Pacific Journal of Mathematics

ON CONSTRAINED EXTREMA

THOMAS VOGEL

Volume 176 No. 2 December 1996

PACIFIC JOURNAL OF MATHEMATICS
Vol. 176, No. 2, 1996

ON CONSTRAINED EXTREMA

THOMAS I. VOGEL

Assume that I and J are smooth functionals defined on a Hilbert space H. We derive sufficient conditions for I to have a local minimum at y subject to the constraint that J is constantly $J(y)$.

The first order necessary condition for I to have a constrained minimum at y is that for some constant λ, $I'_y + \lambda J'_y$ is identically zero. Here I'_y and J'_y are the Fréchet derivatives of I and J at y. For the rest of the paper, we assume that y in H satisfies this necessary condition.

A common misapprehension (upon which much of the stability results for capillary surfaces has been based) is to assume that if the quadratic form $I''_y + \lambda J''_y$ is positive definite on the kernel of J'_y then I has a local constrained minimum at y. This is not correct in a Hilbert space of infinite dimension; Finn [**1**] has supplied a counterexample in the unconstrained case, and the same difficulty will occur in the constrained case. In the unconstrained case, if (as often occurs in practice) the spectrum of I''_y is discrete and 0 is not a cluster point of the spectrum, then I''_y positive definite at a critical point y implies that I''_y is strongly positive, (i.e., there exists $k > 0$ such that $I''_y(x) \geq k\|x\|^2$ holds for all x), and this in turn *does* imply that y is a local minimum (see [**2**]). However, in the constrained case, things are not so easy. Even if $I''_y + \lambda J''_y$ has a nice spectrum (in some sense), it is not clear that $I''_y + \lambda J''_y$ being positive definite on the kernel of J'_y implies that this quadratic form is strongly positive on the kernel, nor that strong positivity implies that y is a local minimum.

In [**3**], Maddocks obtained sufficient conditions for $I''_y + \lambda J''_y$ to be positive definite on the kernel of J'_y. As Maddocks points out, this is not quite enough to say that I has a constrained minimum at y. Remarkably, essentially the same conditions as Maddocks obtained for positive definiteness do in fact imply that I has a strict local minimum at y subject to the constraint $J = J(y)$, as we shall see.

For any $h \in H$ we may say $J(y+h) - J(y) = J'_y(h) + \frac{1}{2}J''_y(h) + \epsilon_1(h)\|h\|^2$, where ϵ_1 goes to zero as $\|h\|$ goes to zero. If we consider an h for which $J(y+h) = J(y)$, then of course $0 = J'_y(h) + \frac{1}{2}J''_y(h) + \epsilon_1(h)\|h\|^2$. Now, for that h we have

$$\begin{aligned}\Delta I = I(y+h) - I(y) &= I_y'(h) + \frac{1}{2}I_y''(h) + \epsilon_2\|h\|^2 \\ &= -\lambda J_y'(h) + \frac{1}{2}I_y''(h) + \epsilon_2\|h\|^2 \\ &= \frac{1}{2}(I_y'' + \lambda J_y'')(h) + (\lambda\epsilon_1 + \epsilon_2)\|h\|^2.\end{aligned} \tag{1}$$

Since $I_y'' + \lambda J_y''$ is a bilinear form, there is a linear operator A defined on H so that $(I_y'' + \lambda J_y'')(u,v) = \langle u, Av\rangle$. Similarly there is some element of H, call it ∇J, so that J_y' applied to a vector h is $\langle h, \nabla J\rangle$. Let $\sigma(A)$ be the spectrum of A. There are three cases which often arise in practice:

Theorem 1. *If $\sigma(A)\cap(-\infty, c] = \emptyset$ for some $c > 0$, then I has a constrained minimum at y.*

Proof. From (1) we may write ΔI as $\langle h, Ah\rangle + (\lambda\epsilon_1 + \epsilon_2)\|h\|^2$. But $\langle h, Ah\rangle \geq c\|h\|^2$ (this is easily verified using the spectral theorem, see [**5**]), so for h sufficiently small, ΔI is positive. □

Theorem 2. *Suppose that $\sigma(A) \cap (-\infty, \epsilon]$ consists of a single negative eigenvalue λ_0 for some $\epsilon > 0$. Let ζ solve $A\zeta = \nabla J$. (A will be invertible.) I has a constrained minimum at y if $J_y'(\zeta) = \langle \zeta, A\zeta\rangle < 0$, and I does not have a constrained minimum at y if $J_y'(\zeta) = \langle\zeta, A\zeta\rangle > 0$.*

The proof of Theorem 2 will proceed in a series of steps.

Step 1. Assume that $\langle\zeta, A\zeta\rangle < 0$. Then $I_y'' + \lambda J_y''$ is strongly positive on the kernel of J_y'.

Proof. Take x in the kernel of J_y'. As in [**4**], x may be written as $v+\alpha\zeta$, where v is perpendicular to φ_0, the eigenfunction corresponding to λ_0. (The key to this calculation is that $\langle\zeta, \varphi_0\rangle \neq 0$. But if ζ is orthogonal to φ_0, it can be shown that $\langle\zeta, A\zeta\rangle > 0$.) One can verify that $\langle x, Ax\rangle = \langle v, Av\rangle - \alpha^2\langle\zeta, A\zeta\rangle$, so that $\langle x, Ax\rangle \geq \langle v, Av\rangle$.

Let $\{E_\lambda\}$ be the spectral family associated with A, so that $A = \int_{-\infty}^{\infty} \lambda dE_\lambda$. By our assumption on $\sigma(A)$, $A = \lambda_0 E_{\lambda_0} + \int_\epsilon^\infty \lambda dE_\lambda$, where E_{λ_0} is orthogonal projection onto φ_0. Therefore,

$$\langle v, Av\rangle = \langle v, \lambda_0 E_{\lambda_0}(v)\rangle + \int_\epsilon^\infty \lambda d\|E_\lambda v\|^2.$$

The first term vanishes, so that

$$\langle v, Av\rangle \geq \epsilon\int_\epsilon^\infty d\|E_\lambda v\|^2 \geq \epsilon\int_{-\infty}^\infty d\|E_\lambda v\|^2 \geq \epsilon\|v\|^2.$$

Therefore, $\langle x, Ax\rangle \geq \epsilon\|v\|^2$.

To conclude the proof that $I''_y + \lambda J''_y$ is strongly positive on the kernel of J'_y, we need to show that $\|v\| \geq k\|x\|$ for some fixed positive constant k. Assume without loss of generality that $\|x\| = 1$. For any fixed x, $\|v\|$ is greater than or equal to the distance from x to the line $\{c\zeta : c \in \mathbf{R}\}$. Consider the projection of x onto ζ. Its length is $|\langle x, \zeta/\|\zeta\|\rangle|$. We may write ζ as $\beta\nabla J + \hat{\zeta}$, where $\hat{\zeta}$ is perpendicular to ∇J. We cannot have β equaling 0, since by assumption, $\langle\zeta, A\zeta\rangle = \langle\zeta, \nabla J\rangle < 0$.

Then the projection has length at most $\|x\|\|\hat{\zeta}\|/\|\zeta\|$. But $\|\hat{\zeta}\| < \|\zeta\|$ (since $\beta \neq 0$). Letting γ equal $\|\hat{\zeta}\|/\|\zeta\|$, we have $\gamma < 1$ and the length of the vector component of x perpendicular to ζ is greater than or equal to $\sqrt{1-\gamma^2}$. But $\|v\|$ is greater than or equal to the length of that component, so we get our k to be $\sqrt{1-\gamma^2}$, concluding step 1.

Step 2. If $\langle\zeta, A\zeta\rangle < 0$, then I has a minimum at y subject to the constraint $J = J(y)$.

Proof. Take an h for which $J(y+h) = J(y)$. Now h need not be in the kernel of J'_y, but we may write h as $h_1 + \alpha\zeta$, where h_1 is in the kernel of J'_y, by taking α to be $\langle h, \nabla J\rangle/\langle\zeta, \nabla J\rangle$. (Note that $\langle\zeta, \nabla J\rangle = \langle\zeta, A\zeta\rangle \neq 0$.) Substituting into equation (1),

$$\text{(2)} \qquad \Delta I = \frac{1}{2}\langle h_1, Ah_1\rangle + \alpha\langle h_1, A\zeta\rangle + \frac{1}{2}\alpha^2\langle\zeta, A\zeta\rangle + (\lambda\epsilon_1 + \epsilon_2)\|h\|^2.$$

However, $\langle h_1, A\zeta\rangle = \langle h_1, \nabla J\rangle = 0$, causing this term to vanish. We have $0 = \Delta J = J'_y(h) + \epsilon_3\|h\|$, where ϵ_3 tends to 0 as $\|h\|$ tends to 0. Thus $\alpha^2 = \epsilon_3^2\|h\|^2$, and we conclude that

$$\Delta I = \frac{1}{2}\langle h_1, Ah_1\rangle + \epsilon\|h\|^2$$

where ϵ tends to zero as $\|h\|$ tends to 0. From Step 1, A is strongly positive on the kernel of J'_y, so

$$\Delta I \geq \frac{k}{2}\|h_1\|^2 + \epsilon\|h\|^2.$$

Since $h = h_1 + \alpha\zeta$, with $\alpha = -\epsilon_3\|h\|$, it is easy to see that for $\|h\|$ sufficiently small there holds $\|h_1\| \geq \frac{1}{2}\|h\|$. Thus

$$\Delta I \geq \|h\|\left(\frac{k}{8} + \epsilon\right)$$

which must be greater than 0 for $\|h\|$ sufficiently small. Therefore I has a minimum at y subject to the constraint $J = J(y)$, concluding the proof of step 2 and the first half of Theorem 2.

Step 3. Suppose that $\langle\zeta, A\zeta\rangle > 0$. Then I does not have a minimum at y subject to the constraint $J = J(y)$.

Proof. First, $I''_y + \lambda J''_y$ is no longer positive definite on the kernel of J'. Indeed, $\eta = \varphi_0 + c\zeta$ is in the kernel of J'_y if $c = -\frac{\langle\varphi_0, \nabla J\rangle}{\langle\zeta, \nabla J\rangle} = -\frac{\langle\varphi_0, \nabla J\rangle}{\langle\zeta, A\zeta\rangle}$, but one can verify that $\langle\eta, A\eta\rangle < 0$.

Now consider $f(r,s) = J(y + r\eta + s\nabla J) - J(y)$, a differentiable function of r and s. Then $\nabla f(0,0) = (0, \|\nabla J\|^2)$, so the zero set of f is tangent to the r axis at the origin. From this we conclude that there is a function $s(r)$ so that $J(y + r\eta + s(r)\nabla J) - J(y) = 0$, with $\lim_{r\to 0} \frac{s(r)}{r} = 0$. From equation (1), for $h = r\eta + s(r)\nabla J$ we have

$$\Delta I = (I'' + \lambda J'')(r\eta + s(r)\nabla J) + (\lambda\epsilon_1 + \epsilon_2)\|r\eta + s(r)\nabla J\|^2$$

so that $\Delta I = r^2\langle\eta, A\eta\rangle + o(r^2)$. Thus, for all r sufficiently small $\Delta I < 0$, indicating that we do not have a constrained minimum, concluding the proof of Theorem 2. □

Theorem 3. *If $\sigma(A) \cap (-\infty, 0)$ consists of more than one point, I does not have a constrained minimum at y.*

Proof. Suppose that ν and μ are in $\sigma(A) \cap (-\infty, 0)$, with $\nu < \mu$. Let E_λ be the spectral decomposition of A, so that E_λ is not constant in any neighborhood of ν nor in any neighborhood containing μ. Take an $\epsilon > 0$ so that the two ϵ neighborhoods around ν and μ are disjoint and contained in $(-\infty, 0)$. Then $E_{\nu+\epsilon} - E_{\nu-\epsilon}$ is nonzero, i.e., is a nontrivial projection. Therefore there is some $\varphi_0 \neq 0$ so that $(E_{\nu+\epsilon} - E_{\nu-\epsilon})\,\varphi_0 = \varphi_0$. I claim that $\langle\varphi_0, A\varphi_0\rangle < 0$.

Indeed, $\langle\varphi_0, A\varphi_0\rangle = \langle\varphi_0, \int_{-\infty}^{\infty} \lambda dE_\lambda(\varphi_0)\rangle$, which is $\int_{-\infty}^{\infty} \lambda d\langle E_\lambda(\varphi_0), \varphi_0\rangle$, where the latter just a Stieljes integral. But beyond $\nu + \epsilon$, $E_\lambda(\varphi_0) = \varphi_0$, so we only get a negative contribution. It is certainly strictly negative, since for $\lambda < \nu - \epsilon$, $E_\lambda(\varphi_0) = 0$.

Now find a φ_1 for μ in the same fashion. We need to show that $\langle\varphi_0, A\varphi_1\rangle = 0$. But $\langle\varphi_0, A\varphi_1\rangle = \int_{-\infty}^{\infty} \lambda d\langle\varphi_0, E_\lambda\varphi_1\rangle$, and it is routine to show that $\langle\varphi_0, E_\lambda\varphi_1\rangle = 0$ for all λ.

We may take c_0 and c_1, not both zero, so that $c_0\varphi_0 + c_1\varphi_1$ is perpendicular to ∇J. Then $\langle c_0\varphi_0 + c_1\varphi_1, Ac_0\varphi_0 + Ac_1\varphi_1\rangle = c_0^2\langle\varphi_0, A\varphi_0\rangle + c_1^2\langle\varphi_1, A\varphi_1\rangle < 0$. The proof now proceeds as in Step 3 of Theorem 2. □

Note. It often occurs in practice that the spectrum of A is discrete and may be written as $\lambda_0 < \lambda_1 \leq \lambda_2 \leq \dots$, with 0 not a cluster point of $\sigma(A)$. In this special case, the parts of the hypotheses of the above theorems which relate to $\sigma(A)$ are as follows. In Theorem 1 we require that $0 < \lambda_0$, in Theorem 2 we require that $\lambda_0 < 0 < \lambda_1$ (in addition to the hypotheses on ζ), and in Theorem 3 we require that $\lambda_0 < \lambda_1 < 0$.

References

[1] R. Finn, *Editorial comment on Stability of a Catenoidal Liquid Bridge*, by Lianmin Zhou, to appear in Pac. J. Math.

[2] I.M. Gelfand and S.V. Fomin, *Calculus of Variations*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1963.

[3] J.H. Maddocks, *Stability and Folds*, Arch. Rat. Mech. Anal., **99** (1987), 301-328.

[4] ______, *Stability of nonlinear elastic rods*, Arch. Rat. Mech. Anal., **85** (1984), 311-354.

[5] M. Schechter, *Spectra of Partial Differential Operators*, North-Holland Publishing Co., Amsterdam, 1971.

Received November 17, 1994.

TEXAS A & M UNIVERSITY
COLLEGE STATION, TX 77840
E-mail address: tvogel@math.tamu.edu

PACIFIC JOURNAL OF MATHEMATICS

Founded in 1951 by

E. F. Beckenbach (1906-1982) F. Wolf (1904-1989)

Manuscripts must be prepared in accordance with the instructions provided on the inside back cover.

The table of contents and the abstracts of the papers in the current issue, as well as other information about the Pacific Journal of Mathematics, may be found on the Internet at http://www.math.uci.edu/pjm.html.

The *Pacific Journal of Mathematics* (ISSN 0030-8730) is published monthly except for July and August. Regular subscription rate: $245.00 a year (10 issues). Special rate: $123.00 a year to individual members of supporting institutions.

Subscriptions, back issues published within the last three years and changes of subscribers address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. Prior back issues are obtainable from Kraus Periodicals Co., Route 100, Millwood, NY 10546.

The Pacific Journal of Mathematics at the University of California, c/o Department of Mathematics, 981 Evans Hall, Berkeley, CA 94720 (ISSN 0030-8730) is published monthly except for July and August. Second-class postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Pacific Journal of Mathematics, P.O. Box 6143, Berkeley, CA 94704-0163.

PUBLISHED BY PACIFIC JOURNAL OF MATHEMATICS at University of California,
Berkeley, CA 94720, A NON-PROFIT CORPORATION
This publication was typeset using AMS-LATEX,
the American Mathematical Society's TEX macro system.

PACIFIC JOURNAL OF MATHEMATICS

Volume 176 No. 2 December 1996

0030-8730(1996)176:2;1-A