is a function W, i.e., no two members of $cl(Q)$ have the same first term. The space $H^{m,2}(\Omega)$ is defined as the set of all first terms of W, with, for $u \in H^{1,2}(\Omega)$,

$$\|u\|_{H^{1,2}}(\Omega) = (\|u\|^2_{L_2(\Omega)} + \|Wu\|^2_K)^{1/2}. \quad (5.24)$$

The orthogonal projection of $H \times K$ onto W will be helpful in later chapters for construction of various Sobolev gradients. Generally, the calculation of such a projection involves the solution of n constant coefficient elliptic equation of order 2m on Ω. L. We will be particularly interested in the numerical solution of such problems.

CHAPTER 6

Introducing Boundary Conditions

In this chapter we give a some simple examples which illustrate how one may deal with boundary conditions in the context of descent processes. We deal here with a Hilbert space H; more general spaces are dealt with in later chapters.

1. Projected Gradients

Many systems of partial differential equations may be expressed as the problem of finding $u \in H$ such that

$$\phi(u) = 0 \tag{6.1}$$

(for example,

$$\phi(u) = \int_0^1 (u' - u)^2/2, \ u \in H^{1,2}([0,1])). \tag{6.2}$$

Boundary conditions may be expressed by means of a function B from H to a second Hilbert space V :

$$B(u) = 0 \tag{6.3}$$

(for example

$$B(u) = u(0) - 1, \ u \in H^{1,2}([0,1])) \tag{6.4}$$

with $V = R$ in this case. Thus if u satisfies (6.1) and (6.3) where ϕ and B are given by (6.2) and (6.4) respectively, then we have

$$u' - u = 0, \ u(0) = 1. \tag{6.5}$$

In general we seek a gradient that takes both ϕ and B into account. Specifically we seek $\nabla^B \phi$ so that if $x \in H$

$$z(0) = x, \ z'(T) = -(\nabla^B \phi)(z(t)), \ t \geq 0, \tag{6.6}$$

then

$$B(z(t)) = B(x), \ t \geq 0, \tag{6.7}$$

i.e., B provides an invariant for solutions z to (6.6). In particular, if $x \in H$, $B(x) = 0$ and (6.7) holds, then $B(u) = 0$, that is, u satisfies the required boundary conditions if $u = lim_{t \to \infty} z(t)$.

How does one find such a gradient? For B differentiable, the problem is partially solved by means of the function P_B so that if $x \in H$, then $P_B(x)$ is the orthogonal projection onto $N(B'(x))$. Then one defines

$$(\nabla^B \phi)(x) = P_B(x)(\nabla\phi)(x), \; x \in H. \tag{6.8}$$

Hence if z satisfies (6.6) and $t \geq 0$,

$$(B(z))'(t) = B'(z(t))z'(t) = -B'(z(t))P_B(z(t))(\nabla\phi)(z(t)) = 0$$

since $P_B(z(t))$ is the orthogonal projection of H onto $N(B'(z(t))$.

So, the main task in constructing gradients which meet our requirement is the task of constructing families of orthogonal projections $P_B(x)$, $x \in H$.

Before starting with an example, we indicate how one might arrive at (6.8). For $y \in H$, define the real-valued function α_y to have domain $N(B'(y))$ so that

$$\alpha_y(v) = \phi(y+v), \; v \in N(B'(y)).$$

Note α_y has a gradient as in Chapter 4; if $v \in N(B'(y))$, then $\nabla\alpha_y(v)$ is the element of $N(B'(y))$ so that

$$\alpha_y'(v)h = \langle h, \nabla\alpha_y(v)\rangle_H, \; h \in N(B'(y)).$$

But note that

$$\alpha_y'(v)h = \phi'(y+v)h, \; h \in N(B'(y))$$

In particular

$$\alpha_y'(0)h = \phi'(y)h, \; h \in N(B'(y))$$

and

$$\phi'(y)h = \langle h, (\nabla\phi)(y)\rangle_H = \langle h, P_B(y)(\nabla\phi)(y)\rangle_H, \; h \in N(B'(y)).$$

Hence

$$(\nabla\alpha_y)(0) = P_B(y)(\nabla\phi)(y)$$

and so

$$(\nabla^B\phi)(y) = (\nabla\alpha_y)(0).$$

We now treat in some detail how such a gradient is calculated in a simple example. A number of essential features of Sobolev gradient construction are reflected in the following calculations.

Example 6.1. Take $H = H^{1,2}([0,1])$ and $K = L_2([0,1])$. Define ϕ and B so that

$$\phi(y) = \int_0^1 (y' - y)^2/2, \; B(y) = y(0) - 1, \; y \in H. \tag{6.9}$$

Then if $y \in H$ is a solution to

$$\phi(y) = 0, \; B(y) = 0,$$

it also satisfies (6.5).

We now get an explicit expression for $\nabla^B\phi$. In treating this example in some detail, we illustrate a number of general aspects of our theory. For ease in writing define the functions C and S by

$$C(t) = \cosh(t),\ S(t) = \sinh(t),\ j(t) = t,\ t \in R.$$

THEOREM 28. *With ϕ, B defined in* (6.9),

$$((\nabla^B\phi)(y))(t) = y(t) - [y(1)S(t) + y(0)C(1-t)]/C(1),\ t \in [0,1].$$

We will break down this calculation in a series of lemmas. The first one calculates $(\nabla\phi)(y),\ y \in H$.

LEMMA 6. *If $y \in H$, then*

$$((\nabla\phi)(y))(t) = y(t) - [y(1)C(t) - y(0)C(1-t)]/S(1),\ t \in [0,1].$$

PROOF. Note that

$$\phi'(y)h = \int_0^1 (h'-h)(y'-y), = \langle \tbinom{h}{h'}, \tbinom{y-y'}{y'-y} \rangle = \langle \tbinom{h}{h'}, P\tbinom{y-y'}{y'-y} \rangle,\ h, y \in H,$$

where P is the orthogonal projection defined in (5.8). Since $P\binom{y}{y'} = \binom{y}{y'}$ we have that

$$(\nabla\phi)(y) = y - u$$

where $\binom{u}{u'} = P\binom{y'}{y}$.

Using (5.7),

$$\begin{aligned} u(t) &= (C(1-t)\int_0^t (Cy' + Sy) + C(t)\int_t^1 (C(1-j)y' - S(1-j)y))/S(1) \\ &= (C(1-t)(C(t)y(t) - C(0)y(0))) + C(t)(C(0)y(1) - C(1-t)y(t)))/S(1) \\ &= (C(t)y(1) - C(1-t)y(0))/S(1),\ t \in [0,1], \end{aligned}$$

since $(Cy)' = Cy' + Sy$ and $(C(1-j)y)' = C(1-j)y' - S(1-j)y$. With this expression for u, (6.8) yields the concrete expression:

$$((\nabla\phi)(y))(t) = y(t) - (C(t)y(1) - C(1-t)y(0))/S(1), t \in [0,1].$$

□

Recall now that

$$(\nabla^B\phi)(y) = P_B(y)(\nabla\phi)(y),\ y \in H.$$

We next seek an explicit expression for $P_B(y)h$ where $h, y \in H$. The next three lemmas will give us this.

LEMMA 7. *Suppose that for Hilbert spaces H, V, $Q \in L(H,V)$ and $(QQ^*)^{-1}$ exists and belongs to $L(H,V)$. Then the orthogonal projection J of V onto $N(Q)$ is given by*

$$J = I - Q^*(QQ^*)^{-1}Q.$$

PROOF. Note that $Q^*(QQ^*)^{-1}Q$ is (1) idempotent, (2) symmetric, (3) its range is a subset of $R(Q^*)$, and (4) it is fixed on $R(Q^*)$. These four properties imply that $Q^*(QQ^*)^{-1}Q$ is the orthogonal projection onto $R(Q^*)$. Thus J is its complementary projection. □

Note that for B as in (6.9), $B'(y)h = h(0)$ for all $h, y \in H$. Let Q be this common value of $B'(y)$, $y \in H$.

The next lemma gives us an expression for Q^* in a special case for which $V = R$.

LEMMA 8. *Suppose $Qh = h(0)$, $h \in H$. Then*

$$(Q^*w)(t) = wC(1-t)/S(1),\ t \in [0,1],\ w \in R. \tag{6.10}$$

Proof. Suppose $w \in R$. Then there is a unique $f \in H$ so that

$$h(0)w = \langle Qh, w\rangle_R = \langle h, f\rangle_H,\ h \in H.$$

We want to determine this element f. To this end note that if $f \in C^2([0,1])$, then

$$\langle h, f\rangle_H = \int_0^1 (hf + h'f') = \int_0^1 h(f - f'') + h(1)f'(1) - h(0)f'(0)$$

and so if

$$wh(0) = \langle h, f\rangle_H, h \in H,$$

it must be that

$$f - f'' = 0, f'(1) = 0 \text{ and } f'(0) = -w. \tag{6.11}$$

It is an exercise in ordinary differential equations to find f satisfying (6.11). Such an f is given by (6.10).

LEMMA 9. *Let J be the orthogonal projection of H onto $N(Q)$. Then*

$$(Jh)(t) = h(t) - h(0)C(1-t)/C(1),\ t \in [0,1].$$

PROOF. We calculate $Q^*(QQ^*)^{-1}Q$. Using Lemma 9, if $w \in R$,

$$QQ^*w = (Q^*w)(0) = wC(1)/S(1)$$

and so

$$(QQ^*)^{-1}w = wS(1)/C(1).$$

Hence if $h \in H$,

$$\begin{aligned}(Q^*(QQ^*)^{-1}Qh)(t)) &= ((QQ^*)^{-1}Q^*Qh)(t)\\ &= h(0)(S(1)/C(1))(C(1-t)/S(1)) = h(0)C(1-t)/C(1),\ t \in [0,1].\end{aligned}$$

The conclusion follows immediately. □

PROOF. (Of Theorem 28). Note that $P_B(y) = J$ (J is as in Lemma 7) for any $y \in H$. A direct calculation using Lemmas 6 and 8 yields

$$((\nabla^B\phi)(y))(t) = (J((\nabla\phi)(y)))(t)$$
$$= y(t) - [y(1)S(t) + y(0)C(1-t)]/C(1),\ t \in [0,1].$$

□

The reader may verify that if $(\nabla^B\phi)(y) = 0$, then

$$y(t) = y(0)\exp(t),\ t \in [0,1].$$

A second example concerns a partial differential equation. The particular equation has a conventional variational principal and is treated again in Chapter 9 from that point of view. It is usually preferable to deal with conventional variational principals if they are available since that typically involves only first derivatives whereas the following procedure involves second derivatives.

Example 6.2. Suppose g is a real-valued $C^{(1)}$ function on R and Ω is a region with piecewise smooth boundary in R^2. We seek $u \in H = H^{2,2}(\Omega)$ so that

$$-\Delta u + g(u) = 0, u - w = 0 \text{ on } \quad \partial\Omega.$$

Define ϕ on H by

$$\phi(u) = (1/2)\int_\Omega (-\Delta u + g(u))^2 u \in H.$$

Denote $\{v \in H : v = 0 \text{ on } \partial\Omega\}$ by H_0. Then

$$\phi'(u)h = \int_\Omega (-\Delta u + g(u))(-\Delta h + g'(u)h), u \in H,\ h \in H_0.$$

Denote by P the orthogonal projection of $L_2(\Omega)^6$ onto

$$\{(h, h_1, h_2, h_{1,1}, h_{1,2}, h_{2,2}) : h \in H_0\}.$$

Accordingly

$$\phi'(u)h = \langle(h, h_1, h_2, h_{1,1}, h_{1,2}, h_{2,2}), (g'(u)m(u), 0, 0, -m(u), 0, -m(u))\rangle^6_{L_2(\Omega)}$$
$$= \langle h, \pi P(g'(u)m(u), 0, 0, -m(u), 0, -m(u))\rangle_H, u \in H, h \in H_0,$$

where $m(u) = -\Delta u + g(u)$, $u \in H$. Hence we have

$$(\nabla\phi)(u) = \pi P(g'(u)m(u), 0, 0, -m(u), 0, -m(u)),\ u \in H.$$

We emphasize that the boundary conditions are incorporated in this gradient. If $u \in H$, $\delta > 0$, $u - w = 0$ on $\partial\Omega$ and $v = u - \delta(\nabla\phi)(u)$, then we also have that $v - w = 0$ on $\partial\Omega$. Observe that if $B(u) = (u - w)|_{\partial\Omega}$, $u \in H$, then the condition

$$B(u) = 0$$

expresses the boundary conditions used in this example. In continuous steepest descent using $\nabla_B\phi$ in (6.2), $z(t) - w = 0$ on $\partial\Omega$, $t \geq 0$. Consequently, if $u = \lim_{t\to\infty} z(t)$ exists, we would have $B(u) = 0$ also, *i.e.*, $u - w = 0$ on $\partial\Omega$.

Alternatively we might have followed a path analagous to that in Example 6.1.

Theorems in the second part of chapter 4 all have appropriate generalizations in which $\nabla_B \phi$ replaces $\nabla \phi$. These generalizations will not be formally stated or proved here.

2. Approximation of Projected Gradients

We turn now to another matter concerning boundary conditions. We give our development only in a finite dimensional setting even though one surely exists also in function spaces. We will illustrate our results in one dimension (see [**57**] for some higher dimensional results). We indicate that laplacians associated with subspaces H_0 are essentially limiting cases of a sequence of laplacians on H taken with a succession of more extremely weighted norms.

Our point of departure is the norm (2.10) which we generalize as follows. For a given positive integer n, denote by

$$\alpha = \{\alpha_i\}_{i=0}^n$$

a sequence of positive numbers and define the weighted norm

$$\|x\|_\alpha = [\sum_{i=1}^n (((x_i - x_{i-1})/\gamma_n)^2 + \alpha_i((x_i + x_{i-1})/2)^2]^{1/2}, \ x \in R^n. \tag{6.12}$$

Denote by $\langle \ , \ \rangle_\alpha$ the corresponding inner product. In what follows, $\| \ \|, \langle \ , \ \rangle$ denote respectively the standard norm and inner product on R^n.

Now the pair of norms $\| \ \|$ and $\| \ \|_\alpha$ on R^{n+1} yield a Dirichlet space in the sense of Theorem 25 of Chapter 5. Denote by M_α the corresponding transformation which results from from Theorem 25 and denote M_α^{-1} by Δ_α, the laplacian for this Dirichlet space.

For each $\lambda > 0$ we choose a weight $\beta(\lambda)$ so that

$$\beta(\lambda)_1 = \lambda, \beta(\lambda)_i = 1, \ i = 2, \ldots, n.$$

We examine the behavior of $M_{\beta(\lambda)}$ as $\lambda \to \infty$. For the space H_0 whose points are those of R^{n+1} whose first term is 0, consider the the Dirichlet space given by H_0 with two norms, the first being the norm restricted to H_0 and the second being unweighted version of (6.12) restricted to H_0. Denote by M_0 the resulting transformation from Theorem 25 and by Δ_0 the inverse of M_0.

THEOREM 29. *If $h \in H_0$ then $M_0 h = \lim_{\lambda \to \infty} M_{\beta(\lambda)} h$.*

PROOF. Denote by P' the orthogonal projection, under the standard Euclidean norm on R^{n+1}, of R^{n+1} onto the points of H_0. Observe that

$$\Delta_{\beta(\lambda)} = (\lambda - 1)P' + \Delta_{\beta(1)}$$

and that if $0 < a < c$, then

$$\langle (cP' + \Delta_{\beta(1)})x, x\rangle \geq \langle (aP' + \Delta_{\beta(1)})x, x\rangle \geq \langle \Delta_{\beta(1)}x, x\rangle \geq \langle x, x\rangle, x \in R^n.$$

Thus

$$I \le M_{\beta(c)} \le M_{\beta(a)}$$

and so

$$J = \lim_{\lambda\to\infty} M_{\beta(\lambda)}$$

exists. Note that J is symmetric and nonnegative.

Suppose that $x \in R^{n+1}, x \ne 0$. For $\lambda > 0$ define

$$y_\lambda = (\lambda P' + \Delta_{\alpha(1)})^{-1}x. \tag{6.13}$$

From the above it is clear that $\lim_{\lambda\to\infty} y_\lambda$ exists. Then for $\lambda > 0, \|y_\lambda\| \le \|x]\|$,

$$x = \lambda P' y_\lambda + \Delta_{\beta(1)} y_\lambda$$

and so

$$\lambda\|P'y_\lambda\| \le \|x\|(1 + |\Delta_{\beta(1)}|).$$

Thus

$$\lim_{\lambda\to\infty} \|P'y_\lambda\| = 0. \tag{6.14}$$

Therefore $P'J = 0$ and consequently $JP' = 0$.

Take C to be the restriction of $(I - P')\Delta_{\beta(1)}$ to the range of $(I - P')$ and note that $C^{-1} = M_0$ (C^{-1} exists since $(\Delta_{\beta(1)})^{-1}$ exists and $\Delta_{\beta(1)}$ is symmetric and positive). For x in the range of $I - P'$ and $\lambda > 0$ and (6.13) holding,

$$\lambda P' y_\lambda + \Delta_{\beta(1)} y_\lambda = x$$

and so

$$\lambda(I - P')P'y_\lambda + (I - P')\Delta_{\beta(1)}P'y_\lambda = x$$

and consequently

$$(I - P')y_\lambda = C^{-1}x + C^{-1}(I - P')\Delta_{\beta(1)}P'y_\lambda.$$

Hence

$$Jx = \lim_{\lambda\to\infty} y_\lambda = C^{-1}x$$

which is what was to be shown. □

3. Singular Boundary Value Problems

Recent work of W.T. Mahavier [**46**] uses weighted Sobolev spaces in an interesting and potentially far reaching way. This work stems from an observation that for the singular problem

$$ty'(t) - y(t) = 0,\ t \in [0, 1]$$

the Sobolev gradient constructed according to previous chapters performs rather poorly. It is shown in [**46**] that a Sobolev gradient taken with respect to a finite dimensional version of

$$\|f\|_W^2 = \int_0^1 [(ty'(t))^2 + (y(t))^2]\ dt$$

gave vastly better numerical performance using discrete steepest descent. In [**46**] it is shown how to choose weights so that the resulting Sobolev gradients have good convergence properties can be chosen which are appropriate to singularities. Generalizations to systems is made in this work, which should be consulted for details are careful numerical comparisons.

4. Dual Steepest Descent

Suppose that each of H, K, S is a Hilbert space, F is a $C^{(1)}$ function from H to K and B is a $C^{(1)}$ function from S to R. Denote by ϕ, ψ, η the functions on H defined by

$$\phi(x) = ||F(x)||_K^2,\ x \in H,$$

$$\psi(x) = ||B(x)||_S^2,\ x \in H.$$

$$\eta = \phi + \psi.$$

Dual steepest, used by Richardson in [**90**], [**92**] descent consists in seeking $u \in H$ such that

$$u = lim_{t\to\infty} z(t) \text{ and} F(u) = 0, B(u) = 0 \tag{6.15}$$

where

$$z(0) = x \in H,\ z'(t) = -(\nabla\eta)(z(t)),\ t \in H. \tag{6.16}$$

The following is a convergence result from [**90**]. All inner products and norms in this section without subscripts are to be taken in H.

THEOREM 30. *Suppose Ω is an open subset of H, z satisfies* (6.16) *and $R(z) \subset \Omega$. Suppose also that there exists $c, d > 0$ such that*

$$||(\nabla\phi)(y)|| \geq ||F(y)||_K, ||(\nabla\psi)(y)|| \geq ||B(y)||_S,\ y \in \Omega.$$

Finally suppose that there is $\alpha \in (-1, 1]$ so that if $y \in \Omega$, then

$$\langle(\nabla\phi)(y), (\nabla\psi)(y)\rangle/||(\nabla\phi)(y)||\ ||(\nabla\psi)(y)|| \geq \alpha.$$

Then (6.16) *holds.*

PROOF. (From [**90**]) Clearly we may assume that $\alpha \in (-1, 0)$. Then

$$\begin{aligned}||(\nabla\eta)(y)||^2 &= ||(\nabla\phi)(y) + (\nabla\psi)(y)||^2 \\ &= ||(\nabla\phi)(y)||^2 + 2\rangle(\nabla\phi)(y), (\nabla\psi)(y)\rangle + ||(\nabla\psi)(y)||^2 \\ &\geq ||(\nabla\phi)(y)|| + 2\alpha||(\nabla\phi)(y)||\ ||(\nabla\psi)(y)|| + ||(\nabla\psi)(y)||^2 \\ &= -\alpha[||(\nabla\phi)(y)||^2 - 2||(\nabla\phi)(y)||||(\nabla\psi)(y)|| + ||(\nabla\psi)(y)||^2] \\ &\quad + (1+\alpha)[||(\nabla\phi)(y)||^2 + ||(\nabla\psi)(y)||^2] \\ &\geq (1+\alpha)\min(c^2, d^2)[||F(y)||^2 + ||B(y)||^2] = (1+\alpha)\min(c^2, d^2)2\eta(y).\end{aligned}$$

The conclusion (6.15) then follows from Theorem 9 of Chapter 4. □

5. Multiple Solutions of Some Elliptic Problems

The following was written by John M. Neuberger and kindly supplied to the present writer. The reader should see [**50**] and [**15**] for more background and for numerical results.

In [**50**] Sobolev gradient descent is used to find sign-changing (as well as one-sign) solutions to elliptic boundary value problems. This numerical algorithm employs gradient descent as described in this text together with projections on to infinite dimensional submanifolds (of finite codimension) of $H = H_0^{1,2}(\Omega)$. The idea was first suggested by an examination of the existence proof for a superlinear problem in [**15**] and later extended to study a more general class of elliptic bvps. Specifically, let Ω be a smooth bounded region in $\mathbf{R}^N$, Δ the Laplacian operator, and $f \in C^1(\mathbf{R}, \mathbf{R})$ such that $f(0) = 0$ (this last condition can be relaxed and the resulting variational structure similarly handled). We seek solutions to the boundary value problem in Ω:

$$\Delta u + f(u) = 0, \ u = 0 hboxon \partial\Omega. \tag{6.17}$$

Under certain technical assumptions on f (e.g., subcritical and superlinear growth), it was proven in [**15**] that (6.17) has at least four solutions, the trivial solution, a pair of one-sign solutions, and a solution which changes sign exactly once. These solutions are critical points of the C^2 action functional

$$J(u) = \int_\Omega \{\frac{1}{2}|\nabla u|^2 - F(u)\}\, dx,$$

where $F(u) = \int_0^u f(s)\, ds$. Necessarily these critical points belong to the zero set of the functional $\gamma(u) = J'(u)(u)$. We denote this zero set and an important subset by

$$\begin{aligned} S &= \{u \in H - \{0\} : \gamma(u) = 0\} \\ S_1 &= \{u \in S : u_+ \neq 0, u_- \neq 0, \gamma(u_+) = 0\}, \end{aligned}$$

where we note that nontrivial solutions to (6.17) are in S (a closed subset of H) and sign-changing solutions are in S_1 (a closed subset of S). In this case, 0 is a local minimum of J, $J(\lambda u) \to -\infty$ as $\lambda \to \infty$ for all $u \in H - \{0\}$, and there exists $\delta, M > 0$ such that $||u|| \geq \delta$ and $J(u) \geq M$ for all $u \in S$. Thus, we have the alternative definitions $S = \{u \in H - \{0\} : J(u) = \max_{\lambda > 0} J(\lambda u)\}$ and $S_1 = \{u \in S : u_+ \in S, u_- \in S\}$. Think of a volcano with the bottom of the crater at the origin; the rim is S. Since S is diffeomorphic to the unit sphere in H and it can be shown that S_1 separates positive from negative elements of S, think of the one-sign elements as north and south poles with S_1 the equator.

For the problem outlined above, steepest descent will converge to the trivial solution if the initial estimate is inside the volcano. The remaining critical points are saddle points and elements of S. The algorithm used to keep iterates on S is steepest *ascent* in the ray direction. Projecting the gradient $\nabla J(u)$ on to the ray $\{\lambda u : \lambda > 0\}$ results in the formula

$$\frac{\langle \nabla J(u), u\rangle}{\langle u, u\rangle} u = \frac{\gamma(u)}{||u||^2} u.$$

Thus, the sign of γ determines the ascent direction towards S. If we denote the unique limit of this convergent sequence for a given initial element u by Pu, then $Pu \in S$ and $P(u_+) + P(u_-) \in S_1$. Standard finite difference methods can be used to compute the integrals $\gamma(u) = ||u||^2 - \int_\Omega uf(u)\,dx$ and $||u||^2 = \int_\Omega |\nabla u|^2$. This leads to satisfactory convergence of the sequence $u_{k+1} = u_k + s\frac{\gamma(u_k)}{||u_k||^2}u_k$, where the stepsize is some $s \in (.1, 1)$. This is a fairly slow part of the algorithm for which more efficient methods might be found.

To find one-sign solutions, let u_0 be any function satisfying the boundary condition and of roughly correct sign. Before each Sobolev gradient descent step, apply the iterative procedure effecting the projection P. To find a sign-changing solution, let u_0 be any function satisfying the boundary condition and of roughly correct nodal structure (eigenfunctions make excellent initial guesses). Before each Sobolev descent step, apply the projection to the positive and negative parts of the iterate and add them together. The result will be an approximation of an element of S_1.

CHAPTER 7

Newton's Method in the Context of Sobolev Gradients

This chapter contains two developments. The first shows how continuous Newton's method arises from a variational problem. It provides an example of what was indicated in (1.3). The second development concerns a version of Newton's method in which the usually required invertability fails to hold.

1. Newton Directions From an Optimization.

Up until now we have concentrated on constructing Sobolev gradients for problems related to differential equations. If

$$\phi(u) = 0 \tag{7.1}$$

represents a systems of differential equations on some linear space X of functions, then a norm for X was always chosen which involved derivatives. For example if

$$\phi(u) = (1/2)\int_0^1 (u' - u)^2$$

for functions u on [0,1], then our chosen norm was $X = H^{1,2}([0,1])$, i.e., we chose

$$\|u\|_X^2 = \int_0^1 (u^2 + {u'}^2).$$

With regard to this norm ϕ is a continuously differentiable function with respect to which a differentiable gradient function may be defined. Finite dimensional emulations of both ϕ and the above norm gave us numerically useful gradients. This is in stark contrast with the choice $X = L_2([0,1]$ for which ϕ would be everywhere discontinuous and only densely defined — an unpromising environment in which to consider gradients. We noted in Chapter 2 that finite dimensional emulations of such situations performed very poorly. Now $X = H^{1,2}([0,1])$ was chosen over $H^{m,2}([0,1])$ for $m > 1$ since a space of minimal complexity (which still leaves ϕ continuously differentiable) seems preferable.

In case (7.1) represents an expression in which derivatives are not in evidence, it would seem likely that no Sobolev metric would present an obvious choice. Many finite dimensional cases that do not emulate derivatives still suffer the same problems that are illustrated by Theorem 2.

We raise the question as to how effective gradients may be defined in such cases. For this we return to (1.3). Suppose that n is a positive integer and ϕ is a $C^{(3)}$ function on R^n. We seek $\beta : R^n \times R^n \to R$, $\beta \in C^{(3)}$, so that if $x \in R^n$, the problem of finding a critical point h of

$$\phi'(x)h \text{ subject to the constraint } \quad \beta(x,h) = c \in R$$

leads us to a numerically sound gradient. What ought we require of β? One criterion (which served us rather well for differential equations even though we did not mention it explicitly) is that the sensitivity of h in $\beta(x,h)$ should somewhat match the sensitivity of $\phi'(x)h$. This suggests a choice of

$$\beta(x,h) = \phi(x+h),\ x,h \in R^n.$$

For $x \in R^n$, define

$$\alpha(h) = \phi'(x)h, \gamma(h) = \beta(x,h),\ h \in R^n. \tag{7.2}$$

If h is an extremum of α subject to the constraint $\gamma(h) = c$ (for some $c \in R$), then, using Lagrange multipliers, we must have that

$$(\nabla\alpha)(h) \text{ and } (\nabla\gamma)(h) \text{ are linearly dependent.}$$

But

$$(\nabla\alpha)(h) = (\nabla\phi)(h) \text{ and } (\nabla\gamma)(h) = (\nabla\phi)(x+h).$$

We summarize some consequences in the following:

THEOREM 31. *Suppose that ϕ is a real-valued $C^{(3)}$ function on R^n, $x \in R^n$, and (7.2) holds. Suppose also that $((\nabla\phi)'(x))^{-1}$ exists. Then there is an open interval J containing 1 such that if $\lambda \in J$, then*

$$\lambda(\nabla\phi)(x) = (\nabla\phi)(x+h)$$

for some $h \in R^n$.

PROOF. Since $((\nabla\phi)'(x))^{-1}$ exists, then $((\nabla\phi)'(y))^{-1}$ exists for all y in some region G containing x. The theorem in the preface gives that there is an open interval J containing 1 on which there is a unique function z so that

$$z(1) = 0, z'(t) = ((\nabla\phi)'(x+z(t)))^{-1}(\nabla\phi)(z(t)),\ t \in J.$$

This is rewritten as

$$((\nabla\phi)'(x+z(t))z'(t) = (\nabla\phi)(z(t)),\ t \in J.$$

Taking anti-derivatives we get

$$(\nabla\phi)(x+z(t)) = t(\nabla\phi)(x) + c_1,\ t \in J.$$

But $c_1 = 0$ since $z(1) = 0$ and the argument is finished. □

Note that

$$z'(1) = ((\nabla\phi)'(x))^{-1}(\nabla\phi)(x)$$

is the Newton direction of $\nabla\phi$ at x. For a given x, the sign of

$$\langle((\nabla\phi)'(x))^{-1}(\nabla\phi)(x), (\nabla\phi)(x)\rangle_{R^n}$$

is important. If this quantity is positive then

$$((\nabla\phi)'(x))^{-1}(\nabla\phi)(x)$$

is an ascent direction; if negative it is a descent direction; if zero, then x is already a critical point of ϕ.

2. Generalized Inverses and Newton's Method.

Suppose that each of H and K is a Hilbert space and $F : H \to K$ is a $C^{(2)}$ function. Suppose also that if $x \in H$, then for some $m > 0$,

$$||F'(x)^* y||_H \geq m||y||_K, \ y \in K. \tag{7.3}$$

This condition allows for $F'(x)$ to have a large null space (corresponding to differential equations for which full boundary conditions are not specified) so long as $F'(x)$ is bounded from below, *i.e.*, that (7.3) holds.

LEMMA 10. *Suppose $T \in L(H, K)$ and for some $m > 0$*

$$||T^* y||_H \geq m||y||_K, \ y \in K.$$

Then

$$(TT^*)^{-1} \text{ exists and is in } L(K, K)$$

and

$$\langle (TT^*)^{-1} y, y\rangle_K \leq ||y||_K^2/m^2, \ y \in K.$$

This is essentially the theorem on page 266 of [**91**].

The following indicates that a condition somewhat stronger than (4.9) gives convergence of continuous Newton's method in cases where there may be a continuum of zeros of F — cases in which $F'(x)$ fail to be invertable for some $x \in H$.

THEOREM 32. *Suppose that F is a C^2 function from $H \to K$ and for each $r > 0$ there is $m > 0$ so that (7.3) holds if $||x||_H \leq r$. Suppose also that $x \in H$ and $w : [0, \infty) \to H$ satisfies*

$$w(0) = x, \ w'(t) = -F'(w(t))^*(F'(w(t))F'(w(t))^*)^{-1}F(w(t)), \ t \geq 0. \tag{7.4}$$

If the range of w is bounded then

$$u = \lim_{t\to\infty} w(t) \text{ exists and } F(u) = 0.$$

PROOF. Suppose w has bounded range. Denote by r a positive number so that

$$||w(t)||_H \leq r, \ t \geq 0,$$

and denote by m a positive number so that (7.3) holds for all x so that $||x||_H \leq r$. By Lemma 10

$$|(F'(x)F'(x)^*)|_{L(K,K)} \leq 1/m^2, \ ||x||_H \leq r.$$

Now suppose that $x \in H$ and w satisfies (7.4). Observe that

$$(F(w))' = F'(w)w' = -F(w)$$

so that

$$F(w(t)) = e^{-t}F(x),\ t \geq 0.$$

It follows that

$$\begin{aligned}\|w'(t)\|_H^2 &= \|F'(w(t))^*(F'(w(t))F'(w(t))^*)^{-1}F(w(t))\|_H^2 \\ &= \langle (F'(w(t))F'(w(t))^*)^{-1}F(w(t)), F(w(t))\rangle_H. \\ &= e^{-2t}\langle (F'(w(t))F'(w(t))^*)^{-1}F(x), F(x)\rangle_K,\ t \geq 0.\end{aligned}$$

Hence

$$\|w'(t)\|_H^2 \leq e^{-2t}M_0^2,\ t \geq 0$$

where

$$M_0 = \|F(x)\|_K/m.$$

Since

$$\|w'(t)\|_H \leq e^{-t}M_0, t \geq 0,$$

it follows that

$$\int_0^\infty \|w'\|_H \ < \infty$$

and hence

$$u = \lim_{t\to\infty} w(t) \text{ exists.}$$

Moreover

$$F(u) = \lim_{t\to\infty} F(w(t)) = 0,$$

and the argument is finished. □

Theorem 19 gives some instances in which (7.3) is satisfied. Many other instances can be found.

A promising group of applications for this chapter appears to be the following: For $\phi :\ H \to R, \phi \in C^{(2)}$, many of the Sobolev gradients in this monograph have the property that $((\nabla\phi)'(u))^{-1}$ exists and is in $L(H,H), u \in H$. Taking $F(u) = (\nabla\phi)(u), u \in H$, there is in such cases the possibility of using either discrete or continuous Newton's method on F. As a start in this direction there are the following two theorems.

THEOREM 33. *Suppose that ϕ is a real valued $C^{(2)}$ function on the Hilbert space H and $\nabla\phi$ has a locally lipschitzian derivative. Suppose furthermore that $x \in H$ and $\alpha, r > 0$ so that $((\nabla\phi)'(u))^{-1}$ exists, is in $L(H,H)$, $u \in H, \|u - x\|_H \leq r$ and*

$$\|((\nabla\phi)'(y))^{-1}(\nabla\phi)(x)\|_H \leq \alpha,\ \|y - x\|_H \leq r.$$

If

$$z(0) = x,\ z'(t) = -((\nabla\phi)'(z(t)))^{-1}(\nabla\phi)(z(t)),\ t \geq 0, \tag{7.5}$$

then

$$u = \lim_{t\to\infty} z(t)$$

exists and

$$(\nabla\phi)(u) = 0 \textit{ provided that } \alpha < r.$$

PROOF. For z as in (7.5),

$$(\nabla\phi)'(z(t))z'(t) = -(\nabla\phi)(z(t)),\ t \geq 0,$$

and so

$$((\nabla\phi)(z))' = -(\nabla\phi)(z)$$

and hence

$$(\nabla\phi)(z(t)) = e^{-t}(\nabla\phi)(z(t)),\ t \geq 0.$$

But this implies that

$$z'(t) = -e^{-t}((\nabla\phi)'(z(t)))^{-1}(\nabla\phi)(x),\ t \geq 0$$

and hence if $t > 0$ and $\|z(s) - x\|_H \leq r,\ 0 \leq s \leq t$ then

$$\|z(t) - x\|_H \leq \int_0^t \|z'\|_H = \int_0^t e^{-s}\|((\nabla\phi)'(z(s))^{-1}(\nabla\phi)(x)\|_H\ ds$$
$$\leq \int_0^t e^{-s}\alpha\ ds = \alpha(1 - e^{-t}) < \alpha.$$

It follows using the fact that $\alpha < r$ that $\|z(t)-x\|_H \leq \alpha$ and also that $\int_0^t \|z'\|_H \leq \alpha,\ t \geq 0$. Consequently,

$$u = \lim_{t\to\infty} z(t)$$

exists and

$$(\nabla\phi)(u) = \lim_{t\to\infty}(\nabla\phi)(z(t)) = \lim_{t\to\infty} e^{-t}(\nabla\phi)(x) = 0.$$

□

THEOREM 34. *Suppose ϕ is a real-valued $C^{(2)}$ function on H and $\nabla\phi$ has a locally lipschitzian derivative and*

$$((\nabla\phi)'(u))^{-1} \in L(H, H),\ u \in H.$$

Suppose also that if $r > 0$, there is $\alpha > 0$ so that

$$|((\nabla\phi)'(u))|_{L(H,H)} \leq \alpha,\ \|u\|_H \leq r. \tag{7.6}$$

Finally suppose that $\nabla\phi$ is coercive in the sense that if $M > 0$, then

$$\{y \in H\ :\ \|(\nabla\phi)(y)\|_H \leq M\}$$

is bounded. If

$$z(0) = x,\ z'(t) = -((\nabla\phi)'(z(t)))^{-1}(\nabla\phi)(z(t)),\ t \geq 0 \tag{7.7}$$

then

$$u = \lim_{t\to\infty} z(t) \textit{ exists and } (\nabla\phi)(u) = 0.$$

PROOF. Suppose $x \in H$ and (7.7) is satisfied. Since, as in the argument for the previous theorem,

$$(\nabla\phi)(z(t)) = e^{-t}(\nabla\phi)(x),\ t \geq 0,$$

it follows, due to (7.6), that $R(z)$ is bounded. Denote by r a positive number such that $r \geq \|z(t)\|_H$, $t \geq 0$ and by α a positive number such that

$$|((\nabla\phi)'(u))^{-1}|_{L(H,H)} \leq \alpha,\ \|u\|_H \leq r.$$

Then as in the above argument,

$$z'(t) = -e^{-t}((\nabla\phi)'(z(t))^{-1}(\nabla\phi)(x),\ t \geq 0,$$

and

$$\int_0^t \|z'\| \leq \int_0^t e^{-s}\|(\nabla\phi)'(z(s))^{-1}(\nabla\phi)x)\|\, ds$$
$$\leq \int_0^t e^{-s}\alpha\|(\nabla\phi)(x)\| \leq \alpha\|(\nabla\phi)(x)\|,\ t \geq 0$$

and hence

$$u = \lim_{t\to\infty} z(t) \text{ exists and } (\nabla\phi)(u) = 0.$$

□

CHAPTER 8

Finite Difference Setting: the Inner Product Case

This chapter is devoted to construction of Sobolev gradients for a variety of finite dimensional approximations to problems in differential equations. To illustrate some intended notation we first reformulate the finite dimensional problem of Chapter 2. These considerations give a guide to writing computer codes for various Sobolev gradients.

Example 1. Suppose n is a positive integer and $D_0, D_1 : R^{n+1} \to R^n$ are defined by

$$D_0 y = ((y_1 + y_0)/2, ..., (y_n + y_{n-1})/2),$$
$$D_1 y = ((y_1 - y_0)/\delta, ..., (y_n - y_{n-1})/\delta), \quad (8.1)$$

where $y = (y_0, y_1, ..., y_n) \in R^{n+1}$ and $\delta = 1/n$. Define

$$D : R^{n+1} \to (R^n)^2$$

so that

$$Dy = \begin{pmatrix} D_0 y \\ D_1 y \end{pmatrix}$$

The norm expressed by (2.10) is also expressed by

$$\|y\|_D = ((\|D_0 y\|_{R^n})^2 + (\|D_1 y\|_{R^n})^2)^{1/2} = \|Dy\|_{R^{2n}},\ y \in R^{n+1}. \quad (8.2)$$

The real-valued function in (2.4) can be expressed as

$$\phi(y) = \|G(Dy)\|^2_{R^n}/2,\ y \in R^{n+1}, \quad (8.3)$$

where $G : R^2 \to R$ is defined by

$$G(r, s) = s - r,\ \binom{r}{s} \in R^2.$$

Then for $u \in R^{n+1}$, the composition $G(Du)$ is to be understood as the member of R^n whose ith component is

$$G((D_0 u)_i, (D_1 u)_i) = (D_1 u)_i - (D_0 u)_i,\ i = 1, \ldots, n. \quad (8.4)$$

If $u \in R^{n+1}$ then $G(Du) = 0$ if and only if u satisfies

$$(u_i - u_{i-1})/\delta - (u_i + u_{i-1})/2 = 0,\ i = 1, \ldots, n.$$

We will use some of this notation to express each of the two gradients that were the object of study in Chapter 2. We first take the Fréchet derivative of ϕ in (8.3):

$$\phi'(u)h = \langle G'(Du)^t Dh, G(Du)\rangle_{R^n}, \ u, h \in R^{n+1}. \tag{8.5}$$

Our gradient construction takes two different paths depending on which norm we take on R^{n+1}. First we take the standard norm on R^{n+1} and calculate

$$\phi'(u)h = \langle Dh, G'(Du)^t G(Du)\rangle_{R^n} \tag{8.6}$$

$$= \langle h, D^t G'(Du)^t G(Du)\rangle_{R^{n+1}}, \ h, u \in R^{n+1}, \tag{8.7}$$

so that

$$(\nabla\phi)(u) = D^t G'(Du)^t G(Du), \ u \in R^{n+1}. \tag{8.8}$$

Note that this $(\nabla\phi)(u)$ is also obtained by making a list of the $n+1$ partial derivatives of ϕ evaluated at u.

Proceeding next using the norm $\| \ \|_D$ in (8.2),

$$\phi'(u)h = \langle G'(Du)Dh, G(Du)\rangle_{R^n}$$

$$= \langle Dh, G'(Du)^t G(Du)\rangle_{R^{2n}} = \langle Dh, PG'(Du)^t G(Du)\rangle_{R^{2n}} \tag{8.9}$$

where P is the orthogonal projection (using the standard norm on $(R^n)^2$) of $(R^n)^2$ onto $R(D) \subset (R^n)^2$. Taking $\pi : R^n \times R^n \to R^n$ defined by

$$\pi\binom{x}{y} = x, \ x, y \in R^n,$$

we have from (8.9)

$$\phi'(u)h = \langle h, \pi PG'(Du)^t G(Du)\rangle_D \tag{8.10}$$

and hence the gradient of ϕ with respect to $\| \ \|_D$ is given by

$$(\nabla_D \phi)(u) = \pi PG'(Du)^t G(Du), u \in R^{n+1} \tag{8.11}$$

where $\langle \ , \ \rangle_D$ is the inner product associated with $\| \ \|_D$. Note that

$$P = D(D^t D)^{-1} D^t \tag{8.12}$$

since $D(D^t D)^{-1} D^t$ is symmetric, idempotent, is fixed on $R(D)$ and has range which is a subset of $R(D)$. Thus $\pi P = (D^t D)^{-1} D^t$ and so

$$(\nabla_D \phi)(u) = (D^t D)^{-1} D^t G'(Du)^t G(Du) \tag{8.13}$$

$$= (D^t D)^{-1} (\nabla\phi)(u), \ u \in R^{n+1}. \tag{8.14}$$

The reader might check that (8.13) is just another way to express (2.11) when $G(r, s) = s - r, \binom{r}{s} \in R^2$, since A_n in (2.11) is the same as $D^t D$. One can see that calculations starting with (8.3), and proceeding to (8.5) - (8.13) hold in considerable generality. In particular they can be extended to the following:

Example 2. The development in the previous example may be duplicated for general $G : R^2 \to R$, $G \in C^{(1)}$. For such a G we seek a gradient in order to obtain approximate solutions to the fully nonlinear equation

$$G(y(t), y'(t)) = 0, \ t \in [0, 1]. \tag{8.15}$$

Then (8.4) may be replaced by

$$G((D_0u)_i, (D_1u)_i) = 0, \ i = 1, ..., n.$$

It is of interest to note that (8.15) may be singular with a singularity depending upon nonlinearity in that it may be that

$$G_2(y(t), y'(t)) = 0$$

for some $t \in [0, 1]$ and some solution y. Our gradient construction is not sensitive to such singularity although numerical performance would likely benefit from the considerations of Section 3 of Chapter 6.

Dirgression on adjoints of difference operators. We pause here to indicate an example of a rather general phenomena which occurs when one emulates a differential operator with a difference operator. Consider for an $n \times (n+1)$ matrix for D_1:

$$\delta(D_1)_{i,i} = -1, \delta(D_1)_{i,i+1} = 1, \ \delta(D_1)_{i,j} = 0, j \neq i, i+1, \ i = 1, \ldots, n.$$

The transpose of D_1 is then denoted by

$$D_1^t \tag{8.16}$$

Commonly the derivative operator (call it L here to not confuse it with D) on $L_2([0, 1])$ is taken to have domain those points in $L_2([0, 1])$ which are also in $H^{1,2}([0, 1])$. Denoting by L^t the adjoint of L (considered as a closed densely defined unbounded operator on $L_2([0, 1])$), we have that the domain of L^t is

$$\{u \in H^{1,2}([0, 1]) : u(0) = 0 = u(1)\}. \tag{8.17}$$

Furthermore,

$$L^t u = -u', u \text{ in the domain of } L^t.$$

From (8.16), if $v = (v_0, v_1, ..., v_n) \in R^{n+1}$,

$$D_1^t v = (-v_1/\delta, -(v_2 - v_1)/\delta, ..., -(v_n - v_{n-1})/\delta, v_n/\delta) \tag{8.18}$$

so that D_1^t on R^n is like $-D_1$ would be on R^n except for the first and last terms of (8.18). Remembering that $\delta = 1/n$, we see that in a sort of limiting case of $L_2([0, 1])$, the condition $u(0) = 0 = u(1)$ in (8.17) seems almost forced by the presence of the first and last terms of the rhs of (8.18). It would be possible to develop the subject of boundary conditions for adjoints of differential operators (particularly as to exactly what is to be in the domain of the adjoint the formal expression for an adjoint is usually clear) by means of limits of difference operators on finite dimensional emulations of $L_2(\Omega), \Omega \subset R^m$ for some positive integer m. We write this knowing that these adjoints are rather well understood. Nevertheless it seems that it might be of interest to see domains of adjoints obtained by something like the above considerations.

Example 3. Before looking at finite dimensional Sobolev gradients for approximating problems to partial differential equations, we see how Examples 1 and 2 in this chapter are changed when we require a gradient which maintains an initial condition at 0. Choose a number α. We will try to maintain the boundary condition $y_0 = \alpha$. Pick

$$y = (y_0, y_1, ..., y_n) \in R^{n+1},\ y_0 = \alpha.$$

Denote by H_0 the subspace of R^{n+1} consisting of all points

$$(x_0, x_1, ..., x_n) \in R^{n+1}$$

so that $x_0 = 0$. We wish to represent the functional on H_0 :

$$h \to \phi'(y)h,\ h \in H_0.$$

Using (8.6),

$$\phi'(u)h = \langle h, D^t G'(Du)^t G(Du)\rangle_{R^{n+1}} = \langle h, \pi_0 D^t G'(Du)^t G(Du)\rangle_{R^{n+1}},$$

$u \in R^{n+1}$, $h \in H_0$, where π_0 is the orthogonal projection of R^{n+1} onto H_0 in the standard inner product for R^{n+1}. Thus

$$\pi_0 D^t G'(Du)^t G(Du)$$

is the gradient of ϕ at u relative to the subspace H_0 and the standard norm on R^{n+1}. To find an gradient of ϕ relative to H_0 and the metric induced by $\|\ \|_D$, we will see essentially that (8.9) holds with P replaced by P_0, the orthogonal projection of $(R^n)^2$ (using the standard metric on $(R^2)^n$) onto

$$\{Dh : h \in H_0\}.$$

Denote $\pi_0 D^t D|_{H_0}$ by E_0.

Theorem 35. $P_0 = DE_0^{-1}\pi_0 D^t$.

Proof. Define $Q_0 = DE_0^{-1}\pi_0 D^t$. First note that $Q_0^2 = Q_0$ since

$$Q_0^2 = (DE_0^{-1}\pi_0 D^t)(DE_0^{-1}\pi_0 D^t) =$$
$$DE_0^{-1}\pi_0 D^t DE_0^{-1}\pi_0 D^t = DE_0^{-1}\pi_0 D^t = Q_0$$

and

$$Q_0^t = (DE_0^{-1}\pi_0 D^t)^t = (D\pi_0 E_0^{-1}\pi_0 D^t)^t =$$
$$D\pi_0^t E_0^{-1}\pi_0^t D^t = DE_0^{-1}\pi_0 D^t = Q_0.$$

Further note that $R(Q_0) \subset R(D|_{H_0})$ and that Q_0 is fixed on $R(D|_{H_0})$. This characterizes the orthogonal projection of $(R^2)^n$ onto

$$\{Dh : h \in H_0\}$$

and thus we have that $P_0 = Q_0$. ☐

As in (8.9), if $h \in H_0, u \in H$,

$$\phi'(u)h = \langle Dh, G'(u)^t G(Du)\rangle_{R^{2n}},$$

but here

$$\langle Dh, G'(u)^t G(Du)\rangle_{R^{2n}} = \langle Dh, P'_G(u)^t G(Du)\rangle_{R^{2n}}$$

and so in place of (8.8) we have

$$(\nabla \phi_B)(u) = \pi P_0 G'(u)^t G(Du) = E_0^{-1} \pi_0 (\nabla \phi)(u)$$

where

$$(\nabla \phi)(u)$$

is the conventional gradient of ϕ (constructed without regard to boundary conditions) and

$$(\nabla \phi_B)(u)$$

is the Sobolev gradient of ϕ constructed with regard to the boundary condition

$$B(u) - \alpha = 0$$

where if $u = (u_0, u_1, ..., u_n) \in R^{n+1}$, then $B(u) = u_0$.

We next give an example of how a finite difference approximation for a partial differential equation fits our scheme.

Example 4. Suppose n is a positive integer and G is the grid composed of the points

$$\{(i/n, j/n)\}_{i,j=0}^{n}.$$

Denote by G_d the subgrid

$$\{(i/n, j/n)\}_{i,j=1}^{n-1}.$$

Denote by H the $(n+1)^2$ dimensional space whose points are the real-valued functions on G and denote by H_d the $(n-1)^2$ dimensional space whose points are the real-valued functions on G. Denote by D_1, D_2 the functions from H to H_d so that if $u \in H$, then

$$D_1 u = \{(u_{i+1,j} - u_{i-1,j})/(2\delta)\}_{i,j=1}^{n-1}$$

and

$$D_2 u = \{(u_{i,j+1} - u_{i,j-1})/(2\delta)\}_{i,j=1,n}^{n-1}.$$

Denote by D the transformation from H to $H \times H_d \times H_d$ so that

$$Du = (u, D_1 u, D_2 u), \ u \in H.$$

Denote by F a $C^{(1)}$ from $R^3 \to R$. For example, if

$$F(r, s, t) = s + rt, \ (r, s, t) \in R^3$$

then the problem of finding $u \in H$ such that

$$F(Du) = 0$$

is a problem of finding a finite difference approximation to a solution z to the viscosity free Burger's equation:

$$z_1 + z z_2 = 0 \tag{8.19}$$

on $[0,1] \times [0,1]$. For a metric on H we choose the finite difference analogue of the norm on $H^{1,2}([0,1] \times [0,1]$, namely,

$$\|u\|_D = (\|u\|^2 + \|D_1 u\|^2 + \|D_2 u\|^2)^{1/2}, \ u \in H, \tag{8.20}$$

All norms and inner products without subscripts in this example are Euclidean.

Define then ϕ with domain H so that

$$\phi(u) = \|F(Du)\|^2/2, u \in H.$$

We compute (with $\pi = D^{-1}$))

$$\begin{aligned}\phi'(u)h &= \langle F'(Du)Dh, F(Du)\rangle \\ &= \langle Dh, (F'(Du))^t F(Du)\rangle = \langle Dh, PF'(Du)^t F(Du)\rangle \\ &= \langle h, \pi PF'(Du)^t F(Du)\rangle_D = \langle h, (D^tD)^{-1}D^tF'(Du)^tF(Du)\rangle_D\end{aligned}$$

so that

$$(\nabla_D\phi)(u) = (D^tD)^{-1}D^tF'(Du)^tF(Du) = (D^tD)^{-1}(\nabla\phi)(u)$$

where $\langle\ ,\ \rangle_D$ the inner product associated with $\|\ \ \|_D$ (in the preceding we use the fact that $P = D(D^tD)^{-1}D)$ and $(\nabla\phi)(u)$ is the ordinary gradient of ϕ, $u \in H$). A gradient which takes into account boundary conditions may be introduced into the present setting much as in Examples 2 and 3. We pause here to recall briefly some known facts about (8.19).

Theorem 36. *Suppose z is a C^1 solution on $\Omega = [0,1] \times [0,1]$ to* (8.19). *Then* $[0,1] \times [0,1]$ *is the union of a collection Q of closed intervals such that*

(i) no two members of Q intersect,

(ii) if $J \in Q$, then each end point of J is in $\partial\Omega$,

(iii) if $J \in Q$ and J is nondegenerate, then the slope m of J is such that

$$m = z(x) \quad x \in J,$$

i.e., the members of Q are characteristic lines for (8.19). Some reflection reveals that no nondegenerate interval S contained in $\partial\Omega$ is small enough so that if arbitrary smooth data is specified on S then there would be a solution z to (8.19) assuming that data on S. This is indicative of the fundamental fact that for many systems of nonlinear partial differential equations the set of all solutions on a given region is not conveniently specified by specifying boundary conditions on some designated boundary. This writer did numerical experiments in 1976-77 (unpublished) using a Sobolev gradient (8.19). It was attractive to have a numerical method which was not boundary condition dependent. It was noted that a limiting function from a steepest descent process had a striking resemblance to the starting function used (recall that for linear homogeneous problems the limiting value is the nearest solution to the starting value). It was then that the idea of a foliation emerged: the relevant function space is divided into leaves in such a way that two functions are in the same leaf provided they lead (via steepest descent) to the same solution. It still remains a research problem to characterize such foliations even in problems such as (8.19). More specifically we say that

$$r, s \in H = H^{1,2}(\Omega)$$

are equivalent provided that if

$$z_r(0) = r, z_s(0) = s,$$

and

$$z_r'(t) = -(\nabla\phi)(z_r(t)), z_s'(t) = -(\nabla\phi)(z_s(t)),\ t \geq 0, \tag{8.21}$$

then

$$\lim_{t\to\infty} z_r(t) = \lim_{t\to\infty} z_s(t),$$

where

$$\phi(u) = (1/2)\int_\Omega (u_1 + uu_2)^2,\ u \in H.$$

We would like to understand the topological, geometrical and algebraic nature of these equivalence classes. Each contains exactly one solution to (8.19). The family of these equivalence classes should characterize the set of all solutions to (8.19) and provide a point of departure for further study of this equation. The gradient to which we refer in (8.21) is the function $\nabla\phi$ so that if $u \in H$,

$$\phi'(u)h = \langle h, (\nabla\phi)(u)\rangle_H,\ h \in H.$$

Chapter 14 deals rather generally with foliations which arise as in the above.

We present now an example of a Sobolev gradient for numerical approximations to a second order problem. At the outset we mention that there are alternatives to treating a second order problem in the way we will indicate.

(1) If a second order problem is an Euler equation for a first order variational principal, we may deal with the underlying problem using only first order derivatives as we indicated in Chapter 9

(2) In any case we can always convert a second order problem into a system of first order equations by introducing one or more unknowns. Frequently there are several ways to accomplish this.

Example 5. In any case if we are determined to treat a second order problem directly, we may proceed as follows. We will suppose that

$$\Omega = [0, 1] \times [0, 1]$$

and that our second order problem is Laplace's equation on Ω. We do not particularly recommend what follows; it is for purposes of illustration only. Take the grid as in Example 4. Take the linear space H to consist of all real-valued functions on this grid. Define D_0, D_1, D_2 on H so that if $u \in H$, then D_0u, D_1u, D_2u are as in Example 4 (all norms and inner products without subscripts in this example are Euclidean) and

$$(D_{11}u)_{i,j} = (u_{i+1,j} - 2u_{i,j} + u_{i-1,j})/\delta^2$$
$$(D_{12}u)_{i,j} = (u_{i+1,j+1} - u_{i-1,j+1} - u_{i+1,j-1} + u_{i-1,j-1})/\delta^2,$$

$$(D_{22}u)_{i,j} = (u_{i,j+1} - 2u_{i,j} + u_{i,j-1})/\delta^2 i, j = 1, 2, ..., n-1.$$

Take

$$Du = (D_0u, D_1u, D_2u, D_{11}u, D_{12}u, D_{22}u),\ u \in H,$$

and take, for $u \in H$, $\|u\|_D^2 =$

$$\|D_0u\|^2 + \|D_0u\|^2 + \|D_1u\|^2 + \|D_2u\|^2 + \|D_{11}u\|^2 + \|D_{12}u\|^2 + \|D_{22}u\|^2 \tag{8.22}$$

where all norms without subscript are Euclidean. Define ϕ on H so that

$$\phi(u) = (1/2)\sum_{i,j=1}^{n-1}((D_{11}u)_{i,j} + (D_{22}u)_{i,j})^2,\ u \in H.$$

Thus if $u, h \in H$,

$$\phi'(u)h = \sum_{i,j=1}^{n-1}((D_{11}u)_{i,j}(D_{11}h)_{i,j} + (D_{22}u)_{i,j}(D_{22}h)_{i,j})$$

and so

$$\begin{aligned}\phi'(u)h &= \langle Dh, (0,0,0,(D_{11}u),0,(D_{11}u))\rangle \\ &= \langle Dh, P(0,0,0,(D_{11}u),0,(D_{22}u))\rangle \\ &= \langle h, \pi P(0,0,0,(D_{11}u),0,(D_{11}u))\rangle_D\end{aligned}$$

and so

$$(\nabla_D\phi)(u) = \pi P(0,0,0,(D_{11}u),0,(D_{11}u))$$

where $\langle\ ,\ \rangle_D$ denotes the inner product derived from (8.22). In this particular case, of course, the work involved in constructing P is at least as much as that of solving Laplace's equation directly.

In [**87**], there is described a computer code which solves a general second order quasi-linear partial differential equation on an arbitrary grid whose intervals are parallel to the axes of R^2.

If boundary conditions on $\partial\Omega$ are required, then the above constructions are modified. The various vectors h appearing above are to be in

$$H_0 = \{u \in H : u_{i,j} = 0 \text{ if one of } i \text{ or } j = 0 \text{ or } n\}.$$

Then the resulting gradient $\nabla_D\phi$ will be a member of H_0.

Example 6. How one might treat a system of differential equations is illustrated by the following:

Consider the pair of equations

$$u'(t) = f(u(t), v(t)),\ v'(t) = g(u(t), v(t)),\ t \in [0,1],$$

where $f, g : R^2 \to R$ are of class C^1. Consider

$$\phi : H = H^{1,2}([0,1]) \to R$$

defined by

$$\phi(u,v) = (1/2)\int_0^1 ((u' - f(u,v)^2 + (v' - g(u,v)^2),\ u,v \in H.$$

Take $Du = \binom{u}{u'},\ u \in H$. Calculate:

$$\phi'(u,v)(h,k) = \int_0^1 ((u' - f(u,v)(h' - f_1(u,v)h - f_2(u,v)k)$$

$$+(v' - g(u,v)(k' - g_1(u,v)h - g_2(u,v)k)) = \langle \binom{Dh}{Dk}, \binom{r}{s} \rangle^2_{L_2([0,1])}, \tag{8.23}$$

where

$$r = \begin{pmatrix} -(u' - f(u,v))f_1(u,v) - ((v' - g(u,v))g_1(u,v) \\ (u' - f(u,v) \end{pmatrix}$$

and

$$s = \begin{pmatrix} -(u' - f(u,v))f_2(u,v) - ((v' - g(u,v))g_2(u,v) \\ v' - g(u,v) \end{pmatrix},$$

$u, v, h, k \in H$. Thus we have from (8.23)

$$\phi'(u,v)(h,k) = \binom{Pr}{Ps}, \tag{8.24}$$

where P is the orthogonal projection of $L_2([0,1])^2$ onto

$$\{\binom{u}{u'} : u \in H\}.$$

Hence we have that

$$(\nabla\phi)(u,v) = \binom{\pi Pr}{\pi Ps}$$

where $\pi\binom{\alpha}{\beta} = \alpha,\ \alpha, \beta \in H$.

Now let us incorporate boundary conditions, say

$$u(0) = 1, v(1) = 1.$$

Then in (8.23),

$$h \in H_1 = \{\alpha \in H : \alpha(0) = 0\},$$

$$k \in H_2 = \{\beta \in H : \beta(1) = 0\},$$

and instead of (8.24) we have

$$\phi'(u,v)(h,k) = \binom{P_0 r}{Q_0 s} \tag{8.25}$$

where here P_0 is the orthogonal projection of $L_2([0,1])^2$ onto

$$\{D\alpha : \alpha \in H_1\}.$$

and Q_0 is the orthogonal projection of $L_2([0,1])^2$ onto

$$\{D\beta : \beta \in H_2\}$$

Hence

$$(\nabla\phi)(u,v) = \binom{\pi P_0 r}{\pi Q_0 s}$$

For mixed boundary conditions, say $u(0) = v(1)$, $u(1) = 2v(0)$, we would have in place of (8.25)

$$\phi'(u,v)(h,k) = P_0\binom{r}{s}$$

where here P_0 is the orthogonal projection of $L_2([0,1])^4$ onto

$$\{\binom{D\alpha}{D\beta} : \alpha(0) - b(1) = 0, \alpha(1) - 2v(0) = 0\}.$$

Hence

$$(\nabla\phi)(u,v) = \pi P_0\binom{r}{s}$$

where here $\pi(\alpha,\beta,\gamma,\delta) = (\alpha,\gamma)$.

In this latter case the boundary conditions are coupled between the two components of the solution and a single orthogonal projection on $L_2([0,1])^4$ is to be calculated rather than two individual projections on $L_2([0,1])^2$. Finite difference considerations are similar to those of other examples in this chapter.

In [**4**] the use of finite elements in construction and computation with Sobolev gradients is developed. Applications are made to Burger's equation and the time-independent Navier-Stokes equations.

CHAPTER 9

Sobolev Gradients for Weak Solutions: Function Space Case

Suppose Ω is a bounded open subset of R^n with piecewise smooth boundary and F is a $C^{(1)}$ function from $R \times R^n$ so that if

$$\phi(u) = \int_\Omega F(u, \nabla u),\ u \in H = H^{1,2}(\Omega), \tag{9.1}$$

then ϕ is a $C^{(1)}$ function on H. For $w \in H$ consider the problem of finding a minimum (or maximum or critical point) of ϕ over all $u \in H$ such that $u - w = 0$ on $\partial\Omega$. Such an extremum u of ϕ has the property that

$$\phi'(u)h = \int_\Omega (F_1(u, \nabla u)h + \langle F_2(u, \nabla u), \nabla h\rangle_{R^n} = 0,\ h \in H_0, \tag{9.2}$$

where $H_0 = \{h \in H : h = 0 \text{ on } \partial\Omega\}$. If $u \in H$ satisfies (9.2) and has the property that $F_2(u, \nabla u) \in C^{(1)}(\Omega)$, then

$$\int_\Omega (F_1(u, \nabla u) - \nabla \cdot F_2(u, \nabla u))h = 0,\ h \in H_0, \tag{9.3}$$

i.e., the Euler equation

$$F_1(u, \nabla u) - \nabla \cdot F_2(u, \nabla u) = 0 \tag{9.4}$$

holds. Even if there is not sufficient differentiability for (9.4) to hold, it is customary to say that if (9.3) holds then u satisfies (9.4) in the weak sense.

What we intend to produce is a Sobolev gradient $\nabla\phi$ for ϕ so that

$$(\nabla\phi)(u) = 0$$

if and only if $u \in H$ satisfies (9.2). More generally, suppose $p > 1$, Ω is an open subset of R^n and F is a $C^{(1)}$ function from $R \times R^n$ to R so that the function ϕ defined by

$$\phi(u) = \int_\Omega F(u, \nabla u),\ u \in H^{1,p}(\Omega)$$

has the property that it is a $C^{(1)}$ function from $H = H^{1,p}(\Omega)$ to R. Denote by H_1 a closed subspace of H. Then $u \in H$ is a critical point of ϕ relative to H_1 provided that

$$\phi'(u)h = \int_\Omega (F_1(u, \nabla u)h + \langle F_2(u, \nabla u), \nabla h\rangle_{R^n} = 0,\ h \in H_1.$$

For $u \in H$, the functional $\phi'(u) \in H_1^*$ and H_1 is a subspace of $H^{1,p}(\Omega)$. Accordingly, since $p > 1$, there is a unique point (cf [**34**]) $h \in H_1$ so that

$$\phi'(u)h \text{ is maximum subject of } \|h\|_{H^{1,p}(\Omega)} = |\phi'(u)|. \tag{9.5}$$

Such a point h is denote by $(\nabla\phi)(u)$ according the criterion (1.3). This will be considered in some detail in chapter 10.

Explicit construction of such gradients in finite dimensional spaces is considered in Chapter 8. Applications including some to transonic flow and minimal surfaces are given in Chapters 11,12,13. In case $p = 2$ we obtain below an explicit expression for gradients in the present function space setting. Before giving this construction we make some general observations.

For $p = 2$ and $\nabla\phi$ a gradient constructed from (9.5), there is the possibility of finding a critical point of ϕ by finding a zero of α :

$$\alpha(u) = \|(\nabla\phi)(u)\|_H^2/2, \ u \in H \tag{9.6}$$

(relative to the subspace H_1 of H) provided that $\nabla\phi \in C^{(1)}$. A gradient $\nabla\alpha$ for α is constructed as in (4.14) with ϕ replaced by α and F replaced by $\nabla\phi$. Perhaps using results from Chapter 4, one may seek a zero of α as

$$u = \lim_{t\to\infty} z(t),$$

where z satisfies

$$z(0) = x \in H, z'(t) = -(\nabla\alpha)(z(t)), \ t \geq 0.$$

Alternatively, discrete steepest descent may be used on α. In this way one may attempt to calculate weak solutions of the variational problem specified by ϕ and H_1.

We now develop an explicit expression for such gradients in case $p = 2$. Our finite dimensional results in Chapters 10,13 will indicate similar results in the general case where $p > 1$.

We use the notation:

$$Du = \binom{u}{\nabla u}, \ u \in H^{1,2}(\Omega).$$

THEOREM 37. *Suppose Ω is an open subset of (or the closure of an open subset of) R^n and F is a $C^{(1)}$ function from $R \times R^n$ so that ϕ is defined by* (9.1) *and is a $C^{(1)}$ function on $H = H^{1,2}(\Omega)$. Suppose furthermore that H_1 is a closed subspace of H. Then*

$$(\nabla\phi)(u) = \pi P((\nabla F)(Du))$$

where P is the orthogonal projection of $L_2(\Omega) \times L_2(\Omega)^n$ onto

$$\{\binom{v}{\nabla v} : v \in H_1\}$$

and $\pi\binom{f}{g} = f$, $f \in L_2(\Omega)$, $g \in L_2(\Omega)^n$.

PROOF. Suppose $u \in H$, $h \in H_1$. Then

$$\begin{aligned}\phi'(u)h &= \int_\Omega F'(Du)Dh \\ &= \langle (\tfrac{h}{\nabla h}), (\nabla F)(Du)\rangle_{L_2(\Omega)^{n+1}} \\ &= \langle (\tfrac{h}{\nabla \phi}), P((\nabla F)(Du))\rangle_{L_2(\Omega)^{n+1}} \\ &= \langle h, \pi P((\nabla F)(Du))\rangle_H\end{aligned}$$

and so

$$(\nabla\phi)(u) = \pi P((\nabla F)(Du)).$$

□

THEOREM 38. *Suppose that in addition to the hypothesis of Theorem 37,*

$$\phi(u) \geq 0,\ u \in H.$$

Suppose also that $x \in H$ *and*

$$z(0) = x,\ z'(t) = -(\nabla\phi)(z(t)),\ t \geq 0.$$

Then there is $t_1, t_2, \ldots . \to \infty$ *so that* $\lim_{n\to\infty}(\nabla\phi)(z(t_n)) = 0$.

PROOF. This follows immediately from the fact that

$$\phi(x) \geq \int_0^t \|(\nabla\phi)(z)\|^2,\ t \geq 0.$$

□

THEOREM 39. *Suppose that in addition to the hypothesis of Theorem 37 that* $\phi \in C^{(2)}$ *and that* α *is defined by* (9.6). *Then*

$$(\nabla\alpha)(u) = (\nabla\phi)'(u)(\nabla\phi)(u), u \in H. \tag{9.7}$$

PROOF. For $u \in H$, $h \in H_1$,

$$\alpha'(u)h = \langle(\nabla\phi)'(u)h, (\nabla\phi)(u)\rangle_H = \langle h, (\nabla\phi)'(u)(\nabla\phi)(u)\rangle_H$$

so that

$$(\nabla\alpha)(u) = (\nabla\phi)'(u)(\nabla\phi)(u),\ u \in H$$

since $(\nabla\phi)'(u)$ is a symmetric member of $L(H, H)$ and $\phi \in C^{(2)}$. □

Note that since

$$(\nabla\phi)(u) = \pi P((\nabla F)(Du)),\ u \in H, \tag{9.8}$$

it follows that

$$(\nabla\phi)'(u)h = \pi P((\nabla F)'(Du)Dh),\ u, h \in H. \tag{9.9}$$

In the case $p = 2$ and $H_1 = H_0$ define W so that

$$Wu = \nabla u,\ u \in H_0.$$

Use formula (5.12) to express the orthogonal projection onto the graph of W. In the construction of P is in terms of inverses of the transformations $(I + W^tW)$ and $(I + WW^t)$ where here W^t denotes the adjoint of W in the usual sense (i.e., as a closed unbounded linear operator on an appropriate L_2 space). But these two operators are positive definite linear elliptic operators of a well understood nature. Thus we might consider the projection P to be understood also. In later chapters we will consider at some length construction of such projections in corresponding finite dimensional settings.

One can construct cases where the minimum of a functional ϕ is not a solution of the corresponding Euler equation but the minimum is a weak solution. Our corresponding Sobolev gradients are zero at precisely such weak solutions. The rather constructively defined gradients of this chapter give another means by which to search for extrema of ϕ.

We close this chapter with an example which will give a construction for $\nabla\phi$ and $\nabla\alpha$ in a specific case. Suppose that Ω is a bounded open set in R^2 with a piecewise smooth boundary (or the closure of such a region). Suppose also that G is a $C^{(2)}$ function from R to R so that the function ϕ defined by

$$\phi(w) = \int_\Omega (\|\nabla w\|^2/2 + G(w)),\ w \in H = H^{1,2}(\Omega) \tag{9.10}$$

is $C^{(1)}$. We propose to calculate both (9.7) and (9.8) in this special case.

Now

$$\phi'(w)h = \int_\Omega (\langle \nabla w, \nabla h\rangle_{R^2} + G'(w)h),\ h \in H.$$

Instead of integrating by parts to obtain an Euler equation we proceed in a way in which various functions are all in H. Note that

$$\phi'(w)h = \langle \binom{h}{h'}, \binom{G'(w)}{\nabla w}\rangle_{L_2(\Omega)^3} = \langle \binom{h}{h'}, P\binom{G'(w)}{\nabla w}\rangle_{L_2(\Omega)^3} = \langle h, \pi P\binom{G'(w)}{\nabla w}\rangle_H$$

where P is the orthogonal projection of $L_2(\Omega)^3$ onto $\{\binom{h}{\nabla h} : h \in H\}$ and $\pi\binom{f}{g} = f, f \in L_2(\Omega), g \in L_2(\Omega)^2$. Calculating further we note that

$$\pi P\binom{G'(w)}{\nabla w} = \pi P\binom{w}{\nabla w} + \pi P\binom{G'(w)-w}{0} = w + \pi P\binom{G'(w)-w}{0}.$$

First note that

$$\{\binom{u}{\nabla u} : u \in H\}^\perp = \{\binom{\nabla \cdot v}{v} : v \in H^{1,2}(\Omega)\}$$

so that if $f \in L_2(\Omega), g \in L_2(\Omega)^2$, there exist uniquely $u \in H, v \in H^{1,2}(\Omega)$ so that

$$u + \nabla\cdot v = f,\ \nabla u + v = g.$$

For fixed $w \in H$, we have that in the current example

$$f = G'(w) - w,\ g = 0$$

so $u \in H^{1,2}(\Omega)$ and will satisfy

$$-\nabla\cdot\nabla u + u = G'(w) - w.$$

We write

$$u = (I - \Delta)^{-1}(G'(w) - w)$$

with the understanding that the above inverse is taken with zero boundary conditions on u. Hence

$$(\nabla\phi)(w) = w + (I - \Delta)^{-1}(G'(w) - w). \tag{9.11}$$

Next we look for $\nabla\alpha$ where

$$\alpha(w) = \|(\nabla\phi)(w)\|_H^2/2, \; w \in H.$$

It follows that

$$(\nabla\phi)'(w)h = h + (I - \Delta)^{-1}(G'(w)h - h), \; w, h \in H. \tag{9.12}$$

We already have that

$$(\nabla\alpha)(w) = (\nabla\phi)'(w))(\nabla\phi)(w), \; w \in H \tag{9.13}$$

so that (9.11) and (9.12) combine to give an expression for $(\nabla\alpha)(w)$ using (9.13).

This problem grew out of discussions with the author's colleagues Alfonso Castro and Hank Warchall concerning the finding of non-zero solutions to the Euler equation associated with (9.10).

Steepest descent with (9.13) seems better than that with (9.11) for the following reason: It is expected that zeros of $\nabla\phi$ are isolated. Steepest descent calculations with $\nabla\alpha$ may take advantage of a basin of attraction associated with a zero of $\nabla\alpha$. On the other hand a steepest descent process with $\nabla\phi$ may overshoot a saddle point of ϕ.

For now we are content to let this example illustrate possible utility for Sobolev gradients for weak formulations. In Chapters 11, 12,13 we will return to the topic in a numerical setting.

CHAPTER 10

Sobolev Gradients in Non-inner Product Spaces: Introduction.

Many problems involving partial differential equations seem not to be placed naturally in Sobolev spaces $H^{m,p}(\Omega)$ for $p = 2$. Some important examples will be seen in Chapter 13 which deals with various transonic flow problems. In the present chapter we first introduce Sobolev gradients in finite dimensional emulations of $H^{1,p}([0,1])$. We seek to construct a gradient with respect to a finite dimensional version of an $H^{m,p}$ space for $p > 2$. As such it provides us with an example of a gradient using the third principle of Chapter 1.

Pick $p > 2$ and denote by n a positive integer. For a finite dimensional version of $H^{1,p}([0,1])$, denote by H the vector space whose points are those of R^{n+1} but with norm

$$\|y\|_H = (\sum_{i=0}^{n} |y_i|^p + \sum_{i=1}^{n} |y_i - y_{i-1}|^p)^{1/p}, \quad y = (y_0, y_1, ..., y_n) \in R^{n+1}. \tag{10.1}$$

This expression differs from (2.9) even if we choose $p = 2$. The present choice gives us somewhat nicer results than the direct analogy of (2.9).

We define a function ϕ as follows. First define D_0, D_1 as in (8.1) and denote by α a real valued $C^{(2)}$ function from $R^n \times R^n$. Define $\phi : R^{n+1} \to R$ by

$$\phi(y) = \alpha(D_0 y, D_1 y) = \alpha((y_1 + y_0)/2, ..., (y_n + y_{n-1})/2; (y_1 - y_0)/\delta, \ldots, (y_n - y_{n-1})/\delta), y = (y_0, y_1, ..., y_n) \in R^{n+1}.$$

Consider the problem of determining $h \in R^{n+1}$ so that

$$\phi'(y)h \text{ is maximum subject to } \|h\|_H^p - |\phi'(y)|^p = 0 \tag{10.2}$$

where

$$|\phi'(y)| = \sup_{g \in R^{n+1}, g \neq 0} |\phi'(y)g| / \|g\|_H.$$

We need some notation. Fix $y \in R^{n+1}$ and choose $A, B \in R^n$ so that the linear functionals

$$\alpha_1(D_0 y, D_1 y), \alpha_2(D_0 y, D_1 y)$$

have the representations

$$\alpha_1(D_0 y, D_1 y)k = \langle k, A \rangle_{R^n}$$
$$\alpha_2(D_0 y, D_1 y)k = \langle k, B \rangle_{R^n}, \; k \in R^n.$$

Then we have that

$$\phi'(y)h = \langle D_0 h, A\rangle_{R^n} + \langle D_1 h, B\rangle_{R^n}$$
$$= \langle h, D_0^t A + D_1^t B\rangle_{R^{n+1}} = \langle h, q\rangle_{R^{n+1}},\ h \in R^{n+1} \quad (10.3)$$

where

$$q = (\nabla\phi)(y) = D_0^t A + D_1^t B.$$

We proceed to find a unique solution to (10.2) and we define a gradient of ϕ at $y, (\nabla_H \phi)(y)$, to be this unique solution. Note that there is some $h \in R^{n+1}$ which satisfies (10.2) since

$$\{h \in R^{n+1} : \|h\|_H^p = |\phi'(y)|^p\}$$

is compact. Define $\beta, \gamma : H \to R$ by

$$\beta(h) = \phi'(y)h,\ \gamma(h) = \|h\|_H^p - |\phi'(h)|^p,\ h \in R^{n+1}.$$

In order to use Lagrange multipliers to solve (10.2) we first calculate conventional (*i.e.*, R^{n+1}) gradients of β, γ. Using (10.3) we have that

$$(\nabla\beta)(h) = (D_0^t A + D_1^t B) = (\nabla\phi)(y) = q, h \in R^{n+1}.$$

Define Q so that $Q(t) = |t|^{p-2}, t \in R$. By direct calculation,

$$(\nabla\gamma)(h) = (\gamma^{(0)}(h), \gamma^{(1)}(h), \ldots, \gamma^{(n)}(h))$$

where

$$\gamma^{(0)}(h) = Q(h_0) - Q((h_1 - h_0)/\delta)/\delta,\ \gamma^{(n)}(h) = Q(h_n) + Q((h_n - h_{n-1})/\delta)/\delta,$$
$$\gamma^{(i)}(h) = Q(h_i) + (Q((h_i - h_{i-1})/\delta) - Q((h_i - h_{i+1})/\delta))/\delta,$$
$$i = 1, \ldots, n-1,\ h \in R^{n+1}. \quad (10.4)$$

This can be written more succinctly as

$$(\nabla\gamma)(h) = E^t(Q(Eh)) \quad (10.5)$$

where

$$E(h) = \binom{h}{D_1 h}, h \in R^{n+1}$$

and the notation $Q(Eh)$ denotes the member of R^{n+1} which is obtained from Eh by taking Q of each of its components. The rhs of expression (10.5) gives a finite dimensional version of the p-Laplacian associated with the embedding of $H^{1,p}([0,1])$ into L_p and is denoted by $\Delta_p(h)$. Some references to p-Laplacians are [**45**].

The theory of Lagrange multipliers asserts that for any solution h to (10.2) (indeed of any critical point associated with that problem), it must be that $(\nabla\gamma)(h)$ and $(\nabla\beta)(h) = (\nabla\phi)(y)$ are linearly dependent. We will see that there are just two critical points for (10.2), one yielding a maximum and the other a minimum.

Lemma 11. *If $h \in R^{n+1}$ the Hessian of γ at $h, (\nabla\gamma)'(h)$ is positive definite unless $h = 0$. Moreover $(\nabla\gamma)'$ is continuous.*

Indication of proof. Using (10.4), we see that for $h \in R^{n+1}$, $h \neq 0, (\nabla\gamma)'(h)$ is symmetric and strictly diagonally dominant. Thus $(\nabla\gamma)'(h)$ must be positive definite. Clearly $(\nabla\gamma)'$ is continuous.

THEOREM 40. *If $g \in R^{n+1}$, there is a unique $h \in R^{n+1}$ so that*

$$(\nabla\gamma)(h) = g. \tag{10.6}$$

Moreover there is a number λ so that if $g = \lambda q$, then the solution h to (10.6) *solves* (10.2).

PROOF. Since $(\nabla\gamma)'(h)$, $h \in R^{n+1}$, $h \neq 0$ is positive definite, it follows that γ is strictly convex. Now pick $g \in R^{n+1}$ and define

$$\eta(h) = \gamma(h) - \langle h, g\rangle_{R^{n+1}},\ h \in R^{n+1}. \tag{10.7}$$

Since γ is convex it follows that η is also convex. Noting that η is bounded below, it is seen that η has a unique minimum, say h. At this element h we have that

$$\eta'(h)k = \gamma'(h)k - \langle k, g\rangle_{R^{n+1}} = 0,\ k \in R^{n+1}. \tag{10.8}$$

Since $\gamma'(h)k = \langle(\nabla\gamma)(h), k\rangle_{R^{n+1}}$, it follows from (10.8) that

$$(\nabla\gamma)(h) = g. \tag{10.9}$$

□

At a critical point h of (10.2),

$$(\nabla\gamma)(h) = \lambda(\nabla\phi)(y)$$

for some $\lambda \in R$ and hence

$$\begin{aligned} h &= (\nabla\gamma)^{-1}(\lambda(\nabla\phi)(y)) \\ &= Q^{-1}(\lambda)(\nabla\gamma)^{-1}((\nabla\phi)(y)). \end{aligned}$$

The condition that $\beta(h) = 0$ determines λ up to sign; one choice indicating a maximum for (10.2) and the other a minimum (pick the one which makes $\phi'(x)h$ positive).

The above demonstrates a special case of the following [**34**] as pointed out in [**104**]:

THEOREM 41. *Suppose X is a uniformly convex Banach space, f is a continuous linear functional on X and $c > 0$. Then there is a unique $h \in X$ so that fh is maximum subject to $\|h\|_X = c$.*

The space H above (R^{n+1} with norm (10.1)) is uniformly convex. Our argument for Theorem 40 gives rise to a constructive procedure for determining the solution of (10.9) in the special case (10.2). Higher dimensional analogues which generalize the material in Chapter 4 follow the lines of the present chapter with no difficulty. In [**104**] (which should be published separately) there are generalizations of a number of the propositions of Chapter 4 to spaces $H^{m,p}, p > 2$. In (10.5), the lhs does not depend on h and so that an effective solution of our maximization problem (10.2) depends on being able to solve, given $g \in R^{n+1}$, for h so that

$$\Delta_p(h) = g. \tag{10.10}$$

To this end we indicate a nonlinear version of the well-known Gauss-Seidel method (for solving symmetric positive definite linear systems) which are to be used to solve (10.10), given $g = (g_0, \ldots, g_n) \in R^{n+1}$. We seek $h \in R^{n+1}$ so that $E^t(Q(E(h))) = g$, that is, so that

$$\gamma^{(0)}(h) = Q(h_0) - Q((h_1 - h_0)/\delta)/\delta = g_0,$$
$$\gamma^{(n)}(h) = Q(h_n) + Q((h_n - h_{n-1})/\delta)/\delta = g_n$$
$$\gamma^{(i)}(h) = Q(h_i) + (Q((h_i - h_{i-1})/\delta) - Q((h_i - h_{i+1})/\delta))/\delta = g_i$$
$$i = 1, \ldots, n-1. \tag{10.11}$$

Now given $a, b, c \in R$, each of the equations individually

$$Q(x) - Q((a - x)/\delta)/\delta = b,$$
$$Q(x) + Q((x - a)/\delta)/\delta = b,$$
$$Q(x) + (Q((x - a)/\delta) - Q((x - c)/\delta))/\delta = b$$

has a unique solution x. Our idea for a nonlinear version of Gauss-Seidel for solving (10.11) consists in making an initial estimate for the vector h and then systematically updating each component in order by solving the relevant equation (using Newton's method), repeating until convergence is observed. Generalization to higher dimensional problems should be clear enough.

CHAPTER 11

The Superconductivity Equations of Ginzburg-Landau

1. Introduction

Sections 2-6 and 9 of this chapter present joint work with Robert Renka. It follows closely [**89**]. There is considerable current interest in finding minima of various forms of the Ginzburg-Landau (GL) functional. Such minima give an indication of electron density and magnetic field associated with superconductors. We are indebted to Jacob Rubinstein for our introduction to this problem. We have relied heavily on [**26**],[**27**],[**95**].

We will present a method for determining such minima numerically.

2. The GL Functional

Following [**26**],[**27**],[**95**] we take the following for the GL functional:

$$E(u,A) = E_0 + \int_\Omega (\|(\nabla - iA)u\|^2/2 + \|\nabla \times A - H_0\|^2/2 + (\kappa^2 V(u)) \tag{11.1}$$

where $V(z) = (1/4)(|z|^2-1)^2, z \in C$. The unknowns are

$$u \in H^{1,2}(\Omega, C), \quad A \in H^{1,2}(\Omega, R^3),$$

and the following are given:

$$E_0 \in R, \quad H_0 \in C(\Omega, R^3), \quad \kappa \in R,$$

with Ω a bounded region in R^3 with regular boundary. H_0 represents an applied magnetic field and κ is a constant determined by material properties.

We seek to minimize (11.1) without imposing boundary conditions or other constraints on u, A in (11.1).

We attempt now to contrast our point of view with previous treatments of the minimization problem for (11.1). In [**26**] a Fréchet derivative for (11.1) is taken:

$$E'(u,A)\begin{pmatrix} v \\ B \end{pmatrix}, \quad u,v \in H^{1,2}(\Omega, C), \quad A, B \in H^{1,2}(\Omega, R^3). \tag{11.2}$$

An integration by parts is performed resulting in the GL **equations** — the Euler equations associated with (11.1) together with the natural boundary conditions

$$(\nabla \times A) \times \nu = H_0 \times \nu, \quad ((\nabla - iA)u) \cdot \nu = 0 \tag{11.3}$$

on $\Gamma = \partial\Omega$ where ν is the outward unit normal function on Γ. (We do not write out the GL equations here since they will not be needed.) Then one seeks a minimum of (11.1) by solving the Euler equation resulting from (11.2) subject to the boundary conditions (11.3).

3. A Simple Example of a Sobolev Gradient

In this section we give an example of a Sobolev gradient for a simple and familiar functional. We hope that this will make our construction for the GL functional easier to follow.

Denote by g a member of $K = L_2([0,1])$ and by ϕ the functional on $H = H^{1,2}([0,1])$:

$$\phi(y) = \int_0^1 (({y'}^2 + y^2)/2 - yg), \quad y \in H. \tag{11.4}$$

Define $D : H \to K \times K$ by $Du = \begin{pmatrix} u \\ u' \end{pmatrix}, \quad u \in H$. Then

$$\begin{aligned} \phi'(y)h &= \int_0^1 (h'y' + hy - hg) = \langle \tbinom{h}{h'}, \tbinom{y-g}{y'} \rangle_{K\times K} \\ &= \langle Dh, P\tbinom{y-g}{y'} \rangle_{K\times K} = \langle h, \pi P\tbinom{y-g}{y'} \rangle_H \end{aligned} \tag{11.5}$$

where P is the orthogonal projection of $K \times K$ onto

$$R(D) = \left\{ \begin{pmatrix} u \\ u' \end{pmatrix} : u \in H \right\}$$

and $\pi\tbinom{f}{k} = f, \quad f, k \in K$.

Define a Sobolev gradient as the function on H so that if $y \in H$, then $(\nabla_S \phi)(y)$ is the element of H which represents the functional $\phi'(y)$:

$$(\nabla_S \phi)(y) = \pi P\tbinom{y-g}{y'}. \tag{11.6}$$

We see $y \in H$ is a critical point of ϕ if and only if

$$(\nabla_S \phi)(y) = 0.$$

Thus a search for a critical point of ϕ is reduced to the problem of finding a zero of $\nabla_S \phi$.

Contrast this with the familiar strategy of writing (11.4) as

$$\phi'(y)h = \int_0^1 h(-y'' + y - g) \quad + \quad hy'|_0^1, \; h \in H \tag{11.7}$$

(making the assumption along the way that $y \in H^{2,2}$) and then seeking a critical point satisfying the Euler equation

$$-y'' + y = g \tag{11.8}$$

together with the natural boundary conditions

$$y'(0) = 0 = y'(1). \tag{11.9}$$

By contrast, using the Sobolev gradient (11.6) a critical point may be sought (successfully) by either continuous steepest descent

$$z(0) = z_0 \in H, \quad z'(t) = -(\nabla_S \phi(z(t)), \quad t \geq 0 \tag{11.10}$$

or discrete steepest descent

$$z_0 \in H, \quad z_n = z_{n-1} - \delta_{n-1}(\nabla_S \phi)(z_{n-1}), \quad n = 1, 2, \ldots \tag{11.11}$$

where $\delta_0, \delta_1, \ldots$ are chosen optimally.

At the risk of running this example into the ground, we express the above in a notation which is similar to what we want to use for the GL functional.

Define $F : R^2 \times R \to R$ by

$$F((x, y), w) = (x^2 + y^2)/2 - xw, \quad x, y, w \in R. \tag{11.12}$$

Then for a given $g \in K$, (11.4) may be written

$$\phi(y) = \int_0^1 F(Dy, g), \quad y \in H. \tag{11.13}$$

The Fréchet derivative ϕ' may then be expressed

$$\begin{aligned} \phi'(y)h &= \int_0^1 F_1(Dy, g)Dh = \langle Dh, (\nabla_1 F)(Dy, g)\rangle_{K\times K} \\ &= \langle Dh, P(\nabla_1 F)(Dy, g)\rangle_{K\times K} \\ &= \langle h, \pi P(\nabla_1 F)(Dy, g)\rangle_H, \quad h, y \in H, \end{aligned} \tag{11.14}$$

where F_1 indicates the partial Fréchet derivative of F in its first argument and, for $p \in R^2, q \in R$,

$$F_1(p, q)s = \langle s, (\nabla_1 F)(p, q)\rangle_{R^2}, \quad s \in R^2. \tag{11.15}$$

Now $\pi P(\nabla_1 F)(Dy, g)$ is just another expression for $(\nabla_S \phi)(y), \quad y \in H$.

4. A Sobolev Gradient for GL.

The boundary conditions (11.3) are relatively complicated and somewhat unusual. Our Sobolev gradient construction avoids explicit consideration of these boundary conditions. Just as they arise naturally in the conventional method of the calculus of variations, they can be avoided naturally in the Sobolev gradient method. We will try to make this clear in what follows. We revert back to the notation of Section 2.

Although it is not necessary to do so, we convert (11.1) into an equivalent real form. This will simplify our explanation and fit more closely with our background reference [**71**].

Define

$$G : H = H^{1,2}(\Omega, R^2) \times H^{1,2}(\Omega, R^3) \to R \tag{11.16}$$

by

$$G(\left(\begin{smallmatrix} r \\ s \end{smallmatrix}\right), A) = E(u, A) - E_0, \quad u = r + is \in H^{1,2}(\Omega, C), \quad A \in H^{1,2}(\Omega, R^3).$$

Then there is

$$F : r^{20} \times R^3 \to R \tag{11.17}$$

(r^{20} denotes $(R^2 \times R^6) \times (R^3 \times R^9)$) so that

$$G(w) = \int_\Omega F(Dw, H_0), \quad w = (\binom{r}{s}, A) \in H \tag{11.18}$$

where

$$Dw = D(\binom{r}{s}, A) = ((\binom{r}{s}, \binom{\nabla r}{\nabla s}), (A, \nabla A)). \tag{11.19}$$

Note that for such $w \in H$, Dw is a function from $\Omega \to r^{20}$. A Fréchet derivative of G is expressed as

$$G'(w)k = \int_\Omega F_1(Dw, H_0)Dk, \quad w, k \in H$$

where the subscript 1 above denotes partial Fréchet differentiation in the first argument of F. Further,

$$G'(w)k = \langle Dk, (\nabla_1 F)(Dw, H_0)\rangle_J, \quad w, k \in H \tag{11.20}$$

where J denotes $L_2(\Omega)^{20}$ and for $p \in r^{20}, q \in R^3$, $(\nabla_1 F)(p, q)$ here represents the element of r^{20} so that

$$F_1(p, q)h = \langle h, (\nabla_1 F)(p, q)\rangle_{r^{20}}, \quad h \in r^{20}.$$

Note that

$$\{Dk : k \in H\}$$

forms a closed subspace of $L_2(\Omega, r^{20})$. Denote by P the orthogonal projection of $L_2(\Omega, r^{20})$ onto $R(D)$. Note that $R(D)$ with the standard norm in $L_2(\Omega, r^{20})$ is isometrically isomorphic with $H' = H^{1,2}(\Omega, R^5)$.

Returning to (11.20) we have that

$$\begin{aligned} G'(w)k &= \langle Dk, (\nabla_1 F)(Dw, H_0)\rangle_J = \langle PDk, (\nabla_1 F)(Dw, H_0)\rangle_J \\ &= \langle Dk, P(\nabla_1 F)(Dw, H_0)\rangle_J = \langle k, \pi P(\nabla_1 F)(Dw, H_0)\rangle_{H'} \end{aligned}$$

where, for $q \in H$, $\pi Dq = q$. (Note that for $w \in H, P((\nabla_1 F)(Dw, H_0))$ is of the form Dq for some $q \in H$.)

We define a Sobolev gradient function for G as

$$\nabla_S G(w) = \pi P(\nabla_1 F)(Dw, H_0)), w \in H.$$

Our descent iteration is

$$w_0 \in H, w_n = w_{n-1} - \delta_{n-1}(\nabla_S G)(w_{n-1}), n = 1, 2, \ldots$$

where $\{\delta_n\}_{n=0}^\infty$ are chosen optimally.

In order to specify a numerical algorithm precisely we need two things: a discretization scheme for G and some illumination concerning P both as defined and in a corresponding discretized version. This is the point of the following section.

5. Finite Dimensional Emulation

We follow the development of Section 4 for a numerical setting for GL. We take Ω to be a square domain in R^2. Choose a positive integer n. Consider the rectangular grid Ω_n on Ω obtained by dividing each side of Ω into n pieces of equal length. Denote by H_n the collection of all functions from Ω_n to R^4. For $v \in H_n$, denote by $D_n v$ the proper analogue of (11.19) (values corresponding to divided differences are considered attached to centers of grid squares and function values at cell centers are obtained by averaging grid-point values). The calculations following (11.19) have their precise analogy in this finite dimensional setting: for F as used in (11.18), there is a function G_n which corresponds to a finite dimensional version of G in (11.18). Thus (11.18) corresponds to

$$G_n(w) = \sum_{i,j=1,\ldots,n} F((D_n w)_{(i,j)}, H_{0(i,j)}), \quad w \in H_n$$

which in turn gives that

$$\begin{aligned} G_n'(w)k &= \langle D_n k, (\nabla_1 F)(D_n w, H_0)\rangle_{J_n} \\ &= \langle k, D_n^t(\nabla_1 F)(D_n w, H_0)\rangle_{H_n'}, \quad w, k \in H_n. \end{aligned} \tag{11.21}$$

From (11.21) it follows that ∇G_n, the conventional gradient function for G_n, is specified by

$$(\nabla G_n)(w) = D_n^t(\nabla_1 F)(D_n w, H_0), \quad w \in H_n,$$

where H_n', J_n, P_n are the appropriate finite dimensional versions of H', J and P, respectively.

From (11.21) it follows that

$$G_n'(w)k = \langle k, \pi P_n(\nabla_1 F)(D_n w, H_0)\rangle_{H_n'}, \quad w, k \in H_n.$$

Hence the Sobolev gradient function $\nabla_S G_n$ is given by

$$(\nabla_S G_n)(w) = \pi P_n(\nabla_1 F)(D_n w, H_0), \quad w \in H_n.$$

To finish a description of how this Sobolev gradient is calculated note first that

$$P_n = D_n(D_n^t D_n)^{-1} D_n^t$$

since the range of P_n is a subset of the range of D_n, P_n is fixed on the range of D_n, P_n is symmetric and idempotent. This is enough to convict P_n of being the orthogonal projection onto the range of D_n. Thus

$$\begin{aligned} (\nabla_S G_n)(w) &= \pi D_n(D_n^t D_n)^{-1} D_n^t(\nabla_1 F)(D_n w, H_0) \\ &= (D_n^t D_n)^{-1}(\nabla G_n)(w) \end{aligned}$$

After computation of the standard gradient $\nabla G_n(w)$, an iterative method is used to solve the symmetric positive definite linear system for the discretized Sobolev gradient.

6. Numerical Results

Here we present some numerical results. We compare results using our Sobolev gradient with those obtained using the ordinary gradient. As indicated in 2, one should expect much better results using the Sobolev gradient. The following table reflects this. Results are for two distinct runs, one using a Sobolev gradient and the second using a conventional gradient, for each of 5 values of n, the number of cells in each direction (Ω is partitioned into n^2 square cells).

	Sobolev gradient			Standard gradient		
n	SD steps	LS iterations	Time	SD steps	Time	Speedup
10	12	209	4	1049	36	9.0
20	16	542	16	3345	413	25.8
30	13	606	35	6579	1913	54.7
40	16	768	77	10591	5666	73.6
50	12	870	122	15275	12632	103.5

The contour plot in Figure 1 matches closely one in Figure 1 of [**27**]. Considering only constant imposed magnetic fields H_0 and restricting oneself to square domains, there are essentially three parameters in (11.1): the value of H_0, the value of κ and the length of one side of the square. The length comes in as a factor by means of derivatives in the first and second terms of (11.1) so it seems to have at least as much influence as H_0 and κ. We do not dwell on physical interpretations of our work but rather stress that we have presented an efficient method for calculating minima of (11.1) that may be extended with no difficulty to a three dimensional setting.

The columns labeled 'SD steps' contain the numbers of steepest descent steps, and the column labeled 'LS iterations' contains the total number of linear solver iterations. Since the condition number of the linear systems increases with n, so does the number of linear solver iterations per descent step. All times in are in seconds on a PC with the Intel 80486-DX2/66 processor. The column labeled 'Speedup' contains the ratios of execution times. Note that the relative advantage of the Sobolev gradient over the standard gradient increases with problem size. Convergence is defined by an upper bound of 10^{-6} on the mean absolute (conventional) gradient component: $(n+1)^{-2}\|\nabla G_n(u, A)\|_{R^{(n+1)^2}}$. Parameter values are $\Omega = [0, 1] \times [0, 1], \kappa = 1., H_0(x, y) = 1., (x, y) \in \Omega$. The initial estimate in all cases was taken from $A = 0, u(x, y) = 1 + i, (x, y) \in \Omega$. Linear systems were solved by a conjugate gradient method in which the convergence tolerance was heuristically chosen to decrease as the descent method approached convergence. The line search consisted of univariate minimization in the search direction (negative gradient direction). The number of evaluations of the functional per descent step averaged 26.2 with the Sobolev gradient and 8.4 with the conventional gradient.

Figures 1-4 are contour plots depicting computed values of electron density $|u|^2$ and the magnetic field $\nabla \times A$ for each of two square domains: $[0, 1] \times [0, 1]$

and $[0, 10] \times [0, 10]$. In both cases, the material number is $\kappa = 1$, and the external magnetic field is $H_0 = 1$.

Plots are given in Section 10

7. A Liquid Crystal Problem

In [**33**], Garza deals with the problem of numerically determining critical points of

$$E(u) = \int_\Omega |\nabla u|^2 \tag{11.22}$$

where for some ρ in $(0, 1)$,

$$\Omega = \{(x, y, z) \in R^3 : \rho^2 \leq x^2 + y^2 \leq 1, 0 \leq z \leq 1\}$$

and $u \in H^{1,2}(\Omega, r^3)$ is subject to the condition that

$$||u(x, y, z)|| = 1, (x, y, z) \in \Omega \tag{11.23}$$

and that $u(x, y, z)$ be normal to the boundary of Ω for all $(x, y, z) \in \partial\Omega$.

According to [**6**], for ρ small enough there are two critical points to (11.22), one being a trivial one and the other being considerably more interesting. In [**33**] there is determined numerically a Sobolev gradient for finding critical points of (11.22). A unique feature of this gradient is that it respects (11.23) in the sense that continuous steepest descent preserves this condition. In [**6**], the full symmetry of Ω is required. On the other hand, Garza's numerical method does not depend on any particular symmetry of Ω. It thus seem appropriate for design purposes on regions of various shapes. See [**33**] for details.

8. An Elasticity Problem

In [**21**] there is considered the problem of finding critical points of

$$\phi(u) = \int_\Omega [||\nabla(u)||^2 + (det(\nabla(u))^{-1/2}], u \in H^{1,2}(\Omega, R^3), \tag{11.24}$$

where Ω is a bounded region in R^3 and $\nabla(u) : \Omega \to L(R^3, R^3)$ is the matrix valued representation of u'. Members $u \in H^{1,2}(\Omega R^3)$ which are a critical point of ϕ are sought so that the determinant in (11.24) are nonnegative. A Sobolev gradient for ϕ is constructed and critical points of it are found numerically. See [**21**] for more references and details.

9. Singularities for a Simpler GL functional

Work in this section is joint with Robert Renka and is taken from [**75**]. Suppose $\epsilon > 0$ and d is a positive integer. Consider the problem of determining critical points of the functional ϕ_ϵ:

$$\phi_\epsilon(u) = \int_\Omega (||\nabla(u)||^2/2 + (|u|^2 - 1)^2/(4\epsilon^2)), \; u \in H^{1,2}(\Omega, C), \; u(z) = z^d, \; z \in \partial\Omega, \tag{11.25}$$

where Ω is the unit closed disk in C, the complex numbers. For each such $\epsilon > 0$, denote by $u_{\epsilon,d}$ a minimizer of (11.25).

In **[7]** it is indicated that for various sequences $\{\epsilon_n\}_{n=1}^{\infty}$ of positive numbers converging to 0, precisely d singularities develop for $u_{\epsilon_n,d}$ as $n \to \infty$. The open problem is raised (Problem 12, page 139 of **[7]**) concerning possible orientation of such singularities. Our calculations suggest that for a given d there are (at least) two resulting families of singularity configurations. Each configuration is formed by vertices of a regular $d-$gon centered at the origin of C, with each corresponding member of one configuration being about .6 times as large as a member of the other. A family of which we speak is obtained by rotating a configuration through some angle α. That this results in another possible configuration follows from the fact (page 88 of **[7]**) that if

$$v_{\epsilon,d}(z) = e^{-id\alpha}u_{\epsilon,d}(e^{i\alpha}z), z \in \Omega,$$

then $\phi_\epsilon(v_{\epsilon,d}) = \phi_\epsilon(u_{\epsilon,d})$ and $v_{\epsilon,d}(z) = z^d, z \in \partial\Omega$.

That there should be singularity patterns formed by vertices of regular $d-$gons has certainly been anticipated although it seems that no proof has been put forward. What we offer here is some numerical support for this proposition. What surprised us in this work is the indication of *two* families for each positive integer d.

We explain how these two families were encountered. Our calculations use steepest descent with numerical Sobolev gradients. One family appears using discrete steepest descent and the other appears when continuous steepest descent is closely tracked numerically. We offer no explanation for this phenomenon but simply report it. For a given d, the family of singularities obtained with discrete steepest descent is closer to the origin (by about a factor of .6) than the corresponding family for continuous steepest descent. In either case, the singularities found are closer to the boundary of Ω for larger d. A source of computational difficulties might be that critical points of ϕ_ϵ are highly singular objects (for small ϵ, a graph of $|u_{\epsilon,d}|^2$ would appear as a plate of height one above Ω with d slim tornadoes coming down to zero). Moreover for each d as indicated above, one expects a continuum of critical points (one obtained from another by rotation) from which to 'choose'.

For calculations the region Ω is broken into pieces using some number (180 to 400, depending on d) of evenly spaced radii together with 40 to 80 concentric circles.

For continuous steepest descent, using $d = 2, \ldots, 10$ we started each steepest descent iteration with a finite dimensional version of $u_{\epsilon,d}(z) = z^d,\ z \in C$. To emulate continuous steepest descent, we used discrete steepest descent with small step size (on the order of .0001) instead of the optimal step size. In all runs reported on here we used $\epsilon = 1/40$ except for the discrete steepest descent run with $d = 2$. In that case $\epsilon = 1/100$ was used (for $\epsilon = 1/40$ convergence seemed not to be forthcoming in the single precision code used - the value .063 given is likely smaller than a successful run with $\epsilon = 1/40$ would give). Runs with somewhat larger ϵ yielded a similar pattern except the corresponding singularities were a little farther from the origin. In all cases we found d singularities arranged on a regular d-gon centered at the origin.

Results for continuous steepest descent are indicated by the following pairs:

$$(2,.15),(3,.25),(4,.4),(5,.56),(6,.63),(7,.65),(8,.7),(9,.75),(10,.775)$$

where a pair (d,r) above indicates that a (near) singularity of $u_{\epsilon,d}$ was found at a distance r from the origin with $\epsilon = 1/40$. In each case the other $d-1$ singularities are located by rotating the first one through an angle that is an integral multiple of $2\pi/d$. Results for discrete steepest descent are indicated by the following pairs:

$$(2,.063),(3,.13),(4,.18),(5,.29),(6,.34),(7,.39),(8,.44),(9,.48),(10,.5)$$

using the same conventions as for continuous steepest descent.

Computations with a finer mesh would surely yeild more precise results.

Some questions. Are there more than two (even infinitely many) families of singularities for each d? Does some other descent method (or some other method entirely) lead one to new configurations? Are there in fact configurations which are not symmetric about the origin?

We thank Pentru Mironescu for his careful description of this problem to JWN in December 1996 at the Technion in Haifa.

10. Some Plots

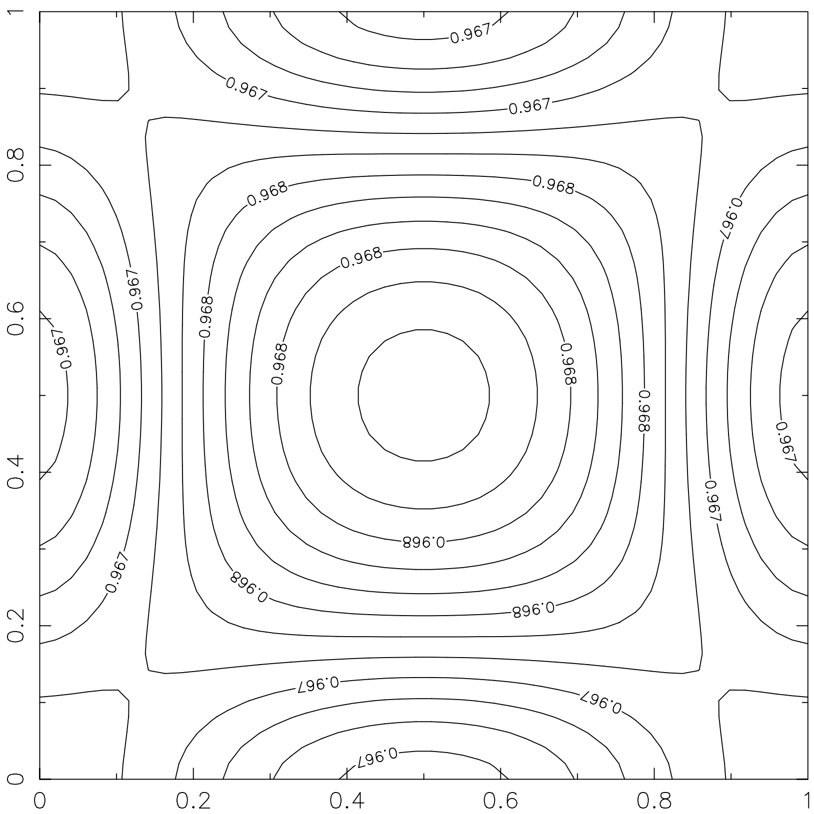

FIGURE 1. Electron density on the unit square

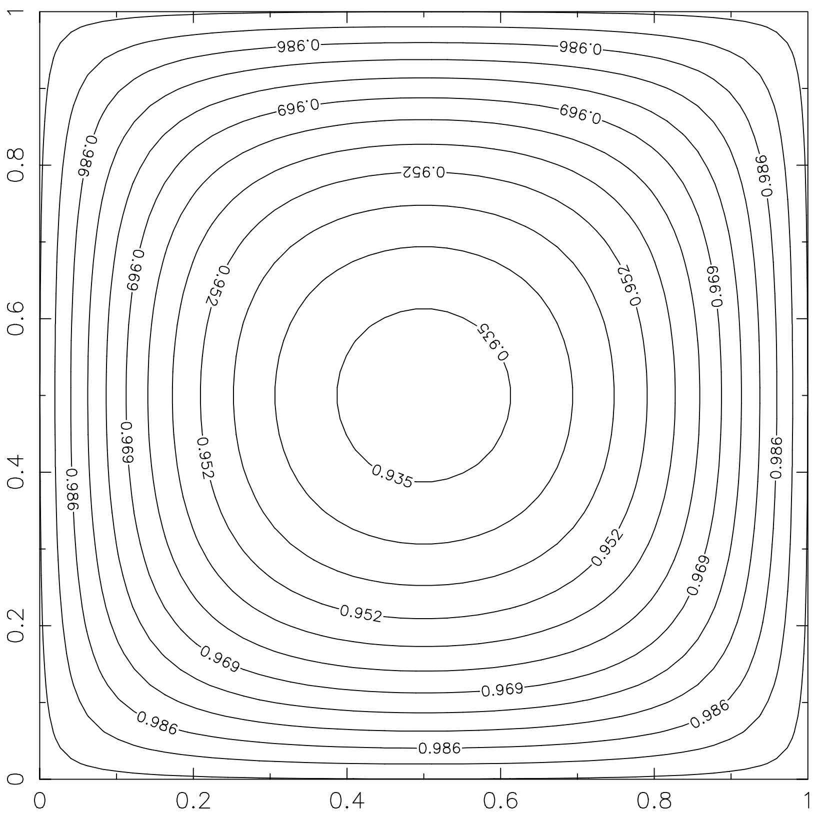

FIGURE 2. Magnetic field on the unit square

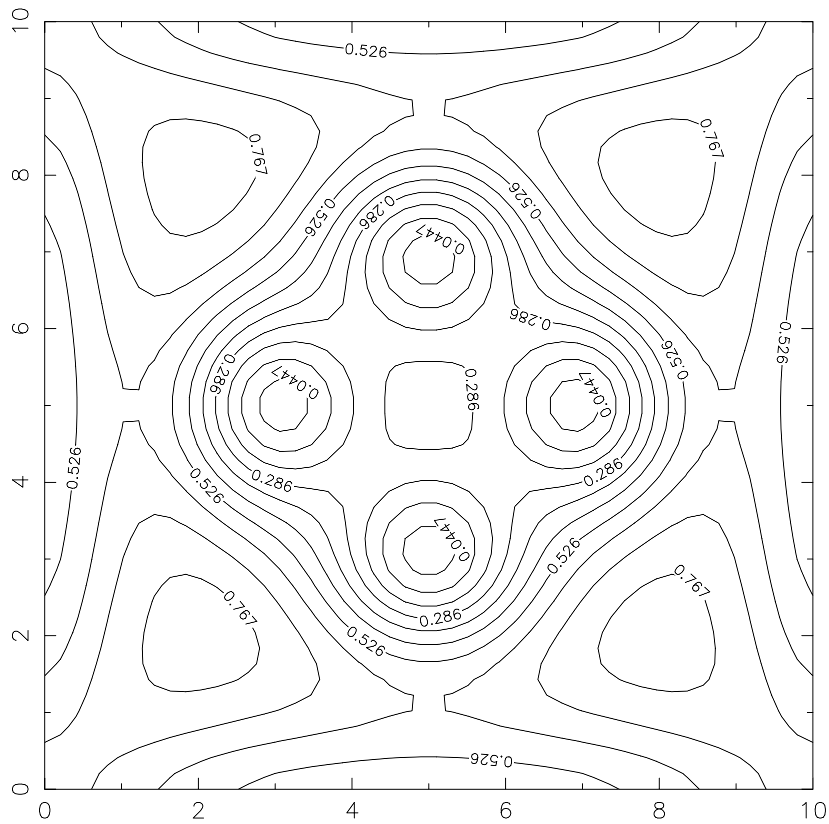

FIGURE 3. Electron density on $[0, 10] \times [0, 10]$

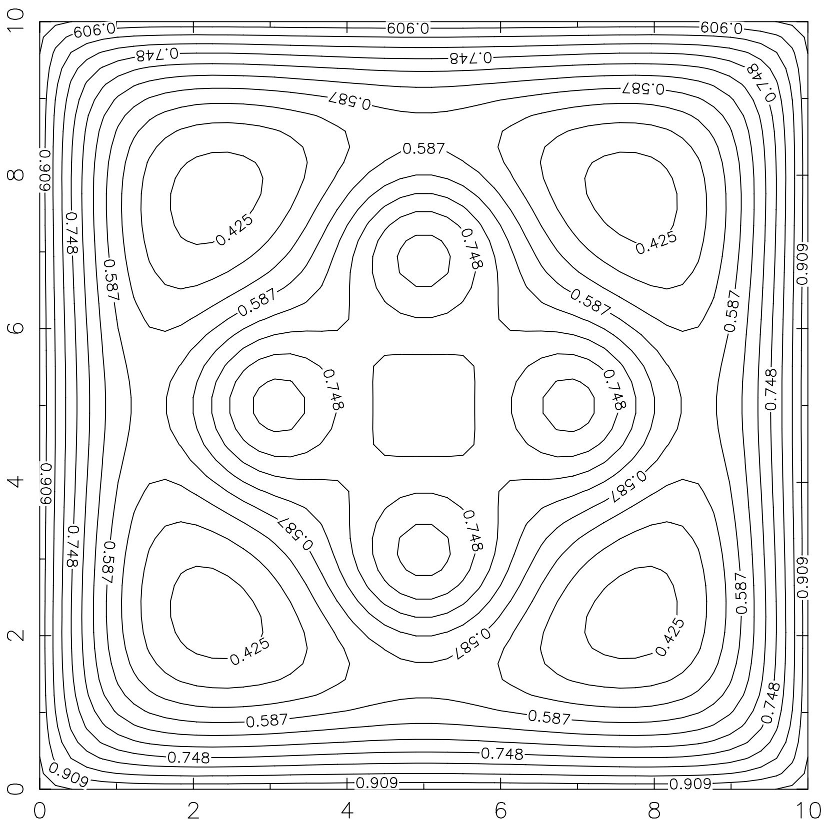

FIGURE 4. Magnetic field on $[0,10]\times[0,10]$

CHAPTER 12

Minimal Surfaces

1. Introduction

In this chapter we discuss an approach to the minimal surface problem by means of a descent method using Sobolev gradients on a structure somewhat similar to a Hilbert manifold. We begin with a rather detailed discussion of the problem of minimal length between two fixed points. This problem, of course, has the obvious solution but we hope that the explicit calculation in this case will reveal some of our ideas. The work of this chapter is recent joint work with Robert Renka and this written in [**88**].

2. Minimum Curve Length

Let S denote the set of smooth regular parametric curves on $[0,1]$; i.e.,

$$S = \{f : f \in C^2([0,1], \mathbf{R}^2) \text{ and } \|f'(t)\| > 0 \quad \forall\, t \in [0,1]\},$$

where $\|\cdot\|$ denotes the Euclidean norm on $\mathbf{R}^2$. Denote curve length $\phi : S \to \mathbf{R}$ by

$$\phi(f) = \int_0^1 \|f'\| = \int_0^1 s',$$

where s is the arc length function associated with f; i.e.,

$$s(t) = \int_0^1 \|f'\| \quad \forall\, t \in [0,1].$$

We now treat $f \in S$ as fixed and, using a steepest descent method with f as the initial estimate, we seek to minimize ϕ over functions that agree with f at the endpoints of $[0,1]$. Variations are taken with functions that satisfy zero end conditions:

$$S_0 = \{h : h \in C^2([0,1], \mathbf{R}^2) \text{ and } h(0) = h(1) = 0\}.$$

The derivative of $\phi(f)$ in the direction h is

$$\begin{aligned}
\phi'(f)h &= \lim_{\alpha\to 0}(1/\alpha)[\phi(f+\alpha h) - \phi(f)] = \lim_{\alpha\to 0}(1/\alpha)\int_0^1 (\|f' + \alpha h'\| - \|f'\|) \\
&= \lim_{\alpha\to 0}(1/\alpha)\int_0^1 \frac{\|f'+\alpha h'\|^2 - \|f'\|^2}{\|f'+\alpha h'\| + \|f'\|} = \lim_{\alpha\to 0}(1/\alpha)\int_0^1 \frac{2\alpha\langle f', h'\rangle + \alpha^2\|h'\|^2}{\|f'+\alpha h'\| + \|f'\|} \\
&= \int_0^1 \langle f', h'\rangle/\|f'\| = \int_0^1 \langle f', h'\rangle/s' \qquad \forall\, h \in S_0. \quad (12.1)
\end{aligned}$$

Note that $\phi'(f)h$ can be rewritten as a Stieltjes integral:

$$\phi'(f)h = \int_0^1 \langle f'/s', h'/s' \rangle\, s' = \int_0^1 \left\langle \frac{df}{ds}, \frac{dh}{ds} \right\rangle ds,$$

where $\frac{df}{ds} = f'/s'$ and $\frac{dh}{ds} = h'/s'$ are the derivatives of f and h with respect to the arc length function s (associated with f). Thus we have expressed $\phi'(f)h$ in a parameter-independent way in the sense that both the Stieltjes integral and the indicated derivatives depend only on the two curves involved and not on their parameterization.

We obtain a Hilbert space by defining an inner product on the linear space S_0. The gradient of ϕ at f depends on the chosen metric. We first consider the standard L_2 norm associated with the inner product

$$\langle g, h \rangle_{(L_2[0,1])^2} = \int_0^1 \langle g, h \rangle \qquad \forall\, g, h \in S_0.$$

Integrating by parts, we have

$$\begin{aligned}\phi'(f)h = \int_0^1 \langle f', h' \rangle / s' &= \int_0^1 \langle f'/s', h' \rangle \\ &= \int_0^1 \langle -(f'/s')', h \rangle = \langle -(f'/s')', h \rangle_{(L_2[0,1])^2} \quad \forall\, h \in S_0.\end{aligned}$$

Thus the representation of the linear functional $\phi'(f)$ in the L_2 metric is

$$\nabla \phi(f) = -(f'/s')'.$$

Note that the negative gradient direction (used by the steepest descent method) is toward the center of curvature; i.e.,

$$-\nabla \phi(f) = (f'/s')' = s' \kappa N$$

for curvature vector

$$\kappa N = \frac{d^2 f}{ds^2} = \frac{d}{ds}\left(\frac{f'}{s'}\right) = \frac{1}{s'}\left(\frac{f'}{s'}\right)' = \frac{f' \times f'' \times f'}{s'^4}.$$

We now consider a variable metric method in which the Sobolev gradient of ϕ at f is defined by an inner product that depends on f (but not the parameterization of f):

$$\langle g, h \rangle_f = \int_0^1 \langle g', h' \rangle / s' = \int_0^1 \left\langle \frac{dg}{ds}, \frac{dh}{ds} \right\rangle ds \quad \forall\, g, h \in S_0. \tag{12.2}$$

Let $g \in S_0$ denote the Sobolev gradient representing $\phi'(f)$ in this metric; i.e.,

$$\phi'(f)h = \langle g, h \rangle_f \quad \forall\, h \in S_0. \tag{12.3}$$

Then, from (12.1), (12.2), and (12.3),

$$\phi'(f)h = \int_0^1 \langle f', h' \rangle / s' = \int_0^1 \langle g', h' \rangle / s' \quad \forall\, h \in S_0$$

$$\Rightarrow \int_0^1 \left\langle \frac{f'-g'}{s'}, h' \right\rangle$$
$$= -\int_0^1 \left\langle \left(\frac{f'-g'}{s'}\right)', h \right\rangle = 0 \quad \forall\, h \in S_0$$
$$\Rightarrow (f'-g')/s' = c \quad \text{for some } c \in \mathbf{R}^2.$$

Hence $g(t) = \int_0^t g' = \int_0^t (f' - cs') = f(t) - f(0) - cs(t)$, where $g(1) = f(1) - f(0) - cs(1) = 0 \Rightarrow c = [f(1) - f(0)]/s(1)$; i.e.,

$$f(t) - g(t) = f(0) + \frac{s(t)}{s(1)}[f(1) - f(0)] \quad \forall\, t \in [0,1]. \tag{12.4}$$

The right hand side of (12.4) is the line segment between $f(0)$ and $f(1)$ parameterized by arc length s. Thus steepest descent with the Sobolev gradient g leads to the solution in a single iteration with step-size 1. While this remarkable result does not appear to extend to the minimal surface problem, our tests show that steepest descent becomes a viable method when the standard gradient is replaced by the (discretized) Sobolev gradient.

3. Minimal Surfaces

Denote the parameter space by $\Omega = [0,1] \times [0,1]$. The minimal surface problem is to find critical points of the surface area functional (subject to Dirichlet boundary conditions)

$$\phi(f) = \int_\Omega \|f_1 \times f_2\|, \quad f \in C^1(\Omega, \mathbf{R}^3), \quad f_1 \times f_2 \neq 0,$$

where f_1 and f_2 denote the first partial derivatives of f. This functional will be approximated by the area of a triangulated surface.

Define a triangulation T of Ω as a set of triangles such that

1. no two triangles of T have intersecting interiors,
2. the union of triangles of T coincides with Ω, and
3. no vertex of a triangle of T is interior to a side of a triangle of T.

Denote by V_T the set of all vertices of triangles of T, and let S_T be the set of all functions f from V_T to $\mathbf{R}^3$ such that, if $q - p$ and $r - p$ are linearly independent, then $f_q - f_p$ and $f_r - f_p$ are linearly independent for all $p, q, r \in V_T$ such that p is adjacent to q and r in the triangulation. Let Q be the set of all triples $\tau = [a,b,c] = [b,c,a] = [c,a,b]$ such that a, b,and c enumerate the vertices of a member of T in counterclockwise order. Denote the normal to a surface triangle by

$$f_\tau = (f_b - f_a) \times (f_c - f_a) = f_a \times f_b + f_b \times f_c + f_c \times f_a \quad \text{for } \tau = [a,b,c].$$

Note that $f_\tau \neq 0$ and the corresponding triangle area $\frac{1}{2}\|f_\tau\|$ is positive, where $\|\cdot\|$ now denotes the Euclidean norm on $\mathbf{R}^3$. Define surface area $\phi_T : S_T \to \mathbf{R}$

by

$$\phi_T(f) = \frac{1}{2} \sum_{\tau \in Q} \|f_\tau\| \,.$$

Now fix $f \in S_T$ and let $S_{0,T}$ denote the linear space of functions from V_T to $\mathbf{R}^3$ that are zero on the boundary nodes $V_T \cap \partial\Omega$. A straightforward calculation results in

$$\phi_T'(f)h = \frac{1}{2} \sum_{\tau \in Q} \langle f_\tau, (f,h)_\tau \rangle \,/\, \|f_\tau\| \quad \forall\, h \in S_{0,T}, \tag{12.5}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product on $\mathbf{R}^3$ and

$$\begin{aligned}(f,h)_\tau &= (f_b - f_a) \times (h_c - h_a) + (h_b - h_a) \times (f_c - f_a) \\ &= f_a \times h_b \,+\, h_a \times f_b \,+\, f_b \times h_c \,+\, h_b \times f_c \,+ \\ &\qquad f_c \times h_a \,+\, h_c \times f_a \quad \text{for } \tau = [a,b,c].\end{aligned}$$

It can be shown that the approximation to the negative L_2-gradient is proportional to the discretized mean curvature vector. Brakke has implemented a descent method based on this gradient [**11**]. However, for a metric on $S_{0,T}$, we choose the one related to the following symmetric bilinear function which depends on f:

$$\langle g, h \rangle_f = \frac{1}{4} \sum_{\tau \in Q} \langle (f,g)_\tau, (f,h)_\tau \rangle \,/\, \|f_\tau\| \quad \forall\, g, h \in S_{0,T}. \tag{12.6}$$

We will show that, at least for a regular triangulation T of Ω, ((12.6) defines a positive definite function and hence an inner product). To this end, let n be a positive integer and consider the uniform rectangular grid with $(n+1)^2$ grid points $\{(i/n, j/n)\}_{i,j=0}^{n}$. Then let T_n denote the triangulation of Ω obtained by using the diagonal with slope -1 to partition each square grid cell into a pair of triangles.

THEOREM 42. *For* $f \in S_T, \quad T = T_n, \quad \langle \cdot, \cdot \rangle_f$ *is positive definite on* $S_{0,T}$.

PROOF. Suppose there exists $h \in S_{0,T}$ such that $\langle h, h \rangle_f = 0$. Then $(f,h)_\tau = 0 \;\forall\, \tau \in Q$. It suffices to show that $h = 0$. Consider a pair of adjacent triangles indexed by $\tau_1 = [a,b,p]$ and $\tau_2 = [b,c,p]$ for which $h_a = h_b = h_c = 0$ so that $(f,h)_{\tau_1} = (f_b - f_a) \times h_p = 0$ and $(f,h)_{\tau_2} = (f_c - f_b) \times h_p = 0$. For every such pair of triangles in T_n, a, b, and c are not collinear, and $f_b - f_a$ and $f_c - f_b$ are therefore linearly independent. Hence, being dependent on both vectors, $h_p = 0$. The set of vertices p for which $h_p = 0$ can thus be extended from boundary nodes into the interior of Ω. More formally, let $B_0 = V_T \cap \partial\Omega$ and denote by B_k the union of B_{k-1} with $\{p \in V_T : \exists\, a,b,c \in B_{k-1} \text{ such that } [a,b,p],[b,c,p] \in Q\}$ for k as large as possible starting with $k = 1$. Then for some k, $B_k = V_T$ and, since $h_p = 0 \;\forall\, p \in B_k$, we have $h = 0$. ∎

Let $g \in S_{0,T}$ denote the Sobolev gradient representing $\phi_T'(f)$ in the metric defined by (12.6); i.e.,

$$\phi_T'(f)h = \langle g, h\rangle_f \quad \forall\, h \in S_{0,T}. \tag{12.7}$$

Then from (12.5), (12.6), and (12.7),

$$\phi_T'(f)h = \frac{1}{2}\sum_{\tau\in Q} \langle f_\tau, (f,h)_\tau\rangle \,/\, \|f_\tau\|$$

$$= \frac{1}{4}\sum_{\tau\in Q} \langle (f,g)_\tau, (f,h)_\tau\rangle \,/\, \|f_\tau\|$$

implying that

$$4\,\langle u, h\rangle_f = \sum_{\tau\in Q} \langle (f,u)_\tau, (f,h)_\tau\rangle \,/\, \|f_\tau\| = 0 \quad \forall\, h \in S_{0,T}, \tag{12.8}$$

where $u = f - g$ since $(f,u)_\tau = (f,f)_\tau - (f,g)_\tau = 2f_\tau - (f,g)_\tau$. For an alternative characterization of u, define $\beta(v) = \frac{1}{2}\|v\|_f^2$ $\forall v \in S_T$, and let v be the minimizer of β over functions in S_T that agree with f on $\partial\Omega$. Then $\beta'(v)h = \langle v, h\rangle_f = 0\;\forall\, h \in S_{0,T}$. This condition is uniquely satisfied by $v = u = f - g$.

The Sobolev gradient g used in the descent iteration is obtained from u which is defined by (12.8). We expand the left hand side of (12.8) as follows. For $\tau = [a,b,c]$,

$$(f,u)_\tau = u_a \times (f_b - f_c) \,+\, u_b \times (f_c - f_a) \,+\, u_c \times (f_a - f_b) \quad \text{and}$$
$$(f,h)_\tau = h_a \times (f_b - f_c) \,+\, h_b \times (f_c - f_a) \,+\, h_c \times (f_a - f_b).$$

Hence

$$\begin{aligned}
\langle (f,u)_\tau, (f,h)_\tau\rangle &= \langle h_a, (f_b - f_c)\times(f,u)_\tau\rangle + \langle h_b, (f_c - f_a)\times(f,u)_\tau\rangle \\
&\quad + \langle h_c, (f_a - f_b)\times(f,u)_\tau\rangle \\
&= \langle h_a, (f_b - f_c)\times u_a \times (f_b - f_c) \,+\, (f_b - f_c)\times \\
&\qquad [u_b \times (f_c - f_a) \,+\, u_c \times (f_a - f_b)]\rangle \,+ \\
&\quad \langle h_b, (f_c - f_a)\times u_b \times (f_c - f_a) \,+\, (f_c - f_a)\times \\
&\qquad [u_c \times (f_a - f_b) \,+\, u_a \times (f_b - f_c)]\rangle \,+ \\
&\quad \langle h_c, (f_a - f_b)\times u_c \times (f_a - f_b) \,+\, (f_a - f_b)\times \\
&\qquad [u_a \times (f_b - f_c) \,+\, u_b \times (f_c - f_a)]\rangle
\end{aligned}$$

For $p \in V_T$, denote $\{\tau \in Q : p \in \tau\}$ by T^p. Then

$$\sum_{\tau\in Q} \langle (f,u)_\tau, (f,h)_\tau\rangle \,/\, \|f_\tau\| =$$

$$\sum_{p\in V_T} \langle h_p, \sum_{\tau=[p,b,c]\in T^p} \{(f_b - f_c)\times u_p \times (f_b - f_c) \,+$$
$$(f_b - f_c)\times[u_b \times (f_c - f_p) \,+\, u_c \times (f_p - f_b)]\} / \|f_\tau\|\rangle.$$

From (12.8), this expression is zero for all $h \in S_{0,T}$. Thus

$$\sum_{\tau=[p,b,c]\in T^p} \{(f_b - f_c) \times u_p \times (f_b - f_c) + (f_b - f_c) \times [u_b \times (f_c - f_p) + u_c \times (f_p - f_b)]\} \|f_\tau\| = 0 \quad \forall p \in V_{I,T}, \quad (12.9)$$

where $V_{I,T}$ denotes the interior members of V_T. Equation (12.9) can also be obtained by setting $\frac{\partial \beta}{\partial u_p}$ to 0. In order to obtain an expression in matrix/vector notation, let $u = v + w$ where $v \in S_{0,T}$ and $w \in S_T$ is zero on $V_{I,T}$ (so that $v = u$ on $V_{I,T}$ and $w = u = f$ on the boundary nodes). Then (12.9) may be written

$$Au = q, \quad (12.10)$$

where

$$(A\mathbf{u})_p = \sum_{\tau=[p,b,c]\in T^p} \{(f_b - f_c) \times v_p \times (f_b - f_c) + (f_b - f_c) \times [v_b \times (f_c - f_p) + v_c \times (f_p - f_b)]\} / \|f_\tau\|$$

and

$$q_p = -\sum_{\tau=[p,b,c]\in T^p} (f_b - f_c) \times [w_b \times (f_c - f_p) + w_c \times (f_p - f_b)] \,/\, \|f_\tau\|$$

for all $p \in V_{I,T}$. For N interior nodes in $V_{I,T}$ and an arbitrarily selected ordering of the members of V_T, $\mathbf{u}$ and $\mathbf{q}$ denote the column vectors of length $3N$ with the components of u_p and q_p stored contiguously for each $p \in V_{I,T}$. Then A is a symmetric positive definite matrix with N^2 order-3 blocks. This follows from Theorem 42 since

$$\begin{aligned}\mathbf{u}^T A\mathbf{u} &= \sum_{p\in V_{I,T}} \left\langle u_p, (A\mathbf{u})_p \right\rangle = \sum_{p\in V_T} \left\langle v_p, (A\mathbf{u})_p \right\rangle \\ &\sum_{p\in V_T} \langle v_p, \sum_{\tau=[p,b,c]\in T^p} \{(f_b - f_c) \times v_p \times (f_b - f_c) + \\ &(f_b - f_c) \times [v_b \times (f_c - f_p) + v_c \times (f_p - f_b)]\} / \|f_\tau\| \rangle \\ &= \sum_{\tau\in Q} \langle (f, v)_\tau, (f, v)_\tau \rangle \,/\, \|f_\tau\| = 4 \langle v, v \rangle_f .\end{aligned}$$

Equation (12.10) may be solved by a block Gauss-Seidel or SOR method using $u = f$ as an initial solution estimate. No additional storage is required for the matrix (thus allowing for a large number of vertices), and convergence is guaranteed since A is positive definite [**18**, p. 72].

If f is sufficiently close to a local minimum of ϕ_T that second derivatives in all directions are positive, we obtain a Hessian inner product

$$\langle g, h \rangle_H = \phi_T''(f)gh = \frac{1}{2} \sum_{\tau\in Q} \frac{\langle (f,g)_\tau, (f,h)_\tau \rangle + \langle f_\tau, (g,h)_\tau \rangle}{\|f_\tau\|} - \frac{\langle f_\tau, (f,g)_\tau \rangle \langle f_\tau, (f,h)_\tau \rangle}{\|f_\tau\|^3},$$

for g, $h \in S_{0,T}$. The Hessian matrix H is defined by $\langle g, h\rangle_H = \langle Hg, h\rangle_{L_2}$ $\forall g, h \in S_{0,T}$, and letting g now denote the H-gradient, g is related to the standard gradient $\nabla\phi_T(f)$ by $\phi_T'(f)h = \langle g, h\rangle_H = \langle Hg, h\rangle_{L_2} = \langle \nabla\phi_T(f), h\rangle_{L_2}$ $\forall h \in S_{0,T}$, implying that $g = H^{-1}\nabla\phi_T(f)$. The displacement $u = f - g$ is obtained by minimizing

$$\langle u, u\rangle_H = \frac{1}{2}\sum_{\tau\in Q} \frac{\|(f,u)_\tau\|^2 + 2\langle f_\tau, u_\tau\rangle}{\|f_\tau\|} - \frac{\langle f_\tau, (f,u)_\tau\rangle^2}{\|f_\tau\|^3}$$

over functions u that agree with f on the boundary. Note that, for $g = 0$, $\frac{1}{2}\langle u, u\rangle_f$ and $\frac{1}{2}\langle u, u\rangle_H$ are both equal to $\phi_T(f)$.

4. Uniformly Parameterized Surfaces

Numerical tests of the method revealed a problem associated with non uniqueness of the parameterization. Recall that, even in the minimum curve length computation, the parameterization of the solution depends on the initial curve. Thus, depending on the initial surface f_0, the method may result in a triangulated surface whose triangular facets vary widely in size and shape. Also, with a tight tolerance on convergence, the method often failed with a nearly null triangle. (Refer to Section 6.) Currently available software packages such as EVOLVER [**11**] treat this problem by periodically retriangulating the surface (by swapping diagonals in quadrilaterals made up of pairs of adjacent triangles) during the descent process. As an alternative we considered adding bounds on $\|f_\tau\|$ to the minimization problem. This finally led us to a new characterization of the problem as described in the following two theorems.

THEOREM 43. *Let* $\phi(f) = \int_\Omega \|f_1 \times f_2\|$ *and* $\gamma(f) = \int_\Omega \|f_1 \times f_2\|^2$ *for* $f \in C^2(\Omega, \mathbf{R}^3)$ *such that* $f_1 \times f_2 \neq 0$. *Then critical points of* γ *are critical points of* ϕ*; i.e., if* $\gamma'(f)h = 0\ \forall\, h \in C_0^2(\Omega, \mathbf{R}^3) = \{h \in C^2(\Omega, \mathbf{R}^3) : h(x) = 0 \quad \forall\, x \in \partial\Omega\}$, *then* $\phi'(f)h = 0\ \forall\, h \in C_0^2(\Omega, \mathbf{R}^3)$. *Furthermore, such critical points* f *are uniformly parameterized:* $\|f_1 \times f_2\|$ *is constant (and hence equal to the surface area* $\phi(f)$ *at every point since* Ω *has unit area).*

PROOF. $\phi'(f)h = 0\ \forall\, h \in C_0^2(\Omega, \mathbf{R}^3)$ if and only if

$$-\nabla\phi(f) = D_1\left(\frac{f_2 \times f_1 \times f_2}{\|f_1 \times f_2\|}\right) + D_2\left(\frac{f_1 \times f_2 \times f_1}{\|f_1 \times f_2\|}\right) = 0,$$

where D_1 and D_2 denote first partial derivative operators. (Note that the L_2-gradient $\nabla\phi(f)$ is proportional to the mean curvature of f.) Also, $\gamma'(f)h = 0\ \forall\, h \in C_0^2(\Omega, \mathbf{R}^3)$ if and only if $Lf = D_1(f_2 \times f_1 \times f_2) + D_2(f_1 \times f_2 \times f_1) = 0$. Thus it suffices to show that $Lf = 0 \Rightarrow \|f_1 \times f_2\|$ is constant. Expanding Lf, we have

$$\begin{aligned} Lf = f_{12} \times (f_1 \times f_2) + f_2 \times D_1(f_1 \times f_2) + D_2(f_1 \times f_2) \times f_1 + \\ (f_1 \times f_2) \times f_{12} \\ f_2 \times D_1(f_1 \times f_2) + D_2(f_1 \times f_2) \times f_1 \\ f_2 \times f_{11} \times f_2 + f_2 \times (f_1 \times f_{12}) + (f_{12} \times f_2) \times f_1 + \\ f_1 \times f_{22} \times f_1 \\ \langle f_2, f_2\rangle f_{11} - \langle f_2, f_{11}\rangle f_2 + \langle f_2, f_{12}\rangle f_1 - \langle f_1, f_2\rangle f_{12} + \\ \langle f_1, f_{12}\rangle f_2 - \langle f_1, f_2\rangle f_{12} + \langle f_1, f_1\rangle f_{22} - \langle f_1, f_{22}\rangle f_1 \quad , \end{aligned}$$

where the last equation follows from the identity $u \times (v \times w) = \langle u, w\rangle v - \langle u, v\rangle w$. Now suppose $Lf = 0$. Then $\langle f_1, Lf\rangle = \langle f_2, Lf\rangle = 0$ and hence

$$\begin{aligned} \langle f_2 \times f_1 \times f_2, f_{11}\rangle + \langle f_1 \times f_2 \times f_1, f_{12}\rangle = \\ \langle f_2, f_2\rangle \langle f_1, f_{11}\rangle - \langle f_1, f_2\rangle \langle f_2, f_{11}\rangle + \\ \langle f_1, f_1\rangle \langle f_2, f_{12}\rangle - \langle f_1, f_2\rangle \langle f_1, f_{12}\rangle = \langle f_1, Lf\rangle = 0 \end{aligned}$$

and

$$\begin{aligned} \langle f_2 \times f_1 \times f_2, f_{12}\rangle + \langle f_1 \times f_2 \times f_1, f_{22}\rangle = \\ \langle f_2, f_2\rangle \langle f_1, f_{12}\rangle - \langle f_1, f_2\rangle \langle f_2, f_{12}\rangle + \\ \langle f_1, f_1\rangle \langle f_2, f_{22}\rangle - \langle f_1, f_2\rangle \langle f_1, f_{22}\rangle = \langle f_2, Lf\rangle = 0 \end{aligned}$$

Hence,

$$\begin{aligned} D_1\left(\|f_1 \times f_2\|\right) = \frac{\langle f_1 \times f_2, D_1\left(f_1 \times f_2\right)\rangle}{\|f_1 \times f_2\|} \\ \frac{\langle f_2 \times f_1 \times f_2, f_{11}\rangle + \langle f_1 \times f_2 \times f_1, f_{12}\rangle}{\|f_1 \times f_2\|} = 0 \end{aligned}$$

and

$$\begin{aligned} D_2\left(\|f_1 \times f_2\|\right) = \frac{\langle f_1 \times f_2, D_2\left(f_1 \times f_2\right)\rangle}{\|f_1 \times f_2\|} \\ \frac{\langle f_2 \times f_1 \times f_2, f_{12}\rangle + \langle f_1 \times f_2 \times f_1, f_{22}\rangle}{\|f_1 \times f_2\|} = 0 \end{aligned}$$

implying that $\|f_1 \times f_2\|$ is constant. □

The following theorem implies the converse of Theorem 43; i.e., critical points of ϕ are (with a change of parameters) critical points of γ. Note that the surface should not be confused with its representation by a parametric function.

THEOREM 44. *Any regular parametric surface $f \in C^1(\Omega, \mathbf{R}^3)$ can be uniformly parameterized.*

PROOF. Let $\alpha(x, y) = \|f_1(x, y) \times f_2(x, y)\|$, $(x, y) \in \Omega$, and define $\beta : \Omega \to \Omega$ by

$$\beta(x, y) = \begin{pmatrix} u(x, y) \\ v(x, y) \end{pmatrix}$$

where

$$u(x,y) = \frac{\int_0^x \alpha(r,y)dr}{\int_0^1 \alpha(r,y)dr}, \qquad v(x,y) = \frac{\int_0^y \int_0^1 \alpha(r,s)drds}{\phi(f)}.$$

Then

$$u_1(x,y) = \frac{\alpha(x,y)}{\int_0^1 \alpha(r,y)dr}, \quad v_1(x,y) = 0, \quad \text{and} \quad v_2(x,y) = \frac{\int_0^1 \alpha(r,y)dr}{\phi(f)}.$$

Note that $\phi(f) = \int_0^1 \int_0^1 \alpha(r,s)drds$, and by regularity of f, $\alpha(x,y) > 0 \quad \forall\ (x,y) \in \Omega$. It is easily verified that β is invertible. Its Jacobian has determinant $u_1 v_2 - u_2 v_1 = \alpha/\phi(f)$. Denote the reparameterized surface by $g(u,v) \equiv f\left(\beta^{-1}(u,v)\right)$. Then $f(x,y) = g\left(\beta(x,y)\right)$ and

$$\begin{aligned} f_1(x,y) \times f_2(x,y) = [g_1(u,v)u_1(x,y) + g_2(u,v)v_1(x,y)] \times \\ [g_1(u,v)u_2(x,y) + g_2(u,v)v_2(x,y)] \\ = (u_1 v_2 - u_2 v_1)\left[g_1(u,v) \times g_2(u,v)\right]. \end{aligned}$$

Hence $\|g_1(u,v) \times g_2(u,v)\| = \phi(f)$. □

Note that, in the analogous minimum curve length problem, the minimizer of $\int_0^1 \|f'\|^2$ satisfies $f'' = 0$ implying constant velocity resulting in a uniformly parameterized line segment, while the minimizer of $\int_0^1 \|f'\|$ satisfies $(f'/\|f'\|)' = 0$ implying zero curvature but not a uniform parameterization.

For the minimum curve length problem the analog of Theorem 43 holds in both the discrete and continuous cases, but this is not true of the minimal surface problem; i.e., the theorem does not apply to the triangulated surface. However, to the extent that a triangulated surface approximates a critical point of γ, its triangle areas are nearly constant. This is verified by our test results.

On the other hand, there are limitations associated with minimizing the discretization of γ. Forcing a uniformly triangulated surface eliminates the potential advantage in efficiency of an adaptive refinement method that adds triangles only where needed — where the curvature is large. Also, in generalizations of the problem, minimizing the discretization of γ can fail to approximate a minimal surface. In the case of three soap films meeting along a triple line, the triangle areas in each film would be nearly constant but the three areas could be different, causing the films to meet at angles other than 120 degrees. Furthermore, it is necessary in some cases of area minimization to allow surface triangles to degenerate and be removed.

It should be noted that, while similar in appearance, γ is not the Dirichlet integral of f, $\delta(f) = \frac{1}{2}\int_\Omega \|f_1\|^2 + \|f_2\|^2$, which is equal to $\phi(f)$ when f is a conformal map (f is parameterized so that $\|f_1\| = \|f_2\|$ and $\langle f_1, f_2\rangle = 0$) [**19**]. Minimizing δ has the advantage that the Euler equation is linear (Laplace's equation) but requires that the nonlinear side conditions be enforced by varying nodes of T.

The discretized functional to be minimized is

$$\gamma_T(f) = \frac{1}{4}\sum_{\tau\in Q} \|f_\tau\|^2, \quad f \in S_T,$$

and the appropriate inner product is

$$\langle g, h\rangle_f = \frac{1}{4}\sum_{\tau\in Q} \langle (f,g)_\tau, (f,h)_\tau\rangle \quad \forall\, g, h \in S_{0,T}.$$

Theorem 42 remains unaltered for this definition of $\langle\cdot,\cdot\rangle_f$. A Sobolev gradient g for γ_T is defined by $u = f - g$, where $\gamma_T'(f)h = \langle g, h\rangle_f$ implying that

$$\sum_{\tau\in Q} \langle (f,u)_\tau, (f,h)_\tau\rangle = 0 \quad \forall\, h \in S_{0,T},$$

and thus u satisfies (12.9) without the denominator $\|f_\tau\|$ (or with $\|f_\tau\|$ taken to be constant). The Hessian inner product associated with γ_T is

$$\langle g, h\rangle_H = \frac{1}{3}\gamma_T''(f)gh = \frac{1}{6}\sum_{\tau\in Q} \langle (f,g)_\tau, (f,h)_\tau\rangle + \langle f_\tau, (g,h)_\tau\rangle,$$

and the displacement is obtained by minimizing

$$\langle u, u\rangle_H = \frac{1}{6}\sum_{\tau\in Q} \|(f,u)_\tau\|^2 + 2\langle f_\tau, u_\tau\rangle.$$

This expression is considerably simpler than the corresponding expression associated with ϕ_T.

5. Numerical Methods and Test Results

We used the Fletcher Reeves nonlinear conjugate gradient method [**29**] with Sobolev gradients and step size obtained by a line search consisting of Brent's one dimensional minimization routine FMIN [**32**]. The number of conjugate gradient steps between restarts with a steepest descent iteration was taken to be 2. At each step, the linear system defining the gradient was solved by a block SOR method with relaxation factor optimal for Laplace's equation and the number of iterations limited to 500. Convergence of the SOR method was defined by a bound on the maximum relative change in a solution component between iterations. This bound was initialized to 10^{-3} and decreased by a factor of 10 (but bounded below by 10^{-13}) after each descent iteration in which the number of SOR iterations was less than 10, thus tightening the tolerance as the initial estimates improved with convergence of the descent method.

The SOR method was also used to solve the linear systems associated with attempted Newton steps. A Newton step was attempted if and only if the root-mean-square norm of the L_2 gradient at the previous iteration fell below a tolerance which was taken to be a decreasing function of n. Failure of the SOR method due to an indefinite Hessian matrix was defined as an increase in the Euclidean norm of the residual between any pair of consecutive iterations. In most cases this required only two wasted SOR iterations before abandoning the attempt and falling back to a conjugate gradient step. In some cases however, the number of wasted SOR iterations was as high as 20, and, more generally, there was considerable inefficiency caused by less than optimal tolerances.

The selection of parameters described above, such as the number of conjugate gradient iterations between restarts, the tolerance defining convergence of SOR, etc., were made on the basis of a small number of test cases and are not necessarily optimal. The Fletcher-Reeves method could be replaced by the Polak-Ribière method at the cost of one additional array of length $3(n+1)^2$. Also, alternative line search methods were not tried, nor was there any attempt to optimize the tolerance for the line search. However, again based on limited testing, conjugate gradient was not found to be substantially faster than steepest descent, and the total cost of the minimization did not appear to be sensitive to the accuracy of the line search. Adding a line search to the Newton iteration (a damped Newton method) was found to be ineffective, actually increasing the number of iterations required for convergence.

We used the regular triangulation $T = T_n$ of $\Omega = [0,1] \times [0,1]$ and took the initial approximation f_0 to be a displacement of a discretized minimal surface f : $f_0 = f + p$ for $f \in S_T$, $p \in S_{0,T}$. Note that f_0 defines the boundary curve as well as the initial value. The following three minimal surfaces $F \in C^\infty(\Omega, \mathbf{R}^3)$ were used to define f: **Catenoid** $F(x,y) = (R\cos\theta, R\sin\theta, y)$ for radius $R = \cosh(y - .5)$ and angle $\theta = 2\pi x$. The surface area is $\phi(F) = \int_\Omega \|F_1 \times F_2\| = \pi(1 + \sinh(1)) \cong 6.8336$.

Right Helicoid $F(x,y) = (x\cos(10y), x\sin(10y), 2y)$ with surface area $\phi(F) = \sqrt{26} + \left[\ln\left(5 + \sqrt{26}\right)\right]/5 \cong 5.5615$.

Enneper's Surface $F(x,y) = \left(\xi - \xi^3/3 + \xi\eta^2, \eta - \eta^3/3 + \xi^2\eta, \xi^2 - \eta^2\right)$ for $\xi = (2x-1)R/\sqrt{2}$ and $\eta = (2y-1)R/\sqrt{2}$, where $R = 1.1$. The surface area is $\phi(F) = 2R^2 + \frac{4}{3}R^4 + \frac{14}{45}R^6 \cong 4.9233$.

For each test function f, all three components were displaced by the discretization p of $P(x,y) = 100x(1-x)y(1-y)$ which is zero on $\partial\Omega$.

Table 1 displays the computed surface areas associated with minimizing $\phi_T(f)$ (method 1) and $\gamma_T(f)$ (method 2), for each test function and each of five values of n. Convergence of the descent method was defined by a bound of $.5 \times 10^{-4}$ on the relative change in the functional between iterations. The number of conjugate gradient iterations and the total number of SOR iterations (in parentheses) for each test case is displayed in the row labeled CG (SOR) iterations. Similarly, the number of Newton iterations is followed by the total number of SOR iterations (for all Newton steps) in parentheses. Note that the cost of each SOR iteration is proportional to n^2 (with a smaller constant for method 2). Although each SOR step has a higher operation count (by a factor of 2 with method 1 and 1.5 with method 2) for a Newton iteration than a conjugate gradient iteration, this is offset by the fact that no line search is required for the Newton iteration. Method 2 is more efficient in all cases except the helicoid with $n = 50$. This is further discussed below. The rows labeled RMS L_2 gradient

	$n = 10$	$n = 20$	$n = 30$	$n = 40$	$n = 50$
Catenoid $\phi_T(f)$	6.6986	6.7995	6.8184	6.8250	6.8281
$\phi(F) = 6.8336$					
Method 1					
Surface area	6.6984	6.8057	8.0053	8.0518	8.0593
CG (SOR) iterations	10(375)	20(996)	17(1538)	30(2648)	26(3766)
Newton iterations	0	0	0	0	0
RMS L_2 gradient	.15E-1	.74E-2	.33E0	.35E0	.28E0
Method 2					
Surface area	6.6964	6.8035	6.8219	6.8289	6.8352
CG (SOR) iterations	4(141)	11(435)	17(852)	15(1265)	21(1920)
Newton iterations	2(17)	0	0	0	0
RMS L_2 gradient	.36E-2	.86E-2	.45E-2	.48E-2	.45E-2
Helicoid $\phi_T(f)$	4.8409	5.3731	5.4771	5.5139	5.5310
$\phi(F) = 5.5615$					
Method 1					
Surface area	4.8139	5.3697	5.4801	5.5214	5.5397
CG (SOR) iterations	28(875)	20(971)	25(2670)	32(3856)	32(5170)
Newton iterations	0	0	0	0	0
RMS L_2 gradient	.36E-1	.76E-2	.15E-1	.12E-1	.15E-1
Method 2					
Surface area	4.8572	5.3820	5.4820	5.5170	5.5344
CG (SOR) iterations	4(137)	8(440)	15(1012)	20(1816)	41(6500)
Newton iterations	2(13)	2(25)	0	0	0
RMS L_2 gradient	.27E-1	.46E-2	.19E-2	.22E-2	.16E-1
Enneper $\phi_T(f)$	4.9077	4.9194	4.9215	4.9223	4.9227
Method 1					
$\phi(F) = 4.9233$					
Surface area	4.9118	4.9308	4.9311	4.9303	4.9581
CG (SOR) iterations	13(325)	12(846)	14(1493)	21(3134)	20(5753)
Newton iterations	0	0	0	0	0
RMS L_2 gradient	.11E-1	.17E-1	.88E-2	.15E-1	.34E0
Method 2					
Surface area	4.9139	4.9216	4.9228	4.9237	4.9245
CG (SOR) iterations	4(133)	15(674)	18(1231)	21(1743)	21(2470)
Newton iterations	2(24)	1(22)	0	0	0
RMS L_2 gradient	.63E-2	.16E-2	.14E-2	.13E-2	.11E-2

TABLE 1. Surface Areas and Iteration Counts, Low Accuracy

display the root-mean-square Euclidean norms of the L_2 gradients of the surface area ϕ_T.

The table also displays the triangulated surface areas $\phi_T(f)$ associated with the undisplaced surfaces. From these values, the discretization error is verified to be of order $\left(\frac{1}{n}\right)^2$; i.e., $n^2\,|\phi(F) - \phi_T(f)|$ approaches a constant with increasing n, where $T = T_n$ and f is the discretization of F. Note, however, that the computed surface areas do not closely match the ϕ_T values and are in some cases smaller because, while both are triangulated surface areas with the same boundary values, the former are minima of the discretized functionals while the latter have nodal function values taken from smooth minimal surfaces.

	$n = 10$	$n = 20$	$n = 30$	$n = 40$	$n = 50$
Catenoid $\phi_T(f)$	6.6986	6.7995	6.8184	6.8250	6.8281
$\phi(F) = 6.8336$					
Method 1					
Surface area	6.6962	6.7993	7.9922	8.0468	8.0435
CG (SOR) iterations	15(441)	498(13532)	104(5467)	82(4746)	95(8992)
Newton iterations	8(118)	8(790)	0	0	0
RMS L_2 gradient	.25E-7	.49E-8	.29E0	.35E0	.27E0
Method 2					
Surface area	6.6964	6.7994	6.8184	6.8251	6.8320
CG (SOR) iterations	4(141)	17(582)	69(2976)	122(5702)	180(7877)
Newton iterations	7(45)	10(138)	8(294)	6(540)	2(1000)
RMS L_2 gradient	.36E-2	.65E-3	.21E-3	.92E-4	.21E-2
Helicoid $\phi_T(f)$	4.8409	5.3731	5.4771	5.5139	5.5310
$\phi(F) = 5.5615$					
Method 1					
Surface area	4.8043	5.3632	5.4731	5.5127	5.5314
CG (SOR) iterations	192(13915)	384(10462)	785(23897)	572(30088)	311(15319)
Newton iterations	0	0	0	0	0
RMS L_2 gradient	.27E0	.26E-1	.29E-1	.66E-1	.87E-2
Method 2					
Surface area	4.8571	5.3821	5.4817	5.5163	5.5333
CG (SOR) iterations	4(137)	8(440)	104(4486)	209(10817)	500(62376)
Newton iterations	7(51)	11(119)	6(356)	5(1035)	0
RMS L_2 gradient	.27E-1	.46E-2	.14E-2	.83E-3	.16E-1
Enneper $\phi_T(f)$	4.9077	4.9194	4.9215	4.9223	4.9227
$\phi(F) = 4.9233$					
Method 1					
Surface area	4.9077	4.9197	4.9244	4.9242	4.9417
CG (SOR) iterations	172(4699)	933(41510)	146(4672)	372(20615)	37(6841)
Newton iterations	13(879)	0	0	0	0
RMS L_2 gradient	.29E-7	.16E-2	.79E-2	.71E-2	.34E0
Method 2					
Surface area	4.9138	4.9216	4.9227	4.9230	4.9231
CG (SOR) iterations	4(133)	15(674)	41(3527)	107(9024)	371(38551)
Newton iterations	12(128)	13(483)	20(3025)	11(3997)	11(4118)
RMS L_2 gradient	.63E-2	.14E-2	.65E-3	.36E-3	.22E-3

TABLE 2. Surface Areas and Iteration Counts, High Accuracy

The tabulated surface areas reveal some anomalies associated with non uniqueness of the solution. In the case of the catenoid, there is a second surface satisfying the same boundary conditions: a pair of parallel disks connected by a curve. This surface has an approximate area of 7.9893. Plots verify that it is this surface that is approximated by method 1 with n = 30, 40, and 50. The curve connecting the disks is approximated by long thin triangles. We assume that the nearly constant triangle area maintained by method 2 prevented it from converging to this solution. Additional tests on Enneper's surface with a larger domain (ξ and η in the range -2 to 2) revealed the apparent existence of a second minimal surface with the same boundary but with smaller surface area. This non-uniqueness of Enneper's surface was noted by Nitsche ([**77**]).

Table 2 displays the same quantities as in Table 1 but with convergence tolerance 1.0×10^{-14}. The computed solutions are not significantly more accurate, but the tests serve to demonstrate the deficiencies of method 1 and the robustness of method 2. Method 1 failed to converge in all cases except the catenoid with $n = 10$ and $n = 20$ and Enneper's surface with $n = 10$. In all other cases the procedure was terminated when the minimum triangle area fell below 100ϵ for machine precision ϵ ($.222 \times 10^{-15}$). (Allowing the procedure to continue would have resulted in failure with a nearly singular linear system or a triangle area of 0.) Method 2, on the other hand, failed to converge only on the helicoid with $n = 50$. Table 2 displays the result of 500 iterations, but another 500 iterations resulted in no improvement. Pictures reveal a uniformly triangulated surface but with long thin triangles apparently caused by the tendency of triangle sides to be aligned with the lines of curvature. Also, for method 2, the ratio of largest to smallest triangle area is at most 5 in all cases other than the helicoid with $n = 50$, in which the ratio is 46 with the larger convergence tolerance and 98 with the smaller tolerance. In no case was a saddle point encountered with either method.

Excluding the cases in which the method failed to converge, the number of descent steps and the number of SOR steps per descent step both increase with n for all test functions and both the conjugate gradient and Newton methods, implying that the Hessian matrices and their approximations become increasingly ill-conditioned with increasing n. This reflects the fact that finite element approximations to second-order elliptic boundary value problems on two-dimensional domains result in condition numbers of O(N) for N nodes.

The small iteration counts demonstrate the effectiveness of the preconditioner for both methods. Additional tests revealed that the standard steepest descent method (using the discretized L_2 gradient) fails to converge unless the initial estimate is close to the solution. Also, the conjugate gradient method without preconditioning is less efficient than preconditioned steepest descent even when starting with a good initial estimate.

6. Conclusion

We have described an efficient method for approximating parametric minimal surfaces. In addition to providing a practical tool for exploring minimal surfaces, the method serves to illustrate the much more generally applicable technique of solving PDE's via a descent method that employs Sobolev gradients, and it demonstrates the effectiveness of such methods. Furthermore, it serves as an example of a variable metric method.

The implementations of method 1 (`MINSURF1`) and method 2 (`MINSURF2`) are available as Fortran software packages which can be obtained from netlib.

CHAPTER 13

Flow Problems and Non-inner Product Sobolev Spaces

1. Full Potential Equation

From [**39**] we have the following one-dimensional flow problem. Consider a horizontal nozzle of length two which has circular cross sections perpendicular to its main axis and is a figure of revolution about its main axis. We suppose that the cross sectional area is given by

$$A(x) = .4[1 + (1 - x^2)],\ 0 \le x \le 2.$$

We suppose that pressure and velocity depend only on the distance along the main axis of the nozzle and that for a given velocity u the pressure is given by

$$p(u) = [1 + ((\gamma - 1)/2)(1 - u^2)]^{\gamma/(\gamma-1)}$$

for all velocities for which $1 + ((\gamma - 1)/2)(1 - u^2) \ge 0$. We choose $\gamma = 1.4$, the specific heat corresponding to air. Define a density function m by $m(u) = -p'(u), u \in R$. Further define

$$J(f) = \int_0^2 Ap(f'),\ f \in H^{1,7}. \tag{13.1}$$

For general $\gamma > 1$ we would choose $H^{1,2\gamma/(\gamma-1)}$ in order that the integrand of (13.1) be in $L_1([0,2])$. Thus the specific heat of the media considered determines the appropriate Sobolev space; for $\gamma = 1.4$ we have $2\gamma/(\gamma - 1) = 7$. Taking a first variation we have

$$J'(f)h = -\int_0^2 Am(f')h',\ f, h \in H^{1,7} \tag{13.2}$$

where the perturbation $h \in H^{1,7}$ is required to satisfy $h(0) = 0 = h(2)$ and f is required to satisfy

$$f(0) = 0, f(2) = c \tag{13.3}$$

for some fixed positive number c. Denote

$$H_0 = H_0^{1,7}([0,2]) = \{h \in H^{1,7}([0,2]) : h(0) = 0 = h(2)\}.$$

Suppose $f \in H^{1,7}([0,2])$. By Theorem 41 there is a unique $h \in H_0$ such that

$$J'(f)h$$

is maximum subject to

$$\|h\| = |J'(f)|$$

where $|J'(f)|$ is the norm of $J'(f)$ considered as a member of the dual of H^*. This maximum $h \in H$ may be denoted by $(\nabla_H J)(f)$, the Sobolev gradient of J at f. Later in this chapter we construct a finite dimensional emulation of this gradient. First we point out some peculiar difficulties that a number of flow problems share with this example. From (13.2), if $Am(f')$ were to be differentiable, we would arrive at an Euler equation

$$(Am(f'))' = 0 \tag{13.4}$$

for f a critical point of J. Furthermore, given sufficient differentiability, we would have

$$A'm(f') + Am'(f')f'' = 0$$

for f a critical point of J. Observe that for some f, the equation (13.4) may be singular if for some

$$x \in [0,2],\, m'(f'(x)) = 0.$$

This simple appearing example leads to a differential equation (13.4) which has the particularly interesting feature that it might be singular depending on the whims of the nonlinear coefficient of f''. Some calculation reveals that this is exactly the case at $x \in [0,2]$ if $f'(x) = 1$ which just happens to be the case at the speed of sound for this problem ($f'(x)$ is interpreted as the velocity corresponding to f at x). It turns out that the choice of c in (13.3) determines the nature of critical points of (13.1) - in particular whether there will be transonic solutions, i.e., solutions which are subsonic for x small enough ($f'(x) < 1$) and then become supersonic ($f'(x) > 1$) for some larger x.

It is common that there are many critical points of (13.1). Suppose we have one, denoted by f. An examination of m yields that for each choice of a value of $y \in [0,1]$, there are precisely two values $x_1, x_2 \in [0, ((\gamma+1)/(\gamma-1))^{1/2})$ so that $x_1 < x_2$ and $m(x_1) = y = m(x_2)$. The value x_1 corresponds to a subsonic velocity and x_2 corresponds to a supersonic velocity. So if f is such that

$$0 < f'(t_1) < 1,\ 0 < f'(t_2) > 1 \text{ for some } t_1, t_2 \in [0,2],$$

and (13.3) holds, then we may construct additional solutions as follows:

Pick two subintervals $[a,b], [c,d]$ of $[0,2]$ so that f is subsonic on [a,b] and supersonic on [c,d]. Define g so that it agrees with f on the complement of the union of these two intervals so that g' on $[a,b]$ is the supersonic value which corresponds as in the preceding paragraph to the subsonic values of f' on $[a,b]$. Similarly, take g' on $[c,d]$ to have the subsonic values of f' on $[c,d]$. Do this in such a way that

$$\int_0^2 f' = \int_0^2 g'.$$

In this way one may construct a large family of critical points g from a particular one f. Now at most one member of this family is a physically correct solution if one imposes the conditions that the derivative of a such a solution does not shock 'up' in the sense that going from left to right (the presumed flow direction) there is no point of discontinuity of the derivative that jumps from subsonic to supersonic. A discontinuity in a derivative which goes from supersonic to subsonic as one moves from left to right is permitted.

We raise the question as to how a descent scheme can pick out a physically correct solution given the possibility of an uncountable collections of non-physical solutions.

A reply is the following: Take take the left hand side of (13.4) and add an artificial dispersion, that is pick $\epsilon > 0$ and consider the problem of finding f so that

$$-(Am(f'))' + \epsilon f''' = 0. \tag{13.5}$$

We now turn to a numerical emulation of this development. Equation (13.5) is no longer singular. A numerical scheme for (13.5) may be constructed using the ideas of Chapter 8. We denote by w a numerical solution to (13.5) (on a uniform grid on $[0,2]$ with n pieces) satisfying the indicated boundary conditions. Denote by J_n the numerical functional corresponding to J on this grid. Denote by H_n the space of real-valued functions on this grid where the expression (10.1) is taken for a norm on H_n. Using the development in Chapter 10, denote the $H^{1,7}([0,2])$ Sobolev gradient of J_n at $v \in H_n$ by $(\nabla J_n)(v)$. Consider the steepest descent process

$$w_{k+1} = w_k - \delta_k(\nabla J_k)(w_k),\ k = 1, 2, \ldots \tag{13.6}$$

where $w(1) = w$, our numerical solution indicated above and for $k = 1, 2, \ldots, \delta_k$ is chosen optimally. We do not have a convergence proof for this iteration but do call attention to Figure 1 at the end of this chapter which shows two graphs superimposed. The smooth curve is a plot of w satisfying (13.5). The graph with the sharp break is the limit of $\{w_k\}_{k=1}^{\infty}$ of the sequence (13.6). The process sharp picks out the one physically viable critical point (13.1). The success of this procedure seems to rest on the fact that the initial value w, the viscous solution, has the correct general shape. The iteration (13.6) then picks out a nearest, in some sense, solution to w which is an actual critical point of (13.1). The reader will recognize the speculative nature of the above 'assertions'; this writer would be quite pleased to be able to offer a complete formulation of this problem together with complete proofs, but we must be content here to raise the technical and mathematical issues concerning this approach to the problem of transonic flow. It might be called a smeared shock. Results as indicated in Figure 1 are in good agreement with those of F. T. Johnson [**39**]. This writer expresses great appreciation to F.T. Johnson of Boeing for his posing this problem and for his considerable help and encouragement.

A similar development has been coded for a two dimensional version:

$$J(u) = \int_\Omega (1 + ((\gamma - 1)/2)|\nabla u|^2)^{\gamma/(\gamma-1)} \tag{13.7}$$

$u \in H^{1,7}(\Omega)$, where Ω is a square region in R^2 with a NACA-12 airfoil removed. Details follow closely those for the nozzle problem outlined above. We present results for two runs, one where air speed 'at infinity' is subsonic (Figure 2) and the second (Figure 3) in which air speed at infinity is supersonic. In the first we see supersonic pockets built up on the top and bottom of the airfoil; in the second we see a subsonic stagnation region on the leading edge of the airfoil. Both are expected by those familiar with transonic flow problems.

Calculations in the two problems were straight 'off the shelf' in that procedures outlined in Chapter 10 were follow closely (with appropriate approximations being made on airfoil to simulate zero Neumann boundary conditions there. It is this writer's belief that the same procedure can be followed in three dimensions. Our point is that procedures of this chapter are straightforward and that there should be no essential difficulties in implementing them in large scale three dimensional problems in which 'shock fitting' procedures would be daunting (we claim the procedure of this chapter is 'shock capturing).

2. A Linear Mixed-Type Problem

In [**41**] Kim studies the problem of finding u such that

$$u_{11}(x, y) + y * u_{22}(x, y) = 0, (x, y) \in \Omega \tag{13.8}$$

where

$$\Omega = (x, y) \in R^2 : -1 \leq y \leq 1, 0 \leq x \leq 1.$$

Problem (13.8) is elliptic in the part of Ω in the upper half plane and hyperbolic in the part of Ω in the lower half plane. The presence of this mixed-type gives a common ground with the preceding section. A numerical solution to this problem is given in (13.8). What boundary conditions for (13.8) yield a unique solution appears to be still unknown but (13.8) provides an experimental tool with which to explore this problem. The code in (13.8) assumes no boundary conditions. It is the transformation T which associates a given $u \in H^{1,2}(\Omega)$ with its limit under continuous steepest descent which is of interest here. Much remains to be done in order to understand this problem on Ω and, of course, on more complicated regions which also intersect both the upper and lower planes.

3. Other Codes for Transonic Flow

From [**14**] one has the problem of determining $u : R^2 \to R$ such that

$$F(u, u_1, u_2)u_{11} - G(u, u_1, u_2)u_{12} + H(u, u_1, u_2)u_{22} = 0$$

where

$$F(u, u_1, u_2) = a^2 + ((\gamma - 1)/2)(u_0^2 - u_1^2 - u_2^2) - u_1^2$$
$$G(u, u_1, u_2) = -2u_1 u_2$$
$$H(u, u_1, u_2) = a^2 + ((\gamma - 1)/2)(u_0^2 - u_1^2 - u_2^2) - u_2^2,$$

a, u_0 being the speed of sound and air velocity at infinity respectively and γ is as in the previous section. For boundary conditions it is required that for each $y \in R$, $lim_{x\to\infty}(u(x, y) - x) = 0$, $lim_{x\to\infty}(u_1(x, y) - u_0) = 0$, and $lim_{x^2+y^2} u_2(x, y) = 0$. Assuming an airfoil as in the previous section, it is also required that the tangential velocity component be zero on the object. In [**44**] a finite element code using Sobolev gradients is presented. In [**60**] a finite difference code is given for this problem.

4. Transonic Flow Plots

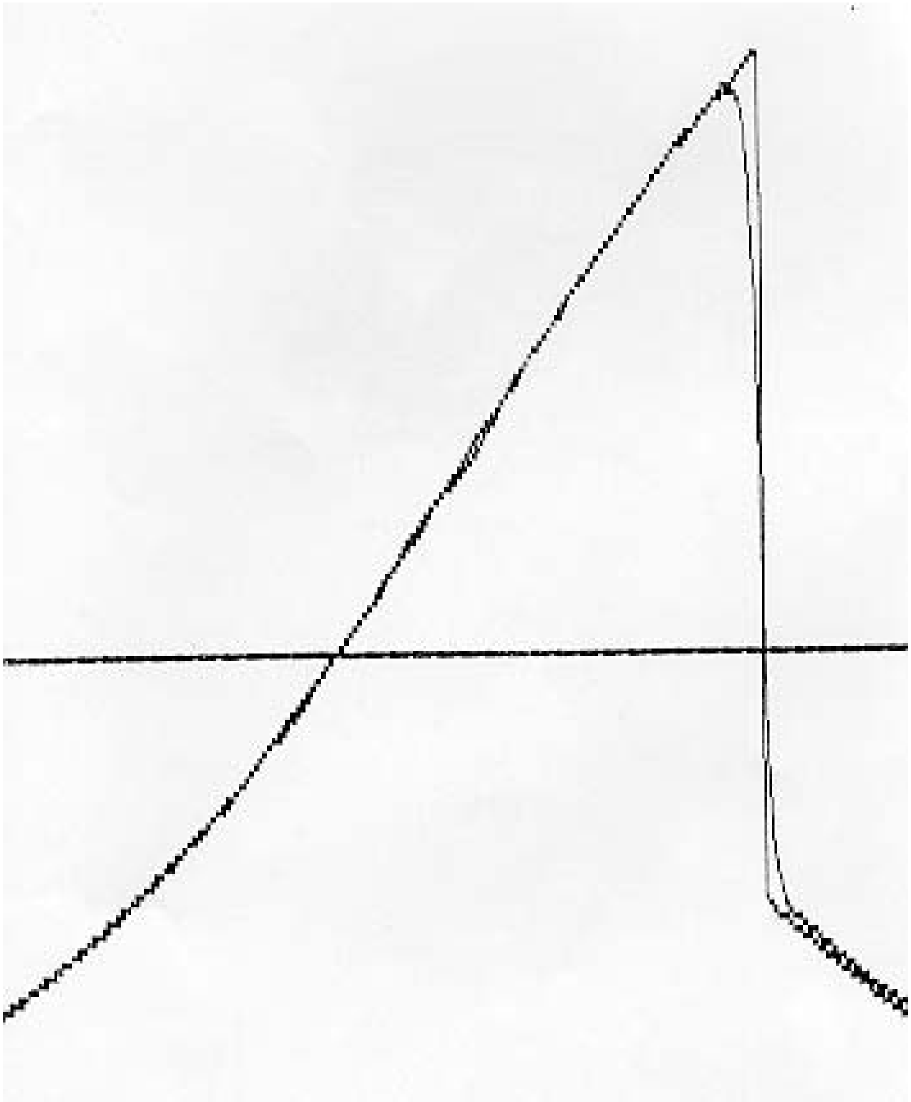

FIGURE 1. Smeared and Sharp Shocks in Nozzle Problem

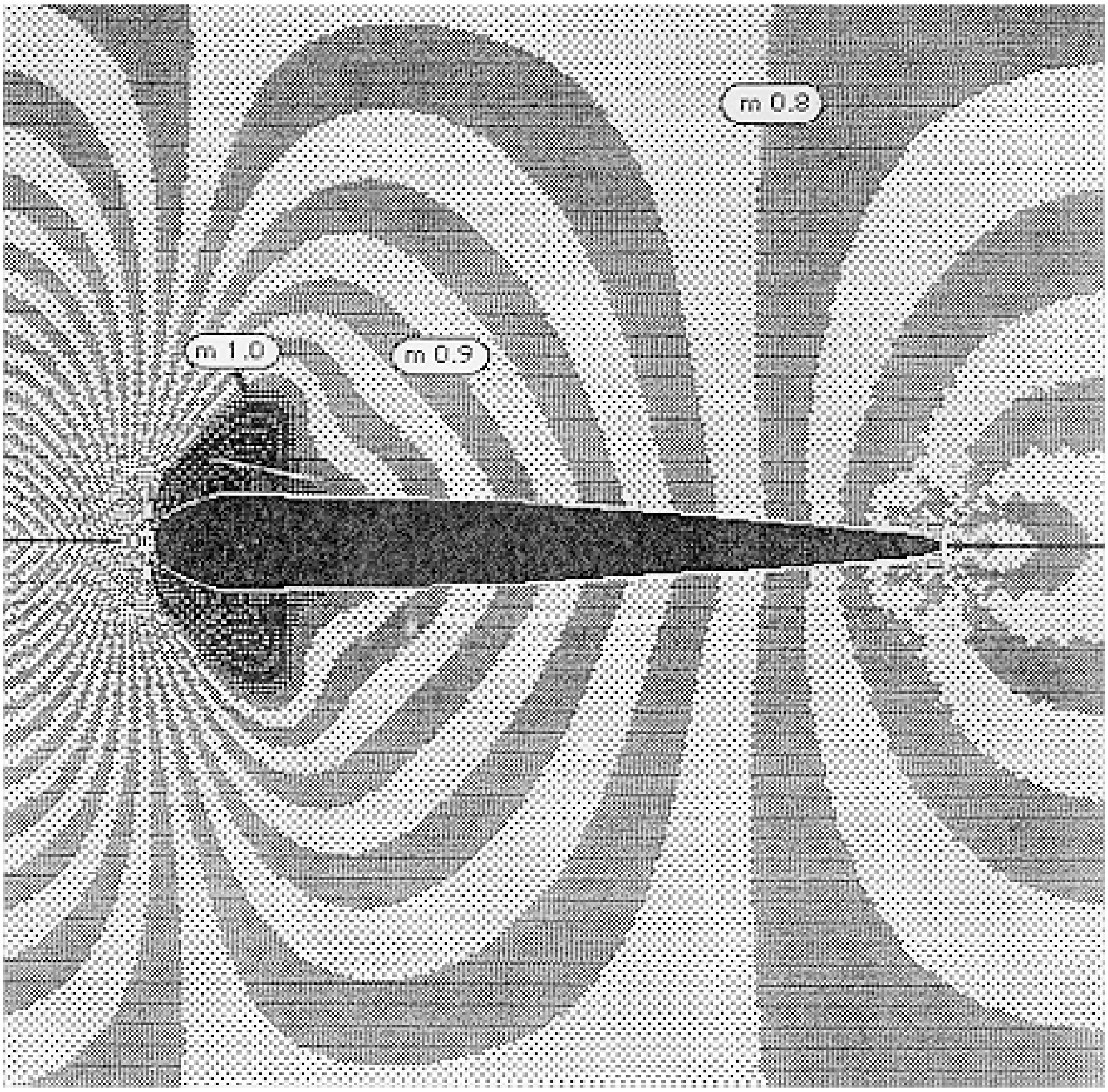

FIGURE 2. Mach .8 Velocity Contours

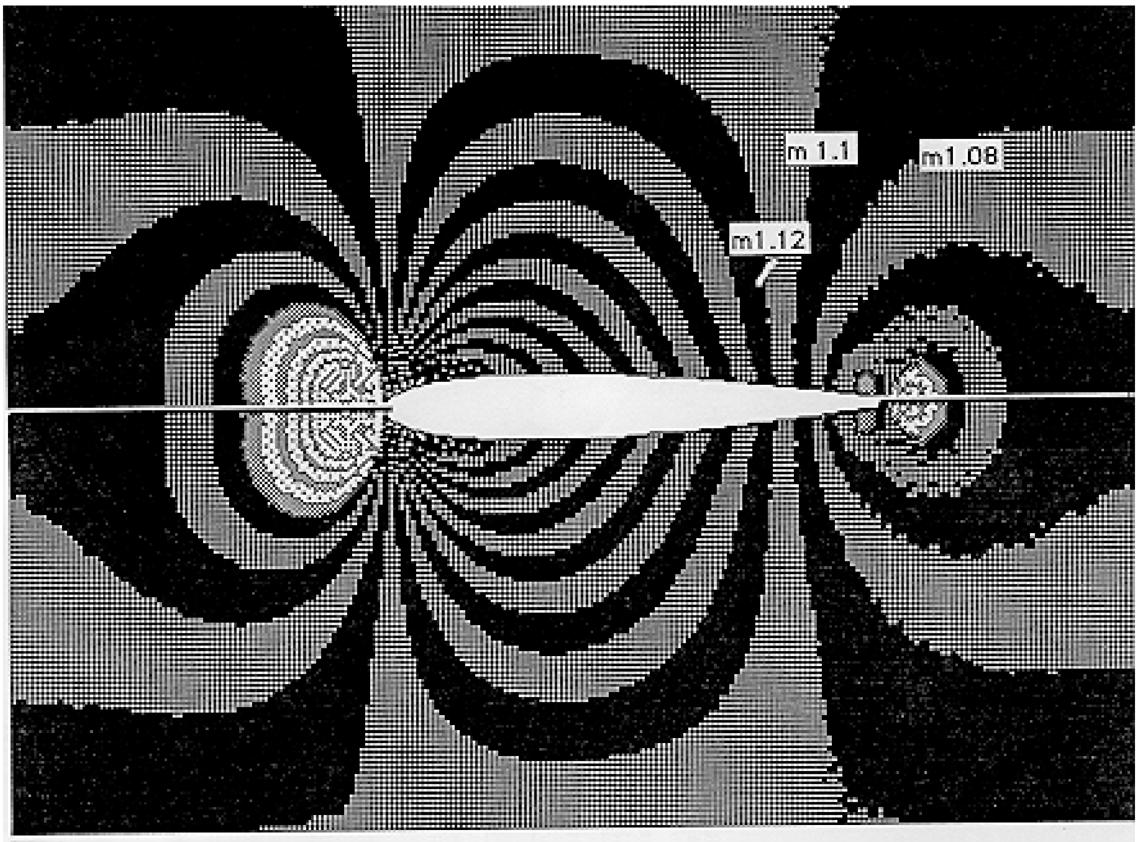

FIGURE 3. Mach 1.1 Velocity Contours

CHAPTER 14

Foliations as a Guide to Boundary Conditions

For a given system of partial differential equations, what side conditions may be imposed in order to specify a unique solution? For various classes of elliptic, parabolic or hyperbolic equations there are, of course, well established criteria in terms of boundary conditions. For many systems, however, there is some mystery concerning characterization of the set of all solutions to the system.

1. A Foliation Theorem

This section is taken largely from [**72**]. Suppose that each of H and K is a Hilbert space and F is a $C^{(3)}$ function from $H \to K$. Define

$$\phi : H \to R$$

by

$$\phi(x) = \|F(x)\|_K^2/2, \ x \in H$$

and note that

$$\phi'(x)h = \langle F'(x)h, F(x)\rangle_K = \langle h, F'(x)^* F(x)\rangle_H, x, h \in H \tag{14.1}$$

where $F'(x)^* \in L(K, H)$ is the Hilbert space adjoint of $F'(x), x \in H$. In view of (14.1), we take $F'(x)^*F(x)$ to be $(\nabla\phi)(x)$, the gradient of ϕ at x.

By Theorem 7 there is a unique function

$$z : [0, \infty) \times H \to H$$

such that

$$z(0, x) = x, \ z_1(t, x) = -(\nabla\phi)(z(t, x)), \ t \geq 0, \ x \in H \tag{14.2}$$

where the subscript in (14.2) indicates the partial derivative of z in its first argument.

In this chapter we have the following standing assumptions on F : If $r > 0$, there is $c > 0$ such that

$$\|F'(x)^* g\|_H \geq c\|g\|_K, \ \|x\| < r, \ g \in K$$

and if

$$x \in H, z(0, x) = x, z_1(t, x) = -(\nabla\phi)(z(t, x))t > 0,$$

then

$$\{z(t, x) : t \geq 0\} \textit{ is bounded.}$$

Using Theorem 9 we have that if $x \in H$, then

$$u = lim_{t\to\infty} z(t,x), \text{ exists and } F(u) = 0. \tag{14.3}$$

Define $G : H \to H$ so that if $x \in H$ then $G(x) = u, u$ as in (14.3). Denote by Q the collection of all $g \in C^{(1)}(H, R)$ so that

$$g'(x)(\nabla\phi)(x) = 0, \ x \in H.$$

THEOREM 45. *(a) G' exists, has range in $L(H, H)$ and*

$$(b) \quad G^{-1}(G(x)) = \cap_{g\in Q}\, g^{-1}(g(x)), \ x \in H.$$

LEMMA 12. *Under the standing hypothesis suppose $x \in H$ and*

$$Q = \{z(t,x), t \geq 0\} \cup \{G(x)\}.$$

There are $\gamma, M, r, T > 0$ so that if

$$Q_\gamma = \cup_{w\in Q} B_\gamma(w),$$

then

$$|(\nabla\phi)'(w)| \leq M, \ |(\nabla\phi)''(w)| \leq \ M, w \in Q$$

and if $y \in H$, $\|y - x\|_H < r$, then

$$[z(t,y), z(t,x)] \subset Q_\gamma, \ t \geq 0 \ and \ [z(t,y), G(y)] \subset Q_{\gamma,} \ t \geq T.$$

For $a, b \in H$, $[a,b] = \{ta + (1-t)b : 0 \leq t \leq 1\}$ and for $w \in H$, $|(\nabla\phi)'(w)|$, $|(\nabla\phi)''(w)|$ denote the norms of $(\nabla\phi)'(w)$, $(\nabla\phi)''(w)$ as linear and bilinear functions on $H \to H$ respectively:

$$|(\nabla\phi)'(w)| = \sup_{h\in H, \|h\|_H=1} |(\nabla\phi)'(w)h|$$

$$|(\nabla\phi)''(w)| = \sup_{h,k\in H, \|h\|_H=1, \|k\|_H=1} |(\nabla\phi)''(w)(h,k)|.$$

PROOF. Since Q is compact and both $(\nabla\phi)', (\nabla\phi)''$ are continuous on H, there is $M > 0$ and an open subset α of H containing Q so that

$$|(\nabla\phi)'(w)|, |(\nabla\phi)''(w)| < M, \ w \in \alpha.$$

Pick $\gamma > 0$ such that $Q_\gamma \subset \alpha$. Then the first part of the conclusion clearly holds.

Note that Q_γ is bounded. Denote by c a positive number so that

$$\|F'(w)^* g\|_H \geq c\|g\|_K, \ g \in K, \ w \in Q_\gamma.$$

Pick $T > 0$ so that $\|z(T,x) - G(x)\|_H < \gamma/4$ and $\|F(z(T,x))\|_K < c\gamma/4$ (this is possible since $\lim_{t\to\infty} F(z(t,x)) = F(u) = 0$). Pick $v > 0$ such that $v \cdot exp(TM) < \gamma/4$. Suppose $y \in B_v(x)$ $(= \{w \in H : \|w - x\|_H < v)\}$. Then

$$z(t,y) - z(t,x) = y - x - \int_0^t ((\nabla\phi)(z(s,y)) - (\nabla\phi)(z(s.x)))ds$$

and so

$$z(t,y) - z(t,x) = y - x$$

$$-\int_0^t \int_0^1 ((\nabla\phi)'((1-\tau)z(s,x)+\tau z(s,y)))d\tau(z(s,y)-z(s,x))ds,\ t \geq 0.$$

Hence there is $T_1 > 0$ such that $[z(s,y), z(s,x)] \subset Q_\gamma,\ 0 \leq s \leq T_1$, and so $|\int_0^1((\nabla\phi)'((1-\tau)z(s,x)+\tau z(s,y)))d\tau| \leq M$ and

$$\|z(t,y)-z(t,x)\|_H \leq \|y-x\|_H + M\int_0^t \|z(s,y)-z(s,x)\|_H ds,\ 0 \leq s \leq T_1.$$

But this implies that

$$\|z(t,y)-z(t,x)\|_H \leq \|y-x\|_H exp(tM) < v \cdot exp(tM)$$
$$\leq v \cdot exp(T_1 M) < \gamma, \|y-x\|_H < r$$

and so

$$[z(t,y), z(t,x)] \subset Q_\gamma, \|y-x\|_H < v,\ 0 \leq t \leq T_1.$$

Supposing that the largest such T_1 is less than T, we get a contradiction and so have that

$$[z(t,y), z(t,x)] \subset Q_\gamma, 0 \leq t \leq T,\ \|y-x\|_H < r.$$

Now choose $r > 0$ such that $r \leq v$ and such that if $\|y-x\|_H < r$, then

$$\|F(y)\|_K \leq 2\|F(x)\|_K, \|z(T,y)-z(T,x)\|_H < \gamma/4$$

and

$$\|F(z(T,y)) - F(z(T,x))\|_K < c\gamma/4.$$

Hence for $\|y-x\|_H < r$,

$$\|z(T,y)-G(x)\|_H \leq \|z(T,y)-z(T,x)\|_H + \|z(T,x)-G(x)\|_H < \gamma/2$$

and

$$\|F(z(T,y))\|_K \leq \|F(z(T,y)) - F(z(T,x))\|_K + \|F(z(T,x))\|_K < c\gamma/2.$$

According to Theorem 14, it must be that

$$\|z(t,y)-z(T,y)\|_H < \gamma/2,\ t \geq T$$

and so

$$\|G(y)-z(T,y)\|_H \leq \gamma/2,\ t \geq T$$

since $G(y) = \lim_{t\to\infty} z(t,y)$. Note also that Theorem 14 gives that

$$\|z(t,x)-z(T,x)\|_H < \gamma/2,\ t \geq T$$

and so we have that the convex hull of

$$G(x), G(y), \{z(t,x) : t \geq T\}, \{z(t,y) : t \geq T\}$$

is a subset of $B_\gamma(G(x)) \subset \alpha$. This gives us the second part of the conclusion since we already have that

$$[z(t,y), z(t,x)] \subset Q_\gamma,\ 0 \leq t \leq T.$$

□

LEMMA 13. *Suppose* $B \in L(H, K)$, $c > 0$ *and*

$$||B^* g||_H \geq c||g||_K, \ g \in K. \tag{14.4}$$

Then

$$|\exp(-tB^*B) - (I - B^*(BB^*)^{-1}B)| \leq \exp(-tc^2), \ t \geq 0.$$

Note that the spectral theorem (df [**91**]) gives that $\exp(-tB^*B)$ converges pointwise on H to $(I - B^*(BB^*)^{-1}B)$, the orthogonal projection of H onto $N(B)$, as $t \to \infty$. What Lemma 13 gives is exponential convergence in operator norm.

PROOF. First note that (14.4) is sufficient for

$$(BB^*)^{-1}$$

to exist and belong to $L(K, K)$. Note next the formula

$$\exp(-tB^*B) = I - B^*(BB^*)^{-1}B + B^*(BB^*)^{-1}\exp(-tBB^*)B \tag{14.5}$$

which is established by expanding $\exp(-tBB^*)$ in its power series and collecting terms, $t \geq 0$. Note also that

$$B^*(BB^*)^{-1}\exp(-tBB^*)B = B^*(BB^*)^{-1/2}\exp(-tBB^*)(BB^*)^{-1/2}B$$

and that

$$|B^*(BB^*)^{-1/2}| = |(BB^*)^{-1/2}B| \leq 1$$

and hence

$$|B^*(BB^*)^{-1}\exp(-tBB^*)B| \leq |\exp(-tBB^*)|.$$

Now denote by ξ a spectral family for BB^*. Since

$$< BB^*g, g >_K = ||B^*g||_K^2 \geq c^2||g||_K^2, \ g \in K$$

it follows that c^2 is a lower bound to the numerical range of BB^*. Denote by b the least upper bound to the numerical range of BB^*. Then

$$BB^* = \int_{c^2}^{b} \lambda \ d\xi(\lambda)$$

and

$$\exp(-tBB^*) = \int_{c^2}^{b} \exp(-t\lambda)d\xi(\lambda), \ t \geq 0.$$

But this implies that $|\exp(-tBB^*)| \leq \exp(-tc^2)$, $t \geq 0$. This fact together with (14.4),(14.5) give the conclusion to the lemma. □

LEMMA 14. *Suppose* x, γ, M, r, T, c *are as in Lemma 12. If* $||x - w|| < r$, *then*

$$||z(t, w) - G(w)|| \leq M_2 \exp(-tc^2), \ t \geq 0$$

where $M_2 = 2^{-1/2}||F(x)||/(1 - \exp(-c^2))$.

PROOF. This follows from the argument for Theorem 9. □

From ([**31**], Theorem (3.10.5)) we have that $z_2(t, w)$ exists for all $t \geq 0, w \in H$ and that z_2 is continuous. Furthermore if $Y(t, w) = z_2(t, w),\ t \geq 0,\ w \in H$, then

$$Y(0, w) = I, Y_1(t, w) = -(\nabla\phi)'(z(t, w))Y(t, w),\ t \geq 0,\ w \in H.$$

Consult [**31**] for background on various techniques with differential inequalities used in this chapter.

LEMMA 15. *Suppose x, γ, M, r, T, c are as in Lemma 12 and $\epsilon > 0$. There is $M_0 > 0$ so that if $t > s > M_0$ and $\|w - x\|_H < r$, then*

$$|Y(t, w) - Y(s, w)| < \epsilon.$$

PROOF. First note that if $\|w - x\|_H < r$ then

$$|Y(t, w)| \leq \exp(Mt),\ t \geq 0 \textit{ since}$$

$$Y(t, w) = I - \int_0^t (\nabla\phi)'(z(s, w))Y(s, w)ds, t \geq 0 \tag{14.6}$$

and $|(\nabla\phi)'(z(s, w))| \leq M,\ 0 \leq s$. In particular,

$$|Y(T, w)| \leq \exp(MT),\ \|w - x\|_H < r.$$

Suppose that $t > s \geq T$ and $\delta = t - s$. Then

$$|Y(t, w) - Y(s, w)| = \lim_{n \to \infty} |(\Pi_{k=1}^n (I - (\delta/n)(\nabla\phi)'(z(s + (k-1)\delta/n, w))))Y(s)|$$

(This is an expression that the Cauchy polygon methods works for solving (14.6) on the interval $[s, t]$). For n a positive integer and $\|w - x\|_H < r$,

$$|(\Pi_{k=1}^n (I - (\delta/n)(\nabla\phi)'(z(s + (k-1)\delta/n, w))))Y(s, w)|$$

$$\begin{aligned} &\leq |(\Pi_{k=1}^n (I - (\delta/n)(\nabla\phi)'(G(w))))Y(s, w)| \\ &\quad + |(\Pi_{k=1}^n (I - (\delta/n)(\nabla\phi)'(z(s + (k-1)\delta/n, w))))Y(s, w) \\ &\qquad - (\Pi_{k=1}^n (I - (\delta/n)(\nabla\phi)'(G(w))))Y(s, w)| \end{aligned}$$

Now by Lemma 13,

$$|(\Pi_{k=1}^n (I - (\delta/n)(\nabla\phi)'(G(w))))Y(s, w)| \leq \exp(-c^2 s)|Y(s, w)|.$$

Define

$$A_k = I - (\delta/n)(\nabla\phi)'(z(s + (k-1)\delta/n, w))$$

and

$$B_k = I - (\delta/n)(\nabla\phi)'(G(w)),$$

$k = 1, 2, ..., n$, and denote $Y(s, w)$ by W. By induction we have that

$$|(\Pi_{k=1}^n A_k)W - (\Pi_{k=1}^n B_k)W| \leq \sum_{k=1}^n |A_n \cdots A_{k+1}(A_k - B_k)B_{k-1} \cdots B_1 W|.$$

Now

$$
\begin{aligned}
|A_j| &\le |I-(\delta/n)(\nabla\phi)'(z(G(w)))| \\
&\quad + (\delta/n)|(\nabla\phi)'(G(w)) - (\nabla\phi)'(z(s+(j-1)\delta/n, w))| \\
&\le 1 + (\delta/n)M|(\nabla\phi)'(G(w)) - (\nabla\phi)'(z(s+(j-1)\delta/n, w))| \\
&\le 1 + (\delta/n)(\int_0^1 |(\nabla\phi)''((1-\tau)z(s+(j-1)\delta/n, w) + \tau G(w)|dr) \\
&\qquad \|G(y) - z(s+(j-1)\delta/n, w)\|_H \\
&\le 1 + (\delta/n)M\|G(y) - z(s+(j-1)\delta/n, w)\|_H \\
&\le 1 + (\delta/n)MM_2\exp(-c^2(s+(j-1)\delta/n) \\
&= 1 + (\delta/n)M_3\exp(-c^2 s)(\exp(-c^2\delta/n))^{j-1},
\end{aligned}
$$

$j = 1, ..., n$. We note that $|B_j| \le 1, j = 1, ..., n$. Note that

$$
\begin{aligned}
|A_n \cdots A_{k+1}| &\le |A_n| \cdots |A_{k+1}| \qquad (14.7) \\
&\le \Pi_{j=k+1}^n (1 + (\delta/n)M_3\exp(-c^2 s)(\exp(-c^2\delta/n))^{j-1}) \\
&\le \Pi_{j=k+1}^n \exp((\delta/n)M_3\exp(-c^2 s)(\exp(-c^2\delta/n))^{j-1}) \\
&\le \exp(M_3\exp(-c^2 s)(\delta/n)\sum_{j=k+1}^n (\exp(-c^2\delta/n))^{j-1}) \\
&\le \exp(M_3\exp(-c^2 s)(\delta/n)/(1-\exp(-c^2\delta/n))) \\
&\le \exp(M_4\exp(-c^2 s))
\end{aligned}
$$

so long as $\delta/n \le 1$ where $M_4 = M_3\sup_{\beta\in(0,1]}\beta/(1-\exp(-c^2\beta))$ and $M_3 = MM_2$. Note that

$$
\begin{aligned}
|A_k - B_k| &= (\delta/n)|(\nabla\phi)'(z(s+(j-1)\delta/n, w)) - (\nabla\phi)'(G(w))| \\
&= (\delta/n)|(\int_0^1 (\nabla\phi)''((1-\tau)z(s+(j-1)\delta/n, w) + \tau G(w))\, d\tau \\
&\qquad (z(s+(j-1)\delta/n, w) - G(w))| \qquad (14.8) \\
&\le (\delta/n)M|(z(s+(j-1)\delta/n, w) - G(w))| \\
&\le (\delta/n)MM_2\exp(-c^2 s)(\exp(-c^2\delta/n)_)^{j-1},\ k = 1, ..., n,
\end{aligned}
$$

and so using (14.7),(14.8), we get that

$$
\begin{aligned}
&|(\Pi_{k=1}^n A_k)W - (\Pi_{k=1}^n B_k)W| \\
&\le \sum_{k=1}^n \exp(M_4\exp(-c^2 s))|A_k - B_k||W| \\
&\le \exp(M_4\exp(-c^2 s))MM_2 exp(-c^2 s)\sum_{k=1}^n \exp(-c^2\delta/n)^{k-1})|W| \\
&\le \exp(M_4\exp(-c^2 s))\exp(-c^2 s))M_4|W|.
\end{aligned}
$$

Thus

$$|(\Pi_{k=1}^n(I-(\delta/n)(\nabla\phi)'(z(s+(k-1)\delta/n,w))))Y(s,w)| \tag{14.9}$$

$$\leq |(\Pi_{k=1}^n(I-(\delta/n)(\nabla\phi)'(G(w))))Y(s,w)| \\ +\exp(M_4\exp(-c^2s))\exp(-c^2s)M_4|Y(s,w)|$$

Taking the limit of both sides of (14.9) as $n\to\infty$, we have

$$|Y(t,w)-Y(s,w)|\leq \exp(-c^2s)|Y(s,w)|(1+\exp(M_4\exp(-c^2s))M_4)$$

$0<T\leq s<t$. Taking for the moment $s=T$ the above yields

$$|Y(t,w)-Y(T,w)|\leq \exp(-c^2T)|Y(T,w)|(1+\exp(M_4\exp(-c^2T))M_4).$$

Note $\{|Y(T,w)|,\ \|y-x\|_H<r\}$ is bounded, say by $M_5>0$. Hence

$$|Y(t,w)-Y(T,w)|\leq$$

$$\exp(-c^2s)M_5(1+\exp(M_4\exp(-c^2T))M_4),\ t>s\geq T,\ \|w-x\|_H<r,$$

and so the conclusion follows. □

Denote by U the function with domain $B_r(x)$ to which $\{Y(t,\cdot)\}_{t\geq 0}$ converges uniformly on $B_r(x)$. Note that $U:B_r(x)\to L(H,H)$. and U is continuous.

LEMMA 16. *Suppose that $x\in H$, $r>0$, each of $v_1,v_2,...$ is a continuous function from $B_r(x)\to H$, q is a continuous function from H to $L(H,H)$ which is the uniform limit of $v_1',v_2',....$ on $B_r(x)$. Then q is continuous and*

$$v'(y)=q(y),\|y-x\|_H<r.$$

PROOF. Suppose $y\in B_r(x)$, $h\in H$, $h\neq 0$ and $y+h\in B_r(x)$. Then

$$\int_0^1 q(y+sh)h\ ds=\lim_{n\to\infty}\int_0^1 v_n'(y+sh)h\ ds$$

$$=\lim_{n\to\infty}(v_n(y+h)-v_n(y))=v(y+h)-v(y).$$

Thus

$$\|v(y+h)-v(y)-q(y)h\|_H/\|h\|_H=\|\int_0^1 q(y+sh)h-q(y)h)ds\|$$

$$\leq\int_0^1|q(y+sh)-q(y)|ds\to 0$$

as $\|h\|\to 0$. Thus v is Frechet differentiable at each $y\in B_r(x)$ and $v'(y)=g(y)$, $y\in B_r(x)$. □

PROOF. To prove the Theorem, note that Lemmas 14,15,16 give the first conclusion. To establish the second conclusion, suppose that $g\in Q$. Suppose $x\in H$ and $\beta(t)=g(z(t,x))$, $t\geq 0$. Then

$$\beta'(t)=g'(z(t,x))z_1(t,x)=-g'(z(t,x))(\nabla\phi)(z(t,x)),\ t\geq 0.$$

Thus β is constant on $R_x = \{z(t,x) : t \geq 0\} \cup \{G(x)\}$. But if y is $\bar{i}n$ $G^{-1}(G(x))$ then $g(\{z(t,y) : t \geq 0\} \cup \{G(y)\})$ must also be in $G^{-1}(G(x))$ since $G(y) = G(x)$. Thus $G^{-1}(G(x))$ is a subset of the level set $g^{-1}(g(x))$ of g. Therefore,

$$G^{-1}(G(x)) \subset \cap_{g\in Q} g^{-1}(g(x)).$$

Suppose now that $x \in H$, $y \in \cap_{g\in Q} g^{-1}(g(x))$ and $y \notin G^{-1}(G(x))$. Denote by f a member of H^* so that $f(G(x)) \neq f(G(y))$. Define $p : H \to R$ by $p(w) = f(G(w)), w \in H$. Then $p'(w)h = f'(G(w))G'(w)$ and so

$$p'(w)(\nabla\phi)(w) = fG'(w)G'(w)(\nabla\phi)(w) = 0,$$

$w \in H$, and hence $p \in Q$, a contradiction since

$$y \in \cap_{q\in Q} g^{-1}(g(x)) and p \in Q$$

together imply that $p(y) = p(x)$. Thus

$$G^{-1}(G(x)) \supset \cap_{q\in Q} g^{-1}(g(x))$$

and the second part of the theorem is established. □

We end this section with an example.

Example. Take H to be the Sobolev space $H^{1,2}([0,1]), K = L_2([0,1])$,

$$F(y) = y' - y,\ y \in H.$$

We claim that the corresponding function G is specified by

$$(G(y))(t) = exp(t)(y(1)e - y(0))/(e^2 - 1),\ t \in [0,1],\ y \in H. \tag{14.10}$$

In this case

$$G^{-1}(G(x)) = \{w \in H : w(1)e - w(0) = y(1)e - y(0)\}.$$

This may be observed by noting that since F is linear,

$$\nabla\phi)(y) = F^*Fy,\ y \in H.$$

The equation

$$z(0) = y \in H, z'(t) = -F^*Fz(t),\ t \geq 0$$

has the solution

$$z(t) = \exp(-tF^*F)y,\ t \geq 0.$$

But $\exp(-tF^*F)y$ converges to

$$(I - F^*(FF^*)^{-1}F)y,$$

the orthogonal projection of y onto $N(F)$, *i.e.*, the solution $w \in H$ to $F(w) = 0$ that is nearest (in H) to y. A little calculation shows that this nearest point is given by $G(y)$ in (14.10). The quantity $(y(1)e - y(0))/(e^2 - 1)$ provides an invariant for steepest descent for F relative to the Sobolev metric $H^{1,2}([0,1])$. Similar reasoning applies to all cases in which F is linear but invariants are naturally much harder to exhibit for more complicated functions F. In summary, for F linear, the corresponding function G is just the orthogonal projection of H onto $N(F)$. In general G is a projection on H in the sense that it is an idempotent transformation. We hope that a study of these idempotents in increasingly involved cases will give information about 'boundary conditions' for significant classes of partial differential equations.

2. Another Solution Giving Nonlinear Projection

Suppose that each of H and K is a Hilbert space and $F : H \to K$ is a $C^{(2)}$ transformation such that

$$(F'(x)F'(x)^*)^{-1} \in L(K, K), \; x \in H$$

and $F(0) = 0$. Denote by P, Q functions on H such that if $x \in H$, then $P(x)$ is the orthogonal projection of H onto the null space of $F'(x)$ and $Q(x) = I - P(x)$. Denote by f, g the functions on $R \times H \times H$ to H such that if $x, \lambda \in H$ then

$$f(0, x, \lambda) = x, \; f_1(t, x, \lambda) = P(f(t, x, \lambda)\lambda, \; t \geq 0,$$

and

$$g(0, x, \lambda) = x, \; g_1(t, x, \lambda) = Q(g(t, x, \lambda)\lambda, \; t \geq 0.$$

Define M from H to H so that if $\lambda \in H$ then

$$M(\lambda) = g(1, f(1, 0, \lambda), \lambda).$$

The following is due to Lee May [**47**].

Theorem 46. *There is an open subset V of H, centered at $0 \in H$ such that the restriction of M to V is a diffeomorphism of V and $M(V)$ is open.*

Denote by S the inverse of the restriction of M to V and denote by by G the function with domain S so that

$$G(w) = f(1, 0, S(w)), \; w \in R(M).$$

The following is also from [**47**]:

Theorem 47. *$R(G) \subset N(F)$.*

The function G is a solution giving nonlinear projection. Two elements $x, y \in D(S)$ are said to be equivalent if $G(x) = G(y)$. Arguments in [**47**] use an implicit function and are essentially constructive. Thus G associates each element near enough to 0 with a solution. Many of the comments about the function G of the preceding section apply as well to the present function G. The reader might consult [**47**] for careful arguments for the two results of this section.

CHAPTER 15

Some Related Iterative Methods for Differential Equations

This chapter describes to developments which had considerable influence on the theory of Sobolev gradients. They both deal with projections. As a point of departure, we indicate a result of Von Neumann ([**100**],[**36**]).

THEOREM 48. *Suppose H is a Hilbert space and each of P and L is an orthogonal projection on H. If Q denotes the orthogonal projection of H onto $R(P) \cap R(L)$, then*

$$Qx = \lim_{n \to \infty} (PLP)^n x, \ x \in H. \tag{15.1}$$

If $T, S \in L(H,H)$, and are symmetric, then $S \leq T$ means

$$\langle Sx, x \rangle \leq \langle Tx, x \rangle, \ x \ inH.$$

PROOF. (Indication) First note that $\{(PLP)^n\}_{n=1}^{\infty}$ is a non-increasing sequence of symmetric members of $L(H,H)$ which is bounded below (each term is non-negative) and hence $\{(PLP)^n\}_{n=1}^{\infty}$ converges pointwise on H to a non-negative symetric member Q of H which is idempotent and which commutes with both P and L. Since Q is also fixed on $R(P), R(L)$ it must be the required orthogonal projection. □

How might this applied to differential equations is illustrated first by a simple example.

Example 15.1. Suppose $H = L_2([0,1])^2$, P is the orthogonal projection of H onto

$$\{(\begin{smallmatrix} u \\ u' \end{smallmatrix})\} : u \in H^{1,2}([0,1])\}$$

and L is the orthogonal projection of H onto

$$\{(\begin{smallmatrix} u \\ u \end{smallmatrix}) : u \in L_2([0,1])\}.$$

Then

$$R(P) \cap R(L) = \{(\begin{smallmatrix} u \\ u' \end{smallmatrix}) : u \in H^{1,2}([0,1]), \ u' = u\}.$$

Thus $R(P) \cap R(L)$ essentially yields solutions u to $u' = u$.

The above example is so simple that a somewhat more complicated one might shed more light.

Example 15.2. Suppose that each of a, b, c, d is a continuous real-valued function on $[0,1]$ and denote by L the orthogonal projection of $H = L_2([0,1])^4$ onto

$$\{(u, au + bv, v, cu + dv) : u, v \in L_2([0,1])\}. \tag{15.2}$$

Denote by P the orthogonal projection of H onto

$$\{(u, u', v, v') : u, v \in H^{1,2}([0,1])\} \tag{15.3}$$

Then

$$(u, f, v, g) \in R(P) \cap R(L)$$

if and only if

$$f = u', g = v', \ f = au + bv, g = cu + dv,$$

that is, the system

$$u' = au + bv, \quad v' = cu + dv \tag{15.4}$$

is satisfied.

In a sense we split up (15.4) into an algebraic part represented by (15.2) and an analytical part (i.e., a part in which derivatives occur) (15.3). Then $R(P) \cap R(L)$ gives us all solutions to (15.4). One may see that the iteration (15.1) provides a constructive way to calculate a solution since we have already seen that the P may be presented constructively. As to a construction for L observe that if

$$t \in [0,1], \ \alpha = a(t), \ \beta = b(t), \ \gamma = c(t), \ \delta = d(t), \ p, q, r, s \in R$$

then the minimum (x, y) to

$$\|(p, q, r, s) - (x, \alpha x + \beta y, y, \gamma x + \delta y)\| \tag{15.5}$$

is given by the unique solution (x, y) to

$$(1 + \alpha^2 + \gamma^2)x + (\alpha\beta + \gamma\delta)y = p + \alpha q + \gamma s$$
$$(\alpha\beta + \gamma\delta)x + (1 + \beta^2 + \delta^2)y = r + \beta q + \delta s.$$

We remark that appropriate boundary conditions could be imposed using the ideas of Chapter 6.

From [**55**] there is a nonlinear generalization of (15.1). Suppose that H is a Hilbert space, P is an orthogonal projection on H and Γ is a strongly continuous function from H to

$$S = \{T \in L(H, H) : T^* = T, 0 \leq \langle Tx, x\rangle \leq \|x\|^2, \ x \in H\}.$$

For T any non-negative symmetric member of $L(H, H), T^{1/2}$ denotes the unique non-negative symmetric square root of T. By Γ being strongly continuous we mean that if $\{x_n\}_{n=0}^{\infty}$ is a sequence in H converging to $x \in H$ and $w \in H$, then

$$lim_{n\to\infty}\Gamma(x_n)w = \Gamma(x)w.$$

THEOREM 49. *Suppose* $w \in H, Q_0 = P$,

$$Q_{n+1} = Q_n^{1/2}\Gamma(Q_n^{1/2}w)Q_n^{1/2},\ n = 1, 2, \ldots.$$

Then

$$\{Q_n^{1/2}w\}_{n=1}^{\infty} \text{ converges to } z \in H$$

such that

$$Pz = z,\ \Gamma(z)z = z.$$

.

PROOF. First note that Q_0 is symmetric and nonnegative. Using the fact that the range of Γ contains only symmetric and nonegative members of $L(H, H)$, one has by induction that each of $\{Q_n\}_{n=1}^{\infty}$ is also symmetric and nonnegative. Moreover for each positive integer n, and each $x \in H$,

$$\begin{aligned}\langle Q_{n+1}x, x\rangle &= \langle Q_n^{1/2}\Gamma(Q_n^{1/2}x)Q_n^{1/2}x, x\rangle \\ &= \langle \Gamma(Q_n^{1/2}x)Q_n^{1/2}x, Q_x^{1/2}\rangle \leq \langle Q_n^{1/2}x, Q_n^{1/2}x\rangle = \langle Q_n x, x\rangle, \quad (15.6)\end{aligned}$$

so that $Q_{n+1} \leq Q_n$. Hence $\{Q_n\}_{n=1}^{\infty}$ converges strongly on H to a symmetric nonnegative transformation Q and so also $\{Q_n^{1/2}\}_{n=1}^{\infty}$ converges strongly to $Q^{1/2}$. Denote by z the limit of $\{Q_n^{1/2}w\}_{n=1}^{\infty}$ and note that then $\{\Gamma(Q_n^{1/2}w)Q_n^{1/2}w\}_{n=1}^{\infty}$ converges to $\Gamma(z)z$. Since for each positive integer $n, Q^{1/2}$ is the strong limit of a sequence of polynomials in Q_n, it follows by induction that $PQ^{1/2} = Q^{1/2}$. Hence $Pz = z$.

Since for each positive integer n and each $x \in H$,

$$\langle Q_{n+1}x, x\rangle = \langle \Gamma(Q_n^{1/2}w)Q^{1/2}x, Q^{1/2}x\rangle,$$

it follows that

$$\langle Qx, x\rangle = \langle Q^{1/2}x, Q^{1/2}x\rangle = \langle \Gamma(z)Q^{1/2}w, Q^{1/2}w\rangle$$

and hence

$$\langle (I - \Gamma(z))x, x\rangle = 0,\ x \in H$$

and so

$$(I - \Gamma(z))z = 0, \text{i.e.,} \Gamma(z)z = z.$$

This together with the already established fact that $Pz = z$ is what was to be shown. ☐

Further examples of a nonliner projection methods are given in [**24**] (particularly Proposition 5) and in references contained therein.

Brown and O'Malley in [**12**] have generalized the above result. The next Lemma and Theorem give their result. Denote by $B_1(H)$ the set of symmetric, bounded members of $L(H, H)$ which have numerical range in $[0, 1]$.

LEMMA 17. *(From* [**12**]*) Let* $Q \in B_1(H)$ *and let* α *be a positive rational number other than* 1. *If* $Q^{\alpha} = Q$, *then* $Q = Q^2$.

PROOF. Let $\alpha = r/s$; the presumed equality is equivalent to $Q^r = Q^s$. Assume that $r < s$ and that r is the minimal positive power of Q which reoccurs in the sequence $\{Q^n\}_{n=1}^{\infty}$. From the fact that powers of an operator descend in the quasi-order mentioned above, together with the limited anti-symmetry of this relation, it follows that $Q^t = Q^r$ for all integral t between r and s. From $Q^r = Q^{r+1}$, it follows that $Q^t = Q^r$ for all $t \geq r$. If r is odd, then

$$(Q^{(r+1)/2})^2 = Q^{r+1} = Q^{2r} = (Q^r)^2$$

By uniqueness of square roots, $Q^r = Q^{(r+1)/2}$ whence $r = (r+1)/2$ and $r = 1$. If r is even, then $(Q^{r/2})^2 = A^r = (Q^r)^2$, whence $r = r/2$, which is impossible for positive r. Thus $r = 1$ and $Q = Q^2$. □

THEOREM 50. *Let $w \in H$, let P be an orthogonal projection on H, and let $L : H \to B_1(H)$ be strongly continuous. Let α, β be positive rational numbers with $\alpha \in [1/2, \infty)$. Set $Q_0 = P$, and let*

$$Q_{n+1} = Q_n^{\alpha} L(Q_n^{\beta} w) Q_n^{\alpha},\ n = 1, 2, \ldots.$$

Then $\{Q_n\}_{n=0}^{\infty}$ is a decreasing sequence of elements of $B_1(H)$ which converges to an element $Q \in B_1(H)$ such that (1) if $\alpha > 1/2$ then Q is idempotent and $z = Qw$ satisfies $L(z)z = z$ and $Pz = z$, and (2) if $\alpha = 1/2, \beta \geq 1/2$, then $z = Q^{\beta} w$ satisfies $L(z)z = z, Pz = z$.

PROOF. (From [**12**].) Fix $\alpha \geq 1/2, \beta > 0$. Since $Q_0 = P \in B_1(H)$ and the range of L is in $B_1(H)$, it follows inductively that $Q_n \in B_1(H)$ for all n. Since $2\alpha \geq 1, Q_n^{2\alpha} \leq Q_n$; moreover,

$$\begin{aligned}\langle (Q_n^{2\alpha} - Q_{n+1})x, x\rangle &= \langle (Q_n^{2\alpha} - Q_n^{\alpha} L(Q_n^{\beta} w) Q_n^{\alpha})x, x\rangle \\ &= \langle Q_n^{\alpha}(I - L(Q_n^{\beta} w)) Q_n^{\alpha} x, x\rangle = \langle (I - L(Q_n^{\beta} w)) Q_n^{\beta} w)) Q_n^{\alpha} x, Q_x^{\alpha}\rangle.\end{aligned}$$

Thus, since $I - L(Q_n^{\beta} x) \geq 0$, if follows that $Q_{n+1} \leq Q_n^{2\alpha}$. Hence we have

$$Q_{n+1} \leq Q_n^{2\alpha} \leq Q_n,\ n = 0, 1, 2, \ldots. \tag{15.7}$$

In particular, the sequence $\{Q_n\}_{n=1}^{\infty}$ is monotonically decreasing in the (operator) interval from 0 to I. Thus we have by [**96**] that the sequence $\{Q_n^{\alpha}\}_{n=1}^{\infty}$ converges to Q^{α} and $\{Q_n^{\beta}\}_{n=1}^{\infty}$ converges to Q^{β}. Since L is continuous and operator multiplication is jointly continuous in the strong topology on $B_1(H)$, we have by uniqueness of limits that

$$Q = Q^{\alpha} L(Q(^{\beta} w) Q^{\alpha}.$$

Also from (15.7) and the closed graph of the relation $\leq$, we have

$$Q \leq Q^{2\alpha} \leq Q.$$

Thus, since $Q, Q^{2\alpha}$ commute, we have that $Q = Q^{2\alpha}$. Moreover, since $P = Q_0$, we have $PQ_n = Q_n$, whence $PQ^{\gamma} = Q^{\gamma}$ for all positive rational γ.

(i) Suppose $\alpha > 0$. By (17) $Q = Q^2$, from which it follows that $Q = Q^{\gamma}$ for all positive rational γ, and in particular $Q = QL(Qw)Q$.

Let $z = Qw$, and fix $x \in H$.

$$\langle Qx, x\rangle = \langle QL(z)Qx, x\rangle = \langle L(z)Qx, Qx\rangle,$$

and since $Q^2 = Q$, it follows that

$$0 = \langle Qx, Qx\rangle - \langle L(z)Qx, Qx\rangle = \langle (I - L(z))Qx, Qx\rangle.$$

Therefore, since $I - L(z)$ and hence $(I - L(z))^{1/2}$ belong to $B_1(H)$, we have that $Q = L(z)Q$. In particular, $z = Qw = L(z)Qw = L(z)z$.

(ii) Suppose $\alpha = 1/2, \beta \geq 1/2$. Let $z = Q^\beta w$; then $Q = Q^{1/2}L(z)q^{1/2}$ from which

$$\langle Qx, x\rangle = \langle Q^{1/2}L(z)Q^{1/2}x, x\rangle = \langle L(z)Q^{1/2}x, Q^{1/2}x\rangle.$$

Since $\langle Qx, x\rangle = \langle Q^{1/2}, Q^{1/2}\rangle$ also we have

$$0 = \langle Q^{1/2}x - L(z)Q^{1/2}x, Q^{1/2}x\rangle = \langle (I - L(z))Q^{1/2}x, Q^{1/2}x\rangle.$$

Now as in (i), it follows that $Q^{1/2} = L(z)Q^{1/2}$. In particular,

$$z = Q^\beta w = Q^{1/2}Q^{\beta-1/2}w = L(z)Q^{1/2}Q^{\beta-1/2}w = L(z)Q^\beta w = L(z)z.$$

That $Pz = z$ in both cases is obvious from the fact that $PQ^\gamma = Q^\gamma$ for all positive rational γ. □

Further applications to differetial equations are illustrated by the following simple case. Compare results of Chapter 3. Take A as a function from $L_2([0,1)])^2$ into $L(L_2([0,1])^2, L_2([0,1]))$ and take $D : H^{1,2}([0,1]) \to L_2([0,1])$ so that $Du = \binom{u}{u'}$, $u \in H^{1,2}([0,1])$. Make the assumption that

$$A(Du)A(Du)^* = I, \text{the identity on} L_2([0,1])$$

a result that can be obtained by normalization granted that $A(Du)A(Du)^*$ is invertable. Define Γ so that

$$\Gamma(u) = u - A(Du)^*A(Du), u \in H^{1,2}([0,1]).$$

Take P to be the orthogonal projection of $L_2([0,1])$ onto $R(D)$. Then if $z \in H^{1,2}([0,1])$ and

$$Pz = z, \Gamma(z)z = z$$

hold, i.e., that the conclusion to Theorem 49 holds, it follows that

$$A(Du)Du = 0. \tag{15.8}$$

Equation (15.8) represents a substantial family of quasilinear differential equations.

The following problem from **[56]** is related to the above projection methods. Suppose we hve two Hilbert spaces H, K, an othogonal projection P on $H, g \in K$ and a continuous linear transformation B from H to K such that $BB^* = I$. Find y in H so that

$$By = g \text{ and } Py = y. \tag{15.9}$$

This is equivalent to finding $x \in H$ so that

$$BPx = g. \tag{15.10}$$

Make the definitions $L = I - B^*B, L_g x = Lx + B^*x, x\ inH$ and note that if $x \in H$, then $L_g x$ is the nearest element z to x so that $Bz = g$. We seek solutions x to (15.10) as

$$x = lim_{n\to\infty}(PL_gP)^n z \text{ for some } z \in H.$$

First note that if $z \in H, g \in K$, by induction we have

$$(PL_gP)^k z = (PLP)^k z + PB^*(I + (I - M) + \cdots + (I - M)^{k-1})g, k = 1, 2, \ldots \tag{15.11}$$

where

$$M = BPB^*$$

Note that M is a symmetric nonnegative linear transformation from K to K and $M \leq I$ so that the numerical range of M is a subset of $[0, 1]$. The next theorem gives a characterization of $R(\bar{B}P)$ and the one after that gives a characterization of $R(BP)$.

Theorem 51. *If $g \in K$ then $g \in N(PB^*)^\perp$ if and only if*

$$lim_{k\to\infty}(I - M)^k g = 0.$$

Proof. Define $z = lim_{k\to\infty}(I - M)^k g$. This limit exists since M is symmetric and $0 \leq I - M \leq I$. Now

$$(I - M)z = lim_{k\to\infty}(I - M)^{k+1}g = z$$

so that $Mz = 0$ and hence $BPB^*z = 0$. Therefore

$$0 = \langle BPB^*z, z\rangle = ||PB^*z||^2.$$

Hence $z \in N(PB^*)$. If in addition, $g \in N(PB^*)^\perp$, then

$$0 = \langle g, z\rangle = \langle g, lim_{k\to\infty}(I - M)^{2k}g\rangle = ||lim_{k\to\infty}(I - M)^k g||^2 = ||z||^2$$

and hence $z = 0$.

Now suppose that $z = 0$ and $w \in N(PB^*)$. Then

$$0 = ||z||^2 = \langle \lim_{k\to\infty}(I - M)^k g, w\rangle = \langle g, \lim_{k\to\infty}(I - M)^k g, w\rangle = \langle g, w\rangle$$

since $(I - M)w = z$. Hence $\langle g, w\rangle = 0$ for all $w \in N(PB^*)$ and so $g \in N(PB^*)^\perp$. □

Theorem 52. *If $g \in R(\bar{B}P)$ then $g \in R(BP)$ if and only if*

$$lim_{k\to\infty}PB^*(I + (I - M) + \cdots + (I - M)^{k-1})g$$

exists.

PROOF. Suppose the limit in the Theorem exists and call it z. Note that $Pz = z$. Then

$$Bz = lim_{k\to\infty} BPB^*(I + (I - M) + \cdots + (I - M)^{k-1})g$$
$$= lim_{k\to\infty} M(I + (I - M) + \cdots + (I - M)^{k-1})g$$
$$= lim_{k\to\infty}(g - (I - M)^k g) = g$$

since $g \in R(\bar{B}P) = N(PB^*)^\perp$ and so by (51) $lim_{k\to\infty}(I - M)^k g = 0$. Therefore $Bz = g$ and so $BPz = g$ since $Pz = z$. □

THEOREM 53. *Suppose* $g \in H, lim_{k\to\infty}(I - M)^k g = 0$ *and*

$$lim_{k\to\infty} PB^*(I + (I - M) + \cdots + (I - M)^{k-1})g$$

exists. If $z \in H$*, then* $x = lim_{k\to\infty}(PL_g P)^k z$ *exists and satisfies* $Px = x, Bx = g$*, and hence* $BPx = g$*. Moreover,* x *is the nearest point* w *to* z *so that* $Pw = w, Bw = g$*.*

PROOF. Suppose $z \in H$. Define $r = lim_{k\to\infty}(PLP)^k z$ and

$$y = PB^*(I + (I - M) + \cdots + (I - M)^{k-1})g.$$

Using (15.11), $x = lim_{k\to\infty}(PL_g P)^k z = r + y$. But $Pr = r, Br = 0$ and, as in the proof of (52), $Py = y, By = g$. Therefore $Px = x, Bx = g$ and consequently $BPx = x$.

Suppose now that $w \neq x, Pw = w$. Then $BP(w - x) = g - g = 0$. Since $r = lim_{k\to\infty}(PLP)^k z$, if follows that r is the nearest point of $R(P) \cap N(B)$ to z. Hence $\langle r - z, w - x\rangle = 0$ since $w - x \in R(P) \cap R(L)$. Also since for each positive integer k,

$$\langle w - x, PB^*(I + (I - M) + \cdots + (I - M)^{k-1})g\rangle$$
$$= \langle BP(w - x), (I + (I - M) + \cdots + (I - M)^{k-1})g\rangle = 0,$$

it follows that $\langle w - x, y\rangle = 0$. Hence x is the closest point w to z such that $Pw = w, Bw = g$. □

We apply the above development to a problem of functional differential equations with both advanced and retarded arguments. If $c \in R, f\ H = H^R$ then f_c denotes the member of H so that $f(t) = f(t + c), t \in R$. Suppose $\alpha, \beta > 0, g \in C^{(1)}(R), r, s \in C(R), r, s$ bounded . We have the problem of finding $f \in H$ so that

$$f' + rf_\alpha + sf_{-\beta} = g. \tag{15.12}$$

. This is a functional differential equation with both advanced and retarded arguments **[35]**,**[23]** . Define $A : L_2(R)^2 \to L_2(R)$ by

$$A\binom{f}{g} = g + rf_\alpha + sf_{-\beta}, f, g \in L_2(R)^2. \tag{15.13}$$

LEMMA 18. *Suppose A satisfies* (15.13). *Then*

$$A^*z = \binom{(rz)_{-\alpha}+(sz)_\beta}{z},\ z \in L_2(R).$$

Moreover, AA^* has a bounded inverse defined on all of $L_2(R)$.

PROOF. For $\binom{f}{g} \in L_2(R)^2, z \in L_2(R)$

$$\langle A\binom{f}{g}, z\rangle_{L_2(R)} = \int_{-\infty}^{\infty} (rf_\alpha + sf_{-\beta} + g)z =$$

$$\int_{-\infty}^{\infty} (f(rz)_{-\alpha} + f(sz)_\beta + gz) = \langle \binom{f}{g}, \binom{((rz)_{-\alpha}+(sz)_\beta}{z})\rangle_{L_2(R)}.$$

The first conclusion then follows. To get the second, compute

$$AA^*z = z + r[(rz)_{-\alpha} + (sz)_\beta]_\alpha + s[(rz)_{-\alpha} + (sz)_\beta]_{-\beta}$$
$$= (1 + r^2 + s^2)z + r(sz)_{\alpha+\beta} + s(rz)_{-\alpha-\beta} = (I + CC^*)z$$

where

$$C\binom{f}{g} = rf_\alpha + sf_{-\beta},\ \binom{f}{g} \in L_2(R)^2.$$

It is then clear that the second conclusion holds. □

Define $B = (AA^*)^{1/2}A$. Then B satisfies the hypothesis of the previous three theorems.

LEMMA 19. *The orthogonal projection Q of $L_2(R)^2$ onto $\{\binom{u}{u'}, u \in H^{1,2}(R)$ is given by $Q\binom{f}{g} = \binom{u}{u'}$ where*

$$u(t) = (e^t/2)\int_t^{\infty} e^{-s}(f(s) - g(s))\ ds + (e^{-t}/2)\int_{-\infty}^{t} e^s(f(s) + g(s))\ ds, t \in R.$$

PROOF. Note that Q is idempotent, symmetric and fixed on all points of $\{\binom{u}{u'} : u \in H_{1,2}(R)\}$ and has range in this set. This is enough to show that Q is the required orthogonal projection. □

The formula in Lemma 19 came from a calculation like that of Problem 5.2 of Chapter 5. On this infinite interval one has the requirement that various functions are in $L_2(R)$ but there are no boundary conditions.

With A as in Lemma 18 and P as in Lemma 19, form B as above and take $M = BPB^*$. If the conditions of Theorem 53 are satisfied one is assured of existence of a solution to (15.12). Linear functional differential equations in a region $\Omega \subset R^m$ for $m > 1$ may be cast similarly. As an alternative to these projection methods, one can use the continuous steepest descent propositions of Chapters 3 and 4 and numerical techniques similar to those of Chapter 8. We do not have concrete results of functional differential equations to offer as of this writing but we suspect that there are possibilities along the lines we have indicated. For references on functional differential equations in addition to **[35]**,**[23]** we mention the references **[3]**, **[93]**, **[101]**. In this last reference there are refinements to some of the propositions of this chapter.

CHAPTER 16

A Related Analytic Iteration Method

This chapter deals with an iteration scheme for partial differential equations which, at least for this writer, provided a predecessor theory to the main topic of this monograph. Investigations leading to the main material of this monograph started after a failed attempt to make numerical the material of the present chapter. The scheme deals with spaces of analytic functions, almost certainly not the best kind of spaces for partial differential equations. To deal with these iterations we will need some notation concerning higher derivatives of functions on a Euclidean space.

Suppose that each of m and n is a positive integer and u is a real-valued $C^{(1)}$ function on an open subset of R^m. For $k \leq n$ and $x \in D(u)$, $u^{(k)}$ denotes the derivative of order k of u at x - it is a symmetric $k-linear$ function on R^m (cf [**76**]).

Denote by $M(m, n)$ the vector space of all real-valued $n-linear$ functions on R^m and for $v \in M(m, n)$ take

$$\|v\| = (\sum_{p_1=1}^{m} \cdots \sum_{p_n=1}^{m} (v(e_{p_1}, \ldots, e_{p_n})))^{1/2} \qquad (16.1)$$

where $e_1, \ldots, e_m$ is an orthonormal basis of R^m. As Weyl points out in [**103**], p 139, the expression in (16.1) does not depend on particular choice of orthonormal basis. Note that the norm in (16.1) carries with it an inner product.

Denote by $S(m, n)$ the subspace of $M(m, n)$ which consists of all symmetric members of $M(m, n)$ and denote by $P_{m,n}$ the orthogonal projection of $M(m, n)$ onto $S(m, n)$. For $y \in R^m, v \in S(m, n)$, vy^n denotes $v(y_1, \ldots, y_n)$ where $y_j = y,\ i = 1, \ldots, n$.

Suppose $r > 0$ and u is a $C^{(1)}$ function whose domain includes $B_r(0)$. We have the Taylor formula

$$u(x) = \sum_{q=0}^{n-1} u^q(0)x^q + \int_0^1 ((1-s)^{n-1}/(n-1)!)u^{(n)}(sx)x^n \, ds. \qquad (16.2)$$

For $A, w \in S(m, n), Aw$ denotes the inner product of A with w taken with respect to (16.1).

Now suppose that k is a positive integer, $k < n$, f is a continuous function from

$$R^m \times R \times S(1,n) \times \cdots \times S(k,n) \to R$$

and A is a member of $S(m,n)$ with $\|A\| = 1$. Given $r > 0$ one has the problem of finding $u \in C(n)$ so that

$$Au^{(n)}(x) = f(x, u(x), u'(x), \ldots, u^{(k)}(x)),\ \|x\| \leq r. \tag{16.3}$$

We abbreviate the rhs of (16.3) by $f_u(x)$.

To introduce our iteration method, note that if $v \in S(m,n), g \in R$ then

$$v - (Av - g)A$$

is the nearest element w of $S(m,n)$ to v so that $Aw = g$. This observation together with (16.2) leads us to define

$$T : C^{(n)}(B_r(0)) \to C(B_r(0))$$

by

$$(Tv)(x) = \sum_{q=0}^{n-1} v^q(0)x^q +$$
$$\int_0^1 ((1-s)^{n-1}/(n-1)!)(v^{(n)}(sx) - (Av^{(n)}(sx) - f_v(sx))A)x^n\, ds,$$
$$\|x\| \leq r,\ v \in C^{(n)}(B_r(0)). \tag{16.4}$$

We are interested in the possible convergence of

$$\{(T^j v)(x)\}_{j=1}^{\infty} \tag{16.5}$$

at least for $\|x\| \leq \rho$ for some $\rho \leq r$. This convergence is a subject of a series of papers **[51]**,**[52]**, **[53]**,**[54]** and **[81]**,**[82]**,**[83]**, **[84]**,**[85]**. Some results from these papers are indicated in what follows.

For $r > 0$ denote by α_r the collection of all real-valued functions v so that

$$v(x) = \sum_{q=0}^{\infty}(1/q!)v^{(q)}x^q, \sum_{q=0}^{\infty}(1/q!)\|v^{(q)}\|x^q < \infty,\ \|x\| \leq r. \tag{16.6}$$

Theorem 54. *If $r > 0, u \in C^{(n)}(B_r(0))$, then $Tu = u$ if and only if* (16.3) *holds.*

THEOREM 55. *Suppose $r > 0$ and h is a real-valued polynomial on R^m, $A \in S(m,n), \|A\| = 1$ and*

$$(Tv)(x) = \sum_{q=0}^{n-1} v^q(0)x^q + \int_0^1 ((1-s)^{n-1}/(n-1)!)(v^{(n)}(sx) - (Av^{(n)}(sx) - h(sx))A)x^n \, ds, \quad \|x\| \le r, v \in \alpha_r. \tag{16.7}$$

Then $\{T^j v\}_{j=0}^{\infty}$ converges uniformly on $B_r(0)$ to $u \in \alpha_r$ such that

$$Au^{(n)}(x) = h(x), \ \|x\| \le r.$$

The next two theorems require a lemma from [**52**]. In order to state the lemma we need some additional notation. There are several kinds of tensor products which we will use. Suppose that each of n, k is a positive integer, $n \le k$, i is a nonnegative integer, $i \le n$. For $A \in S(m,n), D \in S(m,k)$ define $A \otimes_i D$ to be the member $w \in S(m, n+k-2i)$ so that

$$w(y_1, \ldots, y_{n+k-2i}) = \langle A(y_1, \ldots, y_{n-i}), D(y_{n-i+1}, \ldots, y_{n+k-2i}\rangle, \quad y_1, \ldots, y_{n+k-2i} \in R^m,$$

where the above inner product $\langle \ , \ \rangle$ is taken in $S(m,i)$. The case $i = 0$ describes the usual tensor product of A and D; the case $i = n$ describes what we will call the inner product of A and D and in this case we will denote $A \otimes_n D$ as simply $AD \in S(m, k-n)$ and call AD the inner product of A with D. If $k = n$ the this agrees with our previous convention. The symmetric product $A \vee D$ is defined as $P_{m,n}(A \otimes D)$.

The following lemma is crucial to proving convergence of our iteration in nonlinear cases. It took about seven years to find; an argument is found in [**52**].

LEMMA 20. *If $A, C \in S(m,n), B, D \in S(m,k)$ and $n \le k$, then*

$$\binom{n+k}{n}\langle A \vee B, C \vee D\rangle = \sum_{i=0}^{n} \binom{n}{i}\binom{k}{i}\langle A \otimes_i D, C \otimes_i B\rangle. \tag{16.8}$$

Observe that the lemma has the easy consequence

$$\|A \vee B\|^2 \ge \binom{n+k}{n}^{-1} \|A\|^2 \|B\|^2. \tag{16.9}$$

This inequality is crucial to the following (essentially from [**53**],[**54**]):

THEOREM 56. *Suppose that $r > 0, v \in \alpha_r$ and f in (16.3) is real-analytic at*

$$(0, v(0), v'(0), \ldots, v^{(k)}(0)) \in R^m \times R \times S(1,n) \times \cdots \times S(k,n).$$

If $k \leq n/2$ there is $\rho \in (0, r]$ so that $\{T^j v\}_{j=0}^{\infty}$ converges uniformly on $B_\rho(0)$ to $u \in \alpha_\rho$ such that

$$Au^{(n)}(x) = f_u(x), \; \|x\| \leq \rho.$$

In [**82**] Pate generalizes the above in the direction of weakening the requirement $k \leq n/2$. The key is an analysis of higher order terms of

$$\binom{n+k}{n} \|A \vee B\|^2 = \sum_{i=0}^{n} \binom{n}{i} \binom{k}{i} \|A \otimes_i B\|^2$$

which is just (16.8) with $C = A \in S(m, n), D = B \in S(m, k)$. The main fact in this analysis is the result from [**81**] that

$$\binom{n+k}{n} \|A \vee B\|^2 = \sum_{i=0}^{\min(n/2,k)} \binom{n}{i} \binom{k}{i} \|A \otimes B\|^2 \mu_q(A)$$

where for $0 \leq q \leq \min(n/2, k)$, $\mu_q(A)$ is defined to be $\lambda_q(A)/\|A\|^2$ and $\lambda_q(A)$ is defined as follows: For q in this range and A a non zero member of $S(M, n)$ define

$$A_q : S(m, q) \to S(m, n-q)$$

by $A_q u = Au$, the above defined 'inner product' of A and u. Then denote $A_q^t : S(m, n-q) \to S(m, n)$ the corresponding adjoint transformation. Finally, define $T_q : S(m, n) \to S(m, n)$ by $T_q u = A_q^t(A_q u), u \in S(m, n)$. Then $\lambda_q(A)$ is the minimum eigenvalue of T_q.

In [**82**], the condition $k \leq n/2$ of (56) is weakened to state $k \leq (n + j)/2$ where j is the largest integer p such that $\lambda_p(A) \neq 0$. This is a result of some deep and difficult mathematics. The reader is encouraged to see [**81**], [**82**],[**83**],[**84**],[**85**] for details.

CHAPTER 17

Steepest Descent for Conservation Equations

Many systems of conservation equations may be written in the form

$$u_t = \nabla \cdot F(u, \nabla(u)) \tag{17.1}$$

where for some positive integers n, k and $T > 0$ and some region

$$\Omega \subset R^n, u : [0, T] \times \Omega \to R^k.$$

Often, however, a more complicated form is encountered:

$$(Q(u))_t = \nabla \cdot F(u, \nabla(u)),\ S(u) = 0 \tag{17.2}$$

where here for some positive integer q,

$$u : [0, T] \times \Omega \to R^{k+q},$$

$$Q : R^{k+q} \to R^k, S : R^{k+q} \to R^q,$$

$$F : R^{k+q} \times R^{n(k+q)} \to R^{nk}.$$

The condition $S(u) = 0$ in (17.2) is a relationship between the components of the unknowns u. It is often called an equation of state. In $Q(u)$ unknowns may be multiplied as in momentum equations. Sometimes (17.2) can be changed into (17.1) by using the condition $S(u) = 0$ to eliminate q of the unknowns and by using some change of variables to convert the term $(Q(u))_t$ into the form u_t, but it often seems better to treat the system (17.2) numerically just as it is written.

We consider homogeneous boundary conditions on $\partial\Omega$ as well as initial conditions. We describe a strategy for a time-stepping procedure. Take w to be a time-slice of a solution at time t_0. We seek an estimate v at time $t_0 + \delta$ for some time step δ. We seek v as a minimum to ϕ where

$$\phi(v) = [||Q(v) - Q(w) - \delta F((w + v)/2, \nabla((v + w)/2))||^2 + ||S((v + w)/2)||^2$$

for all v $H^{2,2}(\Omega, R^n)$ satifying the homogeneous boundary conditions. If v is found so that $\phi(v) = 0$, then it may be a reasonable approximation of a solution at $t_0 + \delta$. Assuming that F and Q are such that ϕ is a $C^{(1)}$ function with locally lipschitzian derivative, we may apply theory in various preceeding chapters in order to arrive at a zero (or at least a minimum) of ϕ. In particular we may consider continuous steepest descent

$$z(0) = w, z'(t) = -(\nabla\phi)(z(t)),\ t \geq 0 \tag{17.3}$$

where we have started the descent at w, the time slice estimating the solution at t_0. If δ is not too large, then w should be a good place to start. For numerical computations we would be interested in tracking (17.3) numerically with a finite number of steps. This procedure was tested on Burger's Equation

$$u_t + uu_x = \nu \Delta u$$

with periodic boundary conditions. The above development is taken from [**66**]. A more serious application is reported in [**68**] fir a magnetohydrodynamical system consisting of eight equations (essentially a combination of Maxwell's equations and Navier-Stokes equations) together with an equation of state. These equations were taken from [**97**] and the computations was carried out in three space dimensions. The code has never been properly evaluated but it seems to work.

CHAPTER 18

A Sample Computer Code with Notes

Following is a simple FORTRAN code for a finite difference approximation to the problem of finding y on $[0, 1]$ so that $y' - y = 0$. It is a computer implementation of the Sobolev gradient given in Chapter 2.

```
      program sg
      dimension u(0:1000), g(0:1000), w(0:1000)
c See Note A
      f(t) = 1+t
      open(7,file='sg.dat')
      n = 1000
      er = .0001
      delta = 1./float(n)
      p = 1./delta + .5
      q = 1./delta - .5
      sq = p**2 + q**2
c See Note B
      do 1 i = 0,n
      u(i) = f(float(i)/float(n))
1     continue
c See Note C
      kkk = 0
100   continue
c See Note D
      g(0) = -p*((u(1)-u(0))/delta - (u(1)+u(0))/2.)
      do 5 i=1,n-1
      g(i) = q*((u(i) - u(i-1))/delta - (u(i) + u(i-1))/2.)
     .  -p*((u(i+1) - u(i))/delta - (u(i+1) + u(i))/2.)
5     continue
      g(n) = q*((u(n) - u(n-1))/delta - (u(n) + u(n-1))/2.)
      kk = 0
c See Note E
      call ls(g,w,n,p,q)
c See Note F
      x = 0.
```

```
      y = 0.
      do 30 i=1,n
      x = x+((w(i)-w(i-1))/delta - (w(i)+w(i-1))/2.)**2
      y = y+((w(i)-w(i-1))/delta-(w(i)+w(i-1))/2.)
     .   *((u(i)-u(i-1))/delta-(u(i)+u(i-1))/2.)
30    continue
      d = y/x
c See Note G
      x = 0.
      do 50 i=0,n
      u(i) = u(i) - d*w(i)
50    continue
      do 51 i=1,n
      x = ((x(i)-x(i-1))/delta)**2 + ((x(i)+x(i-1))/2.)**2
51    continue
      x = sqrt(x/float(n))
      kkk = kkk + 1
      if (x.gt.er) goto 100
      write(7,*) 'number of iterations = ',kkk
      write(7,99) (u(i), i=0,n)
      write(*,*) 'Solution at intervals of length .1:'
      write(*,99) (u(i), i=0,n,n/10)
99       format(10f7.4)
      stop
      end
c See Note H
      subroutine ls(g,w,n,p,q)
      dimension g(0:1000),w(0:1000),a(0:1000),b(0:1000),c(0:1000)
      do 1 i=0,n
      a(i) = -p*q
      b(i) = p**2 + q**2
      c(i) = -p*q
1     continue
      do 2 i=1,n
      b(i) = b(i) - c(i-1)*a(i)/b(i-1)
      g(i) = g(i) - g(i-1)*a(i)/b(i-1)
2     continue
      w(n) = g(n)/b(n)
      do 3 i=n-1,0,-1
      w(i) = (g(i) - c(i)*w(i+1))/b(i)
3     continue
      return
```

end

Note A. The function f specifies the initial estimate of a solution. The integer n is the number of pieces into which $[0, 1]$ is broken. When the Sobolev norm of the gradient is less than er the interation is terminated.

Note B. The vector u at first takes its values from the function f. It later carries a succession of approximations to a finite difference solution.

Note C. The main iteration loop starts at label 100.

Note D. Construction of the conventional gradient of

$$\phi : \phi(u_0, u_1, \ldots, u_n)$$
$$= (1/2)\sum_{n=1}^{n}((u(i) - u(i-1))/delta - (u(i) + u(i-1))/2)^2, u \in R^{n+1}.$$

Note E. Call of the linear solver with input g a conventional gradient and output w a Sobolev gradient. Here Gaussian elimination is used. For problems in higher dimensions one might use Gauss-Seidel, Jacobi or Succesive Overrelaxation (SOR). In practice for large problems, a conjugate gradient scheme may be superimposed on a discrete steepest descent procedure (H). This was done in (Renka-Neuberger GL,SD). It further enhances numerical performance by cutting down on classic alternating direction hunting common in discrete steepest descent. In the case of Sobolev gradients, the iteration normally converges quite fast but a conjugate gradient procedure might cut the number of interations in half. It give a significant but not compelling saving of time.

Note F. Calculation of optimal step size as the number d which minimizes $\phi(u - d * w)$.

Note G. Updating of the estimate u. Calculation of norm of Sobolev gradient of w. Calculation stops when this norm is smaller than er.

Results from running the above code are found in file 'sg.dat'. In addition one should get printed to the screen the following:

Solution at intervals of length .1:

.7756 .8562 .9452 1.0432 1.1514 1.2708 1.4025 1.5481 1.7088 1.8863 2.0824

Bibliography

[1] R. A. Adams, *Sobolev Spaces*, Academic Press, 1978.

[2] , J. Baillon, R. Bruck, S. Reich, *On the Aysmptotic Behavior of Nonexpansive Mappings and Semigroups in Banach Spaces*, Houston J. Math. 4 (1978), 1-9.

[3] W.C. Bakker, *A Method of Iteration for Solving Functional Differential Equations with Advanced and Retarded Arguments* Dissertation, Emory University (1978).

[4] C. Beasley, *Finite Element Solution to Nonlinear Partial Differential Equations*, Dissertation, North Texas State University (1981).

[5] M.L. Berger, *Nonlinearity and Functional Analysis*, Academic Press, 1977.

[6] F. Béthuel, H.Brezis, B.D. Coleman, F. Hélein, *Bifurcation Analysis of Minimizing Harmonic Maps Describing the Equilibrium of Nematic Phases*, Arch. Rat. Mech. 118 (1992), 149-168.

[7] F. Béthuel, H. Brezis, F. Hélein, *Ginzburg-Landau Vorticies*, Birkhauser, 1994.

[8] A. Beurling and J. Deny, *Dirichlet Spaces*, Proceedings Nat. Acad. Sci. 45 (1959), 208-215.

[9] A. Beurling, *Collected Works, Vol. 2*, Birkhauser, 1989.

[10] H. Brezis, *Opérateurs maximaux monotones et semi-groupes nonlinéares*, Math Studies 5, North Holland, 1973.

[11] *The surface evolver*, K. Brakke, Experimental Math. 1 (1992), 141-165.

[12] D.R. Brown and M.J. O'Malley, *A Fixed-point Theorem for certain Operator Valued Maps*, Proc. Amer. Math. Soc. 74 (1979), 254-258.

[13] R. E. Bruck and S. Reich, *A General Convergence Principle in Nonlinear Functional Analysis*, Nonlinear Analysis 4 (1980), 939-950.

[14] G.F. Carey, *Variational Principles for the Transonic Airfoil Problem*, Computer Methods in Applied Mathematics and Engineering 13 (1978), 129-140.

[15] A. Castro, J. Cossio and John M. Neuberger, *Sign-Changing Solutions for a Superlinear Dirichlet Problem*, Rocky Mountain J. Math. (to appear).

[16] A. Castro and J.W. Neuberger *An Implicit Function Theorem* (preprint).

[17] P.L. Cauchy *Méthode générale pour la reésolution des systemes d'équations simultanées*, C.R. Acad. Sci. Paris, 25 (1847), 536-538.

[18] P. Ciarlet, *Introduction to Numerical Linear Algebra and Optimisation*, Cambridge University Press, Cambridge, 1989.

[19] R. Courant, *Dirichlet's Principle, Conformal Mapping, and Minimal Surfaces*, Interscience, 1950.

[20] H.B. Curry, *The Method of Steepest Descent for Nonlinear Minimization Problems*, Quart. Appl Math 2 (1944), 258-261.

[21] J.G. Dix and T.W. McCabe, *On Finding Equilibria for Isotropic Hyperelastic Materials*, Nonlinear Analysis 15 (1990), 437-444.

[22] J. R. Dorroh and J. W. Neuberger, *A Theory of Strongly Continuous Semigroups in Terms of Lie Generators*, J. Functional Analysis 136 (1996), 114-126.

[23] R. Driver, *Ordinary and Delay Differential Equations*, Springer-Verlag, 1977.

[24] J.M. Dye and S. Reich, *Unrestricted Iteration of Nonexpansive Mappings in Hilbert Space*, Trans. Am. Math. Soc. 325 (1991), 87-99.

[25] L.C. Evans, *Weak Convergence Methods for Nonlinear Partial Differential Equations*, CBMS Series in Mathematics 74, 1990.

[26] Qiang Du, M. Gunzburger, and Janet S. Peterson, *Analysis and Approximation of the Ginzburg- Landau Model of Superconductivity*, SIAM Review 34 (1992), 54-81.

[27] Qiang Du, M. Gunzburger, and Janet S. Peterson, *Computational simulation of type-II superconductivity including pinning phenomena*, Physical Review B 51, 22 (June 1995), 16194-16203.

[28] N. Dunford and J. Swartz, *Linear Operators, Part II*, Wiley-Interscience, NY, 1958.

[29] R. Fletcher and C.M. Reeves, *Function Minimisation by Conjugate Gradients*, Computer J. 7 (1964), 149-154.

[30] R. Fletcher, *Unconstrained Optimization*, John Wiley, 1980.

[31] T. M. Flett, *Differential Analysis*, Cambridge University Press, Cambridge, 1980.

[32] G.E. Forsythe, M.A. Malcolm adn C.B. Moler, *Computer Methods for Mathematical Computations*, Prentice-Hall, 1977.

[33] J. Garza, *Using Steepest Descent to Find Energy Minimizing Maps Satisfying Nonlinear Constraints*, Dissertation, Univ. of North Texas (1994).

[34] R.C. James, *Weak Convergence and Reflexivity*, Trans. Amer. Math. Soc. 113 (1964), 129-140.

[35] J.K. Hale, *Functional Differential Equations*, Springer-Verlag, 1971.

[36] P.R. Halmos, *A Hilbert Space Problem Book*, Van Nostrand, Princeton, 1967.

[37] M. Hestenes, *Conjugate Direction Methods in Optimization*, Springer-Verlag, 1980.

[38] L. Hormander, *The Analysis of Linear Partial Differential Operators*, Springer-Verlag, 1990.

[39] F.J. Johnson, *Application of a Poisson Solver to the Transonic Full Potential Equations*, Boeing Note AERO-B8113-C81-004.

[40] N.A. Karmarkar, *A New Polynomial Time Algorithm for Linear Programming*, Combinatorica 4 (1984), 373-395.

[41] K. Kim *Steepest Descent for Partial Differential Equations*, Dissertation, Univ. of North Texas (1992)

[42] J.C. Lagarias, *The Nonlinear Geometry of Linear Programming III. Projective Legendre Transform Co-ordinates and Hilbert Geometry*, Trans. Amer. Math. Soc 320 (1990), 193-225.

[43] P. D. Lax, *Parabolic Equations*, Annls. Math. Stud. V, 167-190.

[44] M. Liaw, *The Steepest Descent Method Using Finite Elements for Systems of Nonlinear Partial Differential Equations*, Dissertation, North Texas State University, (1983).

[45] J.-L. Lions, *Equations Differentielles Operationnelles et Prlblemes aux Limites*, Springer-Verlag, 1961.

[46] W.T. Mahavier, *Steepest Descent in Weight Sobolev Spaces*, J. Abst. Appl. Analysis (to appear).

[47] L.C. May, *A Solution-Giving Transformation for Systems of Differential Equations*, Dissertation, Univ. of North Texas (1990).

[48] Jurgen Moser, *A Rapidly Convergent Iteration Method and Non-linear Partial Differential Equations*, Annali Scuola Norm. Sup. Pisa 20 (1966), 265-315.

[49] M.Z. Nashed, *Aspects of Generalized Inverses in Analysis and Regularization*, Generalized Inverses and Applications, Academic Press, 1976.

[50] John M. Neuberger, *A Numerical Method for Finding Sign-Changing Solutions of Superlinear Dirichlet Problems*, Nonlinear World (to appear).

[51] J.W. Neuberger, *Tensor Products and Successive Approximations for Partial Differential Equations*, Isr. J. Math., 6 (1968), 121-132.

[52] ______, *Norm of Symmetric Product Compared with Norm of Tensor Product*, J. Linear Multi-Linear Algebra, 2 (1974), 115-121.

[53] ______, *An Iterative Method for Solving Nonlinear Partial Differential Equations*, Advances in Math., 19 (1976), 245-265.

[54] ______, *A Resolvent for an Iteration Method for Nonlinear Partial Differential Equations*, Trans. Amer. Math. Soc., 226 (1977), 231-343.

[55] ______, *Projection Methods for Linear and Nonlinear Systems of Partial Differential Equations*, Proceedings 1976 Dundee Conference on DifferentialEquations, Springer-Verlag Lecture Notes 564 (1976), 341-349.

[56] ______, *Square Integrable Solutions to Linear Inhomogenous Systems*, J. Differential Equations, 27 (1978), 144-152.

[57] ______, *Boundary Value Problems for Systems of Nonlinear Partial Differential Equations*, Functional Analysis and Numerical Analysis, Springer-Verlag Lecture Notes 701 (1979), 196-208.

[58] ______, *Boundary Value Problems for Linear Systems*, Proc. Royal Soc., Edinburgh, 83A (1979), 297-302.

[59] ______, *Finite Dimensional Potential Theory Applied to Numerical Analysis*, Linear Alg. Appl., 35 (1981), 193-202.

[60] ______, *A Type-Independent Iterative Method for Systems of Non-Linear Partial Differential Equations: Application to the Problem of Transonic Flow*, Computers and Mathematics With Applications, 7 (1980), 67-78.

[61] ______, *A Solvability Condition for Systems of Functional Differential Equations*, Proceedings Conf. Functional and Integral Equations, Marcel Dekker, Inc., (1981).

[62] ______, *A Type-Independent Iterative Method Using Finite Elements*, Proceedings Third International Conf. on Finite Elements and Flow Theory, University of Calgary, 1980.

[63] ______, *A Type-Independent Numerical Method with Applications to Navier-Stokes Equations*, Proc. Ninth Conference on Numerical Simulation of Plasmas, Northwestern University, 1980.

[64] ______, *Steepest Descent for General Systems of Linear Differential Equations in Hilbert Space*, Springer Lecture Notes 1032 (1983), 390-406.

[65] ______, *Some Global Steepest Descent Results for Nonlinear Systems*, Trends Theory Applications Diff. Eq., Marcel Dekker (1984), 413-418.

[66] ______, *Use of Steepest Descent for Systems of Conservation Equations*, Research Notes in Mathematics, Pitman (1984).

[67] ______, *Steepest Descent for General Systems of Linear Differential Equations in Hilbert Space*, Springer-Verlag Lecture Notes in Mathematics 1032 (1982), 390-406.

[68] ______, *Steepest Descent and Time-Stepping for the Numerical Solution of Conservation Equations*, Proc. Conf. on Physical Math and Nonlinear Differential Equations, Marcel Dekker, 1984.

[69] ______, *Steepest Descent and Differential Equations*, J. Math. Soc. Japan 37 (1985), 187-195.

[70] ______, *Calculation of Sharp Shocks Using Sobolev Gradients*, Comtemporary Mathematics, American Mathematical Society 108 (1990), 111-118.

[71] ______, *Constructive Variational Methods for Differential Equations*, Nonlinear Analysis, Theory, Meth. & Appl. 13 (1988), 413-428.

[72] ______, *Sobolev Gradients and Boundary Conditions for Partial Differential Equations*, Contemporary Mathematics 204 (1997), 171-181.

[73] ______, *Continuous Steepest Descent in Higher Order Sobolev Spaces*, Proceedings of Second World Congress of Nonlinear Analysts (to appear).

[74] ______, *Laplacians and Sobolev Gradients*, Proc. Amer. Math. Soc. (to appear).

[75] J.W. Neuberger and R.R. Renka, *Numerical Calculation of Singularities for Ginzburg-Landau Functionals* (preprint)

[76] F. und R. Nevanlinna, Absolute Analysis, Springer-Verlag, 1959.

[77] J.C.C. Nitsche, *Non-uniqueness for Plateau's Problem. A Bifurcation Process*, Ann. Acad. Sci. Fenn. Ser. A I Math. 2 (1976), 361-373.

[78] F. Olver, *Applications of Lie Groups to Differential Equations, 2nd Ed.*, Springer-Verlag, 1993.

[79] J. Ortega and W. Rheinboldt, *Iterative Solutions of Nonlinear Equations in Several Variables*, Academic Press (1970).

[80] R.S. Palais and S. Smale, *A Generalized Morse Theory*, Bull. Amer. Math. Soc. 70 (1964), 165-172.

[81] T. Pate, *Lower Bounds for the Norm of the Symmetric Product*, Linear Alg. App. 14 (1976), 285-292.

[82] ______, *Some Improvements on Neuberger's Iteration Procedure for Solving Partial Differential Equations*, J. Diff. Eqns. 34 (1979), 261-272.

[83] ______, *An Application of the Neuberger Iteration Procedure to the Constant Coefficient Linear Partial Differential Equation*, J. Math. Anal. Appl. 72 (1979), 771-782.

[84] ______, *A Characterization of a Neuberger type Iteration Procedure that Leads to Solutions of Classical Boundary Value Problems*, Pac. J. Math. 104 (1983), 191-204.

[85] ______, *A Unifying Context for the Classical Cauchy Initial Value Problem and the Type of Problem Solvable Unisng an Iteration Procedure Suggested by Neuberger*, Houston J. Math. 8 (1982), 273-284.

[86] S. Reich, *A Nonlinear Hille-Yosida Theorem in Banach Spaces*, J. Math. Analy. Appl. 84 (1981), 1-5.

[87] R.J. Renka and J.W. Neuberger, *A Portable Software Package for Non-Linear Partial Differential Equations*, Oak Ridge National Laboratory, CSD-102 (1982).

[88] ______, *Minimal Surfaces and Sobolev Gradients*, Siam J. Sci. Comp 16 (1995), 1412-1427.

[89] ______, Sobolev Gradients and the Ginzburg-Landau Equations, Siam J. Sci. Comp (to appear).

[90] W.B. Richardson, Jr., *Nonlinear Boundary Conditions in Sobolev Spaces*, Dissertation, Univ. of North Texas (1984).

[91] F. Riesz and B. Sz.-Nagy, *Functional Analysis*, Ungar, 1955.

[92] W.B. Richardson, Jr., *Steepest Descent and the Least C for Sobolev's Inequality*, Bull. London Math. Soc. 18 (1986), 478-484.

[93] M.L. Robertson, *The Equation* $y'(t) = F(t, y(g(t)))$, Pacific J. Math. 43 (1972), 483-492.

[94] P. L. Rosenblum, *The Method of Steepest Descent*, Proc. Symp. Appl. Math. VI, Amer. Math. Soc., 1961.

[95] J. Rubinstein, Six Lectures on Superconductivity, CRM Summer School, 6-18 August 1995, Banff, Canada.

[96] M. Schecter, *Principles of Functional Analysis*, Academic Press, NY, 1971.

[97] D. Schnack and J. Killeen, *Two-Dimensional Magnetohydrodynamic Calculation*, J. Comp. Physics 35 (1980), 110-145.

[98] F. Treves, *Locally Convex Spaces and Linear Partial Differential Equations*, Springer-Verlag, 1967.

[99] M. M. Vainberg, *On the Convergence of the Method of Steepest Descent for Nonlinear Equations*, Sibirsk. Math. A. 2 (1961), 201-220.

[100] J. VonNeumann, *Functional Operators II*, Annls. Math. Stud. 22, 1940.

[101] J.B. Walsh, *Iterative Solution of Linear Boundary Value Problems*, Dissertation, Univ. of North Texas (1983).

[102] H. Weyl, *The method of orthogonal projections in potential theory*, Duke Math.J. 7 (1940), 411-444.

[103] ———, *The Theory of Groups and Quantum Mechanics*, Dover, 1950.

[104] M. Zahran, *Steepest Descent on a Uniformly Convex Space*, Dissertation, Univ. of North Texas (1995).

Index