35

ON THE APPROXIMATE MINIMIZATION OF FUNCTIONALS*

by

James W. Daniel

Technical Report #42

September 1968

*Prepared under Contract Number N00014-67-A-0128-0004 at the University of Wisconsin.

1. INTRODUCTION

Many theoretical and computational problems either arise or can be formulated as one of locating a minimizing point of some real valued (nonlinear) functional over a certain set; such variational settings often lead to existence theorems as well as to computational methods for solving the problems in question. Computationally, however, one is generally forced to deal with discrete data in place of the original functional; it is therefore necessary to analyze the relationships between variational problems and their discretized analogues.

In [7, Section 4], we first studied under certain equicontinuity assumptions the question of approximately minimizing one functional by minimizing a sequence of nearby functionals. In this present note we state the problem generally, give some convergence theorems, and describe some particular examples.

2. MINIMIZATION OVER W-COMPACT SETS

Let E be a normed linear space and let f be a real valued (nonlinear functional on E. Let there be another notion of convergence (i.e., a topology) in E in addition to norm-convergence, W-continuity, W-compactness, etc. For example, if E is a reflexive Banach space, the W-topology might be the weak-topology. We wish to minimize f over a W-compact set B.

Definition 2.1 A discretization for the functional f on E consists of a family of normed linear spaces E_n, a family of real valued functionals f_n on E_n, a family of mappings p_n of E_n into E, and a family of mappings r_n of E into E_n.

Definition 2.2 A discretization for f on E is consistent if:

1) $\lim_{n \to \infty} \sup f_n(r_n \hat{u}) \leq f(\hat{u})$ if $\hat{u}$ minimizes f over the set in question,

2) $\lim_{n \to \infty} \sup [f(p_n u_n) - f_n(u_n)] \leq 0$ for any sequence $u_n \in E_n$ such that $p_n u_n$ remains in the set B over which we wish to minimize f.

Remark. Generally one would demonstrate that 1) is valid for $\hat{u}$ by proving its validity for all u in E.

We now state the basic minimization problem over B(MPB): Let f be W-lower semicontinuous and bounded below on a W-compact set B; find $\hat{u} \in B$ such that $f(\hat{u}) \leq f(u)$ for all $u \in B$.

Remark. At least one such $\hat{u}$ exists [17]. Recall that lower semicontinuity means $f(u) \leq \liminf f(u_n)$ whenever u_n W-converges to u.

We wish to find $\hat{u}$ solving MPB by solving similar problems for f_n on E_n. Thus consider the basic approximating minimization problem over B_n (MPB_n): Let f_n be bounded below on sets B_n satisfying $p_n B_n \subset B$ and $r_n B \subset B_n$; find $\hat{u}_n \in B_n$ such that $f_n(\hat{u}_n) \leq f_n(u_n) + \varepsilon_n$ for all $u_n \in B_n$, $\varepsilon_n > 0$ converging to zero.

We can now prove the following fundamental theorem on the approximate minimization of functionals.

Theorem 2.1 Let f be W-lower semicontinuous and bounded below on a W-compact set B and suppose f_n, E_n, p_n, r_n gives a consistent discretization for f on E with sets $B_n \subset E_n$ satisfying $p_n B_n \subset B$, $r_n B \subset B_n$. Then there exists $\hat{u}$ and $\hat{u}_n$ solving MPB and MPB_n. For any such solutions, $\lim_{n\to\infty} f(p_n \hat{u}_n) = \lim_{n\to\infty} f_n(\hat{u}_n) = f(\hat{u})$ and all W-limit points of $p_u \hat{u}_n$, at least one of which exists, solve MPB. If MPB has a unique solution, then $p_n \hat{u}_n$ W-converges to it.

Proof. The existence of $\hat{u}$ was noted earlier; the existence of $\hat{u}_n$ is obvious. Since $\hat{u}$ solves MPB, for all n we have $f(\hat{u}) \le f(p_n \hat{u}_n) \equiv f_n(\hat{u}_n) + \eta_n$ where $\eta_n \equiv f(p_n \hat{u}_n) - f_n(\hat{u}_n)$ satisfies $\lim_{n\to\infty} \sup \eta_n \le 0$ by part two of the consistency assumption. On the other hand, $\hat{u}_n$ solves MPB_n in the sense that $f_n(\hat{u}_n) \le \inf_{B_n} f_n(u_n) + \varepsilon_n$; therefore we have $f(\hat{u}) \le f(p_n \hat{u}_n) = f_n(\hat{u}_n) + \eta_n \le \inf_{B_n} f_n(u_n) + \varepsilon_n + \eta_n \le f_n(r_n \hat{u}) + \varepsilon_n + \eta_n$, the latter inequality resulting from the fact that $r_n \hat{u} \in B_n$. By part one of the consistency assumption $\lim_{n\to\infty} \sup f_n(r_n \hat{u}) \le f(\hat{u})$; therefore by letting n tend to infinity in the last string of inequalities we obtain $\lim_{n\to\infty} f(p_n \hat{u}_n) = \lim_{n\to\infty} f_n(\hat{u}_n) = f(\hat{u})$. Since B is W-compact and $p_n \hat{u}_n \in B$, there exists at least one W-limit point of $p_n \hat{u}_n$ in B; let u' be any such point associated with a subsequence $p_{n_j} \hat{u}_{n_j}$. The W-lower semi-continuity of f yields $f(\hat{u}) \le f(u') \le \lim_{n\to\infty} \inf f(p_{n_j} u_{n_j}) = f(\hat{u})$, so u' solves MPB; if the solution to MPB is unique, the sequence $p_n \hat{u}_n$ clearly is W-convergent to it. Q.E.D.

This theorem is a strengthening of Proposition 1, Section 4, in [7], which is concerned with collective compactness and thereby with approximate

solutions of integral equations. Somewhat similar results, in essence directed toward Hammerstein equations, were obtained in unpublished notes by J.-P. Aubin and J. L. Lions [4]; their results, other than those giving explicit conditions on the Hammerstein operators which guarantee the satisfaction of the assumptions of our theorems, are contained in Theorem 2.1. The general ideas concerning discretization schemes are those of [2, 3], in which examples of p_n and r_n are given.

In practice it is often necessary to minimize a given functional over the entire space E; under further restrictions on the discretization scheme, we can handle this case also.

3. MINIMIZATION OVER E

Consider the following global minimization problem over E (MPE): Let f be W-lower semicontinuous and bounded below on E; find $\hat{u}$ such that $f(\hat{u}) \leq f(u)$ for all $u \in E$.

Remark A solution to MPE need not exist.

We approximate this problem by the following approximating global minimization problem over E_n (MPE_n): Let f_n be bounded below on E_n. Find $\hat{u}_n$ such that $f_n(\hat{u}_n) \leq f_n(u_n) + \varepsilon_n$ for all $u_n \in E_n$, where $\varepsilon_n \geq 0$ converges to zero.

Remark MPE_n always has a solution.

We now wish to study situations in which solutions to MPE exist and can be obtained via solutions to MPE_n. The type of condition that usually

is imposed to guarantee the existence of a solution to MPE is a growth condition enabling us to restrict ourselves to bounded sets and then use results for that case. Therefore we now assume that solid spheres in E are W-compact, i.e., for all real R, $S_R \equiv \{u_j \| u \| \leq R\}$ is W-compact.

Definition 3.1 The functional f is said to satisfy a T-condition if there exists an $R > 0$ and u_0 in E with $\|u_0\| \leq R$ such that $\|u\| > R$ implies $f(u) > f(u_0)$.

Remark If there exists a real number b and a monotone function $s(t)$ such that $\lim_{n \to \infty} s(t) = \infty$ and such that $f(u) \geq b + s(\|u\|)$, then f satisfies a T-condition.

Thus, for a functional satisfying a T-condition, problem MPE can be reduced to problem MPB over the W-compact set $B \equiv S_R$, for which we know a solution exists. It is possible, however, to have a consistent discretization for a functional f satisfying a T-condition but such that the points $\hat{u}_n$ satisfy $\lim_{n \to \infty} \|p_n \hat{u}_n\| = \infty$; thus we need further conditions in order to solve MPE via MPE_n. What we need is a type of uniform growth condition on the functionals f_n and a stability condition on the discretization.

Definition 3.2 A discretization for f on E satisfies a uniform growth condition if $\lim_{n \to \infty} \sup f_n(u_n) = \infty$ whenever $\lim_{n \to \infty} \sup \|u_n\|_n = \infty$.

Definition 3.3 A discretization for f on E is stable if there exists a constant A such that $\|p_n u_n\| \leq A \|u_n\|_n$ for all $u_n \in E_n$ for all n.

Now we can prove the following theorem on solving MPE via MPE_n.

Theorem 3.1 Let solid spheres in E be W-compact. Let the W-lower semicontinuous functional f be bounded below on E and satisfy a T-condition with $R = R_0$. Let the given discretization for f on E be stable and consistent (condition 2 must hold if $\|p_n u_n\|$ is bounded) and satisfy a uniform growth condition; suppose each f_n is bounded below on E_n. Then solutions $\hat{u}$ and $\hat{u}_n$ exist solving MPE and MPE_n respectively. For any such solutions $\lim_{n\to\infty} f(p_n\hat{u}_n) = \lim_{n\to\infty} f_n(\hat{u}_n) = f(\hat{u})$, and all W-limit points of $p_n\hat{u}_n$, at least one of which exists, solve MPE. If the solution to MPE is unique, then $p_n\hat{u}_n$ W-converges to it.

Proof: By our assumptions, $\hat{u}_n$ and $\hat{u}$ exist. Since $r_n\hat{u} \in E_n$, we have $f_n(\hat{u}_n) \le f_n(r_n\hat{u})$ and $\lim_{n\to\infty} \sup f_n(r_n\hat{u}) \le f(\hat{u})$ by the consistency assumption; therefore there exists a constant C such that $f_n(\hat{u}_n) \le C$ for all u and hence, because of the uniform growth condition, there exists a constant D such that $\|\hat{u}_n\|_n \le D$. Therefore, by stability, $\|p_n\hat{u}_n\| \le A\|u_n\|_n \le AD$. Let $R = \max(R_0, AD)$. Let $B_n = \{u_n; \|u_n\|_n \le D\}$, $B = S_R = \{u; \|u\| \le R\}$. The theorem now follows by applying Theorem 2.1 with the sets B_n and B as defined above, recalling that B is W-compact, Q.E.D.

It is quite straightforward to apply our two main theorems to generate results concerning, for example, the approximate solution of nonlinear operator equations (by looking at conditions guaranteeing that $\hat{u}$ is an interior point of B or of E and deducing that the derivative of f must vanish there) and the approximate solution of nonlinear eigenvalue-eigenvector problems (by guaranteeing $\hat{u}$ to be on the boundary of B for certain types of sets B

and applying the extended Lagrange multiplier theorem). Results of this type are contained in Sections 4 and 5 of [7]; we pursue this no further here. Instead we look briefly at a number of methods currently in use for solving certain types of problems to see how they fit into the above theory and how the theory indicates the necessary characteristics of the particular methods.

4. EXAMPLES

I. Rayleigh-Ritz.

Perhaps the simplest and oldest example is that of the Rayleigh-Ritz procedure. Here we let E be a Hilbert space with complete basis $\varphi_1, \cdots \varphi_n, \cdots$ and for each n let E_n be the subspace of E spanned by $\varphi_1, \cdots \varphi_n$, with $p_n u_n = u_n$ considered as an element of E, and $r_n u = \sum_{i=1}^{n} \langle u, \varphi_i \rangle \varphi_i$ where $\langle u, v \rangle$ is the inner product on E; the inner product on E_n is the one induced by $\langle \cdot, \cdot \rangle$ on E. Given a functional f, let $f_n = f$ for all n. Such a discretization is stable and, if f is norm continuous, it is consistent. A great deal of interest has arisen lately in the application of this method to the numerical solution of differential equations using various bases $\varphi_1, \cdots$, that is, different finite dimensional subspaces E_n. In particular, in [5] it is shown how very good error bounds can be computed; essentially this approach assumes that f has a differentiable gradient $J = \nabla f$ with uniformly positive definite derivative J'. If $\hat{u}(\hat{u}_n)$ minimizes $f(f_n)$ on $E(E_n)$ then $J(\hat{u}) = 0$ and $\langle J(\hat{u}_n), \varphi_i \rangle = 0$, $i = 1, \cdots n$; hence

$0 = \langle J(\hat{u}) - J(\hat{u}_n), \varphi_i \rangle = \langle J'(\hat{u}_n + \vartheta(\hat{u} - \hat{u}_n))(\hat{u} - \hat{u}_n), \varphi_i \rangle$ which states that $\hat{u}_n$ is the projection under the inner product $[u, v] \equiv \langle J'(\hat{u}_n + \vartheta(\hat{u} - \hat{u}_n))u, v \rangle$ of $\hat{u}$ onto E_n. Thus $[\hat{u} - \hat{u}_n, \hat{u} - \hat{u}_n] \leq [\hat{u} - u_n, \hat{u} - u_n]$ for all $u_n \in E_n$ which provides an error estimate. Details and extensions of this approach are contained in [5, 6].

One difficulty not mentioned in the above papers is that of minimizing f_n over E_n; in this case f_n is an integral of a complicated differential form, often nonlinear. In practice this is discretized by a quadrature sum so that we do not in fact have $f_n = f$. The difficulties therein created are treated in [9].

II. Regularization.

Often one seeks not just any arbitrary minimizing point to a functional but one which, in some sense, is smoothest or most regular. For example, let f and g be weakly lower semicontinuous functionals on a reflexive Banach space E, and let g be nonnegative. Let $E_n = E$, $p_n = r_n =$ the identity, $f_n = f + a_n g$ where $a_n > 0$ converges to zero. Let B be a weakly compact set in E with f bounded below on B. The consistency condition 1) becomes $\lim_{n \to \infty} \sup |f(\hat{u}) + a_n g(\hat{u})| \leq f(u)$ which clearly is satisfied. Condition 2) becomes $\lim_{n \to \infty} \sup -a_n g(u_n) \leq 0$ which follows from the nonnegativity of a_n and g. For future reference we note that $f(\hat{u}_n) + a_n g(\hat{u}_n) \leq f(\hat{u}) + a_n g(\hat{u}) \leq f(\hat{u}_n) + a_n g(\hat{u})$, implying $g(\hat{u}_n) \leq g(\hat{u})$. It follows then from Theorem 2.1 that any weak limit point u' of $\{u_n\}$ minimizes f over B; moreover, since

$g(u') \leq \lim_{n \to \infty} \inf\ g(u_n) \leq g(u)$, $g(u')$ is the minimum value of g over the set of minimizing points of f in B. One says that the minimizing point has been regularized by g[11, 16]; if B is convex and g is uniformly convex, then u' is in fact a norm limit point and moreover if f is quasi-convex [12], the entire sequence $\{u_n\}$ converges in norm to u' [11].

III. Chebyshev Solution of Differential Equations.

Suppose one seeks to solve $Au = b$ where A is a uniformly elliptic linear differential operator in two variables over a bounded domain D, under the condition $u = 0$ on Γ, the boundary of D, assumed to be sufficiently smooth; more general types of equations may also be treated by the method to be presented. A numerical method of recent popularity [10, 14] given a sequence of functions $\{\varphi_i\}$ satisfying the boundary data, consists in choosing numbers $a_{n,1}, \cdots a_{n,n}$ to minimize $\max_{1 \leq j \leq M} |[A(\sum_{i=1}^{n} a_{n,i}\, \varphi_i)](x_j) - b(x_j)|$ where the M points $\{x_j\}$ form a "grid" over D. Strictly for convenience we take $M = 4n$ and suppose that the grid is such that any point in D is at a distance of at most h_n from a grid point x_j. We wish to find conditions under which $\hat{u}_n \equiv \sum_{i=1}^{n} a_{n,i}\, \varphi_i$ will converge, in some sense, to the solution $\hat{u}$ of our problem.

Since we seek to minimize a supremum norm, the norm must be defined, so let $E = \{u;\ u = 0$ on Γ, all partial derivatives of u through second order are continuous on $\bar{D} = D \cup \Gamma\}$; for $u \in E$, let $\|u\| = \|u\|_\infty = \max_{\bar{D}} |u(x)|$. Let $f(u) = \|Au - b\|_\infty$, where we now need to assume that b

is bounded on $\bar{D}$. Let E_n be that subset of E spanned by the functions $\varphi_1, \ldots \varphi_n$, assumed to lie in E; let p_n be the identity mapping, and r_n be at the moment undefined. Define $f_n(u_n) = \|Au_n - b\|_{4n,\infty} = \max_{1 \le i \le 4n} |[Au_n](x_i) - b(x_i)|$; we now seek conditions for consistency. Consider 2) of our Definition 2.2. $f(p_n u_n) - f_n(u_n) = \|Au_n - b\|_\infty - \|Au_n - b\|_{4n,\infty}$. Since this quantity is always non-negative, the requirement $\lim_{n \to \infty} \sup [f(p_n u_n) - f_n(u_n)] \le 0$ in fact demands convergence; in order to conpare suprema over discrete and continuous sets, we need to know something about the growth of the functions $Au_n - b$ between grid points. Hence we now need to assume that $A\varphi_i$ satisfies a Lipschitz condition with Lipschitz constant λ_i (this restricts A somewhat also) and that b satisfies one with a constant λ_0. From this it follows that $|f(p_n u_n) - f_n(u_n)| \le h_n \sum_{i=0}^{n} |a_{n,i}| \, |\lambda_i|$, where $a_{n,0} = 1$. Thus we need next a growth condition on $\sum_{i=0}^{n} |a_{n,i}| \, |\lambda_i|$. For example, the conditions that i) there exists a constant C such that $\sum_{i=0}^{n} |a_{n,i}| \le C$ for all n, and ii) $h_n \Lambda_n$ tends to zero, where $\Lambda_n = \max_{1 \le i \le n} \lambda_i$, would be sufficient; in practice the Λ_n do in fact become large, while the restriction on the $a_{n,i}$ is easy to implement. In essence, the above restrictions are defining B_n, i.e., $B_n = \{u_n; \sum_{i=0}^{n} |a_{n,i}| \le C\}$.

Next consider condition 1) of Definition 2.2, where r_n is to be defined. We require $\lim_{n \to \infty} \sup f_n(r_n u) \le f(u)$. Now $f_n(r_n u) \le f(r_n u)$, so we need only require that $\lim_{n \to \infty} \sup f(r_n u) \le f(u)$; this is certainly true if $r_n u$ is an approximation method in which $Ar_n u$ converges uniformly to Au, for

example, if $r_n u$ and all its partial derivatives through second order converge uniformly to those for u. We remark that it is necessary to have $r_n u$ in B_n.

Under the above conditions, it follows from Theorem 2.1 and its proof that $\lim_{n\to\infty} f_n(\hat{u}_n) = \lim_{n\to\infty} f(p_n \hat{u}_n) = f(\hat{u}) = 0$, where $\hat{u}$ solves $A\hat{u} = b$ and lies in E; the conditions on W-compactness and W-lower semicontinuity are needed only to prove convergence for $p_n u_n$, a problem easily handled differently here. We know that $\|A\hat{u}_n - b\|_\infty = f(p_n u_n)$ converges to zero. By a simple use of the maximum principle [13], we deduce $\|u_n - \hat{u}\|_\infty \leq \|A\hat{u}_n - b\|_\infty \|w\|_\infty$ where w solves $Aw = -1$ in D, $w = 0$ on Γ; therefore $\hat{u}_n$ converges uniformly to the solution $\hat{u}$.

The application of the theory in Section 2 to this problem indicates the type of approach necessary to prove convergence for this numerical method. One requires: 1) smooth functions φ_i with Lipschitz constants λ_i for $A\varphi_i$ that do not grow too rapidly, 2) results from approximation theory that state that <u>if</u> one approximates functions b by combinations of functions $A\varphi_i$, the sums $\sum_{i=1}^{n} |a_{n,i}|$ reamin bounded, and 3) functions b <u>can</u> be approximated by functions $A\varphi_i$. The requirements 1) and 3) here are probably less difficult; generalized Bessel inequality results like 2), however, are not known to this author for general cases. While numerical work with this method proceeds, theoretical results of the type suggested by our theorem should be sought.

IV. Calculus of Variations

Consider the problem of minimizing the functional $f(u) = \int_0^1 g(t, u, \dot{u})dt$, $u(0) = u(1) = 0$, where $\dot{u} = \frac{du}{dt}$. The following simple case of a general numerical method has been suggested [8]: minimize $f_n(u_n) = \sum_{i=1}^{n} h_i\, g(t_{i-1}, u_{n,i-1}, \frac{u_{n,i} - u_{n,i-1}}{h_i})$, $u_{n,0} = u_{n,n} = 0$, $h_i = t_i - t_{i-1}$, where the minimization is over the set of values of $u_{n,1}, \cdots u_{n,n-1}$; this method can be fit neatly into the theory of Theorem 2.1. In [8], under the assumption that there exist unique minimizing points for f (in $C'[0,1]$) and f_n satisfying the "spike" condition $|\frac{\hat{u}_{n,i} - \hat{u}_{n,i-1}}{h_i}| \le A$ for some constant A independent of n, it was purportedly proved that $p_n \hat{u}_n$, the piecewise linear interpolation to $\hat{u}_n$, converges uniformly to $\hat{u}$; because the authors inadvertently left out an assumption guaranteeing a lower semi-continuity property for the functional f, the proof is in fact incorrect. However, as we shall show below by use of Theorem 2.1, the usual assumptions guaranteeing a unique minimizing point for f, in conjunction with an assumption guaranteeing the satisfaction of a type of spike condition, yield a convergence proof.

For convenience let us take $h_i = h = \frac{1}{n}$ for all i. For a fixed $p > 1$, let $E = \{u; u(0) = u(1) = 0$, u is absolutely continuous on $[0,1]$, $\dot{u} \in L_p[0,1]\}$; for $u \in E$, let $\|u\| = \|\dot{u}\|_p = \{\int_0^1 |\dot{u}(t)|^p dt\}^{1/p}$. For each n, let E_n be $(n-1)$-dimensional Euclidean space, where $u_n \in E_n$ has the norm $\|u_n\|_n = \{h \sum_{i=1}^{n} [|u_{n,i} - u_{n,i-1}|]^p\}^{1/p}$, where $u_{n,0} = u_{n,n} = 0$ by

definition. Let p_n be the mapping defined by piecewise linear joining of the values $u_{n,i}$ at the points $t_i = ih$, thus $p_n u_n \in E$. Define the mapping r_n via $(r_n u)_i = u(t_i)$, $i = 1, \cdots n-1$. Define W-convergence as follows: a sequence $u^{(n)}$ in E W-converges to u in E if $\int_0^1 (\dot{u}^{(n)} - \dot{u})v \, dt$ converges to zero for every function $v(t) \in L_q(0, 1)$, $\frac{1}{p} + \frac{1}{q} = 1$.

Proposition 4.1 Solid spheres in E are W-compact. If $u^{(n)}$ W-converges to u, then $u^{(n)}$ converges to u uniformly, i.e., in the supremum norm.

Proof: If we take any sequence $u^{(n)}$ in E with $\|u^{(n)}\| = \|\dot{u}^{(n)}\|_p \leq R$, then, by Alaoglu's Theorem [15, p. 228], there exists a subsequence $u^{(n_j)}$ and a function s in L_p such that for all v in L_q, $\int_0^t (\dot{u}^{(n_j)} - s)v \, dt$ tends to zero. Let $u(t) = \int_0^t s(t) \, dt$; it follows that $u \in E$ and $u^{(n_j)}$ W-converges to u. For the second part of the proof, we note first that $\|u^{(n)}\|$ is bounded, by the uniform boundedness principle [15, p. 202]. Thus $|u^{(n)}(t_1) - u^{(n)}(t_2)| \leq \int_{t_1}^{t_2} |\dot{u}^{(n)}| \, dt \leq [\int_{t_1}^{t_2} |\dot{u}^{(n)}|^p \, dt]^{1/p} |t_2 - t_1|^{1/q} \leq \|u^{(n)}\| \, |t_2 - t_1|^{1/q}$, which implies that the sequence $u^{(n)}$ is uniformly bounded and equicontinuous. If $u^{(n)}$ does not converge uniformly to u, then there exists $\varepsilon > 0$ and a sequence t_j such that $|u^{(n_j)}(t_j) - u(t_j)| \geq \varepsilon$; applying Ascoli's theorem [15, p. 276] to the bounded, equicontinuous sequence $u^{(n_j)}$ yields a contradiction. Q.E.D.

We now make the standard type of assumption in the calculus of variations in order to guarantee the existence of a minimizing point for f(u).

Assumptions: i) $g(t, u, w)$ is jointly continuous in its variables for $0 \le t \le 1$ and $-\infty \le u, w \le \infty$. ii) There exist constants a, b with $b > 0$ such that $g(t, u, w) \ge a + b\,|w|^p$ for all t in $[0,1]$, u finite. iii) g is differentiably convex in w, i.e., $g(t, u, w_1) - g(t, u, w_2) \ge (w_1 - w_2)\, g_w(t, u, w_2)$.

Proposition 4.2 The functional f is W-lower semicontinuous on E, bounded below, and satisfies a T-condition.

Proof: For the last two statements, note that $f(u) = \int_0^1 g(t, u, \dot{u})\, dt \ge \int_0^1 [a + b\,|\dot{u}|^p]\, dt = a + b\,\|u\|^p$. The proof of the W-lower semicontinuity is straightforward using the convexity of g; details may be found in [1, p. 137-139]. Q.E.D.

Proposition 4.3 The discretization scheme defined above is stable and satisfies a uniform growth condition.

Proof: $\|p_n u_n\|^p = \int_0^1 |(p_n u_n)^{\cdot}|^p\, dt = \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} |(p_n u_n)^{\cdot}|^p\, dt = h \sum_{i=1}^{n} \left[\frac{|u_{n,i} - u_{n,i-1}|}{h}\right]^p = \|u_n\|_n^p$. $f_n(u_n) = h \sum_{i=1}^{n} g(t_i, u_{n,i}, \frac{u_{n,i} - u_{n,i-1}}{h})$ $\ge h \sum_{i=1}^{n} \left[a + b\left|\frac{u_{n,i} - u_{n,i-1}}{h}\right|^p\right] = a + b\,\|u_n\|_n^p$. Q.E.D.

The only remaining ingredient for application of Theorem 3.1 is the consistency; in [8], the spike condition was needed for this. In our case, we must make the following assumptions.

Assumptions iv) Some solution $\hat{u}$ minimizing $f(u)$ lies in $C^1[0,1]$, i.e., $\hat{u}^{\cdot}$ is continuous, v) There exist constants c and d and a continuous function $s(t, v)$ such that $|g(t_1, v_1, z) - g(t_2, v_2, z)| \le (c + d\,|z|^p)$

$|s(t_1, v_1) - s(t_2, v_2)|$ where t_1, t_2 are arbitrary points in $[0, 1]$ and v_1, v_2, z are arbitrary real numbers.

Remark. If $g(t, u, w) \equiv \frac{w^2}{2} + r(t, u)$, then (v) is satisfied with $s \equiv r$. If $g(t, u, w) \equiv \ell(w)\, m(t, u)$ with $|\ell(w)| \le c + d\,|w|^p$, then (v) is satisfied with $r \equiv m$; many actual problems are of the above types.

Proposition 4.4 The discretization is consistent.

Proof: For condition 1), we prove $\lim_{n \to \infty} |f_n(r_n \hat{u}) - f(\hat{u})| = 0$. Since, by assumption, $\hat{u}$ is in $C^1[0, 1]$, given ε, for sufficiently large n, $|\hat{u}(t_{i-1}) - \hat{u}(t)| < \varepsilon$ and $|\hat{u}'(t) - \frac{\hat{u}_i - \hat{u}_{i-1}}{h}| < \varepsilon$ for $t_{i-1} \le t \le t_i$. For convenience, we write merely u for $\hat{u}$. $|f(u) - f_n(r_n u)| \le \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} |g(t, u, \dot{u}) - g(t_{i-1}, u_{i-1}, \frac{u_i - u_{i-1}}{h})|\, dt$. But, by uniform continuity of g, given $\delta > 0$ there exists $\varepsilon > 0$ and then N such that $n > N$ implies $|f(u) - f_n(r_n u)| \le \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \delta\, dt = \delta$. Since $\delta > 0$ was arbitrary, condition 1) is proved. For condition 2), we show that $\lim_{n \to \infty} |f(p_n u_n) - f_n(u_n)| = 0$ if $\|p_n u_n\|$ is bounded. $|f(p_n u_n) - f_n(u_n)| \le h \sum_{i=1}^{n} \int_0^1 |g(t_{i-1} + \alpha h, (1-\alpha)\, u_{n,i-1} + \alpha u_{n,i}, \frac{u_{n,i} - u_{n,i-1}}{h}) - g(t_{i-1}, u_{n,i-1}, \frac{u_{n,i} - u_{n,i-1}}{h})|\, d\alpha \le h \sum_{i=1}^{n} \int_0^1 (c + d\,|\frac{u_{n,i} - u_{n,i-1}}{h}|^p)\, |s(t_{i-1} + \alpha h, (1-\alpha)\, u_{n,i-1} + u_{n,i}) - s(t_{i-1}, u_{n,i-1})|\, d\alpha$. Now $\|u_n\|_n = \|p_n u_n\|$ is bounded, $|u_{n,i}|$ is bounded, and $|u_{n,i} - u_{n,i-1}| \le h^{1-1/p}\, \|u_n\|_n$. Thus, using the uniform continuity of $s(t, u)$, given $\varepsilon > 0$ there exists N such that $n > N$ implies

$$|f(p_n u_n) - f_n(u_n)| \leq h \sum_{i=1}^{n} \int_0^1 (c + d|\frac{u_{n,i} - u_{n,i-1}}{h}|^p) \varepsilon \, d\alpha = \varepsilon[c + d\|u_n\|_n^p];$$

since $\varepsilon > 0$ was arbitrary, condition 2) follows. Q.E.D.

We now can state the following theorem which follows immediately from Theorem 3.1 and the above propositions.

Theorem 4.1 Let Assumptions (i) - (v) be valid and let the discretization method described above be used. Then all W-limit points of $p_n \hat{u}_n$ (at least one of which exists) minimize f. If the solution $\hat{u}$ is unique, then in particular $p_n \hat{u}_n$ converges uniformly to $\hat{u}$ and the derivatives converge L_p-weakly.

REFERENCES

1. Akhiezer, N. I., The calculus of variations, Blaisdell, New York (1962).

2. Aubin, J.-P., "Approximation des espaces de distribution et des operateurs differentiels," Bull. Soc. Math. France Memoire No. 12, (1967).

3. _____, "Evaluation des erreurs de trancature des approximations des espaces de Sobolev," J. Math. Anal. Appl. Vol 21 (1968), 356-368.

4. _____, Lions, J. L., unpublished notes on minimization.

5. Ciarlet, P. G., Schultz, M. H., Varga, R. S., "Numerical methods of high-order accuracy for nonlinear boundary value problems I. One dimensional problem," Num. Mathematik, Vol. 9 (1967), 394-430.

6. _____, "Numerical methods of high order accuracy for nonlinear boundary value problems. Monotone operators," to appear.

7. Daniel, J. W., "Collectively compact sets of gradient mappings," Koninkl. Nederl. Akad. Weten., Proceedings Vol. 71, (1968), 270-279.

8. Greenspan, D., "On approximating extremals of functionals, II," Int. J. Engng. Sci. Vol. 5, (1967), 571-588.

9. Herbold, R. J., "Constant quadrature schemes for the numerical solution of boundary value problems by variational techniques," Ph.D. Thesis, Case Western Reserve Univ., Cleveland, Ohio, 1968.

10. Krabs, W., "Einige Methoden zur Lösung des diskreten lineareu Tschebyscheff-Problems," Dissertation, University of Hamburg (1963).

11. Levitin, E. S., Polyak, B. T., "Convergence of minimizing sequences in conditional extremum problems," Soviet Math Sokl., Vol 7 (1966), 764-767.

12. Mangasarian, O. L., "Pseudo-convex functions," J. SIAM Control, Vol. 3 (1965), 281-290.

13. Protter, M. H., Weinberger, H. F., Maximum principles in differential equations, Prentice Hall, (1967), 61.

14. Rosen, J. B., "Approximate solution and error bounds for quasi-linear elliptic boundary value problems," Univ. of Wisconsin, Computer Sciences Technical Report No. 30 (1968).

15. Taylor, A. E., Introduction to functional analysis, Wiley, New York, (1958).

16. Tihonov, A. N., "Methods for the regularization of optimal control problems," Soviet Math. Dokl., Vol. 6 (1965), 761-763.

17. Vainberg, M. M., Variational methods for the study of nonlinear operators, Holden-Day, San Francisco (1964), p. 78.