A NUMERICAL METHOD UTILIZING WEIGHTED SOBOLEV DESCENT TO SOLVE SINGULAR DIFFERENTIAL EQUATIONS

W. T. MAHAVIER

Abbreviated title: Weighted Sobolev Descent for Singular ODE.

Keywords: Steepest descent, weighted Sobolev space, ODE, finite difference, boundary value problem, differential equation, numerical, pre-conditioning, least squares.

AMS subject classifications: 65K05, 65K10, 65L05, 65L10, 65L12

Abstract. A numerical method is developed for solving singular differential equations using steepest descent based on weighted Sobolev gradients. The method is demonstrated on a variety of first and second order problems, including linear constrained, unconstrained, and partially constrained first order problems, a nonlinear first order problem with irregular singularity, and two second order variational problems.

The method is an extension of steepest descent in Sobolev spaces which is a variation of descent based on the Euclidean gradient. The differential equation is cast as a least-squares problem yielding a functional representing the equation. A weighted Sobolev space for the problem is chosen where the weights are based on the functional. The gradients associated with the functional take into account both the weights and the boundary conditions for the given equation.

Results are presented which demonstrate the improvements obtained by computing based on weighted Sobolev gradients rather than computing based on either unweighted Sobolev gradients or on the Euclidean gradient.

1. Introduction. A numerical method for the study of differential equations which have linear singularities is presented. The method extends the work pioneered by Neuberger [N5], where steepest descent based on Sobolev gradients is introduced. Gradients arise from weighted Sobolev spaces such as those considered by Kufner [KA] and Elschner [E] where weights are determined by the singularity of the differential equation under consideration and by the functional representing the equation. Neuberger and his students have demonstrated the power of Sobolev descent for specific problems. See [K], [G], [DM] for examples of Sobolev descent on non-singular problems and see [N2] for an example of a nonlinear second order differential equation with nonlinear singularity arising from a problem concerning transonic flow. Existence and uniqueness arguments for singular problems in Sobolev spaces have been given by Schuchman in [S] and by Canic and Keyfitz in [CK]. For a paper concerning Sobolev gradients which are constructed based on the problem at hand, consider the paper [RN]. See [A] for a general reference on Sobolev spaces.

Three types of descent are addressed: L^2 descent, Sobolev descent, and weighted Sobolev descent. For the problems studied, weighted Sobolev descent outperforms Sobolev descent, which in turn outperforms L^2 descent. The motivational problem,

$$\begin{aligned} &2ty'(t) = y(t) \\ &y(1) = 1, \end{aligned} \tag{1}$$

is introduced in § 2. In § 3 the method is exhibited for solving the general class of singular first order ordinary differential equations,

$$\begin{aligned} &q(t)y'(t) = f(t, y(t)) \\ &k_1 y(a) + k_2 y(b) = k_3 \end{aligned} \tag{2}$$

where $a, b \in I$, $f \in C_I^2$, $q \in C_I^1$, and $q(t) = 0$ for some $t \in I = [0,1]$. In § 5 an alternative to traditional variational methods is offered for the second order boundary value problems,

$$(t^2y')' - u = 0 \\ u(0) = 0 \\ u(1) = 1 \tag{3}$$

and

$$((1-t^2)y')' + 2u = 0 \\ u(0) = 0 \\ u(1) = 1. \tag{4}$$

2. Theory and Motivational Example. Consider $2ty' = y$ on $I = [0,1]$ with final condition $y(1) = 1$ and cast the equation as a least squares minimization problem, setting

$$J(u) = \int_I (2ju' - u)^2$$

for every $u \in C_I^1$ where j denotes the identity on I. If there existed a C^1 solution, then a zero of J would indicate this solution. The fact that there is no C^1 solution motivates the development of the spaces which follow. Let $L = L_I^2$ and $\langle \cdot, \cdot \rangle_L$ denote the L inner product. Define for any $w \in C_I^1$ which is positive almost everywhere,

$$D_1^w = \overline{\left\{ \begin{pmatrix} u \\ wu' \end{pmatrix} : u \in C_I^1 \right\}}^{L \times L},$$

and let H_w be the space of all elements which arise as first components of elements of D_1^w. The inner product for H_w is given by,

$$\langle u, v \rangle_{H_w} = \langle u, v \rangle_L + \langle D_1^w(u), D_1^w(v) \rangle_L.$$

In the unweighted case ($w \equiv 1$), we denote D_1^w by D_1 and H_w by H. It is known, [N3] that D_1 is a function and H is the Hilbert space, $H_{[0,1]}^{1,2}$. An argument follows, showing that D_1^w is a function and H_w is a Hilbert space for appropriate w.

THEOREM 2.1. *If $w \in C_I^1$ is positive almost everywhere and w vanishes at some point, $\hat{t}$, then D_1^w is a function in the sense that no two elements of D_1^w have the same first coordinates and distinct second coordinates.*

Proof. Let

$$A = \left\{ \begin{pmatrix} u \\ wu' \end{pmatrix} : u \in C_I^1 \right\}$$

so that $\overline{A} = D_1^w$ and suppose $\begin{pmatrix} f \\ g \end{pmatrix}$, $\begin{pmatrix} f \\ h \end{pmatrix}$ are elements of $\overline{A}$. Let $\begin{pmatrix} a_n \\ wa_n' \end{pmatrix}$ be a sequence in A converging to $\begin{pmatrix} f \\ g \end{pmatrix}$ in $L^2 \times L^2$ and let $\begin{pmatrix} b_n \\ wb_n' \end{pmatrix}$ be a sequence in A converging to $\begin{pmatrix} f \\ h \end{pmatrix}$ in $L^2 \times L^2$. Three lemmas are proved, the last resulting in the proof of the theorem. □

LEMMA 2.2. *$(wa_n)_{\mathbf{N}}$ is uniformly Cauchy on I.*

Proof. If $a_n \to f$ in L^2 then $w'a_n \to w'f$ in L^1, thus $(w'a_n)_{\mathbf{N}}$ is L^1 Cauchy. Similarly, $(wa'_n)_{\mathbf{N}}$ is L^1 Cauchy. Let $\epsilon > 0$ and N_1 in $\mathbf{N}$ such that for all $n, m \geq N_1$, $\int_I |w'(a_n - a_m)| < \frac{\epsilon}{2}$. Let N_2 in $\mathbf{N}$ such that for all $n, m \geq N_2$, $\int_I |w(a_n - a_m)'| < \frac{\epsilon}{2}$. Let $N = max\{N_1, N_2\}$, $t \in I$, and $n, m \geq N$. Since w arises from the singularity of the differential equation, choose $\hat{t}$ such that $w(\hat{t}) = 0$. The inequality, $|w(t)(a_n(t) - a_m(t))| = |\int_{\hat{t}}^t (wa_n)' - (wa_m)'| \leq \int_{\hat{t}}^t |w(a'_n - a'_m)| + \int_{\hat{t}}^t |w'(a_n - a_m)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$ concludes the proof of the lemma. □

LEMMA 2.3. *$|\int_s^t (wa'_n - wb'_n)|$ converges to zero uniformly on I.*

Proof. Let $\epsilon > 0$. Since $(wa_n)_{\mathbf{N}}$ and $(wb_n)_{\mathbf{N}}$ are uniformly Cauchy, $wa_n \to wf$ uniformly and $wb_n \to wf$ uniformly. Hence, $(wa_n - wb_n) \to wf - wf = 0$ and there exists $N_1 \in \mathbf{N}$ such that for all $n \geq N_1$, $|w(x)a_n(x) - w(x)b_n(x)| < \frac{\epsilon}{3}$ for every $x \in I$. Let $N_2 \in \mathbf{N}$ such that for all $n \geq N_2$, $\|w'(a_n - b_n)\| < \frac{\epsilon}{3}$. Let $n \geq max\{N_1, N_2\}$ and $s, t \in I$. $|\int_s^t wa'_n - wb'_n| = |\int_s^t (wb'_n + w'b_n - w'b_n + w'a_n - w'a_n - wa'_n)| = |\int_s^t ((wb_n)' - (wa_n)') + \int_s^t (w'a_n - w'b_n)| \leq |\int_s^t ((wb_n)' - (wa_n)')| + \int_s^t |w'(a_n - b_n)| \leq |w(t)b_n(t) - w(t)a_n(t) - w(s)b_n(s) + w(s)a_n(s)| + \|w'(a_n - b_n)\| \leq |w(t)b_n(t) - w(t)a_n(t)| + |w(s)b_n(s) - w(s)a_n(s)| + \|w'(a_n - b_n)\| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon$. □

LEMMA 2.4. *$|\int_s^t (g - h)| = 0$ for every $s, t \in I$.*

Proof. Recall, $wa'_n \to g$ in L^2 implies $wa'_n \to g$ in L^1 and $wb'_n \to h$ in L^2 implies $wb'_n \to h$ in L^1. Pick $s, t \in I$ and we have $|\int_s^t g - \int_s^t wa'_n| \leq \int_s^t |g - wa'_n| \to 0$ and $|\int_s^t h - \int_s^t wb'_n| \leq \int_s^t |h - wb'_n| \to 0$. Therefore, $\lim_{n\to\infty} \int_s^t wa'_n = \int_s^t g$ and $\lim_{n\to\infty} \int_s^t wb'_n = \int_s^t h$. We conclude $|\int_s^t (g - h)| = |\lim_{n\to\infty} \int_s^t wa'_n - \lim_{n\to\infty} \int_s^t wb'_n| = |\lim_{n\to\infty} (\int_s^t (wa'_n - wb'_n))| = \lim_{n\to\infty} |\int_s^t (wa'_n - wb'_n)|$. Thus, $|\int_s^t (g - h)| = \lim_{n\to\infty} |\int_s^t (wa'_n - wb'_n)|$. Let $\epsilon > 0$ and from lemma 2 choose $N \in \mathbf{N}$ such that for all $n \geq N$, $|\int_s^t (wa'_n - wb'_n)| < \epsilon$ for every $s, t \in I$. Then, $\lim_{n\to\infty} |\int_s^t (wa'_n - wb'_n)| \leq \epsilon$ for every $s, t \in I$. Hence for every $\epsilon > 0$ and for every $s, t \in I$ we have $|\int_s^t (g - h)| \leq \epsilon$ thus $\int_s^t (g - h) = 0$. This implies that $g - h = 0$ almost everywhere, since if $g - h \neq 0$ almost everywhere then there exist $x \in \Re$ and $\epsilon > 0$ such that, without loss of generality, $g - f > 0$ on the interval $(x - \epsilon, x + \epsilon)$. Therefore, $\int_{x-\epsilon}^{x+\epsilon} (f - g) > 0$, a contradiction. Conclude, $|\int_s^t (g - h)| = 0$ in L^2 and $g = h$ almost everywhere. □

THEOREM 2.5. *D_1^w is a non-expansive, closed, bounded, densely defined, linear operator.*

Proof. D_1^w is closed by definition. Recall that D_1^w is densely defined iff the domain of D_1^w is dense in L. Since polynomials on I are dense in L, and H_w is a superset of the polynomials on I and a subset of L, D_1^w is densely defined. For any $u \in H_w$,

$$\frac{\|D_1^w u\|_L}{\|u\|_{H_w}} = \frac{\|D_1^w u\|_L}{\|u\|_L + \|D_1^w u\|_L} \leq 1,$$

thus D_1^w is bounded and non-expansive as an operator from H_w to L. □

For the motivational problem, the weight will be the identity, $j(t) = t$. That H is a proper subset of H_j may be shown by considering $u(t) = \sqrt{t} \in H_j \backslash H$.

THEOREM 2.6. *H_w is a Hilbert space.*

Proof. Certainly, $\langle u, v \rangle_{H_w} = \langle u, v \rangle_L + \langle D_1^w u, D_1^w v \rangle_L$ is an inner product, thus it suffices to show H_w is complete. If (u_n) is a Cauchy sequence in H_w then there exists $\begin{pmatrix} u_n \\ v_n \end{pmatrix} \in D_w$ such that $u_n = \pi_1 \begin{pmatrix} u_n \\ v_n \end{pmatrix}$. Since (u_n) is Cauchy in H_w, $\|u_n - u\|_L \to 0$

Table 1
Motivational Problem

2ty′ − y = 0			**y(1) = 2**		**N = 100**
Gradient	Iterations	Seconds	Residual	Avg. Abs. Err.	Max. Abs. Err.
L	21518	208	10^{-5}	10^{-2}	1.3×10^{-1}
H	830	10	10^{-5}	10^{-3}	7.1×10^{-2}
H_w	11	1	10^{-5}	10^{-5}	2.7×10^{-3}

2ty′ − y = 0			**y(1) = 2**		**N = 1,000**
Gradient	Iterations	Seconds	Residual	Avg. Abs. Err.	Max. Abs. Err.
H	10,000	1290	10^{-10}	10^{-4}	4.1×10^{-3}
H_w	26	3	10^{-10}	10^{-6}	8.7×10^{-4}

2ty′ − y = 0			**y(1) = 2**		**N = 10,000**
Gradient	Iterations	Seconds	Residual	Avg. Abs. Err.	Max. Abs. Err.
H_w	41	53	10^{-10}	10^{-8}	2.8×10^{-4}

and $\|v_n - v\|_L \to 0$ for some $\begin{pmatrix} u \\ v \end{pmatrix} \in L \times L$. Conclude that $\|u_n - u\|_{H_w} = \|u_n - u\|_L + \|v_n - v\|_L \to 0$ and thus $\|u_n - u\|_{H_w} \to 0$ in H_w. □

The following theorem, [N3], guarantees convergence in the continuous case for each problem in the paper. Convergence for the descrete case constitutes work in progress. Results for singular problems have been proved for specific problems and a generalization is forthcoming. Two theorems are available in the literature, [N5], which apply to the ordinary differential equation, y'=y, y(0)=1.

Theorem 2.7. *(Neuberger) Suppose $\mathcal{H}$ and $\mathcal{K}$ are Hilbert spaces and $G \in L(\mathcal{H}, \mathcal{K})$. Suppose $g \in \mathcal{K}$, $v \in \mathcal{H}$, $Gv = g$, and $\phi(u) = \frac{1}{2}\|Gu - g\|^2$ for every $u \in \mathcal{H}$. If $x \in \mathcal{H}$ and z is the function on $[0, \infty)$ so that*

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

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

D_1^w was shown to be bounded on H_w in Theorem 2. Put $v(t) = \sqrt{t}$, $g = 0$, $\mathcal{H} = H_w$, $\mathcal{K} = \Re$, $Gu = 2D_1^w(u) - u$ and

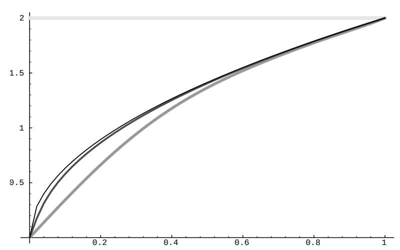
$$J(u) = \int_I (2D_1^j u - u)^2$$

For any $u \in H_w$, $J'(u)$ is a linear functional, hence there esists a uniqe element which we denote by $(\nabla_H J)$, satisfying, $J'(u)(h) = \langle (\nabla_{H_w} J)(u), h \rangle_{H_w}$ for all $u, h \in H_w$. Putting $\nabla = (\nabla_H J)$ satisfies the hypothesis of the theorem, guaranteeing convergence in the weighted spaces.

Note that we have just redefined J. Prior to the redefinition of J, the domain of J was C_I^1 which is unacceptable, as the solution to the problem is $y(t) = \sqrt{t}$ which is not C^1. Now the domain is J is H_j.

Having defined the spaces and continuous theory, results are presented for the motivational example and the numerical method is postponed for the general case in § 3. Table 1 illustrates that the number of iterations and the time required to solve the problem decrease while the obtained accuracy increases for L, H, and H_j descent

Fig. 1. *Motivational Problem*



respectively. As the number of divisions is increased, H_j descent performs well, yet the desired accuracy was unobtainable using L descent. For all cases considered, L and H descent do not converge for too tight a stopping criteria or for too large a number of divisions while H_w descent does. The numerical results indicate that the order of magnitude of the gradient for H_w descent is typically on the order of the stopping criteria, another quality of H_w descent which is not shared by L or H descent for singular problems. Figure 1 shows the difference between the two descent processes. The graph shows four lines shaded from light to dark and varying from thick to thin. Respectively they represent the initial estimate, the Sobolev approximation to the solution after 100 iterations, the weighted Sobolev approximation to the solution after three iterations, and the solution itself. The advantage of the weight near the singularity is evident from the graph. The results in Figure 1 are supported by the following reasoning. L descent does not take the derivative of the function into consideration, hence H descent outperforms L descent. While H descent considers the derivative, the solution does not belong to H. Since $\sqrt{t} \in H_j \backslash H$, H_j descent outperforms H descent. These difficulties in the continuous setting carry over to the numerical settings. Examples in later sections will demonstrate machine precision results for H_w descent which L and H descent are unable to obtain.

3. Numerical Method for First Order Problems. Let $k_1, k_2, k_3, a, b \in \Re$ with $a < b$ and $q \in C_I^1$. Let $f : [a,b] \times \Re \to \Re$ be differentiable with respect to the second variable. The problem is

$$(5) \qquad \begin{aligned} q(t)y'(t) &= f(t, y(t)) \\ k_1 y(a) + k_2 y(b) &= k_3. \end{aligned}$$

Descent based on each of the three gradients is considered for unconstrained, constrained, and partially constrained problems. Much of the notation extends and compliments the work of Neuberger in [N1] and [N5].

Denote the Euclidean norm by $\|\cdot\|_L$ and $x \in \Re^m$ by $x = (x_1, \ldots, x_m)$. Suppose n

is the number of divisions into which the interval $[a,b]$ is partitioned and $\delta = (b-a)/n$. Let ϵ be the stopping criteria; stop when $\|y^{new} - y\|_L < \epsilon$ where y and y^{new} denote successive approximations to the solution.

Define discretized versions of the identity and derivative operators, $D_0 : \Re^{n+1} \to \Re^n$, $D_1^w : \Re^{n+1} \to \Re^n$, and $D_w : \Re^{n+1} \to \Re^{2n}$.

$$D_0(x) = \begin{pmatrix} \frac{x_1+x_2}{2} \\ \vdots \\ \frac{x_n+x_{n+1}}{2} \end{pmatrix} \quad D_1^w(x) = \begin{pmatrix} \left(\frac{w_2+w_1}{2}\right)\left(\frac{x_2-x_1}{\delta}\right) \\ \vdots \\ \left(\frac{w_n+w_{n+1}}{2}\right)\left(\frac{x_{n+1}-x_n}{\delta}\right) \end{pmatrix} \quad D_w(x) = \begin{pmatrix} D_0(x) \\ D_1^w(x) \end{pmatrix}$$

The three discretized versions of the spaces L, H, and H_w are $\left(\Re^{n+1}, \langle\cdot,\cdot\rangle_L\right)$, $\left(\Re^{n+1}, \langle\cdot,\cdot\rangle_H\right)$, and $\left(\Re^{n+1}, \langle\cdot,\cdot\rangle_{H_w}\right)$ where

$$\langle u,v\rangle_{H_w} = \langle D_0(u), D_0(v)\rangle_L + \langle D_1^w(u), D_1^w(v)\rangle_L =$$

$$\sum_{k=1}^{n} \left(\frac{u_{k+1}+u_k}{2}\right)\left(\frac{v_{k+1}+v_k}{2}\right) + \left(\frac{w_{k+1}+w_k}{2}\right)^2 \left(\frac{u_{k+1}-u_k}{\delta}\right)\left(\frac{v_{k+1}-v_k}{\delta}\right)$$

for all $u, v \in \Re^{n+1}$.

D_w relates the Euclidean and Sobolev norms by $\|\cdot\|_{H_w} = \|D_w(\cdot)\|_L$. Let $y \in \Re^{n+1}$ and for all $k = 1, 2, \ldots, n+1$, let $t_k = a + (k-1)\delta$ and $f_k = f(t_k, y_k)$. Define $J : (\Re^{n+1}, \|\cdot\|_{H_w}) \to \Re$ by

$$\begin{aligned} J(y) &= \frac{1}{2}\|D_1^w y - D_0 f\|_L^2 \\ &= \frac{1}{2}\sum_{k=1}^{n}\left(\frac{q_{k+1}+q_k}{2}\,\frac{y_{k+1}-y_k}{\delta} - \frac{f_k+f_{k+1}}{2}\right)^2. \end{aligned}$$

For $x \in \Re^{n+1}$, $J'(x)$ is a bounded linear functional. Let $(\nabla_{H_w} J)(x)$ denote the unique element in $\Re^{n+1}$ satisfying $J'(x)(y) = \langle (\nabla_{H_w} J)(x), y\rangle_{H_w}$ for all $y \in \Re^{n+1}$.

One well known theorem is proved in the case of interest and determines the the matrix, A_w. The reader is referred to [RSN] to verify the non-singular nature of the matrix.

Theorem 3.1. *If $\langle\cdot,\cdot\rangle_{H_w}$ denotes the discretized Sobolev inner product on $\Re^{n+1}$ and $\langle\cdot,\cdot\rangle$ represents the standard inner product on $\Re^{n+1}$ then there exists a matrix A_w in $L(\Re^{n+1}, \Re^{n+1})$ such that $\langle x,y\rangle_{H_w} = \langle A_w x, y\rangle = \langle x, A_w y\rangle$ for every $x, y \in \Re^{n+1}$. Moreover, $A_w (\nabla_{H_w} J)(x) = (\nabla_L J)(x)$ for every $x \in \Re^{n+1}$.*

Proof. Since for every $x, y \in \Re^{n+1}$,

$$\begin{aligned} \langle x,y\rangle_{H_w} &= \langle D_0 x, D_0 y\rangle + \langle D_1^w x, D_1^w y\rangle \\ &= \left\langle D_0^t D_0 x, y\right\rangle + \left\langle (D_1^w)^t D_1^w x, y\right\rangle \\ &= \left\langle (D_0^t D_0 + (D_1^w)^t D_1^w) x, y\right\rangle \\ &= \left\langle D_w^t D_w x, y\right\rangle \end{aligned}$$

we have, $A_w = D_w^t D_w$. Also, for every $x, y \in \Re^{n+1}$,

$$\begin{aligned} \langle (\nabla_L J)(x), y\rangle &= J'(x)(y) \\ &= \langle (\nabla_{H_w} J)(x), y\rangle_{H_w} \\ &= \langle A_w (\nabla_{H_w} J)(x), y\rangle. \end{aligned}$$

Consequently, $A_w \left(\nabla_{H_w} J\right)(x) = \left(\nabla_L J\right)(x)$ for every $x \in \Re^{n+1}$. □

The boundary conditions are $k_1 y(a) + k_2 y(b) = k_3$ and the canonical perturbation space is $\Re_0^{n+1} = \{x \in \Re^{n+1} : k_1 x_1 + k_2 x_{n+1} = 0\}$. Let π_L denote the orthogonal projection of $\Re^{n+1}$ onto $\Re_0^{n+1}$ under the Euclidean inner product and π_{H_w} denote the orthogonal projection of $\Re^{n+1}$ onto $\Re_0^{n+1}$ under the Sobolev inner product. For $x \in \Re^{n+1}$,$J'(x)\big|_{\Re_0^{n+1}}$ is a bounded linear functional. Let $\left(\nabla_{H_w^0} J\right)(x)$ denote the unique element in $\Re_0^{n+1}$ satisfying $J'(x)(y) = \left\langle \left(\nabla_{H_w^0} J\right)(x), y\right\rangle$ for all $y \in \Re_0^{n+1}$. For all $x \in \Re^{n+1}, y \in \Re_0^{n+1}$ this yields,

$$\begin{aligned}
\left\langle \left(\nabla_{H_w^0} J\right)(x), y\right\rangle_{H_w} &= J'(x)(y) \\
&= \left\langle \left(\nabla_{H_w} J\right)(x), y\right\rangle_{H_w} \\
&= \left\langle \pi_{H_w} \left(\nabla_{H_w} J\right)(x), y\right\rangle_{H_w}
\end{aligned}$$

and thus $\left(\nabla_{H_w^0} J\right)(x) = \pi_{H_w} \left(\nabla_H J\right)(x)$ for all $x \in \Re^{n+1}$. Applying the Reisz representation theorem twice and using the self-adjoint property of projections repeatedly we have for every $x \in \Re^{n+1}, y \in \Re_0^{n+1}$

$$\begin{aligned}
\langle \pi_L((\nabla_L J)\, J)(x), y\rangle &= \langle((\nabla_L J)\, J)(x), y\rangle \\
&= J'(x)(y) \\
&= \left\langle \left(\nabla_{H_w^0} J\right)(x), y\right\rangle_{H_w} \\
&= \left\langle A_w \left(\nabla_{H_w^0} J\right)(x), y\right\rangle \\
&= \left\langle \pi_L A_w \left(\nabla_{H_w^0} J\right)(x), y\right\rangle .
\end{aligned}$$

This defines the linear system, $\pi_L A_w \left(\nabla_{H_w^0} J\right)(x) = \pi_L \left(\nabla_L J\right)(x)$, while allowing us to solve for $\left(\nabla_{H_w^0} J\right)$ without computing the projection, π_{H_w}. One observation is in order; the system must be solved over the subspace, $\Re_0^{n+1}$ in orderThe projection π_L must still be determined. Compute π_L by defining $\psi(x) = \|x - u\|_L^{\,2}/2$ and minimizing ψ over $\Re_0^{n+1}$ via Lagrange Multipliers to obtain

$$\pi_L(x) = \left(\frac{k_2(k_2 x_1 - k_1 x_{n+1})}{k_1^2 + k_2^2}, x_2, x_3, \ldots, x_n, \frac{-k_1(k_2 x_1 - k_1 x_{n+1})}{k_1^2 + k_2^2}\right)$$

The boundary conditions are handled as four separate cases. If $k_1 = k_2 = 0$ no boundary conditions are given. If both k_1 and k_2 are non-zero then the last row is replaced by the boundary data, $(k_1, 0, ..., 0, k_2)$ and the last entry of the gradient vector, $\left(\nabla_L J\right)(y)$, is set to zero. Initial and final value problems are handled similarly.

All the codes in the paper use optimal step size which is given by the real number h that minimizes $\alpha(h) = J\big(y - h\left(\nabla_{H_w} J\right)(y)\big)$. If f is linear, h is given by

$$h = \frac{\|\left(\nabla_{H_w} J\right)(y)\|_{H_w}^2}{\left\langle \left(\nabla_{H_w} J\right)^2(y), \left(\nabla_{H_w} J\right)(y)\right\rangle_{H_w}},$$

else, h is computed by applying a linear search to the function, α.

Algorithm

1. Compute the matrix, A_w, and the projection, π_L.
2. Choose $y \in \Re^{n+1}$ satisfying the boundary conditions.
3. Compute the gradient of J at y, $\left(\nabla_L J\right)(y)$.

TABLE 2
Unconstrained Singular Problem

$\mathbf{ty' - y = 0}$	$y_0(t) = t^2$	**No Boundary Conditions**			$N = 100$
Gradient	Iterations	Seconds	Residual	Avg. Abs. Err.	Max. Abs. Err.
L	5419	50	10^{-5}	10^{-2}	4.1×10^{-2}
H	2161	28	10^{-5}	10^{-3}	5.3×10^{-3}
H_w	8	1	10^{-5}	10^{-6}	9.8×10^{-6}

$\mathbf{ty' - y = 0}$	$y_0(t) = t^2$	**No Boundary Conditions**			$N = 10,000$
Gradient	Iterations	Seconds	Residual	Avg. Abs. Err.	Max. Abs. Err.
H_w	36	45	10^{-15}	10^{-10}	8.7×10^{-10}

4. Apply π_L to the matrix, A_w, and the gradient, $(\nabla_L J)(y)$.
5. Make A_w nonsingular by replacing the necessary rows.
6. Solve $\pi_L A_w x = \pi_L (\nabla_L J)(y)$ for $x = \pi_{H_w} (\nabla_{H_w} J)(y)$.
7. Determine minimal $h > 0$ which minimizes $J\big(y - h\pi_{H_w} (\nabla_{H_w} J)(y)\big)$.
8. Let $y^{new} = y - h\pi_{H_w} (\nabla_{H_w} J)(y)$.
9. If $\|y^{new} - y\|_L < \epsilon$ stop; else, put $y = y^{new}$ and repeat steps 3 through 8.

4. Results for First Order Problems. Results were presented in § 2 for a constrained problem. Results follow for an unconstrained problem, a partially constrained problem, and a nonlinear problem with irregular singularity.

Consider $ty' = y$ on I with no boundary conditions. The numerical results are in Table 2. An initial condition $y(0) = 0$ is forced by the singularity, and the one parameter family of solutions is given by $z(t) = kt$. The functional chosen is $J(u) = \int_I (D_j^1 u - u)^2$ for every $u \in H_j$.

In the case where boundary conditions are not sufficient to guarantee uniqueness, the solution to which the algorithm will converge may be predicted and depends on both the chosen gradient and the given initial estimate.

THEOREM 4.1. *If y_0 is the initial estimate, steepest descent will converge to $z(t) = kt$ where $k_L = 3\int_I jy_0$, $k_H = \frac{1}{3}\int_I (jy_0 + y_0')$, and $k_{H_j} = \frac{3}{2}\int_I j(y_0 + jy_0')$ for L descent, H descent, and H_j descent respectively.*

Proof. Only the statement associated with weighted descent is proved. Suppose J is as stated above and $\alpha(z) = \|y_0 - z\|^2_{H_w}$. Observe that $\alpha(z) = \|y_0 - z\|^2_{H_w} = \|y_0\|^2_{H_w} + \|z\|^2_{H_w} - 2\langle z, y_0\rangle_{H_w}$. Minimizing α over $S = \{z : z(t) = kt\}$ yields the closest element in $H_w \cap S$. This is a quadratic equation yielding k_{H_j} as stated. □

Choosing the initial function $y_0(t) = t^2$, the resulting solutions are $z_L(t) = \frac{3}{4}t$, $z_H(t) = \frac{15}{16}t$, and $z_{H_w}(t) = \frac{9}{8}t$. The number of divisions is small so that Sobolev descent results may be compared with the L descent results. L descent is outperformed by Sobolev descent, thus L and H results are then omitted so that the number of divisions and accuracy desired may be increased.

Consider the partially constrained problem, $(t - \frac{1}{2})y' = y$ with $y(0) = -\frac{1}{2}$. Solutions are given by,

$$z(t) = \begin{cases} c_1(t - \frac{1}{2}) & \text{if } x \in [0, \frac{1}{2}] \\ c_2(t - \frac{1}{2}) & \text{if } x \in [\frac{1}{2}, 1] \end{cases}$$

Having specified only an initial condition, the value for c_2 is not unique. As in the last example, the solution may be determined. If y_0 is the initial guess, solution, z, will be

Table 3
Partially Constrained Singular Problem

$(\mathbf{t}-\frac{1}{2})\mathbf{y}'-\mathbf{y}=\mathbf{0}$		$\mathbf{y(0)}=-\frac{1}{2}$		$\mathbf{N}=\mathbf{1000}$	
Gradient	Iterations	Seconds	Residual	Avg. Abs. Err.	Max. Abs. Err.
H	526	66	10^{-5}	10^{-2}	1.5×10^{-1}
H_w	7	1	10^{-5}	10^{-7}	4.3×10^{-6}

$(\mathbf{t}-\frac{1}{2})\mathbf{y}'-\mathbf{y}=\mathbf{0}$		$\mathbf{y(0)}=-\frac{1}{2}$		$\mathbf{N}=\mathbf{10,000}$	
Gradient	Iterations	Seconds	Residual	Avg. Abs. Err.	Max. Abs. Err.
H	764	497	10^{-6}	10^{-3}	3.4×10^{-2}
H_w	13	17	10^{-10}	10^{-10}	1.8×10^{-9}

the function which minimizes $\|y_0-z\|$ in whichever norm is chosen. L descent is again outperformed by H descent and H_w descent so these results are omitted here allowing an increase in the number of divisions and an increase in the desired accuracy. H descent yields the solution above with $c_1=1$ and $c_2=\frac{3}{2}$ while H_w descent yields $c_1=1$ and $c_2=\frac{15}{4}$.

H_w descent outperforms H descent by a factor of 66 in time and by 10^5 in accuracy. After increasing the number of divisions, the time factor remains 66, however, the accuracy is improved to 10^7. This trend persists in all examples considered: *as the number of divisions is increased, the differential between the obtainable accuracy between H_w and H descent increases.*

Observe in Table 3 that a less strict stopping criteria is used for H descent than for H_w descent. This is the 'best' result obtainable for the H descent. Superior results to the ones listed were unobtainable since the order of magnitude of $\nabla_H J$ is 10^{-16} or machine precision.

Consider the nonlinear problem with irregular singularity,

$$\begin{aligned} t^2y' &= 2ty+y^2 \\ y(1) &= 1 \end{aligned} \tag{6}$$

which has the solution, $y(t)=t^2/(2-t)$. Results are given in Table 4 which shows the marked improvements obtained by considering the weighted spaces.

We conclude this section by observing that similar results are obtained for problems where series solutions are not obtainable such as $(t-\frac{1}{4})(t-\frac{3}{4})y'=y$ with an initial condition at any one of the interior points $t=0$, $t=\frac{1}{4}$, $t=\frac{3}{4}$, or $t=1$. Hence the algorithm applies where algorithms based on expansion arguments do not.

5. Second Order Problems. Two approaches to this problem were implemented. The method used was to apply steepest descent directly to J. The alternative approach [N1], [N2] was to form the functional $\phi(u)=\frac{1}{2}\|(\nabla_{H_w}J)(u)\|^2$ whose zeroes are critical points of J. Both methods were successful, but the latter requires solving two systems of equations per iteration. Since neither had superior accuracy results and the alternative approach was computationally inferior, only the former approach is presented. For problems where the first method tends to 'fall off' the critical points, the latter method is appropriate and, surprisingly, requires minimal alteration (about 3 lines) of the code.

The first problem considered is to solve $Ku=0$ where K is defined by $Ku=(t^2u')'-u$. Using the method of series solutions to seek $u\in C_I^0\cap C_{(0,1]}^2$ such that $u(0)=$

Table 4
Nonlinear Problem with Irregular Singularity

$\mathbf{t^2y' - 2ty + y^2 = 0}$	$y_0(t) = t$		$\mathbf{y(1) = 1}$		$N = 100$
Gradient	Iterations	Seconds	Residual	Avg. Abs. Err.	Max. Abs. Err.
L	4799	56	10^{-5}	10^{-2}	1.4×10^{-1}
H	402	4	10^{-5}	10^{-3}	1.0×10^{-1}
H_w	23	1	10^{-5}	10^{-4}	1.7×10^{-2}

$\mathbf{t^2y' - 2ty + y^2 = 0}$	$y_0(t) = t$		$\mathbf{y(1) = 1}$		$N = 1,000$
Gradient	Iterations	Seconds	Residual	Avg. Abs. Err.	Max. Abs. Err.
H	5000	780	10^{-8}	10^{-3}	5.0×10^{-2}
H_w	353	56	10^{-8}	10^{-5}	4.3×10^{-3}

$\mathbf{t^2y' - 2ty + y^2 = 0}$	$y_0(t) = t$		$\mathbf{y(1) = 1}$		$N = 10,000$
Gradient	Iterations	Seconds	Residual	Avg. Abs. Err.	Max. Abs. Err.
H_w	331	409	10^{-5}	10^{-5}	5.4×10^{-3}

0 and $u(1) = 1$ yields $u(t) = c_1 t^{\frac{-1+\sqrt{5}}{2}} + c_2 t^{\frac{-1-\sqrt{5}}{2}}$ where only the first summand satisfies the equation, boundary conditions, and space limitations. The solution $t^{\frac{-1+\sqrt{5}}{2}}$ is in $H_j \backslash H$. Descent is based on subspaces of L, H, and H_j and the three subspaces based on the boundary conditions are $L^0 := \{h \in L : h(0) = 0 = h(1)\}$, $H^0 = H \cap L^0$, and $H_j^0 = H_j \cap L^0$. All three functionals agree on the space $C := C_I^0 \cap C^2_{(0,1]}$, and we abuse the notation labeling them all J and setting

$$J(u) = \frac{1}{2} \int_I j^2 (u')^2 + u^2.$$

Ignoring boundary conditions for the moment, the motivation may be summarized in one sentence. *If $j(t) = t$ and $u \in H_j$ then $J'(u)(h) = \int_I j^2 u'h' + uh = \langle u, h \rangle_{H_j}$, and we naturally seek a critical point of J in the space $\left(H_j, \langle \cdot, \cdot \rangle_{H_j}\right)$.* In practice, the gradient takes into consideration both the weight and the boundary conditions as outlined in § 3. Let $u \in C \subset L^0$ and $J'(u)$ is a bounded linear operator thus, there exists a unique element $(\nabla_{L^0} J)$ satisfying

$$\begin{aligned} \langle (\nabla_{L^0} J)(u), h \rangle_{L^0} &= J'(u)(h) \\ &= \int_I j^2 u'h' + uh \\ &= \int_I ((-j^2 u')' + u)h \\ &= -\int_I hKu \\ &= \langle h, -P_L K u \rangle_{L^0}, \end{aligned}$$

for every $h \in L^0$, where $P_L : L \to L^0$ is the orthogonal projection. The parallel in the Hilbert space H^0 is given by

$$J(u) = \frac{1}{2} \int_I j^2 (D_1 u)^2 + u^2. \tag{7}$$

TABLE 5
Variational Problem

$\mathbf{t^2y'' + 2ty' - y = 0}$		$\mathbf{y(0) = 0}$		$\mathbf{y(1) = 1}$	$\mathbf{N = 100}$
Gradient	Iterations	Seconds	Residual	Avg. Abs. Err.	Max. Abs. Err.
L	$10,000$	41	10^{-6}	10^{-2}	1.4×10^{-1}
H	538	2	10^{-6}	10^{-4}	1.8×10^{-2}
H_w	1	1	10^{-6}	10^{-5}	1.9×10^{-3}

$\mathbf{t^2y'' + 2ty' - y = 0}$		$\mathbf{y(0) = 0}$		$\mathbf{y(1) = 1}$	$\mathbf{N = 1,000}$
Gradient	Iterations	Seconds	Residual	Avg. Abs. Err.	Max. Abs. Err.
H	151	4000	10^{-8}	10^{-4}	1.3×10^{-2}
H_w	1	2	10^{-8}	10^{-6}	4.8×10^{-4}

$\mathbf{t^2y'' + 2ty' - y = 0}$		$\mathbf{y(0) = 0}$		$\mathbf{y(1) = 1}$	$\mathbf{N = 100,000}$
Gradient	Iterations	Seconds	Residual	Avg. Abs. Err.	Max. Abs. Err.
H_w	14	3	10^{-16}	10^{-9}	2.8×10^{-6}

For $u \in C$ and $h \in H^0$,

$$\begin{aligned}
\langle (\nabla_{H^0} J)(u), h \rangle_{H^0} &= J'(u)(h) \\
&= \int_I j^2 D_1 u D_1 h + uh \\
&= \left\langle \begin{pmatrix} h \\ D_1 h \end{pmatrix}, \begin{pmatrix} u \\ j^2 D_1 u \end{pmatrix} \right\rangle_{L \times L} \\
&= \left\langle \begin{pmatrix} h \\ D_1 h \end{pmatrix}, P_H \begin{pmatrix} u \\ j^2 D_1 u \end{pmatrix} \right\rangle_{L \times L} \\
&= \left\langle h, \pi_1 P_H \begin{pmatrix} u \\ j^2 D_1 u \end{pmatrix} \right\rangle_{H^0}
\end{aligned}$$

where

$$P_H : L \times L \to \left\{ \begin{pmatrix} u \\ D_1 u \end{pmatrix} : u \in H^0 \right\}$$

is the orthogonal projection and $\pi_1 : \Re \times \Re \to \Re$ such that $\pi_1 \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \alpha$. As in § 3, the chosen weight is the square root of the function in the integrand of the functional resulting from the singularity in the differential equation. In this case the singularity is t^2 which appears again in the functional.

The parallel in the Hilbert space H_j^0 is given by

$$J(u) = \frac{1}{2} \int_I \left(D_1^j u\right)^2 + u^2. \tag{8}$$

For $u \in C$ and $h \in L^0$,

$$\left\langle \left(\nabla_{H_j^0} J\right)(u), h \right\rangle_{H_j} = J'(u)(h)$$

$$
\begin{aligned}
&= \int_I D_1^j u D_1^j h + uh \\
&= \left\langle \begin{pmatrix} h \\ D_1^j h \end{pmatrix}, \begin{pmatrix} u \\ D_1^j u \end{pmatrix} \right\rangle_{L\times L} \\
&= \left\langle \begin{pmatrix} h \\ D_1^j h \end{pmatrix}, P_{H_j} \begin{pmatrix} u \\ D_1^j u \end{pmatrix} \right\rangle_{L\times L} \\
&= \left\langle h, \pi_1 P_{H_j} \begin{pmatrix} u \\ D_1^j u \end{pmatrix} \right\rangle_{H_j^0}
\end{aligned}
$$

where

$$
P_{H_j} : L \times L \to \left\{ \begin{pmatrix} u \\ D_1^j u \end{pmatrix} : u \in H_j^0 \right\}
$$

is the orthogonal projection. This exposition is summarized in the following theorem.

THEOREM 5.1. *For all $u \in C$, the gradients with respect to the Hilbert spaces L^0, H^0, and H_j^0 are* $(\nabla_{L^0} J)(u) = -P_L K u$, $(\nabla_{H^0} J)(u) = \pi_1 P_H \begin{pmatrix} u \\ j^2 D_1 u \end{pmatrix}$, *and* $\left(\nabla_{H_w^0} J\right)(u) = \pi_1 P_{H_j} \begin{pmatrix} u \\ D_1^j u \end{pmatrix}$. The question is: Which of the equations $(\nabla_L J)(u) = 0$ (Euler's equation), $(\nabla_H J)(u) = 0$, or $(\nabla_{H_w} J)(u) = 0$ is the appropriate equation to consider for computing on variational problems concerning singular differential equations? The results in Table 5 and 6 indicate that the latter is the superior choice. Discretizing the functional yields,

$$
J(u) = \frac{1}{2} \sum_{k=1}^{n} \left(\frac{t_{k+1} + t_k}{2} \right)^2 \left(\frac{u_{k+1} - u_k}{\delta} \right)^2 + \left(\frac{u_{k+1} + u_k}{2} \right)^2 .
$$

Since $\pi_L(x) = (0, x_1, \ldots, x_n, 0)$ and $(\nabla_L J)(u) = -Ku$, we have

$$
(\nabla_{L^0} J)(u) = \left(0, \ldots, -t_k^2 \frac{u_{k-1} - 2u_k + u_{k+1}}{\delta^2} - 2t_k \frac{u_{k+1} - u_{k-1}}{\delta} + u_k, \ldots, 0\right).
$$

The algorithm from § 3 may now be implemented.

Figure 2 exhibits the difference between the weighted and non-weighted descent processes. The graph shows four curves, shaded from light to dark and varying from thick to thin. Respectively they represent the initial estimate, the Sobolev approximation to the solution after three iterations, the weighted Sobolev approximation to the solution after three iterations, and the solution itself. The advantage of the weight near the singularity is clear from the graph. The solution and the weighted Sobolev approximation to the solution are already indistinguishable by three iterations. Table 5 represents the numerical results obtained using each of the above methods. Observe the decrease in both time and iterations required and the increase in both average absolute accuracy and maximum absolute accuracy.

The improved results were expected and a defense of the reasoning follows. Necessary conditions are given in [CH] in order that satisfying Euler's equation be a necessary condition for existence of an extremal point; however, this problem does not satisfy these conditions. The difficulty in the continuous case translates over to the poor numerical performance in solving $(\nabla_L J) = 0$. Similarly, seeking the solution, $t^{\frac{-1+\sqrt{5}}{2}}$ which does *not* belong to the space H, makes solving $(\nabla_H J) = 0$ an unpromising task. This leaves the equation $(\nabla_{H_w} J) = 0$ which indeed performs the best.

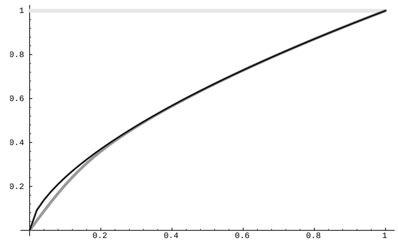


FIG. 2. *Variational Problem*

The second problem is to solve, $Ku = 0$ where $Ku = ((1-t^2)u')' + 2u$ on I with $u(0) = 0$ (forced initial condition), $u(1) = 1$, and $u \in C_I^2$. General solutions are $u(t) = c_1 t + \frac{c_2}{2}\ t\ ln(\frac{1+t}{1-t})$ and only $u(t) = t$ satisfies the boundary conditions. To obtain this solution, consider the functional

$$J(u) = \frac{1}{2}\int_I (1-j^2)(u')^2 + u^2$$

and define the three distinct functionals which parallel those from the previous section. Let $L^0 := \{h \in L : h(0) = 0 = h(1)\}$, $H^0 = H \cap L^0$, and $H_w^0 = H_w \cap L^0$. For $u \in C_I^2$ and $h \in L^0$,

$$\begin{aligned}\langle (\nabla_{L^0} J)(u), h\rangle_{L^0} &= J'(u)(h) \\ &= \int_I (1-j^2)u'h' + uh \\ &= \int_I ((-(1-j^2)u')' + u)h \\ &= -\int_I hKu \\ &= \langle h, -P_L Ku\rangle_{L^0},\end{aligned}$$

where $P_L : L \to L^0$ is the orthogonal projection.

The parallel in H is given by

$$J(u) = \frac{1}{2}\int_I (1-j^2)(D_1 u)^2 + u^2$$

and for $u \in C_I^2$ and $h \in H^0$,

$$\langle (\nabla_{H^0} J)(u), h\rangle_{H^0} = J'(u)(h)$$

Table 6

Legendre's Equation

$\mathbf{(1-t^2)y''-2ty+2y=0}$	$\mathbf{y(0)=0}$	$\mathbf{y(1)=1}$	$\mathbf{N=100}$		
Gradient	Iterations	Seconds	Residual	Avg. Abs. Err.	Max. Abs. Err.
L	5948	24	10^{-6}	10^{-1}	6.6×10^{-1}
H	1998	7	10^{-6}	10^{-6}	3.7×10^{-5}
H_w	64	1	10^{-6}	10^{-7}	8.0×10^{-6}

$\mathbf{(1-t^2)y''-2ty+2y=0}$	$\mathbf{y(0)=0}$	$\mathbf{y(1)=1}$	$\mathbf{N=10,000}$		
Gradient	Iterations	Seconds	Residual	Avg. Abs. Err.	Max. Abs. Err.
H	2142	82	10^{-6}	10^{-6}	3.4×10^{-5}
H_w	85	3	10^{-6}	10^{-6}	1.2×10^{-5}

$\mathbf{(1-t^2)y''-2ty+2y=0}$	$\mathbf{y(0)=0}$	$\mathbf{y(1)=1}$	$\mathbf{N=100,000}$		
Gradient	Iterations	Seconds	Residual	Avg. Abs. Err.	Max. Abs. Err.
H_w	325	125	10^{-15}	10^{-14}	1.7×10^{-14}

$$\begin{aligned}
&= \int_I (1-j^2)D_1uD_1h + uh \\
&= \left\langle \begin{pmatrix} h \\ D_1h \end{pmatrix}, \begin{pmatrix} u \\ (1-j^2)D_1u \end{pmatrix} \right\rangle_{L\times L} \\
&= \left\langle \begin{pmatrix} h \\ D_1h \end{pmatrix}, P_H \begin{pmatrix} u \\ (1-j^2)D_1u \end{pmatrix} \right\rangle_{L\times L} \\
&= \left\langle h, \pi_1 P_H \begin{pmatrix} u \\ (1-j^2)D_1u \end{pmatrix} \right\rangle_{H^0}
\end{aligned}$$

where

$$P_H : L\times L \to \left\{ \begin{pmatrix} u \\ D_1u \end{pmatrix} : u\in H^0 \right\}$$

is the orthogonal projection. As in the previous section the weight chosen is the square root of the function in the functional which results from the singularity in the differential equation. In this case $w(t)=\sqrt{1-t^2}$.

The parallel in $H_{\sqrt{1-j^2}}$ is given by

$$J(u) = \frac{1}{2}\int_I \left(D_1^{\sqrt{1-j^2}}u\right)^2 + u^2 \tag{9}$$

and for $u\in H_{\sqrt{1-j^2}}, h\in H^0_{\sqrt{1-j^2}}$,

$$\begin{aligned}
\left\langle \left(\nabla_{H^0_w} J\right)(u), h\right\rangle &= J'(u)(h) \\
&= \int_I D_1^{\sqrt{1-j^2}}uD_1^{\sqrt{1-j^2}}h + uh \\
&= \left\langle \begin{pmatrix} h \\ D_1^{\sqrt{1-j^2}}h \end{pmatrix}, \begin{pmatrix} u \\ D_1^{\sqrt{1-j^2}}u \end{pmatrix} \right\rangle_{L\times L} \\
&= \left\langle \begin{pmatrix} h \\ D_1^{\sqrt{1-j^2}}h \end{pmatrix}, P_{H_{\sqrt{1-j^2}}} \begin{pmatrix} u \\ D_1^{\sqrt{1-j^2}}u \end{pmatrix} \right\rangle_{L\times L}
\end{aligned}$$

$$= \left\langle h, P_{H_{\sqrt{1-j^2}}} \pi_1 \begin{pmatrix} u \\ D_1^{\sqrt{1-j^2}} u \end{pmatrix} \right\rangle_{H^0_{\sqrt{1-j^2}}}$$

where

$$P_{H_{\sqrt{1-j^2}}} : L \times L \to \left\{ \begin{pmatrix} u \\ D_1^{\sqrt{1-j^2}} u \end{pmatrix} : u \in H^0_{\sqrt{1-j^2}} \right\}$$

is the orthogonal projection. Discretizing the functional,

$$J(u) = \frac{1}{2} \sum_{k=1}^{n} \left(1 - \left(\frac{t_{k+1} + t_k}{2} \right)^2 \right) \left(\frac{u_{k+1} - u_k}{\delta} \right)^2 + \left(\frac{u_{k+1} + u_k}{2} \right)^2$$

Since $\pi_L(x) = (0, x_1, \ldots, x_n, 0)$ and $(\nabla_L J)(u) = -Ku$, the gradient depending on the space and boundary conditions is

$$(\nabla_{L^0} J)(u) = \left(0, \ldots, -(1 - t_k^2) \frac{u_{k-1} - 2u_k + u_{k+1}}{\delta^2} + 2t_k \frac{u_{k+1} - u_{k-1}}{\delta} - 2u_k, \ldots, 0 \right).$$

Table 6 demonstrates the success associated with these problems. The algorithm is parallel to the one from the preceding section. Note the machine precision results.

6. Conclusions. Mathematicians and scientists have oft sought solutions to differential equations using descent based on the Euclidean gradient. The numerical work in this paper indicates that the choice of the underlying space and gradient are crucial for developing efficient numerical methods.

Throughout the paper, weighted descent outperforms both Sobolev descent and Euclidean descent for singular problems. Weighted descent is an extension of the standard descent, thus once the effort has been put forth to implement the non-weighted descent process, little extra effort is required to implement the weighted descent and superior results can be expected.

The versatility of the algorithm has been demonstrated by considering linear constrained, unconstrained, partially constrained first order problems, a nonlinear first order problem with irregular singularity, as well as two variational problems. A report applying the method to singular partial differential equations is forthcoming.

Boundary conditions are maintained at each step of the descent process guaranteeing exact boundary conditions for the solution and the method gives results on a small number of divisions which are representative of the results obtained on a large number of divisions making the method a candidate for multigrid problems.

Convergence results for the discrete case have been shown for specific problems and a general result is forthcoming.

All work was performed on a NeXTstation 33 MHz 68040 Unix platform using the GNU C compiler. Codes for the problems and *Mathematica* codes for computing the necessary matrices are available from the author by e-mail at math-wtm@nich-nsunet.nich.edu.

REFERENCES

[1] R. Seroul and S. Levy, *A Beginner's Book of* TEX, Springer-Verlag, Berlin, New York, 1991.

[A] Adams, R. A., *Sobolev Spaces.* Academic Press, New York, NY, 1975.

[CH] Courant R. and Hilbert D., *Methods of Mathematical Physics, Volume I.* John Wiley & Sons, New York, NY, 1989.

[CK] Canic, S. and Keyfitz, B. L., *An elliptic problem arising from the unsteady transonic small disturbance equation*, J. of Differential Equations, to appear.

[DM] Dix, J. G. and McCabe, T. W., *On finding equilibria for isotropic hyperelastic materials*, Nonlinear Anal., 15 (1990) 437-444.

[E] Elschner, J., *Singular Ordinary Differential Operators and Pseudodifferential Equations.* Springer-Verlag Lecture Notes In Mathematics, 1128, New York, NY, 1980.

[G] Garza, J., *Using steepest descent to find energy-minimizing maps satisfying nonlinear constraints*, Dissertation, University of North Texas (1994).

[K] Kim, K., *Steepest descent for partial differential equations of mixed type*, Dissertation, University of North Texas (1992).

[KA] Kufner, A., *Weighted Sobolev Spaces.* John Wiley & Sons, New York, NY, 1985.

[N1] Neuberger, J. W., *Constructive variational methods for differential equations*, Nonlinear Anal., 13 (1988), no. 4, 413-428.

[N2] Neuberger, J. W., *Calculation of sharp shocks using Sobolev gradients*, Contemp. Math., 108 (1990) 111-118.

[N3] Neuberger, J. W., *Sobolev Gradients and Differential Equations.* In preparation.

[N4] Neuberger, J. W., *Steepest descent and differential equations*, J. Math. Soc. Japan, 37 (1985) 187-195.

[N5] Neuberger, J. W., *Steepest descent for general systems of linear differential equations in Hilbert space*, in Ordinary Differential Equations and Operators, Springer-Verlag Lecture Notes in Mathematics, 1032 (1982) 390-406.

[RN] Renka, R. J. and Neuberger, J. W., *Minimal surfaces and Sobolev gradients*, to appear, SIAM J. of Sci. Comput.

[RSN] Riesz, F. and Sz.-Nagy B., *Functional Analysis.* Dover Publications, New York, NY, 1990.

[S] Schuchman, V., *On behavior of nonlinear differential equations in Hilbert space*, Internat. J. of Math. Math. Sci., 11 (1988) 143-165.