ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 33, Number 4, Winter 2003

STEEPEST DESCENT ON A UNIFORMLY CONVEX SPACE

MOHAMAD M. ZAHRAN

1. Introduction. The idea of steepest descent and how it is of use to find zeros or critical points of nonnegative C^2 functions defined on Hilbert spaces is extensively presented in [**7, 8, 9**] and [**14**]. The main objective of this paper is to generalize parts of these references to many problems of interest or set naturally in the uniformly convex space setting. We are concerned here with numerical solutions of differential equations that fit into the uniformly convex Sobolev spaces $H^{1,p}(\Omega)$ for $p > 2$ and $\Omega \subset R^n$ and do not fit conveniently into the Hilbert space $H^{1,2}(\Omega)$. A good example of this situation is the diffusion problem of the form

$$-\Delta y + F'(y) = 0, \quad y \in H^{1,p}(\Omega)$$

where F is a polynomial function. Let

$$\varphi(y) = \frac{1}{2}\int_{\Omega} y_1^2 + y_2^2 + y_3^2 + F(y), \tag{1}$$

where y_i is the partial derivative with respect to the ith variable.

We seek $y \in H^{1,p}(\Omega)$ so that $\varphi'(y)h = 0$ for all $h \in H^{1,p}(\Omega)$. To do this, note that if $y \in H^{2,p}(\Omega)$,

$$\begin{aligned}\varphi'(y)h &= \int_{\Omega} h_1 y_1 + h_2 y_2 + h_3 y_3 + F'(y)h \\ &= \int_{\partial\Omega} h\frac{\partial y}{\partial n} + \int_{\Omega}(-(y_{11} + y_{22} + y_{33}) + F'(y))h = 0.\end{aligned}$$

This implies that $-\Delta y + F'(y) = 0$ with $\partial y/\partial n = 0$ on the boundary $\partial\Omega$, where n is the outward normal of Ω.

φ in (1) will be well defined if p is chosen so that $F(y) = y^8 \in L_1(\Omega)$ for $y \in H^{1,p}(\Omega)$ and $\Omega \subset R^3$. By the Sobolev embedding theorem in [**1**], it is sufficient to choose p so that $8 \leq 3p/3 - p$. The best choice

Received by the editors on May 21, 2001.

for p is 24/11. We will come back to this example in the numerical experiment part.

2. Gradients.

Definition 1. A Banach space X is said to be uniformly convex if and only if, for every $\varepsilon > 0$, there exists $0 < \delta < 1$ such that whenever x and y are elements of X with $\|x\| = \|y\| = 1$ and $\|x - y\| \geq 2\varepsilon$ we have $\|x + y\| \leq 2\delta$.

Roughly speaking, this means that if two points on the unit sphere of a uniformly convex Banach space are far apart, then their midpoint must be well inside the unit ball. The Lebesgue spaces $L^p(\Omega)$ and the Sobolev spaces $H^{m,p}(\Omega)$ where $\Omega \subset R^n$, m is a nonnegative integer and $1 < p < \infty$, are uniformly convex. For a broad study on Sobolev spaces, see [**1**]. Hilbert spaces are uniformly convex; this can be shown by using the parallelogram law. In [**4**] there is a precise computation of δ to show that the L_p spaces are uniformly convex.

Consider the fact in [**6**] that if X is a uniformly convex Banach space, f is a function in the dual space X^* of X and $c > 0$, then there exists a unique x in X such that $f(x)$ is maximum subject to $\|x\| = c$. We present next the following definition.

Definition 2. Suppose φ is a C^1 function on a uniformly convex Banach space X. There exists a unique h in X such that $\varphi'(x)h$ is a maximum subject to $\|h\|_X = |\varphi'(x)|_{X^*}$. This h is called the gradient of φ at x and is denoted by $(\nabla\varphi)(x)$. And

$$\varphi'(x)(\nabla\varphi)(x) = \sup_{\|t\|_X = |\varphi'(x)|_{X^*}} \varphi'(x)t$$

where $\varphi'(x)$ denotes the Frechet derivative of φ at x.

This definition agrees with the definition of a gradient on a Hilbert space. Recall that if φ is a C^1 function on a Hilbert space X and $x \in X$, then

$$\begin{aligned}\varphi'(x)k &= \langle k, (\nabla\varphi)(x)\rangle \leq \|k\|_X \|(\nabla\varphi)(x)\|_X \\ &= \|(\nabla\varphi)(x)\|_X^2 = \varphi'(x)(\nabla\varphi)(x)\end{aligned}$$

for every k in X such that $\|k\|_X = |\varphi'(x)|_{X^*}$.

For n a positive integer and φ a real-valued C^1 function on R^{n+1} with the usual norm, it is customary to define $\nabla\varphi$ as the function on R^{n+1} so that

$$(2)\qquad \begin{gathered}(\nabla\varphi)(x_0,\dots,x_n) = (\varphi_0(x_0,\dots,x_n),\dots,\varphi_n(x_0,\dots,x_n)),\\ x=(x_0,\dots,x_n)\in R^{n+1}\end{gathered}$$

where φ_i denotes the partial derivative of φ in its ith argument. $(\nabla\varphi)(x)$ is called the ordinary gradient at x.

One space of interest here is the uniformly convex Banach space $X = R^{n+1}$ with the p-norm

$$(3)\qquad \begin{gathered}\|h\|_X = \left(\sum_{i=1}^{n}\left(\left|\frac{h_i-h_{i-1}}{\delta}\right|^p + \left|\frac{h_i+h_{i-1}}{2}\right|^p\right)\right)^{1/p},\\ h=(h_0,h_1,\dots,h_n)\in X,\quad \delta=\frac{1}{n}.\end{gathered}$$

We denote the gradient of the function φ at $x\in X$ by $(\nabla_p\varphi)(x)$ and we call it the p-gradient. Next we illustrate how the ordinary gradient and the p-gradient are related.

3. Relationship between two gradients.

In this section we illustrate how to calculate the p-gradient $\nabla_p\varphi$ of a C^2 function φ on the uniformly convex Banach space $X = R^{n+1}$ with the p-norm (3) by using the ordinary gradient $\nabla\varphi$ constructed with respect to the usual norm.

There exists a unique $h\in X$ such that $\varphi'(x)h = \langle h, (\nabla\varphi)(x)\rangle_{R^{n+1}}$ is maximum subject to $\|h\|_X^p = |\varphi'(x)|_{X^*}^p$. This h, by definition, is the p-gradient $(\nabla_p\varphi)(x)$ at x.

Define the function β from R^{n+1} to R so that

$$\beta(h) = \|h\|_X^p - |\varphi'(x)|_{X^*}^p,\quad \forall\, h\in R^{n+1}.$$

Using Lagrange multipliers, we get $(\nabla\varphi)(x) = \alpha(\nabla\beta)((\nabla_p\varphi)(x))$. Without loss of generality, assume $\alpha = 1$. Denote by D_0, D_1 the functions form R^{n+1} to R^n such that

$$D_0 h = \begin{pmatrix} \frac{h_1+h_0}{2} \\ \frac{h_2+h_1}{2} \\ \vdots \\ \frac{h_n+h_{n-1}}{2} \end{pmatrix}; \quad D_1 h = \begin{pmatrix} \frac{h_1-h_0}{\delta} \\ \frac{h_2-h_1}{\delta} \\ \vdots \\ \frac{h_n-h_{n-1}}{\delta} \end{pmatrix}.$$

Denote by D the function from R^{n+1} to $R^n \times R^n$ such that

$$Dh = \begin{pmatrix} D_0 h \\ D_1 h \end{pmatrix}, \quad \forall\, h \in R^{n+1}.$$

Denote also by D^t the adjoint of D as defined in [**13**].

Define the function Q so that $Q(t) = pt|t|^{(p-2)}$ for all $t \in R$. Note that

$$Q(Dh) = \begin{pmatrix} Q(\frac{h_1+h_0}{2}) \\ \vdots \\ Q(\frac{h_n+h_{n-1}}{2}) \\ Q(\frac{h_1-h_0}{\delta}) \\ \vdots \\ Q(\frac{h_n-h_{n-1}}{\delta}) \end{pmatrix}$$

$$\begin{aligned}
\beta'(h)l &= \lim_{t\to 0} \tfrac{\|h+tl\|^p - \|h\|^p}{t} \\
&= \lim_{t\to 0} \frac{1}{t} \sum_{i=1}^{n} \Bigg(\left| \tfrac{h_i+tl_i-h_{i-1}-tl_{i-1}}{\delta} \right|^p + \left| \tfrac{h_i+tl_i+h_{i-1}+tl_{i-1}}{2} \right|^p \\
&\quad - \left| \tfrac{h_i-h_{i-1}}{\delta} \right|^p - \left| \tfrac{h_i+h_{i-1}}{2} \right|^p \Bigg) \\
&= \sum_{i=1}^{n} \left(Q\left(\tfrac{h_i-h_{i-1}}{\delta}\right) \left(\tfrac{l_i-l_{i-1}}{\delta}\right) + Q\left(\tfrac{h_i+h_{i+1}}{2}\right) \left(\tfrac{l_i+l_{i-1}}{2}\right) \right) \\
&= \langle Q(Dh), Dl \rangle_{R^{n+1}} \\
&= \langle l, D^t Q(Dh) \rangle_{R^{n+1}}.
\end{aligned}$$

Hence,

$$(\nabla\beta)(h) = D^t Q(Dh).$$

Therefore,

$$(\nabla\varphi)(x) = D^t Q(D(\nabla_p \varphi)(x)).$$

The nonlinear operator $D^tQ(Dh)$ is known as the p-Laplacian of h and is often denoted by $\Delta_p h$.

If $p=2$, $X=R^{n+1}$, with the p-norm, is a Hilbert space. There is an argument in [**12**, p. 24] that shows

$$(\nabla\varphi)(x) = (D^tD)(\nabla_2\varphi)(x), \quad \forall\, x\in R^{n+1}$$

where $(\nabla_2\varphi)(x)$ is called the Sobolev gradient of φ at x.

4. Continuous steepest descent. In this section we will seek zeros of a C^2 function φ, defined on a uniformly convex Banach space X, by means of steepest descent, i.e., we seek $u\in X$ so that

$$u=\lim_{t\to\infty} z(t) \quad \text{exists and} \quad \varphi(u)=0$$

where $z:[0,\infty)\to X$ such that $z'(t) = -(\nabla\varphi(z(t)))$.

We first establish global existence for steepest descent in the uniformly convex setting in a way close to the one done in [**10**] by Neuberger in the Hilbert space setting.

Lemma 3. *Suppose φ is a C^1 function on a uniformly convex Banach space X, then*

$$\varphi'(x)(\nabla\varphi)(x) = \|(\nabla\varphi)(x)\|_X^2, \quad \forall\, x\in X.$$

Proof. Since $|\varphi'(x)|_{X^*} = \sup_{\|t\|_X=1}\varphi'(x)t$, $|\varphi'(x)|^2_{X^*} = \sup_{\|t\|_X=|\varphi'(x)|_{X^*}}$ $\varphi'(x)t$. Hence

$$|\varphi'(x)|^2_{X^*} = \varphi'(x)(\nabla\varphi)(x).$$

Therefore

$$\|(\nabla\varphi)(x)\|_X^2 = \varphi'(x)(\nabla\varphi)(x). \qquad \square$$

Theorem 4. *Suppose φ is a C^2 function on X, a uniformly convex space Banach. If $x\in X$, there is a unique function z from $[0,\infty)$ to X such that*

$$z(0)=x, \quad z'(t) = -(\nabla\varphi)(z(t)), \quad t\ge 0. \tag{4}$$

You may refer to [**10**, p. 14] for an argument in the Hilbert space setting, using Lemma 3 whenever necessary to write a proof of Theorem 4 in the uniformly convex setting.

Lemma 5. *Suppose X and Y are two Banach spaces, f is a C^2 function from X to Y. Define φ by*

$$\varphi(x) = \frac{\|f(x)\|_Y^p}{p}, \quad x \in X, \quad p \geq 2. \tag{5}$$

If $a \in X$ such that $f(a) = 0$, then $\varphi'(a) = 0$ where $\varphi'(a)$ is the Frechet derivative of φ at a and $\varphi'(a) \in L(X, R) = X^*$ the dual space of X.

Proof. Suppose $a \in X$ such that $f(a) = 0$. We need to show that $\varphi'(a) = 0$. Let $\varepsilon > 0$. Since f is differentiable at a and $f(a) = 0$, there exists $\delta_1 > 0$ such that if $0 < \|x - a\| < \delta_1$ and $x \in X$, then

$$\frac{\|f(x) - f'(a)(x-a)\|_Y}{\|x-a\|_X} < 1.$$

This implies that

$$\frac{\|f(x)\|_Y}{\|x-a\|_X} - \frac{\|f'(a)(x-a)\|_Y}{\|x-a\|_X} < 1.$$

Therefore,

$$\frac{\|f(x)\|_Y}{\|x-a\|_X} < 1 + \|f'(a)\|_{L(X,Y)}.$$

Since f is continuous at a and $f(a) = 0$, there exists δ_2 such that if $0 < \|x - a\|_X < \delta_2$, then

$$\|f(x)\|_Y < \left(\frac{p\varepsilon}{1 + \|f'(a)\|_{L(X,Y)}}\right)^{1/(p-1)}.$$

Let $\delta = \min(\delta_1, \delta_2)$. If $0 < \|x - a\|_X < \delta$, then

$$\frac{|\alpha(x) - \varphi(a)|}{\|x-a\|_X} = \frac{\|f(x)\|_Y^p}{p\|x-a\|_X} < \varepsilon.$$

Therefore $\varphi'(a) = 0$. $\square$

If $p = 2$ and X is a Hilbert space, the proof of Lemma 5 is straightforward since $\varphi'(a)h = \langle f'(a)h, f(a)\rangle$, for all $h \in X$.

Definition 6. Suppose φ is a C^2 function defined on a subset Ω of a uniformly convex Banach space X. φ is said to satisfy a gradient inequality on Ω provided that there is a $c > 0$ such that

$$\|(\nabla\varphi)(x)\|_X \geq c\sqrt{\varphi(x)}, \quad x \in \Omega. \tag{6}$$

Following are two propositions that will lead to finding zeros of φ by means of steepest descent using the gradient inequality (6).

If we use Lemmas 3 and 5, the proofs of Theorems 7 and 8 are very similar to those done in [**10**, Theorems 4.3 and 4.8] where $p = 2$ and X is a Hilbert space. So proofs for the two theorems will be omitted.

Theorem 7. *Suppose φ is a nonnegative C^2 function on Ω, a subset of X, X is a uniformly convex Banach space and φ satisfies the gradient inequality* (6).

If $x \in \Omega$ and z satisfies (4), *then $u = \lim_{t\to\infty} z(t)$ exists and $\varphi(u) = 0$ provided that $R(z) \subset \Omega$, where $R(z)$ is the range of the function z.*

Theorem 8. *Suppose X and Y are two uniformly convex Banach spaces and f is a C^2 function from X to Y. Suppose also $x \in X$ and r, c are two positive real numbers such that*

$$\|(\nabla\varphi)(y)\|_X \geq c\|f(y)\|_Y, \quad \|y - x\|_X \leq r$$

where φ is defined by (5).

If z satisfies (4), *then $u = \lim_{t\to\infty} z(t)$ exists and $\varphi(u) = 0$ if $\|f(x)\| \leq (cr)^{1/(p-1)}$.*

5. Critical points of convex functions. The convexity condition of nonnegative C^2 functions defined on a Hilbert space is of great importance in the study of steepest descent. In [**10**, Theorem 4.10], the convexity condition is used to find a zero for the gradient of a nonnegative C^2 function defined on a Hilbert space. In this paper we restrict our study to finding critical points for convex nonnegative C^2 functions defined on the uniformly convex space R^{n+1} with the p-norm (3).

Definition 9. A C^2 function φ on a uniformly convex space X is said to be convex in the direction of the gradient $\nabla\varphi$ if and only if there exists $\varepsilon > 0$ such that

$$\varphi''(x)((\nabla\varphi)(x), (\nabla\varphi)(x)) \geq \varepsilon\|(\nabla\varphi)(x)\|_X^2, \quad x \in X.$$

Lemma 10. *Suppose φ is a nonnegative C^2 function on a uniformly convex space X such that $\nabla\varphi \neq 0$ and $c > 0$. Then there exists $y \in X$ such that $(\nabla\varphi)(y) \neq 0$ and $\|(\nabla\varphi)(y)\|_X^2 \leq c$.*

Proof. Let $x \in X$ be such that $(\nabla\varphi)(x) \neq 0$. By Theorem 4, there exists a function $z : [0, \infty) \to X$ such that

$$z(0) = x, \quad z'(t) = -(\nabla\varphi)(z(t)), \quad \forall\, t \geq 0.$$

Let g be a nonnegative function such that $g(t) = \varphi(z(t))$, for all $t \geq 0$.

$$g'(t) = \varphi(z)'(t) = \varphi'(z(t))z'(t) = -\varphi'(z(t))(\nabla\varphi)(z(t)).$$

By Lemma 3,

$$g'(t) = -\|(\nabla\varphi)(z(t))\|_X^2, \quad \forall\, t \geq 0.$$

We intend to show that there is a $t_0 \geq 0$ such that $\|(\nabla\varphi)(z(t_0))\|_X^2 \leq c$. Suppose that $\|(\nabla\varphi)(z(t))\|_X^2 > c$ for all $t \geq 0$. Then $-g'(t) > c$ for all $t \geq 0$. Hence,

$$\int_0^s -g'(t)\, dt > cs, \quad \forall\, s > 0.$$

Thus

$$g(0) - g(s) > cs, \quad \forall\, s > 0.$$

So

$$g(0) - cs > g(s) \geq 0, \quad \forall\, s > 0.$$

Therefore,

$$g(0) > cs, \quad \forall\, s > 0.$$

This cannot happen. Hence there is a $t_0 \geq 0$ such that $\|(\nabla\varphi)(z(t_0))\|_X^2 \leq c$. The proof of the lemma is now complete. $\square$

Lemma 11. *Suppose φ is a C^2 function, the uniformly convex Banach space $X = R^{n+1}$ with the p-norm* (3). *Define the functions $g = \varphi(z)$ where z satisfies* (4) *for $x \in R^{n+1}$. Then*

$$\varphi''(z(t))((\nabla_p\varphi)(z(t)), (\nabla_p\varphi)(z(t))) = g''(t) - \frac{p}{2}(-g'(t))^{\frac{p}{2}-1} g''(t), \qquad \forall\, t \geq 0.$$

Proof. By Lemma 3,

$$g'(t) = \varphi'(z(t))z'(t) = -\varphi'(z(t))(\nabla_p\varphi)(z(t)) = -\|(\nabla_p\varphi)(z(t))\|_X^2.$$

So

$$\begin{aligned} g''(t) &= -\varphi''(z(t))((\nabla_p\varphi)(z(t)), z'(t)) - \varphi'(z(t))(\nabla_p\varphi)'(z(t))z'(t) \\ &= \varphi''(z(t))((\nabla_p\varphi)(z(t)), (\nabla_p\varphi)(z(t))) - \varphi'(z(t))(\nabla_p\varphi)'(z(t))z'(t). \end{aligned}$$

Define the function s so that $s(t) = -(-g'(t))^{p/2}$

$$s(t) = -\|(\nabla_p\varphi)(z(t))\|_X^p = -\beta((\nabla_p\varphi)(z(t))) - |\varphi'(x)|_{X*}^p.$$

where β is the function defined in Section 3.

$$\begin{aligned} s'(t) &= -\beta'((\nabla_p\varphi)(z(t)))(\nabla_p\varphi)'(z(t))z'(t) \\ &= -\langle D^t Q(D(\nabla_p\varphi)(z(t))), (\nabla_p\varphi)'(z(t))z'(t)\rangle_{R^{n+1}} \\ &= -\langle(\nabla\varphi)(z(t)), (\nabla_p\varphi)'(z(t))z'(t)\rangle_{R^{n+1}} \\ &= -\varphi'(z(t))(\nabla_p\varphi)'(z(t))z'(t). \end{aligned}$$

Hence,

$$g''(t) = \varphi''(z(t))((\nabla_p\varphi)(z(t)), (\nabla_p\varphi)(z(t))) + s'(t).$$

This implies that

$$\varphi''(z(t))((\nabla_p\varphi)(z(t)),(\nabla_p\varphi)(z(t)))=g''(t)-s'(t).$$

Note also that

$$s'(t)=\frac{p}{2}(-g'(t))^{\frac{p}{2}-1}g''(t).$$

Thus

$$\varphi''(z(t))((\nabla_p\varphi)(z(t)),(\nabla_p\varphi)(z(t)))=g''(t)-\frac{p}{2}(-g'(t))^{\frac{p}{2}-1}g''(t). \qquad \square$$

Theorem 12. *Suppose φ is a nonnegative C^2 function on the uniformly convex Banach space $X=R^{n+1}$ with the p-norm* (3), *and $\nabla_p\varphi\neq 0$. Suppose also that φ is convex in the sense of Definition* 9. *There exists $x\in X$ such that $(\nabla_p\varphi)(x)\neq 0$ and if z satisfies* (4), *then $u=\lim_{t\to\infty}z(t)$ exists and $(\nabla_p\varphi)(u)=0$.*

Proof. If $g=\varphi(z)$ where z satisfies (4) for $x\in R^{n+1}$ such that $(\nabla_p\varphi)(x)\neq 0$, then by Lemma 11 and Definition 9,

$$g''(t)-\frac{p}{2}(-g'(t))^{\frac{p}{2}-1}g''(t)\geq\varepsilon\|(\nabla_p\varphi)(z(t))\|_X^2, \quad \forall\, t\geq 0.$$

Since $g'(t)=-\|(\nabla_p\varphi)(z(t))\|_X^2$,

$$g''(t)-\frac{p}{2}(-g'(t))^{\frac{p}{2}-1}g''(t)\geq-\varepsilon g'(t).$$

So

$$\frac{g''(t)}{g'(t)}+\frac{p}{2}(-g'(t))^{\frac{p}{2}-2}g''(t)\leq-\varepsilon.$$

Integrating both sides, we get

$$\ln\left(\frac{g'(t)}{g'(0)}\right)-\frac{p}{p-2}(-g'(t))^{\frac{p}{2}-1}+\frac{p}{p-2}(-g'(0))^{\frac{p}{2}-1}\leq-\varepsilon t.$$

Hence

$$\frac{g'(t)}{g'(0)}\leq\exp(-\varepsilon t)\exp\left(\frac{p}{p-2}(-g'(t))^{\frac{p}{2}-1}\right)$$
$$\times\exp\left(-\frac{p}{p-2}(-g'(0))^{\frac{p}{2}-1}\right).$$

Define the function $f(t) = -g'(t)$. Then

$$0 \le f(t) \le f(0)\exp(-\varepsilon t)\exp\left(\frac{p}{p-2}(f(t))^{\frac{p}{2}-1}\right) \times \exp\left(-\frac{p}{p-2}(f(0))^{\frac{p}{2}-1}\right). \tag{7}$$

Hence

$$\lim_{t\to 0}\frac{f(t)-f(0)}{t} \le \lim_{t\to 0}\frac{f(0)[\exp(-\varepsilon t)\exp[p/(p-2)((f(t))^{\frac{p}{2}-1}-(f(0))^{\frac{p}{2}-1}))-1]}{t}.$$

Using L'Hopital's rule on the righthand side, we get

$$f'(0) \le \lim_{t\to 0} f(0)\Bigg[-\varepsilon\exp(-\varepsilon t)\exp\left(\frac{p}{p-2}((f(t))^{\frac{p}{2}-1}-(f(0))^{\frac{p}{2}-1})\right) + \exp(-\varepsilon t)\frac{p}{2}((f(t))^{\frac{p}{2}-2}f'(t)\exp\left(\frac{p}{p-2}((f(t)^{\frac{p}{2}-1}-(f(0))^{\frac{p}{2}-1}))\right)\Bigg].$$

Thus

$$f'(0) \le f(0)\left(-\varepsilon + \frac{p}{2}(f(0))^{\frac{p}{2}-2}f'(0)\right)$$

and $f'(0) - \frac{p}{2}(f(0))^{\frac{p}{2}-1}f'(0) \le -\varepsilon f(0) < 0$ since $f(0) = -g'(0) > 0$. Therefore,

$$f'(0)\left(1-\frac{p}{2}(f(0))^{\frac{p}{2}-1}\right) < 0. \tag{8}$$

We intend to show that $f'(0) < 0$. By Lemma 10 there exists $x \in X$ such that

$$0 < \|(\nabla_p\varphi)(x)\|_X^2 \le \left(\frac{2}{p}\right)^{2/(p-2)}.$$

Note that we can consider such an x from the very beginning. Now

$$f(0) = -g'(0) = \|(\nabla_p\varphi)(z(0))\|_X^2 = \|(\nabla_p\varphi)(x)\|_X^2 \le \left(\frac{2}{p}\right)^{2/(p-2)}.$$

Hence,

$$\frac{p}{2}(f(0))^{(p-2)/2} \le 1.$$

Then (8) implies that $f'(0) < 0$.

Now we intend to show that $f(t) \le f(0)$, for all $t \ge 0$.

Suppose there is a $q > 0$ such that $f(q) > f(0)$. Since $f'(0) < 0$, there is a $t_0 > 0$ such that $f(t_0) < f(0)$. Hence $f(t_0) < f(0) < f(q)$. So there is a b such that $t_0 < b < q$ and $f(b) = f(0)$.

Now by (7), we have

$$f(b) \le f(0)\exp(-\varepsilon b)\exp\left(\frac{p}{p-2}(f(b))^{\frac{p}{2}-1}\right)\exp\left(-\frac{p}{p-2}(f(0))^{\frac{p}{2}-1}\right).$$

Hence,

$$f(b) \le f(0)\exp(-\varepsilon b),$$

and so $f(b) < f(0)$, which is a contradiction, since $f(b) = f(0)$. Therefore,

$$f(t) \le f(0), \quad \forall\, t \ge 0.$$

This implies that

$$\frac{p}{p-2}(f(t))^{\frac{p}{2}-1} \le \frac{p}{p-2}(f(0))^{\frac{p}{2}-1}.$$

Hence

$$\exp\left(\frac{p}{p-2}(f(t))^{\frac{p}{2}-1} - \frac{p}{p-2}(f(0))^{\frac{p}{2}-1}\right) \le 1.$$

Thus (7) implies that

$$f(t) \le f(0)\exp(-\varepsilon t).$$

So

$$0 \le -g'(t) \le -g'(0)\exp(-\varepsilon t).$$

This implies that

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

Now since $g'(t) = -\|(\nabla_p\varphi)(z(t))\|_X^2 = -\|z'\|_X^2$, we have

$$\int_a^{a+1} \|z'\|_X^2 \le -f'(0)\int_a^{a+1} \exp(-\varepsilon t)\,dt, \quad \forall\, a \ge 0.$$

Hence,

$$\int_a^{a+1} \|z'\|_X^2 \leq \frac{g'(0)\exp(-\varepsilon(a+1))}{\varepsilon} - \frac{g'(0)\exp(-\varepsilon a)}{\varepsilon}, \quad \forall\, a \geq 0.$$

Since $g'(0) < 0$,

$$\left(\int_a^{a+1} \|z'\|_X\right)^2 \leq \int_a^{a+1} \|z'\|_X^2 \leq \frac{-g'(0)\exp(-\varepsilon a)}{\varepsilon}, \quad \forall\, a \geq 0.$$

So

$$\int_a^{a+1} \|z'\|_X \leq \left(\frac{-g'(0)}{\varepsilon}\right)^{1/2} \exp\left(\frac{-\varepsilon a}{2}\right), \quad \forall\, a \geq 0.$$

Now

$$\begin{aligned}
\int_0^{\infty} \|z'\|_X &= \sum_{n=0}^{\infty} \int_n^{n+1} \|z'\|_X \\
&\leq \sum_{n=0}^{\infty} \left(\frac{-g'(0)}{\varepsilon}\right)^{1/2} \exp\left(\frac{-\varepsilon n}{2}\right) \\
&= \left(\frac{-g'(0)}{\varepsilon}\right)^{1/2} \sum_{n=0}^{\infty} \exp\left(\frac{-\varepsilon n}{2}\right) \\
&= \left(\frac{-g'(0)}{\varepsilon}\right)^{1/2} \frac{1}{1-\exp(\frac{-\varepsilon}{2})}.
\end{aligned}$$

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

Since $\lim_{t\to\infty} g'(t) = 0$, $\lim_{t\to\infty} \|(\nabla_p\varphi)(z(t))\| = 0$. So $\lim_{t\to\infty}(\nabla_p\varphi)(z(t)) = 0$. Hence $(\nabla_p\varphi)(\lim_{t\to\infty} z(t)) = 0$. Therefore, $(\nabla_p\varphi)(u) = 0$ and the proof of the theorem is now complete. $\square$

The p-norm (3) in $X = R^{(n+1)}$ is a finite-dimensional emulation of the norm

$$\|f\| = \left(\int_0^1 |f|^p + |f'|^p\right)^{1/p}, \quad f \in H^{1,p}[0,1]$$

in the Sobolev space $H^{1,p}[0,1]$.

The work is still underway to extend the above theorem to the infinite dimensional case in the space $H^{1,p}[0,1]$ for $p > 2$, with the above norm.

6. Numerical experiments. If $\Omega = [0,1] \times [0,1] \times [0,1]$, the domain of the function $\varphi(y) = (1/2)\int_\Omega y_1^2 + y_2^2 + y_3^2 + F(y)$ with $F(y) = y^8 \in L_1(\Omega)$ and $y \in H^{1,p}(\Omega)$ which we mentioned in the introduction will be $(n+1)^3$-dimensional space whose points are real-valued functions defined on the grid $\{[(i/n),(j/n),(k/n)]\}_{i,j,k=1}^n$. So one should expect computer memory trouble while running a code to find numerically critical points for the function φ.

Next we present two problems that can be easily computed.

Problem 1. Consider the real-valued C^1 function

$$\varphi(y) = \frac{1}{p}\int_0^1 (y'-y)^p, \quad y \in H^{1,p}[0,1], \quad p > 2. \tag{9}$$

Assume $p = 4$. Then φ is a nonnegative C^1 function on $H^{1,4}[0,1]$. We seek $y \in H^{1,4}[0,1]$ so that $\varphi'(y)h = 0$, for all $h \in H^{1,4}[0,1]$. To do this, note that

$$\begin{aligned}
\varphi'(y)h &= \int_0^1 (y'-y)^3(h'-h)\\
&= \int_0^1 (y'-y)^3h' - (y'-y)^3h\\
&= [h(y'-y)^3]_0^1 + \int_0^1 (-3(y'-y)^2(y''-y')h - (y'-y)^3h)\\
&= [h(y'-y)^3]_0^1 + \int_0^1 h(y'-y)^2(-3y''+3y'-y'+y)\\
&= [h(y'-y)^3]_0^1 + \int_0^1 h(y'-y)^2(-3y''+2y'+y) = 0,\\
&\qquad \forall\, h \in H^{1,4}[0,1],
\end{aligned}$$

which yields the following equations

$$(y'-y)^2(-3y''+2y'+y) = 0, \quad y \in H^{2,4}[0,1], \tag{10}$$

with the boundary conditions $y'(0) = y(0)$ and $y'(1) = y(1)$. It is very important to mention that the solutions for (1) are not necessarily $y = ce^t$. The following is also a solution for (10). Some special initial condition might lead to an approximation to it.

$$y(t) = \begin{cases} e^t & t \in \left[0, \dfrac{1}{3}\right] \\ \Big(\dfrac{e^{1/9} - e^{-1/9}}{e^{1/9} - e^{5/9}}\Big)e^t + \Big(\dfrac{e^{1/3} - e}{e^{1/9} - e^{5/9}}\Big)e^{-t/3} & t \in \left(\dfrac{1}{3}, \dfrac{2}{3}\right), \\ e^{-2/3}e^t & t \in \left[\dfrac{2}{3}, 1\right]. \end{cases} \tag{11}$$

When we seek numerically critical points for (9) using its finite-dimensional emulation

$$\varphi(y) = \frac{1}{4}\sum_{i=1}^{n}\left(\frac{y_i - y_{i-1}}{\delta} - \frac{y_i + y_{i-1}}{2}\right)^4,$$

which is defined on the space R^{n+1} with the norm

$$\|y\| = \left(\sum_{i=1}^{n}\left(\left|\frac{y_i - y_{i-1}}{\delta}\right|^4 + \left|\frac{y_i + y_{i-1}}{2}\right|^4\right)\right)^{1/4},$$

$$y = (y_0, y_1, \dots, y_n) \in R^{n+1}, \quad \delta = \frac{1}{n},$$

the computer shows only $y = ce^t$ as critical points. This is due to the fact that (11) might be a saddle point. To write a computer code, we consider the equation $(\nabla\varphi)(y) = D^tQ(Dh)$ from Section 3, where h is the p-gradient $(\nabla_p\varphi)(y)$ and $(\nabla\varphi)(y)$ is the ordinary gradient (2).

To solve for h, let $(\nabla\varphi)(y) = [c_i]$, $0 \le i \le n$. Given h_{i-1} and h_{i+1} we need to solve for h_i the equation

$$\begin{aligned} F(h_i) = {} & \frac{1}{2}\left[Q\left(\frac{h_{i+1} + h_i}{2}\right) + Q\left(\frac{h_i + h_{i-1}}{2}\right)\right] \\ & + \frac{1}{\delta}\left[Q\left(\frac{h_i - h_{i-1}}{\delta}\right) - Q\left(\frac{h_{i+1} - h_i}{\delta}\right)\right] - c_i = 0. \end{aligned}$$

To do this we use Newton's iteration

$$h_i \longrightarrow h_i - \frac{F(y)}{F'(y)}.$$

Then we consider the following iteration to find the critical point y

$$y \longrightarrow y - \alpha h$$

where α is chosen optimally.

Problem 2. Consider the nonnegative real-valued C^1 function

$$\varphi(y) = \frac{1}{p}\int_0^1 |y'|^p, \quad y \in H^{1,p}[0,1], \quad p > 2. \tag{12}$$

Assume $p = 3$. We seek $y \in H^{1,3}[0,1]$ so that $\varphi'(y)h = 0$ for all $h \in H^{1,3}[0,1]$. To do this note that

$$\begin{aligned}\varphi'(y)h &= \int_0^1 y'|y'|h' \\ &= [y' \mid y']h_0^1 - \int_0^1 (y'|y'|)'h = 0, \quad \forall\, h \in H^{1,3}[0,1],\end{aligned}$$

which yields the following equation

$$(y'|y'|)' = 0, \quad \forall\, y \in H^{1,3}, \tag{13}$$

with the boundary conditions $y'(0) = y(1) = 0$. The solutions for (13) are not necessarily horizontal lines. The following is also a solution for (13). Some special initial condition might lead to an approximation to it.

$$y(t) = \begin{cases} 1 & t \in \left[0, \dfrac{1}{3}\right] \\ 3t & t \in \left(\dfrac{1}{3}, \dfrac{2}{3}\right) \\ 2 & t \in \left[\dfrac{2}{3}, 1\right]. \end{cases} \tag{14}$$

When we seek numerically critical points for (12) using its finite-dimensional emulation

$$\varphi(y) = \frac{1}{3}\sum_{i=1}^{n}\left|\frac{y_i - y_{i-1}}{\delta}\right|^3,$$

the computer shows only horizontal lines as critical points. This is also due to the fact that (14) might be a saddle point. In this code we use the same algorithm as in Problem 1.

REFERENCES

1. R. Adams, *Sobolev spaces*, Academic Press, San Diego, 1978.

2. M.S. Berger, *Nonlinearity and functional analysis*, Academic Press, New York, 1977.

3. J. Diestel, *Sequences and series in Banach spaces*, Springer-Verlag, Berlin, 1992.

4. O. Hanner, *On the uniform convexity of L^p and l^p*, Ark. Mat. **3** (1956), 239–244.

5. V. Huston and J.S. Pym, *Application of functional analysis and operator theory*, Academic Press, New York, 1980.

6. R.C. James, *Weak convergence and reflexivity*, Trans. Amer. Math. Soc. **113** (1964), 129–140.

7. J. Neuberger, *Steepest descent for general systems of linear differential equation in Hilbert space*, in *Ordinary differential equations and operators*, Lecture Notes in Math., vol. 1032, Springer-Verlag, New York, 1983, pp. 390–406.

8. ———, *Steepest descent and differential equations*, J. Math. Soc. Japan **37** (1985), 187–195.

9. ———, *Construction variational methods for differential equations*, Nonlinear Analysis Th. Meth. Appl. **13** (1988), 413–428.

10. ———, *Sobolev gradients and differential equations*, Lecture Notes in Math., vol. 1670, Springer-Verlag, New York, 1997.

11. ———, *Steepest descent for systems of nonlinear partial differential equations*, Oak Ridge National Laboratory/CSD, TM-161 (1982).

12. ———, *Calculation of sharp shocks using Sobolev gradients*, Contemp. Math., vol. 108, Amer. Math. Soc., Providence, RI, 1990.

13. F. Riesz and B. Sz.-Nagy, *Functional analysis*, Ungar, 1955.

14. P. Rosenblum, *The method of steepest descent*, Proc. Symp. Appl. Math. **6**, Amer. Math. Soc., Providence, RI.

Hariri Canadian University, P.O. Box 13/5171, Mechref, Lebanon
E-mail address: mmz20@hotmail.com