

## MULTICONSTRAINED VARIATIONAL PROBLEMS OF NONLINEAR EIGENVALUE TYPE: NEW FORMULATIONS AND ALGORITHMS

ALEXANDER EYDELAND, JOEL SPRUCK, AND BRUCE TURKINGTON

**ABSTRACT.** A new variational approach is proposed for a class of semilinear elliptic eigenvalue problems involving many eigenvalue parameters. These problems arise, for instance, in the modelling of magnetohydrodynamic equilibria with one spatial symmetry. In this case, the physical variational principle imposes a continuously infinite family of constraints, which prescribes the mass and helicity within every flux tube. The equilibrium equations therefore contain unspecified profile functions that are determined along with the solution as multipliers for those constraints. A prototype problem for this general class is formulated, and a natural discretization of its constraint family is introduced. The resulting multiconstrained minimization problem is solved by an iterative algorithm, which is based on relaxation of the given nonlinear equality constraints to linearized inequalities at each iteration. By appealing to convexity properties, the monotonicity and global convergence of this algorithm is proved. The explicit construction of the iterative sequence is obtained by a dual variational characterization.

### 1. INTRODUCTION

Many equilibrium problems in mathematical physics can be formulated as nonlinear elliptic eigenvalue problems. A mathematical prototype for this wide class of problems is expressible as

$$(1.1) \quad -\Delta u = \Lambda(u) \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial\Omega,$$

where  $u = u(x)$  is a real-valued function defined in a domain  $\Omega \subseteq \mathbb{R}^N$ . The term  $\Lambda(u)$ , which we shall call the profile function, is traditionally assumed to have the form  $\Lambda(u) = \lambda f'(u)$ , where  $f'(u)$  is a specified function and  $\lambda$  is an eigenvalue parameter. Then, well-understood analytical and numerical methods apply to (1.1). In particular, solution pairs  $(u, \lambda)$  can be constructed by solving

---

Received November 8, 1988; revised June 8, 1989.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 49H05, 49G05, 76W05.

*Key words and phrases.* Semilinear elliptic eigenvalue problems, constrained minimization, iterative procedure, magnetohydrodynamics.

The research described in this paper was supported by the National Science Foundation under grants DMS-8602316, DMS-8501952, DMS-8601795, and by the U.S. Department of Energy under grant DE-F602-86ER250125 of the Applied Mathematical Science Subprogram of the Office of Energy Research.

the constrained minimization problem

$$(1.2) \quad E(u) := \frac{1}{2} \int_{\Omega} |\nabla u|^2 dx \rightarrow \min \quad \text{over} \quad \int_{\Omega} f(u) dx = \gamma$$

for various constraint values  $\gamma$ . The existence of a family of solutions parameterized by  $\gamma$  can be proved, using standard variational techniques, whenever (say)  $f(0) = f'(0) = 0$ ,  $f$  is convex, and  $f$  satisfies some growth conditions as  $|u| \rightarrow \infty$ . Moreover, effective numerical methods are available for computing these solutions (see, for instance, Glowinski [11]).

In certain physical problems, however, it is not natural to prescribe the form of the profile functions that occur in the equilibrium equations. Rather, it is necessary to characterize an equilibrium solution as an energy minimizer subject to a possibly infinite *family* of constraints derived from the conserved quantities associated with the governing evolution equations. The prescribed data in such a formulation consists of the family of constraint values, while the profile functions are determined implicitly by the solution. This sort of situation is met, for example, in ideal magnetohydrodynamics (MHD), where the natural constraints are defined by mass and helicity integrals over (nested) families of flux tubes. In particular, magnetostatic equilibrium problems with one spatial symmetry have been shown to admit a formulation as multiconstrained minimization problems by Woltjer [21, 22] and Kruskal and Kulsrud [15]. Also, related variational principles for steady flows of an ideal fluid have been formulated by Arnold [1, 2] and Benjamin [4]. These classical variational principles provide the main physical motivation for the present work, which introduces and justifies an algorithm for solving mathematical model problems of this type.

Alternative approaches to these MHD equilibrium problems have been proposed in the context of plasma confinement. Grad [12], who has emphasized the importance of posing these problems in terms of the quantities conserved by the ideal MHD equations, has introduced his so-called generalized (or queer) differential equations as the appropriate equilibrium equations. A computational technique for solving these equations has been devised by Grad et al. [13], and some related analytical studies have been made by Temam [20], Mossino and Temam [18], and Laurence and Stredulinsky [16]. All of these investigations confront the central difficulty (and novelty) inherent in generalized differential equations—namely, their highly implicit definition involving differentiation of the distribution function of the solution. A different approach has been developed by Bauer, Betancourt, and Garabedian [5], in which the family of constraints is eliminated by reformulating the standard equilibrium equations in a special coordinate system derived from the (unknown) flux functions. This leads to an unconstrained minimization problem which can be treated computationally even without a symmetry assumption, but which involves an objective (energy) functional with a rather complicated analytical structure. A similar approach using flux coordinates has been applied to astrophysical problems by Mouschovias [17].

In contrast to these works, our approach is based on a direct treatment of the nonlinear constrained optimization problems resulting from the classical variational principles for equilibrium MHD. As a prototype for this class of problems, we adopt the variational problem  $(P_\infty)$

$$(1.3) \quad E(u) \rightarrow \min \quad \text{over} \quad \int_{\Omega} (u - \sigma)_+ dx = \beta(\sigma), \quad 0 \leq \sigma < +\infty,$$

where  $\beta(\sigma)$  is suitable prescribed data. A formal calculation suggests that the variational equation satisfied by a solution  $u$  of (1.3) has the form (1.1) with a profile function (or distribution) determined by the usual (but formal) Lagrange multiplier rule. In order to justify such a calculation, we find it necessary to replace the infinite family of constraints in  $(P_\infty)$  by an appropriate finite family that approximates it. Thus, we introduce the variational problem  $(P_n)$

$$(1.4) \quad E(u) \rightarrow \min \quad \text{over} \quad F_i(u) := \int_{\Omega} f_i(u) dx = \gamma_i \quad (i = 1, \dots, n),$$

where the functions  $f_i(s)$  and the constraint values  $\gamma_i$  are chosen relative to a given partition of the interval  $0 \leq \sigma < +\infty$  into  $n$  subintervals  $0 = \sigma_0 < \sigma_1 < \dots < \sigma_{n-1} < \sigma_n = +\infty$  according to

$$(1.5) \quad f_i(s) := \int_{\sigma_{i-1}}^{\sigma_i} (s - \sigma)_+ d\sigma, \quad \gamma_i := \int_{\sigma_{i-1}}^{\sigma_i} \beta(\sigma) d\sigma.$$

If the prescribed data satisfy  $\gamma_i > 0$  for every  $i$ , then a nonnegative solution  $u \in H_0^1(\Omega)$  of  $(P_n)$  exists, is smooth, and satisfies (1.1) with

$$(1.6) \quad \Lambda(u) = \sum_{i=1}^n \lambda_i f'_i(u),$$

where  $\lambda_i \in \mathbb{R}$  are precisely the Lagrange multipliers (Theorem 2.2).

It should be emphasized that the finitely-constrained problem  $(P_n)$  yields exact solutions of the relevant equations (say the ideal MHD equations) and is an approximation only in the sense that the family of constraints in (1.3) is discretized. In this discretization, the functions  $f'_i(u)$  ( $i = 1, \dots, n$ ) form a basis for the space of all profile functions  $\Lambda(u)$  which, according to (1.6), are piecewise linear relative to the given partition. This particular choice of basis functions plays a key role in our formulation of  $(P_n)$ .

Besides being analytically tractable, the problem  $(P_n)$  is amenable to a numerical method of solution, the development of which is the principal goal of this paper. To this end, we construct an iterative procedure which is designed to converge to the solutions of  $(P_n)$ . This procedure, which we term the algorithm  $(A_n)$ , may be stated succinctly as follows: the iterate  $u^{k+1}$  is found from the previous iterate  $u^k$  by solving the convex optimization problem

$$(1.7) \quad \begin{aligned} E_\tau(u) := E(u) + \frac{\tau}{2} \langle u, u \rangle &\rightarrow \min \\ \text{over } F_i(u^k) + \langle F'_i(u^k), u - u^k \rangle &\geq \gamma_i, \quad i = 1, \dots, n. \end{aligned}$$

Thus, each iterative step is defined by solving a variational problem in which the objective functional is modified by a quadratic term involving a positive constant  $\tau$ , and the  $n$  constraints are replaced by  $n$  linearized inequalities. As a consequence, each such step admits a simple and efficient numerical implementation.

Perhaps our main result (Theorem 4.1) is the justification of this algorithm. This result may be expressed as follows: given any initialization  $u^0 \in H_0^1(\Omega)$  satisfying the constraints for  $(P_n)$ , there exists a (sufficiently large) constant  $\tau$  depending on  $E(u^0)$  and  $\gamma_i, \sigma_i$  so that the sequence  $\{u^k\}$  generated by the algorithm  $(A_n)$  converges strongly in  $H_0^1(\Omega)$  to the set of critical points of the variational problem  $(P_n)$ , in the sense that the (minimum) distance from  $u^k$  to that set tends to zero as  $k$  tends to infinity. This generalized convergence theorem holds even in the case when solutions of  $(P_n)$  are not unique.

In a forthcoming paper [8], we will discuss the implementation of the algorithm  $(A_n)$  and the results of numerical experiments. We will also extend the algorithm so as to apply to the more general problems discussed in Appendix 1 (§6). Suffice it to say here that for the model problem  $(P_n)$  the entire iterative sequence  $\{u^k\}$  is observed (at least typically) to converge with a linear rate to a unique solution.

We now give the organization of the paper. In §2, we formulate the prototype finite constraint problem  $(P_n)$  and establish its relation to the problem  $(P_\infty)$ . Section 2 also contains the proof of Theorem 2.2, which illustrates in a very clear way some of the fundamental ideas of classical optimization theory; specifically, Lemma 2.3 and its relation to the linear independence of the constraints is of independent interest. In §3, we describe the algorithm  $(A_n)$  for solving  $(P_n)$  and derive its basic convergence properties. In addition, we give a more concrete description of the algorithm in terms of the so-called dual variational problem. Section 4 completes the proof of the convergence of the algorithm (Theorem 4.1). The key step in the proof of convergence is an a priori estimate on Lagrange multipliers (Lemma 4.3) which is of independent interest. In §5, we briefly indicate how to extend the development of §§2–4 to include a free boundary interface, to allow for distributions which change sign, and finally to apply to a more general class of variational problems. Appendix 1 contains a brief discussion of the fundamental variational problems in MHD equilibrium theory which partly motivate our prototype problems. In future work we plan to elaborate on the methods and ideas of the present paper in order to study the more general problems posed in Appendix 1.

The following conventional notations will be used in the sequel. Let  $\Omega \subseteq \mathbb{R}^N$  ( $N \geq 2$ ) be a bounded open set whose boundary,  $\partial\Omega$ , is smooth enough (say  $C^{2,\theta}$ ). We denote the Lebesgue measure on  $\Omega$  by  $dx = dx_1 \cdots dx_N$ ,  $x = (x_1, \dots, x_N) \in \Omega$ , and for any measurable subset  $A \subseteq \Omega$  we write  $|A| = \int_A dx$  for its measure. The inner product and norm on  $L^2(\Omega)$  are denoted by

$$\langle u, v \rangle = \int_{\Omega} u(x)v(x) dx, \quad \|u\| = \langle u, u \rangle^{1/2}.$$

The Sobolev space  $H^m(\Omega)$  ( $m = 1, 2, \dots$ ) is identified with the space of those functions in  $L^2(\Omega)$  whose (weak) partial derivatives of order  $\leq m$  are in  $L^2(\Omega)$ . The space  $H_0^1(\Omega)$  consists of those functions which vanish in the weak sense on  $\partial\Omega$ , and  $H^{-1}(\Omega)$  is its dual with respect to the pairing  $\langle \cdot, \cdot \rangle$ . The class of continuous functions on  $\bar{\Omega} = \Omega \cup \partial\Omega$  is denoted by  $C^0(\bar{\Omega})$ , and  $C^m(\bar{\Omega})$  ( $m = 1, 2, \dots$ ) denotes the class of  $m$  times continuously differentiable functions. Also,  $C^{m,\theta}(\bar{\Omega})$  ( $0 < \theta \leq 1$ ) is the subclass of  $C^m(\bar{\Omega})$  consisting of those functions whose derivatives of order  $m$  are Hölder continuous with exponent  $\theta$ . Throughout the paper, we make use of the function  $\phi(u) = (u - \sigma)_+$ , where  $s_+ = \max(s, 0)$  for  $s \in \mathbb{R}$ .

## 2. FORMULATION OF THE PROBLEMS $(P_n)$ AND $(P_\infty)$

We begin by posing the problem  $(P_\infty)$ , which we adopt as a model for the type of variational problems presented in Appendix 1, where we refer the reader for an indication of the source of these problems in mathematical physics. For a given function  $\bar{u} \in H_0^1(\Omega)$ , we consider the class of functions which are equimeasurable with  $\bar{u}$ , namely,

$$(2.1) \quad M_\infty(\bar{u}) = \{u \in H_0^1(\Omega) : |\{u > \sigma\}| = |\{\bar{u} > \sigma\}| \text{ for all } \sigma \in \mathbb{R}\}.$$

The class  $M_\infty(\bar{u})$  consists of all functions  $u$  in  $H_0^1(\Omega)$  which are rearrangements of  $\bar{u}$  with respect to Lebesgue measure  $dx$  on  $\Omega$ . This characterization can be phrased in terms of the standard rearrangement mapping [18]  $u \rightarrow u_*$  defined by

$$(2.2) \quad u_*(\alpha) := \inf\{\sigma \in \mathbb{R} : |\{u > \sigma\}| > \alpha\} \quad (0 \leq \alpha \leq |\Omega|).$$

Then we see that

$$(2.3) \quad M_\infty(\bar{u}) = \{u \in H_0^1(\Omega) : u_*(\alpha) = \bar{u}_*(\alpha) \text{ for all } 0 \leq \alpha \leq |\Omega|\}.$$

For the purposes of the present study, however, we prefer to define the class  $M_\infty(\bar{u})$  in terms of a certain infinite family of integral constraints.

**Lemma 2.1.** *The following statements are equivalent:*

- (a)  $u \in M_\infty(\bar{u})$ ,
- (b)  $\int_{\Omega} \phi(u) dx = \int_{\Omega} \phi(\bar{u}) dx$  for all  $0 \leq \phi \in C^0(\mathbb{R})$ ,
- (c)  $\int_{\Omega} (u - \sigma)_+ dx = \int_{\Omega} (\bar{u} - \sigma)_+ dx$  for all  $\sigma \in \mathbb{R}$ .

*Proof.* That (a) implies (b) is standard; (a) is equivalent to (b) for functions of the form

$$\phi(s) = \begin{cases} 0, & s \leq \sigma, \\ 1, & s > \sigma, \end{cases}$$

and so (b) holds for any function  $\phi(s)$  which is an increasing limit of step functions (in fact, for any nonnegative Borel measurable function  $\phi(s)$ ). That (b) implies (c) is trivial. That (c) implies (a) follows from differentiation of (c) with respect to the variable  $\sigma$ . Precisely, the right derivative with respect to  $\sigma$  yields the desired result, since

$$\tau^{-1}[(s - \sigma - \tau)_+ - (s - \sigma)_+] \uparrow \begin{cases} 0, & s \leq \sigma, \\ 1, & s > \sigma, \end{cases} \text{ as } \tau \downarrow 0;$$

then, by virtue of monotone convergence, it follows that

$$\lim_{\tau \rightarrow 0^+} \tau^{-1} \left\{ \int_{\Omega} (u - \sigma - \tau)_+ dx - \int_{\Omega} (u - \sigma)_+ dx \right\} = |\{u > \sigma\}|$$

for every  $\sigma$ , and similarly with  $\bar{u}$  replacing  $u$ . This proves the claimed equivalence.  $\square$

The problem  $(P_{\infty})$  is a constrained minimization problem whose objective functional is

$$(2.5) \quad E(u) := \frac{1}{2} \int_{\Omega} |\nabla u|^2 dx,$$

and whose family of constraints is defined by (2.4). For a given  $\bar{u} \in H_0^1(\Omega)$  with  $\bar{u} \geq 0$  a.e. in  $\Omega$ , we pose the variational problem  $(P_{\infty})$

$$(2.6) \quad E(u) \rightarrow \min \quad \text{over } u \in M_{\infty}(\bar{u}).$$

We assume that  $\bar{u}$  is nonnegative for the sake of simplicity in the exposition here, deferring the general case until §5. Equivalently,  $(P_{\infty})$  may be written explicitly in terms of its constraints as

$$(2.7) \quad E(u) \rightarrow \min \quad \text{subject to } \int_{\Omega} (u - \sigma)_+ dx = \beta(\sigma), \quad 0 \leq \sigma < +\infty,$$

where  $\beta(\sigma) := \int_{\Omega} (\bar{u} - \sigma)_+ dx$  may be viewed as given data.

The existence of a solution  $u \in H_0^1(\Omega)$  of  $(P_{\infty})$  is straightforward (using the method of Theorem 2.2). On the other hand, the construction of an appropriate variational equation (a Lagrange multiplier rule) satisfied by such a solution is not routine, and indeed remains an unanswered question. Equally unclear is the question of the (optimal) regularity of the solution. The main source of difficulty in investigating these questions is, of course, the nature of the infinite family of constraints. With this in mind, we therefore proceed to formulate the problem  $(P_n)$ , which replaces the family of constraints in  $(P_{\infty})$  by a finite family of simpler constraints, and which approximates  $(P_{\infty})$  when  $n$  is large. Thus,  $(P_n)$  may be viewed as the natural discretization of  $(P_{\infty})$ .

We introduce a partition  $\{\sigma_i\}_{i=0}^n$  of the interval  $0 \leq \sigma < +\infty$  into  $n$  subintervals  $0 = \sigma_0 < \sigma_1 < \dots < \sigma_{n-1} < \sigma_n = +\infty$ . For a partition of this kind, we define the constraint functionals

$$(2.8) \quad F_i(u) := \int_{\Omega} f_i(u) dx \quad (i = 1, \dots, n),$$

where

$$(2.9) \quad f_i(s) := \int_{\sigma_{i-1}}^{\sigma_i} (s - \sigma)_+ d\sigma = \frac{1}{2}(s - \sigma_{i-1})_+^2 - \frac{1}{2}(s - \sigma_i)_+^2$$

for  $s \in \mathbb{R}$ . Then we let

$$(2.10) \quad M_n(\bar{u}) = \{u \in H_0^1(\Omega) : F_i(u) = F_i(\bar{u}) \text{ for all } i = 1, \dots, n\}.$$

It is easy to verify that  $M_\infty(\bar{u}) \subseteq M_n(\bar{u})$  for any  $n$ , and that if a sequence of partitions is taken so that  $\max_{1 \leq i \leq n-1} (\sigma_i - \sigma_{i-1}) \rightarrow 0$  and  $\sigma_{n-1} \rightarrow +\infty$  as  $n \rightarrow +\infty$ , then  $M_\infty(\bar{u}) = \bigcap_n M_n(\bar{u})$ , the class  $M_n(\bar{u})$  being defined by the  $n$ th partition in the sequence. If  $\bar{u} \in C^0(\bar{\Omega})$ , then it is not necessary that  $\sigma_{n-1} \rightarrow +\infty$  for this conclusion to hold, but rather it suffices that  $\sigma_{n-2} \leq \sup_\Omega \bar{u} < \sigma_{n-1}$ .

The problem  $(P_n)$  is the following multiconstrained minimization problem:

$$(2.11) \quad E(u) \rightarrow \min \quad \text{over } u \in M_n(\bar{u}).$$

Equivalently,  $(P_n)$  may be expressed as

$$(2.12) \quad E(u) \rightarrow \min \quad \text{subject to } F_i(u) = \gamma_i \quad (i = 1, \dots, n),$$

where  $\gamma_i := F_i(\bar{u}) = \int_{\sigma_{i-1}}^{\sigma_i} \beta(\sigma) d\sigma$  may be viewed as given data. The existence of a solution of  $(P_n)$  and the form of the variational equation that it satisfies are given in the next theorem.

**Theorem 2.2.** *If  $\bar{u}$  and  $\{\sigma_i\}_{i=0}^n$  are given so that  $\gamma_i > 0$  for every  $i$ , then there exists a minimizer  $u$  for  $(P_n)$  satisfying*

$$(2.13) \quad \begin{aligned} (a) \quad & u \in C^{2,\theta}(\bar{\Omega}) \text{ for every } 0 < \theta < 1, \\ (b) \quad & u \geq 0 \text{ in } \Omega, \\ (c) \quad & -\Delta u = \sum_{i=1}^n \lambda_i f'_i(u) \text{ for some } \lambda_i \in \mathbb{R}. \end{aligned}$$

*Proof.* We invoke a standard argument to establish the existence of  $u$ . Let  $u_j \in M_n(\bar{u})$  be a minimizing sequence in the sense that

$$\lim_{j \rightarrow \infty} E(u_j) = \inf\{E(\tilde{u}) : \tilde{u} \in M_n(\bar{u})\}.$$

The sequence  $u_j$ , being bounded in  $H_0^1(\Omega)$ , has a subsequence which converges weakly in  $H_0^1(\Omega)$  to a limit  $u$ . We write this subsequence as  $u_j$ , after reindexing, and we have

$$\begin{aligned} E(u) &\leq \lim_{j \rightarrow \infty} E(u_j) = \inf\{E(\tilde{u}) : \tilde{u} \in M_n(\bar{u})\}, \\ F_i(u) &= \lim_{j \rightarrow \infty} F_i(u_j) = \gamma_i \quad \text{for each } i = 1, \dots, n, \end{aligned}$$

using, respectively, the lower semicontinuity of  $E$  with respect to weak convergence in  $H^1$ , and the continuity of each  $F_i$  with respect to strong convergence in  $L^2$ . Thus, the limit  $u \in H_0^1(\Omega)$  solves  $(P_n)$ .

Property (2.13)(b) follows from the observation that  $u_j$  can be replaced by  $|u_j|$  in the above argument without changing the conclusion. (In fact, it is not difficult to see that any solution of (2.13)(a),(c) must be nonnegative.)

The variational equation (2.13)(c) satisfied by  $u$  is the standard Lagrange multiplier rule. As is shown in the reference text [14, Chapter 1], it suffices that the functionals  $E, F_i: H_0^1(\Omega) \rightarrow \mathbb{R}$  be  $C^1$  and that the gradients  $F_1'(u), \dots, F_n'(u)$  be linearly independent. It is immediate from their definitions that the functionals involved are Fréchet differentiable with

$$\langle E'(u), v \rangle = \int_{\Omega} \nabla u \cdot \nabla v \, dx, \quad \langle F_i'(u), v \rangle = \int_{\Omega} f_i'(u) v \, dx,$$

where

$$(2.14) \quad f_i'(s) = (s - \sigma_{i-1})_+ - (s - \sigma_i)_+ \quad (s \in \mathbb{R}).$$

Thus, interpreting these derivatives in the sense of distributions, we have

$$(2.15) \quad E'(u) = -\Delta u \in H^{-1}(\Omega), \quad F_i'(u) = f_i'(u) \in H_0^1(\Omega).$$

Moreover, it is now clear that  $E$  and  $F_i$  are  $C^1$  functionals. The crucial linear independence of the gradients  $f_i'(u)$  ( $i = 1, \dots, n$ ) will be deduced from the following

**Lemma 2.3.** *Let  $f \in C^1(\mathbb{R})$  satisfy  $f(0) = f'(0) = 0$ , and  $f'' \in L^\infty(\mathbb{R})$  with  $f''(s) \geq 0$  for almost all  $s \in \mathbb{R}$ . Then for any  $u \in H_0^1(\Omega)$*

$$(2.16) \quad \int_{\Omega} f(u) \, dx \leq c_N |\Omega|^{1/N} \|f''(u)\|_{L^N} \int_{\Omega} |\nabla u|^2 \, dx.$$

*Proof.* The convexity of  $f(s)$  implies that  $0 = f(0) \geq f(s) - sf'(s)$  for all  $s \in \mathbb{R}$ . Hence

$$(2.17) \quad \int_{\Omega} f(u) \, dx \leq \int_{\Omega} u f'(u) \, dx \leq \|u\|_{L^2} \|f'(u)\|_{L^2}.$$

The Sobolev inequality then yields

$$(2.18) \quad \|f'(u)\|_{L^2} \leq c_N \|\nabla f'(u)\|_{L^p} \leq c_N \|f''(u)\|_{L^N} \|\nabla u\|_{L^2},$$

where  $1 \leq p < 2$  is determined by  $1/2 = 1/p - 1/N$ . The following Poincaré type inequality results from the combination of (2.17) and (2.18):

$$(2.19) \quad \int_{\Omega} f(u) \, dx \leq c_N \|u\|_{L^2} \|f''(u)\|_{L^N} \|\nabla u\|_{L^2}.$$

When this inequality is specialized to the case  $f(s) = \frac{1}{2}s^2$ , the standard Poincaré inequality ensues, namely,

$$\|u\|_{L^2} \leq 2c_N |\Omega|^{1/N} \|\nabla u\|_{L^2}.$$

The claimed inequality (2.16) now follows upon substituting the latter in (2.19).  $\square$



Indeed, now applying Lemma 2.3 to  $f_i(u)$ , we find that each of the sets  $\{\sigma_{i-1} < u < \sigma_i\}$  has strictly positive measure (with a lower bound depending only on upper bounds for  $\gamma_i^{-1}$ ,  $E(u)$ , and  $|\Omega|$ ), since  $f_i''(u) = 1$  in  $\{\sigma_{i-1} < u < \sigma_i\}$  and  $f_i''(u) = 0$  elsewhere. Thus, if for some  $c_i \in \mathbb{R}$  the identity  $0 = \sum_{i=1}^n c_i f_i'(u)$  holds a.e. in  $\Omega$ , then by restricting this identity to the sets  $\{\sigma_{i-1} < u < \sigma_i\}$  we find successively that  $c_1 = 0, c_2 = 0, \dots, c_n = 0$ . Hence, the required linear independence of the gradients of the constraint functionals is proved, and so (2.13)(c) follows for some Lagrange multipliers  $\lambda_i$  ( $i = 1, \dots, n$ ) uniquely determined by  $u$ .

Finally, the smoothness of  $u$  stated in (2.13)(a) can be derived from the standard regularity theory for elliptic partial differential equations [10]. In terms of the Green function  $g(x, x')$ ,  $x, x' \in \Omega$ , defined by  $-\Delta_x g(x, x') = \delta(x - x')$  for  $x \in \Omega$ ,  $g(x, x') = 0$  for  $x \in \partial\Omega$ , we have

$$(2.20) \quad u(x) = \int_{\Omega} g(x, x') \Lambda(u(x')) dx' \quad \text{with } \Lambda(s) := \sum_{i=1}^n \lambda_i f_i'(s).$$

Since  $|\Lambda(u)| \leq (\max_{1 \leq i \leq n} |\lambda_i|) |u|$ , we obtain  $u, \Lambda(u) \in L^\infty(\Omega)$ , using a standard argument. Now differentiating (2.20) in  $x$ , we find that  $u \in C^1(\bar{\Omega})$ . Thus,  $\Lambda(u) \in C^{0,1}(\bar{\Omega})$ , and this is optimal since the Lipschitz function  $\Lambda(s)$  is piecewise linear. The statement (2.13)(a) that  $u \in C^{2,\theta}(\bar{\Omega})$  is therefore a consequence of the global Schauder estimates applied to the equation (2.13)(c). (It is assumed here that  $\partial\Omega$  is of class  $C^{2,\theta}$ , at least.) This completes the proof of Theorem 2.2.  $\square$

The function  $\Lambda(s)$  occurring in (2.20) will play a key role in our subsequent analysis and will be referred to as the *profile function* associated with a solution  $u$ . For later reference we record its definition:

$$(2.21) \quad \Lambda(s) = \begin{cases} 0 & \text{if } s < 0, \\ \sum_{i=1}^n \lambda_i f_i'(s) & \text{if } s \geq 0, \end{cases}$$

where  $\lambda_1, \dots, \lambda_n$  are real constants (the multipliers in Theorem 2.2). The basis functions  $f_i'(s)$ ,  $i = 1, \dots, n$ , each member of which is monotone, is related to the standard basis consisting of finite elements  $\phi_i(s)$  for the partition (or grid)  $\{\sigma_i\}_{i=0}^n$  by the formulas

$$\begin{cases} \phi_i(s) = f_i'(s)/\Delta\sigma_i - f_{i+1}'(s)/\Delta\sigma_{i+1} & (i = 1, \dots, n-2), \\ \phi_{n-1}(s) = f_{n-1}'(s)/\Delta\sigma_{n-1}, & \phi_n(s) = f_n'(s), \end{cases}$$

where  $\Delta\sigma_i = \sigma_i - \sigma_{i-1}$  denotes the increments. At least when  $1 \leq i \leq n-2$ , each piecewise linear  $\phi_i(s)$  is supported on the interval  $\sigma_{i-1} \leq s \leq \sigma_{i+1}$  and is normalized by  $\phi_i(\sigma_i) = 1$ . In terms of these finite elements, the profile function is represented as

$$\Lambda(s) = \begin{cases} 0 & \text{if } s < 0, \\ \sum_{i=1}^{n-1} \Lambda_i \phi_i(s) + \lambda_n \phi_n(s) & \text{if } s \geq 0, \end{cases}$$

where  $\Lambda_i = \Lambda(\sigma_i)$ ; in other words,  $\Lambda(s)$  on  $0 \leq s \leq \sigma_{n-1}$  is identified with the linear interpolant of its values  $\Lambda_i$  at the partition points (or grid nodes)  $\sigma_i$ . The representation (2.21) in terms of the monotone basis functions  $f_i'(s)$  will be used throughout the sequel, however. Its importance stems from the fact that the constraint functionals  $F_i$ , like the objective functional  $E$ , are *convex* for this choice of the basis functions  $f_i(s)$ .

Furthermore, it is a consequence of the definition of  $f_i(s)$  in (2.9) that

$$(2.22) \quad \sum_{i=1}^n f_i(s) = \frac{1}{2}s_+^2 \quad (s \in \mathbb{R}).$$

Therefore, any admissible function  $u \in M_n(\bar{u})$  satisfies

$$(2.23) \quad \frac{1}{2} \int_{\Omega} u^2 dx = \sum_{i=1}^n \gamma_i = \frac{1}{2} \int_{\Omega} \bar{u}^2 dx.$$

Indeed, if  $n = 1$  (and so  $\sigma_0 = 0$  and  $\sigma_1 = +\infty$ ), then the class  $M_1(\bar{u})$  is characterized precisely by the single constraint (2.23). In view of this fact, we may say that the problem  $(P_n)$  constitutes a generalization of the classical Rayleigh principle [6, Chapter VI] characterizing variationally the first eigenvalue-eigenfunction pair for the Laplacian operator  $-\Delta$  on  $H_0^1(\Omega)$ . In going from the linear problem  $(P_1)$  to the general (and nonlinear) problem  $(P_n)$ , the single  $L^2(\Omega)$ -normalization  $\|u\|_{L^2}^2 = 2\gamma_1$  ( $= 1$ , say) is replaced by the family of constraints  $F_i(u) = \gamma_i$  ( $i = 1, \dots, n$ ), which in effect amounts to constraining  $n$   $L^2(\Omega)$ -expressions involving the lower truncations  $(u - \sigma_i)_+$  of  $u$ ; specifically, the constraints for  $(P_n)$  can be written as

$$(2.24) \quad \|(u - \sigma_{i-1})_+\|_{L^2}^2 - \|(u - \sigma_i)_+\|_{L^2}^2 = 2\gamma_i \quad (i = 1, \dots, n).$$

The family of constraints (2.24) is therefore seen to be a strengthening of the single constraint (2.23) in the usual Rayleigh principle  $(P_1)$ .

The sense in which a solution  $u^{(n)}$  of  $(P_n)$  approximates a solution  $u^{(\infty)}$  (say) of  $(P_{\infty})$  as  $n \rightarrow +\infty$  can now be determined. As mentioned above, we suppose that a sequence of partitions is taken with  $\max_{1 \leq i \leq n-1} (\sigma_i - \sigma_{i-1}) \rightarrow 0$  and  $\sigma_{n-1} \rightarrow +\infty$  as  $n \rightarrow \infty$ , so that we are assured that  $M_{\infty}(\bar{u}) = \bigcap_n M_n(\bar{u})$ ; if  $\bar{u} \in C^0(\bar{\Omega})$ , then we suppose instead that  $\sigma_{n-2} \leq \sup_{\Omega} u < \sigma_{n-1}$ , with  $\max_{1 \leq i \leq n-1} (\sigma_i - \sigma_{i-1}) \rightarrow 0$  as  $n \rightarrow +\infty$ . By virtue of the bound  $E(u^{(n)}) \leq E(\bar{u}) < +\infty$ , every subsequence of  $\{u^{(n)}\}_{n=1}^{\infty}$  has a further subsequence that converges weakly in  $H_0^1$  and strongly in  $L^2$ . If we let  $u^{(\infty)} \in H_0^1(\Omega)$  denote a limit point of the solution sequence  $u^{(n)}$  for the multiconstrained problems  $(P_n)$ , then we claim that  $u^{(\infty)}$  is a solution of the (infinitely-constrained) problem  $(P_{\infty})$ . Indeed, any admissible function  $\hat{u} \in M_{\infty}(\bar{u})$  for  $(P_{\infty})$  is admissible for each problem  $(P_n)$ , and hence, using the weak  $H_0^1$ -convergence,  $E(u^{(\infty)}) \leq \liminf E(u^{(n)}) \leq E(\hat{u})$ , where  $n$  is understood to tend to infinity along the subsequence. Also, using the strong  $L^2$ -convergence, it is straightforward to verify

that  $u^{(\infty)} \in M_\infty(\bar{u})$ , because it is evident that for any  $\sigma \in (0, +\infty)$

$$\int_\Omega (u - \sigma)_+ dx = \lim_{n \rightarrow \infty} \frac{1}{\Delta\sigma_i} \int_\Omega f_i(u) dx,$$

where  $i = i(n, \sigma)$  is chosen so that  $\sigma \in [\sigma_{i-1}, \sigma_i)$  for each sufficiently large  $n$ . (We recall that  $\Delta\sigma_i := \sigma_i - \sigma_{i-1}$  and  $f_i$  is given by (2.9).) Thus the claim that  $u^{(\infty)}$  solves  $(P_\infty)$  follows.

This result can be stated in the form

$$(2.25) \quad \text{dist}_{L^2}(u^{(n)}, S_\infty) \rightarrow 0 \quad \text{as } n \rightarrow +\infty,$$

where  $S_\infty$  is defined to be the set of solutions of  $(P_\infty)$  for a given function  $\bar{u}$ ; the distance function  $\text{dist}_{L^2}(u, \mathcal{S})$  from a point  $u$  to a set  $\mathcal{S} \subseteq L^2(\Omega)$  is defined by  $\text{dist}_{L^2}(u, \mathcal{S}) = \inf\{\|u - v\|_{L^2} : v \in \mathcal{S}\}$ . That (2.25) is true is obvious from the above discussion. We remark that it is necessary to consider the set of solutions of  $(P_\infty)$ , since the uniqueness of its solutions has not been established and, in fact, may not hold in general. The statement (2.25) allows us to conclude, at least, that for large  $n$  the solution  $u^{(n)}$  of  $(P_n)$  approximates some solution of  $(P_\infty)$  in the  $L^2(\Omega)$ -norm. This conclusion justifies in part the discretization of the constraints for  $(P_\infty)$  resulting in the finitely many constraints for  $(P_n)$ .

Before passing to the discussion of the algorithm for solving  $(P_n)$ , we note that a variant of this problem can also be employed. We now assume that  $\bar{u} \in H_0^1(\Omega) \cap C^0(\bar{\Omega})$  and let  $\sigma_n = \sup_\Omega \bar{u}$  (rather than  $\sigma_n = +\infty$ ). With respect to a partition  $0 = \sigma_0 < \sigma_1 < \dots < \sigma_{n-1} < \sigma_n$ , we consider the problem  $(P'_n)$

$$E(u) \rightarrow \min \quad \text{over } u \in M_n(\bar{u}), \quad u \leq \sigma_n \text{ a.e. in } \Omega.$$

This problem is of variational inequality type [9, Chapters 1 and 2], being an obstacle problem with  $n$  additional constraints. It can be shown that the analogue of Theorem 2.2 holds in the sense that there is a solution  $u \in H_0^1(\Omega) \cap C^{1,1}(\bar{\Omega})$  satisfying the equation

$$-\Delta u = \sum_{i=1}^n \lambda_i f'_i(u) \quad \text{in } \{u < \sigma_n\} \subseteq \Omega.$$

(If  $\{u = \sigma_n\}$  has positive measure, then  $u$  is not necessarily twice continuously differentiable, of course.) However, we shall restrict our further discussion to  $(P_n)$ , leaving the parallel development for  $(P'_n)$  to the reader.

### 3. DESCRIPTION OF THE ALGORITHM $(A_n)$

We now proceed to define an iterative procedure which is designed to converge to the solutions of  $(P_n)$ .

For a (fixed) positive constant  $\tau$ , let  $E_\tau$  be the modified objective functional

$$(3.1) \quad E_\tau(u) = E(u) + \frac{\tau}{2} \|u\|^2 = \frac{1}{2} \int_\Omega [|\nabla u|^2 + \tau u^2] dx.$$

Also, for any  $v \in H_0^1(\Omega)$ , let  $L_n(v)$  be the class of admissible functions

$$(3.2) \quad L_n(v) = \{u \in H_0^1(\Omega) : F_i(v) + \langle F'_i(v), u - v \rangle \geq \gamma_i \text{ for all } i\}.$$

We notice that in contrast to the class  $M_n(\bar{u})$ , which is defined by  $n$  nonlinear equality constraints, the class  $L_n(v)$  is defined by  $n$  linear (actually affine) inequality constraints. ( $L_n(v)$  depends upon  $\bar{u}$  through the constants  $\gamma_i = F_i(\bar{u})$ , but this dependence will not be emphasized in the notation.) In terms of these definitions we construct a sequence of approximations  $(u^k, \lambda^k) \in H_0^1(\Omega) \times \mathbb{R}^n$  to the solutions of the variational problem  $(P_n)$  as follows.

**Algorithm  $(A_n)$ .** Given  $u^0 \in M_n(\bar{u})$ , let  $(u^k, \mu^k) \in H_0^1(\Omega) \times [0, +\infty)^n$  be defined iteratively by solving the sequence of convex optimization problems

$$(3.3) \quad E_\tau(u) \rightarrow \min \quad \text{over } u \in L_n(u^k),$$

where  $u^{k+1}$  is the (unique) solution, and  $\mu^{k+1}$  is the corresponding  $n$ -vector of nonnegative multipliers (uniquely determined by  $u^{k+1}$ ); then let

$$(3.4) \quad \lambda^{k+1} = \mu^{k+1} - \tau.$$

The algorithm  $(A_n)$  produces a well-defined sequence of approximations once an initialization  $u^0$  and a constant  $\tau$  are chosen. It suffices to take  $u^0 = \bar{u}$ . It will be shown in §4 that  $\tau$  can be chosen sufficiently large to ensure the convergence of the iterative sequence  $\{(u^k, \lambda^k)\}$  to the set  $S_n^*$  of solutions of the problem

$$(3.5) \quad \begin{aligned} E'(u) &= \sum_{j=1}^n \lambda_j F'_j(u) \quad \text{for some } \lambda_1, \dots, \lambda_n \in \mathbb{R}, \\ F_i(u) &= \gamma_i \quad (i = 1, \dots, n), \end{aligned}$$

namely, the set of all critical points for the problem  $(P_n)$  which satisfy the *given* constraints (or, equivalently, belong to  $M_n(\bar{u})$ ). In the present discussion,  $\tau$  will be viewed as a (fixed) positive parameter.

The pair  $(u^{k+1}, \mu^{k+1})$  defined in the algorithm  $(A_n)$  is characterized as the solution of the equations

$$(3.6) \quad \begin{aligned} E'(u^{k+1}) + \tau u^{k+1} &= \sum_{j=1}^n \lambda_j^{k+1} F'_j(u^k), \\ \mu_i^{k+1} [F_i(u^k) + \langle F'_i(u^k), u^{k+1} - u^k \rangle - \gamma_i] &= 0 \quad (i = 1, \dots, n). \end{aligned}$$

Indeed, these equations are precisely the Kuhn-Tucker conditions associated with the convex optimization problem (3.3) which defines the iterative step of  $(A_n)$ . The reader is referred to [14, Chapter 1] for a proof that (3.3) and (3.6) are equivalent, provided that there exists some  $\tilde{u} \in L_n(u^k)$  for which

$$F_i(u^k) + \langle F'_i(u^k), \tilde{u} - u^k \rangle > \gamma_i \quad \text{for every } i = 1, \dots, n.$$

This so-called Slater condition follows from the convexity of the functional  $F_i(u)$ , which implies that

$$(3.7) \quad F_i(u^k) \geq F_i(u^{k-1}) + \langle F'_i(u^{k-1}), u^k - u^{k-1} \rangle \geq \gamma_i,$$

and that  $F_i(u^k) \leq \langle F'_i(u^k), u^k \rangle$ . Thus, it is readily verified that the Slater condition is satisfied by  $\tilde{u} = (1 + \varepsilon)u^k$  with  $\varepsilon > 0$ . It should be emphasized that the equations (3.6) characterize the pair  $(u^{k+1}, \mu^{k+1}) \in L_n(u^k) \times [0, +\infty)^n$  when supplemented by the inequalities

$$(3.8) \quad F_i(u^k) + \langle F'_i(u^k), u^{k+1} - u^k \rangle \geq \gamma_i,$$

$$(3.9) \quad \mu_i^{k+1} \geq 0,$$

for every  $i = 1, \dots, n$ . The vector  $\mu^{k+1} = (\mu_1^{k+1}, \dots, \mu_n^{k+1})$  is the Kuhn-Tucker vector corresponding to the solution  $u^{k+1}$  of (3.3).

The above construction of the algorithm  $(A_n)$  depends fundamentally on the convexity of the optimization problem (3.3); namely, it relies on the strict convexity of the objective functional  $E_\tau$  and the convexity of the class of admissible functions  $L_n(u^k)$ , which is defined by  $n$  affine inequality constraints. The uniqueness of a minimizer  $u^{k+1}$  for (3.3) follows immediately from these attributes. In turn, the uniqueness of the Kuhn-Tucker vector  $\mu^{k+1}$  follows from the smoothness of the objective and constraint functionals and the linear independence of the gradients  $F'_i(u^k)$ . The latter property can be proved exactly as in the proof of Theorem 2.2, now using (3.7) with  $\gamma_i > 0$ .

Next we describe a more explicit construction of the iterative procedure  $(A_n)$  which furnishes us with a concrete numerical implementation of the algorithm. Let  $G_\tau$  denote the Green operator for the elliptic boundary value problem

$$(3.10) \quad -\Delta w + \tau w = h \quad \text{in } \Omega, \quad w = 0 \quad \text{on } \partial\Omega;$$

that is, the solution is represented as  $w = G_\tau h$ , where  $G_\tau: L^2(\Omega) \rightarrow H_0^1(\Omega) \cap H^2(\Omega)$ . For any  $v \in H_0^1(\Omega)$ , let (recall (2.14), (2.15))

$$(3.11) \quad a_{ij}(v) := \langle f'_i(v), G_\tau f'_j(v) \rangle,$$

$$(3.12) \quad c_i(v) := \gamma_i - F_i(v) + \langle F'_i(v), v \rangle$$

for  $i, j = 1, \dots, n$ . (The dependence of these expressions on  $\tau$  is left implicit, for the sake of simplicity in the notation.) We consider the quadratic form

$$(3.13) \quad Q(\mu; v) := \frac{1}{2} \sum_{i,j=1}^n a_{ij}(v) \mu_i \mu_j - \sum_{i=1}^n c_i(v) \mu_i \quad (\mu \in \mathbb{R}^n).$$

This defines a *positive definite* quadratic form on  $\mathbb{R}^n$  whenever  $|\{x \in \Omega: \sigma_{i-1} < v(x) < \sigma_i\}| > 0$  for every  $i = 1, \dots, n$ . We check this fact by calculating

$$\sum_{i,j=1}^n a_{ij}(v) \mu_i \mu_j = \langle \mu \cdot f'(v), G_\tau[\mu \cdot f'(v)] \rangle,$$

where we write  $\mu \cdot f'(v) = \sum_{i=1}^n \mu_i f'_i(v)$ . We then notice that the latter expression is zero if and only if  $\mu \cdot f'(v) = 0$  a.e. in  $\Omega$ . Arguing as in the proof of Theorem 2.2, we deduce that  $\mu_1 = 0, \mu_2 = 0, \dots, \mu_n = 0$  successively, as required.

We now claim that the algorithm  $(A_n)$  is equivalent to the following explicit iterative procedure: for  $u^0 \in M_n(\bar{u})$ , let  $(u^k, \mu^k) \in H_0^1(\Omega) \times [0, +\infty)^n$  be defined iteratively by

$$(3.14) \quad u^{k+1} = \sum_{j=1}^n \mu_j^{k+1} w_j,$$

where

$$(3.15) \quad w_j = G_\tau f'_j(u^k),$$

$$(3.16) \quad \mu^{k+1} = \arg \min \{Q(\mu; u^k) : \mu_i \geq 0, i = 1, \dots, n\}.$$

In other words,  $(u^{k+1}, \mu^{k+1})$  is constructed from  $u^k$  in a three-step process: (1) each  $w_j$  is found by solving (3.10) with  $h = f'_j(u^k)$ ; (2)  $\mu^{k+1}$  is taken to be the unique solution of the quadratic programming problem (3.16); and (3)  $u^{k+1}$  is assembled from  $\mu_j^{k+1}$  and  $w_j$  according to (3.14). The verification of the claim that this procedure is equivalent to  $(A_n)$ , as stated earlier, is standard. Calculating the variational inequalities satisfied by the minimizer  $\mu^{k+1}$  of (3.16), we have for each  $i = 1, \dots, n$

$$(3.17) \quad \sum_{j=1}^n a_{ij}(u^k) \mu_j^{k+1} - c_i(u^k) \begin{cases} \geq 0 & \text{if } \mu_i^{k+1} = 0, \\ = 0 & \text{if } \mu_i^{k+1} > 0. \end{cases}$$

Also, combining (3.14) and (3.15), we have

$$(3.18) \quad -\Delta u^{k+1} + \tau u^{k+1} = \sum_{j=1}^n \mu_j^{k+1} f'_j(u^k) \quad \text{in } \Omega, \quad u^{k+1} = 0 \quad \text{on } \partial\Omega.$$

It is clear that (3.18) is a restatement of the first equation in (3.6); in turn, it is evident from the definitions of  $a_{ij}$  and  $c_i$  that (3.17) is equivalent to the second equation in (3.6) together with the inequalities (3.8) and (3.9), which supplement it. Thus the claimed equivalence of the two forms of the algorithm is verified.

It is interesting to note that the optimization problems (3.16) and (3.3) are dual problems in the sense of convex analysis [19, §30]. Indeed, we can directly verify that

$$-Q(\mu; u^k) = \min \left\{ E_\tau(u) - \sum_{i=1}^n \mu_i [\langle F'_i(u^k), u \rangle - c_i(u^k)] : u \in H_0^1(\Omega) \right\}.$$

The general theory of convex analysis informs us that the Kuhn-Tucker vector  $\mu^{k+1}$  corresponding to the minimizer  $u^{k+1}$  for (3.3) is itself the maximizer for

the dual problem:

$$-Q(\mu; u^k) \rightarrow \max \quad \text{over } \mu_i \geq 0 \quad (i = 1, \dots, n).$$

The preceding construction of the explicit form of the iterative procedure  $(A_n)$  may therefore be viewed as a specific case of the general duality theory. We shall not need this level of generality in the sequel, however.

The algorithm  $(A_n)$  is devised to exploit as much as possible the important fact that the objective and constraint functionals for the variational problem  $(P_n)$  are convex. Even though  $(P_n)$  itself is *not* a convex optimization problem, because its constraints are nonlinear equalities, its convexity attributes are enough to imply that the iterative sequence defined by  $(A_n)$  has very special monotonicity and convergence properties. These properties are the focus of our attention in the next two lemmas.

**Lemma 3.1.** *For every  $k$  we have*

$$(3.19) \quad E_\tau(u^{k+1} - u^k) \leq E_\tau(u^k) - E_\tau(u^{k+1}),$$

$$(3.20) \quad E(u^{k+1}) \leq E(u^0).$$

*Proof.* To prove (3.19), we use the identity

$$(3.21) \quad E_\tau(u^k) - E_\tau(u^{k+1}) = \langle E'_\tau(u^{k+1}), u^k - u^{k+1} \rangle + E_\tau(u^{k+1} - u^k).$$

By virtue of (3.7) we have  $u^k \in L_n(u^k)$ . Therefore,  $(1-t)u^{k+1} + tu^k \in L_n(u^k)$  for all  $0 \leq t \leq 1$ , as this class of functions is convex. Consequently, since  $u^{k+1}$  solves (3.3), we get

$$\begin{aligned} E_\tau(u^{k+1}) &\leq E_\tau(u^{k+1} + t(u^k - u^{k+1})) \\ &= E_\tau(u^{k+1}) + t \langle E'_\tau(u^{k+1}), u^k - u^{k+1} \rangle + O(t^2) \end{aligned}$$

as  $t \rightarrow 0^+$ . Hence,  $\langle E'_\tau(u^{k+1}), u^k - u^{k+1} \rangle \geq 0$ , and so (3.19) follows from (3.21).

To prove (3.20), we note that

$$\|u^0\|^2 = \sum_{i=1}^n \gamma_i \leq \sum_{i=1}^n F_i(u^{k+1}) = \|u^{k+1}\|^2.$$

This yields the desired inequality, since

$$\begin{aligned} E(u^{k+1}) &= E_\tau(u^{k+1}) - \frac{\tau}{2} \|u^{k+1}\|^2 \\ &\leq E_\tau(u^0) - \frac{\tau}{2} \|u^0\|^2 = E(u^0). \quad \square \end{aligned}$$

The above lemma establishes that for any initialization  $u^0 \in M_n(\bar{u})$  the sequence  $E_\tau(u^k)$  is nonincreasing. This monotonicity property of the algorithm  $(A_n)$  leads directly to the following (partial) convergence property of the iterative sequence  $\{u^k\}$ .

**Lemma 3.2.** *Let  $u \in H_0^1(\Omega)$  be any  $H_0^1$ -weak limit point of the sequence  $\{u^k\}$ . Then there exists a unique vector  $\mu = (\mu_1, \dots, \mu_n) \in \mathbb{R}^n$  such that*

$$(3.22) \quad E'_\tau(u) = \sum_{j=1}^n \mu_j F'_j(u),$$

$$(3.23) \quad \mu_i [F_i(u) - \gamma_i] = 0 \quad (i = 1, \dots, n),$$

$$(3.24) \quad \mu_i \geq 0, \quad F_i(u) \geq \gamma_i \quad (i = 1, \dots, n).$$

*Proof.* We recognize (3.22)–(3.24) as precisely the Kuhn-Tucker conditions for the convex minimization problem

$$(3.25) \quad E_\tau(\tilde{u}) \rightarrow \min \quad \text{over } \tilde{u} \in L_n(u),$$

and  $\mu$  as the Kuhn-Tucker vector corresponding to the minimizer  $u$ . Therefore, it suffices to show that  $u$  is indeed the solution of (3.25). Let  $u^{k_p} \rightarrow u$  weakly in  $H_0^1(\Omega)$ , and hence strongly in  $L^2(\Omega)$ , as  $p \rightarrow +\infty$ . Given an arbitrary  $\tilde{u} \in L_n(u)$ , we consider the perturbation  $\tilde{u} + \varepsilon u$  with  $\varepsilon > 0$ . Recalling the definition of  $L_n(u)$ , we find that for each  $i = 1, \dots, n$

$$F_i(u) + \langle F'_i(u), (\tilde{u} + \varepsilon u) - u \rangle \geq \gamma_i + \varepsilon \langle F'_i(u), u \rangle \geq (1 + \varepsilon)\gamma_i.$$

This implies that for sufficiently large  $p$

$$F_i(u^{k_p}) + \langle F'_i(u^{k_p}), (\tilde{u} + \varepsilon u) - u^{k_p} \rangle \geq (1 + \varepsilon/2)\gamma_i,$$

by virtue of the continuity of the terms on the left-hand side of the latter inequality with respect to strong  $L^2$  convergence. Consequently,  $\tilde{u} + \varepsilon u \in L_n(u^{k_p})$  for sufficiently large  $p$ , and so we obtain

$$E_\tau(u) \leq E_\tau(u^{k_p+1}) \leq E_\tau(\tilde{u} + \varepsilon u),$$

where in the first inequality we invoke the monotonicity property (3.19) and the lower semicontinuity of  $E_\tau$  with respect to weak  $H_0^1$ -convergence. Since this inequality holds for arbitrarily small positive  $\varepsilon$ , we conclude that the limit point  $u$  solves (3.25), as required.  $\square$

The partial convergence result given in the above lemma provides the first step in the proof that the iterates  $(u^k, \lambda^k)$  defined by the algorithm  $(A_n)$  converge in an appropriate sense to the critical points of  $(P_n)$ . Indeed, Lemmas 3.1 and 3.2 together permit us to conclude that any subsequence of the iterative sequence  $\{u^k\} \subseteq H_0^1(\Omega)$  has a further subsequence which converges (weakly in  $H_0^1$  and strongly in  $L^2$ ) to a solution  $u$  of (3.22)–(3.24). But (3.22) can be written equivalently as

$$(3.26) \quad E'(u) = \sum_{j=1}^n \lambda_j F'_j(u), \quad \text{where } \lambda_j := \mu_j - \tau,$$

in view of the basic identity (2.23), which implies that

$$(3.27) \quad \sum_{i=1}^n f'_i(s) = s_+ \quad (s \in \mathbb{R}).$$



(Here we also use the fact that  $u^k \geq 0$  for every  $k$ , and hence  $u \geq 0$ , as follows immediately from the explicit form of  $(A_n)$ .) Thus, the addition of the term  $\frac{\tau}{2}\|u\|^2$  to the objective functional  $E(u)$  has the effect of shifting the multipliers by  $\tau$ , sending  $\lambda_i$  into  $\mu_i = \lambda_i + \tau$ . Accordingly, a solution  $u$  of (3.22)–(3.24) is a critical point of the variational problem  $(P_n)$  provided that  $\mu_i > 0$  for every  $i = 1, \dots, n$ , since then the constraints must be equalities  $F_i(u) = \gamma_i$  for every  $i = 1, \dots, n$ . The latter condition can be ensured by choosing  $\tau$  large enough ( $\tau > \max\{0, -\lambda_1, \dots, -\lambda_n\}$ ) depending on the multipliers  $\lambda_i$  corresponding to any solution  $u$  of (3.26). The a priori estimate on  $\max_{1 \leq i \leq n} |\lambda_i|$  needed to complete this argument is given in Lemma 4.3; it provides the second step in the proof of convergence of the algorithm  $(A_n)$ . The complete convergence result is contained in Theorem 4.1.

#### 4. CONVERGENCE THEOREMS

In the previous section, we have studied the algorithm  $(A_n)$  defined by (3.1)–(3.3). There, the positive constant  $\tau$  is introduced in order to allow us to impose linearized inequality constraints at each iteration. In this section, we will show that  $\tau$  can be chosen a priori so that the limit points of the algorithm  $(A_n)$  are critical points of the problem  $(P_n)$  and satisfy its nonlinear equality constraints. More precisely, we have the following convergence theorem.

**Theorem 4.1.** *Given any initialization  $u^0 \in M_n(\bar{u})$ , there exists a constant  $C = C(\Omega, E(u^0), \gamma_1^{-1} \cdots \gamma_n^{-1}, (\Delta\sigma)^{-1})$  such that if  $\tau > C$ , then the iterative sequence  $\{u^k\}$  defined by the algorithm  $(A_n)$  converges strongly in  $H_0^1(\Omega)$  to the set  $S_n^*$  of critical points of the problem  $(P_n)$ , in the sense that*

$$\text{dist}_{H_0^1}(u^k, S_n^*) \rightarrow 0 \quad \text{as } k \rightarrow \infty,$$

where

$$\text{dist}_{H_0^1}(u, S_n^*) = \inf\{\|u - v\|_{H_0^1} : v \in S_n^*\}.$$

Theorem 4.1 says that any subsequence of  $\{u^k\}$  has a further subsequence which converges strongly in  $H_0^1(\Omega)$  to an element of  $S_n^*$  (that is, a solution of (3.5)). In numerical experiments (typically) the entire sequence  $\{u^k\}$  converges. Although we do not have a convergence result of this kind, we do have the following result which indicates that generically there is only one limit.

**Corollary 4.2.** *Let  $A(u^0)$  be the set of limit points of algorithm  $(A_n)$  for a given  $u^0$ . Then either (i)  $A(u^0)$  contains exactly one point, or (ii)  $A(u^0)$  contains infinitely many points, none of which is isolated.*

The key point in the proof of Theorem 4.1 is the following a priori estimate, which is of independent interest.

**Lemma 4.3.** *Let  $u \in H_0^1(\Omega)$  be a (weak) solution of*

$$(4.1) \quad \begin{aligned} -\Delta u &= \sum_{i=1}^n \lambda_i f'_i(u), \\ F_i(u) &= \int_{\Omega} f_i(u) dx \geq \gamma_i > 0 \quad (i = 1, \dots, n) \end{aligned}$$

with  $E(u) \leq E_0 < +\infty$ . Then

$$\max_i |\lambda_i| \leq C = C(\Omega, E_0, \gamma_1^{-1}, \dots, \gamma_n^{-1}, (\Delta\sigma)^{-1}),$$

where  $\Delta\sigma = \min_i \Delta\sigma_i$ .

Assuming the truth of Lemma 4.3 for the moment, let us give the

*Proof of Theorem 4.1.* As we indicated in the previous section, Lemmas 3.1 and 3.2 permit us to conclude that any subsequence of  $\{u^k\}$  has a further subsequence which converges weakly in  $H_0^1(\Omega)$  and strongly in  $L^2(\Omega)$  to a solution of (4.1) with

$$\lambda_i = \mu_i - \tau.$$

By Lemma 4.3 we have that  $\lambda_i > -C$ , which means that  $\mu_i > \tau - C$ . Hence, if a priori we choose  $\tau > C$ , then we ensure that the multipliers  $\mu_i > 0$  for all  $i = 1, \dots, n$ . This implies, by Lemma 3.2, that  $F_i(u) = \gamma_i$ ; that is,  $u$  is a critical point of the problem  $(P_n)$ .

In order to complete the proof of Theorem 4.1, we need to show the strong  $H_0^1(\Omega)$ -convergence of the subsequence, which we still call  $u^k$  for convenience. To see this, we observe that

$$(4.2) \quad \begin{aligned} \|u^{k+1} - u\|_{H_0^1}^2 &= \int_{\Omega} \nabla u^{k+1} \cdot \nabla(u^{k+1} - u) dx - \int_{\Omega} \nabla u \cdot \nabla(u^{k+1} - u) dx \\ &= \sum_{i=1}^n \lambda_i^{k+1} \int_{\Omega} (u^{k+1} - u) f'_i(u^k) dx - \int_{\Omega} \nabla u \cdot \nabla(u^{k+1} - u) dx \\ &= \sum_{i=1}^n \mu_i^{k+1} \int_{\Omega} (u^{k+1} - u) f'_i(u^k) dx + o(1), \end{aligned}$$

as  $k \rightarrow \infty$ . But we claim that the  $\mu_i^{k+1}$  are uniformly bounded. Indeed,

$$\begin{aligned} \sum_{i=1}^n \mu_i^{k+1} \int_{\Omega} u^k f'_i(u^k) dx &= \int_{\Omega} [\nabla u^{k+1} \cdot \nabla u^k + \tau u^{k+1} u^k] dx \\ &\leq E_{\tau}(u^{k+1}) + E_{\tau}(u^k) \leq 2E_{\tau}(u^0), \end{aligned}$$

while  $\int_{\Omega} u^k f'_i(u^k) dx \geq \int_{\Omega} f_i(u^k) dx \geq \gamma_i$ . Hence,  $\sum_{i=1}^n \mu_i^{k+1} \gamma_i \leq 2E_{\tau}(u^0)$ . Since  $\gamma_i > 0$  for every  $i$ , this gives the claimed bound for the  $\mu_i^{k+1}$ . Now (4.2) combined with the strong  $L^2(\Omega)$ -convergence of  $u^{k+1} - u \rightarrow 0$  and the  $L^2(\Omega)$ -boundedness of  $f'_i(u^k)$  yields  $\|u^{k+1} - u\|_{H_0^1} \rightarrow 0$  as  $k \rightarrow \infty$ . The proof of Theorem 4.1 is therefore complete.  $\square$

We can also give the

*Proof of Corollary 4.2.* We recall from Lemma 3.1 that

$$E_\tau(u^{k+1} - u^k) \leq E_\tau(u^k) - E_\tau(u^{k+1}),$$

and consequently that  $\|u^{k+1} - u^k\|_{H_0^1} \rightarrow 0$  as  $k \rightarrow \infty$ . Therefore, using Theorem 4.1, we can invoke the argument of [7, Proposition 2]. For the convenience of the reader, we will repeat this simple argument.

It suffices to show that if  $A(u^0)$  contains one isolated point  $u$ , then the entire sequence converges to  $u$ . If  $u$  is isolated, then we can find disjoint neighborhoods  $N_1$  containing  $u$ , and  $N_2$  containing all other elements of  $A(u^0)$ . Let  $2\varepsilon = \text{dist}(N_1, N_2)$ . There is an integer  $k_1$  so that  $\|u^{k+1} - u^k\|_{H_0^1} \leq \varepsilon$  for all  $k \geq k_1$ . Also, there is an integer  $k_2$  so that for  $k \geq k_2$ ,  $u^k$  belongs to either  $N_1$  or  $N_2$ . For otherwise, we can find a subsequence of  $u^k$  lying in the complement of  $N_1 \cup N_2$ , contradicting Theorem 4.1. Since  $u \in A(u^0)$ , we can find  $u^k \in N_1$  with  $k \geq \max(k_1, k_2)$ . But then  $u^{k+1} \in N_1$ , since  $\text{dist}(u^{k+1}, N_2) \geq \varepsilon$ . By induction,  $u^j \in N_1$  for  $j \geq k$ , and the proposition is proved.  $\square$

The remainder of the section will be devoted to the proof of Lemma 4.3.

*Proof of Lemma 4.3.* We first estimate  $|\lambda_1|, \dots, |\lambda_{n-1}|$ . Recalling that  $\Lambda(s) = \sum_{i=1}^n \lambda_i f'_i(s)$  and  $\Lambda_i = \Lambda(\sigma_i)$ ,  $i = 0, \dots, n - 1$ , we have

$$(4.3) \quad \begin{aligned} \Lambda_0 &= 0 \quad \text{and} \quad \Lambda_i = \Lambda_{i-1} + \lambda_i \Delta \sigma_i, \\ \Lambda(s) &= \Lambda_{i-1} + \lambda_i (s - \sigma_{i-1}) \quad \text{on} \quad [\sigma_{i-1}, \sigma_i]. \end{aligned}$$

Since  $\lambda_i = (\Lambda_i - \Lambda_{i-1})/\Delta \sigma_i$ , it suffices to bound  $|\Lambda_1|, \dots, |\Lambda_{n-1}|$ . We fix the index  $i$  so that

$$(4.4) \quad |\Lambda_i| = \max_{1 \leq j \leq n-1} |\Lambda_j|.$$

Then

$$(4.5) \quad |\Lambda(s)| \geq \frac{1}{3} |\Lambda_i| \quad \text{for} \quad s \in I := [\sigma_i - \frac{1}{3} \Delta \sigma_i, \sigma_i - \frac{1}{6} \Delta \sigma_i],$$

since  $|\lambda_j| \leq 2|\Lambda_j|/\Delta \sigma_j$ .

Let  $\phi(s)$  be a Lipschitz cutoff function with  $\phi(s) = 1$  for  $s \in I$ ,  $\phi(s) = 0$  for  $s \leq \bar{\sigma}_i = \frac{1}{2}(\sigma_{i-1} + \sigma_i)$  and  $s \geq \sigma_i$ , and  $0 \leq \phi(s) \leq 1$ . Using  $\phi(u)$  as a test function in (4.1), we have the identity

$$(4.6) \quad \int_{\Omega} \phi'(u) |\nabla u|^2 dx = \int_{\Omega} \phi(u) \Lambda(u) dx.$$

Since  $\Lambda(s)$  has one sign on the support of  $\phi(s)$ , it follows from (4.6) that

$$(4.7) \quad \begin{aligned} \int_{\{u \in I\}} |\Lambda(u)| dx &\leq \int_{\Omega} \phi(u) |\Lambda(u)| dx \leq \int_{\Omega} |\phi'(u)| |\nabla u|^2 dx \\ &\leq \frac{6}{\Delta \sigma_i} \int_{\Omega} |\nabla u|^2 dx = 12(\Delta \sigma_i)^{-1} E(u). \end{aligned}$$

Combining (4.7) and (4.5), we have

$$(4.8) \quad |\Lambda_i| \leq \frac{36E(u)}{(\Delta\sigma_i)|\{u \in I\}|}.$$

Thus, to estimate  $|\Lambda_i|$ , we need a positive lower bound for  $|\{u \in I\}|$ .

Let  $\sigma = \sigma_i - \frac{1}{3}\Delta\sigma_i$  and  $\sigma' = \sigma_i - \frac{1}{6}\Delta\sigma_i$ , so that  $I = [\sigma, \sigma'] \subset (\sigma_{i-1}, \sigma_i)$  with  $|I| = \frac{1}{6}\Delta\sigma_i$  for some  $1 \leq i \leq n - 1$ . Define

$$f_I(s) = \frac{1}{2}(s - \sigma)_+^2 - \frac{1}{2}(s - \sigma')_+^2.$$

Then

$$(4.9) \quad \|f'_I(u)\|_{L^{N/(N-1)}} \geq |I|\{u > \sigma_{n-1}\}^{(N-1)/N}.$$

On the other hand, by the Sobolev inequality,

$$(4.10) \quad \|f'_I(u)\|_{L^{N/(N-1)}} \leq C_N \int_{\Omega} f''_I(u)|\nabla u| dx \leq C_N |\{u \in I\}|^{1/2} E(u)^{1/2},$$

and by the usual Poincaré inequality applied to  $(u - \sigma_{n-1})_+$ ,

$$(4.11) \quad 2F_n(u) = \|(u - \sigma_{n-1})_+\|_{L^2}^2 \leq C_N |\{u > \sigma_{n-1}\}|^{2/N} E(u).$$

Combining (4.9)–(4.11) gives the required lower bound

$$(4.12) \quad |\{u \in I\}| \geq C_N |I|^2 F_n(u)^{N-1} E(u)^{-N}.$$

Together with (4.11), this gives the desired estimate

$$(4.13) \quad |\Lambda_i| \leq C_N E(u)^{N+1} F_n(u)^{1-N} (\Delta\sigma_i)^{-3} \leq C_N E_0^{N+1} \gamma_n^{1-N} (\Delta\sigma)^{-3}.$$

To complete the proof, we need only to bound  $|\lambda_n|$ . We note that

$$-\Delta u = \Lambda_{n-1} + \lambda_n(u - \sigma_{n-1}) \quad \text{in } \{u > \sigma_{n-1}\}.$$

Thus, using  $(u - \sigma_{n-1})_+$  as a test function, we obtain the identity

$$(4.14) \quad \int_{\Omega} |\nabla(u - \sigma_{n-1})_+|^2 dx = \Lambda_{n-1} \int_{\Omega} (u - \sigma_{n-1})_+ dx + 2\lambda_n F_n(u).$$

Therefore,

$$(4.15) \quad \begin{aligned} \gamma_n |\lambda_n| &\leq E_0 + \frac{1}{2} |\Lambda_{n-1}| \int_{\Omega} u dx \\ &\leq E_0 + C_N |\Lambda_{n-1}| |\Omega|^{1/2+1/N} E_0^{1/2}. \end{aligned}$$

Recalling (4.13), this gives the desired estimate

$$(4.16) \quad |\lambda_n| \leq C(\Omega, E_0) \gamma_n^{-N} (\Delta\sigma)^{-3}.$$

This completes the proof of Lemma 4.3.  $\square$

### 5. GENERALIZATIONS

In this section we give three separate extensions of the prototype problems examined in the preceding sections. Since each extension is fairly straightforward, our discussion here will be brief.

First, we introduce a *free boundary* into problems  $(P_\infty)$  and  $(P_n)$ . This is accomplished simply by imposing the constraints (2.4)(c) only on the interval  $\sigma_0 \leq \sigma < +\infty$  for some given  $\sigma_0 > 0$ . Thus, the problem  $(P_\infty)$  becomes

$$(5.1) \quad E(u) \rightarrow \min \quad \text{subject to} \quad \int_{\Omega} (u - \sigma)_+ dx = \beta(\sigma), \quad \sigma_0 \leq \sigma < +\infty.$$

Similarly, the problem  $(P_n)$  is defined with respect to a partition  $0 < \sigma_0 < \sigma_1 < \dots < \sigma_{n-1} < \sigma_n = +\infty$ , and the constraints for  $(P_n)$  are defined by (2.8), (2.9) as before. Now, however,  $f_i(s) = 0$  for all  $s \leq \sigma_0$  and  $i = 1, \dots, n$ . The validity of Theorem 2.2 remains unaltered in this extension, and a solution  $u$  of  $(P_n)$  satisfies  $-\Delta u = \Lambda(u)$  in  $\Omega$ ,  $u = 0$  on  $\partial\Omega$ , where

$$\Lambda(s) = \begin{cases} 0 & \text{if } s < \sigma_0, \\ \sum_{i=1}^n \lambda_i f'_i(s) & \text{if } s \geq \sigma_0. \end{cases}$$

This may be interpreted as the equivalent free-boundary problem

$$(5.2) \quad \begin{aligned} -\Delta u &= \begin{cases} 0 & \text{in } \{0 < u < \sigma_0\}, \\ \Lambda(u) & \text{in } \{u > \sigma_0\}, \end{cases} \\ |\nabla u| &\text{ is continuous across } \{u = \sigma_0\}, \end{aligned}$$

the latter being the free-boundary condition. As a simple illustration, we note that the case of one constraint ( $n = 1$ ) yields

$$\begin{cases} -\Delta u = \lambda_1(u - \sigma_0)_+ & \text{in } \Omega, & u = 0 & \text{on } \partial\Omega, \\ \int_{\Omega} (u - \sigma_0)_+^2 dx = 2\gamma_1, \end{cases}$$

which is a familiar free-boundary problem.

The algorithm  $(A_n)$  constructed in §3 and its convergence properties established in §4 admit immediate generalizations in the extension just described. Indeed, the only essential changes in the above development needed to include the free boundary into  $(P_n)$  can be summarized as follows. A term  $\tau G_\tau(u^k - \sigma_0)_-$  is subtracted from the right-hand side of (3.14), and a corresponding term  $\tau(f'_i(v), G_\tau(v - \sigma_0)_-)$  is added to  $c_i(v)$  in (3.12);  $a_{ij}$  is unchanged. With these modifications in place, the theory described above extends to the free-boundary case.

Next, we consider an extension of  $(P_\infty)$  and  $(P_n)$  (without free boundaries) which allows  $u$  and  $\bar{u}$  to change sign in  $\Omega$ . The requirement that  $\bar{u} \geq 0$  in  $\Omega$  is now relaxed, and the constraints in  $(P_\infty)$  are replaced by

$$\begin{cases} \int_{\Omega} (u - \sigma)_+ dx = \beta^+(\sigma) & \text{for } 0 \leq \sigma < +\infty, \\ \int_{\Omega} (u - \sigma)_- dx = \beta^-(\sigma) & \text{for } -\infty < \sigma \leq 0, \end{cases}$$

where  $\beta^+$ ,  $\beta^-$  are defined by  $\bar{u}$  (so that the above holds with  $\bar{u}$  substituted for  $u$ ). The extension of  $(P_n)$  is then defined with respect to two partitions  $0 = \sigma_0^+ < \sigma_1^+ < \dots < \sigma_{n-1}^+ < \sigma_n^+ = +\infty$  and  $0 = \sigma_0^- > \sigma_1^- > \dots > \sigma_{m-1}^- > \sigma_m^- = -\infty$  for some  $n$  and  $m$ . The functionals  $F_i^+$  ( $i = 1, \dots, n$ ) and  $F_j^-$

( $j = 1, \dots, m$ ) corresponding to these partitions are defined as before by the convex functions

$$f_i^+(s) = \frac{1}{2}(s - \sigma_{i-1}^+)^2 - \frac{1}{2}(s - \sigma_i^+)^2,$$

$$f_j^-(s) = \frac{1}{2}(s - \sigma_{j-1}^-)^2 - \frac{1}{2}(s - \sigma_j^-)^2;$$

also, the constraint values are taken to be

$$\gamma_i^+ = \int_{\sigma_{i-1}^+}^{\sigma_i^+} \beta^+(\sigma) d\sigma, \quad \gamma_j^- = \int_{\sigma_j^-}^{\sigma_{j-1}^-} \beta^-(\sigma) d\sigma.$$

The constraints for  $(P_n)$  are therefore replaced by the  $n + m$  constraints  $F_i^+(u) = \gamma_i^+$ ,  $F_j^-(u) = \gamma_j^-$ , and so an analogous variational problem is defined, which we shall call  $(P_{n,m})$ . If each  $\gamma_i^+ > 0$  and each  $\gamma_j^- > 0$ , then the analogue of Theorem 2.2 holds, and a solution of  $(P_{n,m})$  satisfies  $-\Delta u = \Lambda(u)$  in  $\Omega$ ,  $u = 0$  on  $\partial\Omega$ , where now

$$\Lambda(s) = \sum_{i=1}^n \lambda_i^+ (f_i^+)'(s) + \sum_{j=1}^m \lambda_j^- (f_j^-)'(s),$$

a piecewise linear function on  $-\infty < s < +\infty$  which is increasing whenever  $\lambda_i^+ \geq 0$  and  $\lambda_j^- \geq 0$  for every  $i$  and  $j$ .

Again, the algorithm and its convergence properties require no essential changes to handle the problem  $(P_{n,m})$ .

Finally, we remark that a variable-coefficient version of  $(P_n)$  can be treated as a straightforward extension of the prototype problem. In this version the objective and constraint functionals are taken to be

$$E(u) = \frac{1}{2} \int_{\Omega} \left\{ \sum_{p,q=1}^N a_{pq}(x) u_{x_p} u_{x_q} + a(x) u^2 \right\} dx,$$

$$F_i(u) = \int_{\Omega} b(x) f_i(u) dx,$$

where the functions  $a_{pq}(x)$ ,  $a(x)$ ,  $b(x)$  belong to  $C^\alpha(\bar{\Omega})$  ( $0 < \alpha < 1$ ) and satisfy the conditions

$$\sum_{p,q=1}^N a_{pq}(x) \xi_p \xi_q \geq \theta |\xi|^2 \quad \text{for all } \xi \in \mathbb{R}^N,$$

$$a(x) \geq 0, \quad b(x) \geq \theta',$$

uniformly for  $x \in \Omega$  for some positive constants  $\theta$  and  $\theta'$ . Then a solution of this version of  $(P_n)$  solves the nonlinear elliptic eigenvalue problem

$$-\sum_{p,q} \frac{\partial}{\partial x_p} \left( a_{pq}(x) \frac{\partial u}{\partial x_q} \right) + a(x)u = b(x)\Lambda(u) \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial\Omega,$$

where  $\Lambda(s)$  is as before.

It is straightforward to check that our development in §§2–4, as well as the two preceding extensions, can be carried out in this more general context.

6. APPENDIX 1. VARIATIONAL PROBLEMS IN MAGNETOHYDRODYNAMICS

Equilibrium problems in magnetohydrodynamics (MHD) supply some of the main examples of multiconstrained variational problems of the kind that we study in this paper. These model problems in plasma physics arise both in controlled thermonuclear fusion research and in astrophysics. The classical approach to solving these problems makes use of a general variational principle which characterizes equilibrium configurations of a plasma assumed to be governed by ideal MHD. This principle depends upon determining the complete family of quantities (expressed as volume integrals) which are conserved under the governing evolution equations. Then it identifies an equilibrium as a minimizer of the total energy over the class of configurations which maintain given values of all of the other conserved quantities. Although this general characterization can be stated formally for fully three-dimensional configurations, we shall consider it only under an assumption of spatial symmetry—either two-dimensionality, axial symmetry, or helical symmetry—since then it takes a simpler and more tractable form. This particular form of the variational principle for symmetric equilibrium configurations has been given by Woltjer [21, 22], who has also determined the complete family of conserved quantities. (He emphasizes the axially symmetric case, but his analysis can be modified to apply as well to the other symmetric cases.) We shall briefly describe the case of magneto-static equilibrium (for which there is no mass flow) in two dimensions; we shall leave aside the minor technical modifications needed to treat axial or helical symmetry, since, even though these cases are important in real applications, they are the same conceptually.

We begin by posing an abstract variational problem similar to  $(P_\infty)$  which encompasses the static physical problems of interest as special cases. Let  $\Omega \subseteq \mathbb{R}^2$  be the cross-section of a cylindrical domain  $\Omega \times \mathbb{R}$ , let  $(x_1, x_2)$  denote the variable point in  $\Omega$ , and let  $x_3$  be the ignorable coordinate. Consider the minimization problem

$$(6.1) \quad \begin{cases} \int_{\Omega} [\frac{1}{2} |\nabla u|^2 + h_1(v_1) + h_2(v_2)] dx \rightarrow \min & \text{over} \\ \int_{\Omega} v_1 (u - \sigma)_+ dx = \beta_1(\sigma), \\ \int_{\Omega} v_2 (u - \sigma)_+ dx = \beta_2(\sigma), \end{cases} \quad (\sigma_0 \leq \sigma < +\infty),$$

where the admissible triple  $(u, v_1, v_2)$  belongs to  $H_0^1(\Omega) \times L^{r_1}(\Omega) \times L^{r_2}(\Omega)$  for some  $1 < r_1, r_2 < +\infty$ . The given functions  $h_1$  and  $h_2$  are assumed to be smooth and strictly convex with  $h_l(0) = h_l'(0) = 0$  and  $h_l(z) = O(|z|^{r_l})$  as  $|z| \rightarrow \infty$  ( $l = 1, 2$ ). As in  $(P_\infty)$ , the two infinite families of constraints are parametrized by  $\sigma \in [\sigma_0, +\infty)$ , and  $\beta_1(\sigma)$  and  $\beta_2(\sigma)$  are given data.

The physical interpretation of (6.1) is as follows. The magnetic field  $\mathbf{B} = (B^1, B^2, B^3)$ , which is independent of  $x_3$ , satisfies  $\nabla \cdot \mathbf{B} = 0$  in  $\Omega \times \mathbb{R}$  and

hence admits a representation  $\mathbf{B} = (u_{x_2}, -u_{x_1}, v_1)$ , where  $u$  is the flux function (or stream function) for its poloidal part and  $v_1$  is its toroidal part. The magnetic energy density (per unit volume) is then  $\frac{1}{2}|\mathbf{B}|^2 = \frac{1}{2}|\nabla u|^2 + \frac{1}{2}v_1^2$ . The mass density  $\rho$  of the plasma is represented by  $v_2$ . The internal energy density (per unit volume) is given by  $\rho^\gamma/(\gamma - 1)$ , in accordance with the polytropic law  $p = \rho^\gamma$  with  $p$  denoting pressure. Therefore, the objective functional represents total (potential) energy when we put

$$(6.2) \quad h_1(v_1) = \frac{1}{2}v_1^2, \quad h_2(v_2) = \frac{v_2^\gamma}{\gamma - 1}.$$

The interpretation of the constraints relies on differentiating them with respect to the parameter  $\sigma$ , for then there results

$$\int_{\{u>\sigma\}} v_1 dx = -\beta'_1(\sigma), \quad \int_{\{u>\sigma\}} v_2 dx = -\beta'_2(\sigma).$$

All of these integrals are extended over the interior of a (cylindrical) flux surface  $\{u = \sigma\}$ —that is, a flux tube  $\{u > \sigma\}$ . But ideal MHD requires that, in evolution, each flux tube must move with the flow preserving its strength and mass. It is readily verified that the above integrals are, respectively, the toroidal flux and mass of the flux tube  $\{u > \sigma\}$ , and hence they are conserved quantities. (The conservation of poloidal fluxes is implicit in the parametrization which uses the values of the flux function  $u$ .)

As is shown in Lemma 2.1, the constraints in (6.1) imply corresponding constraints on all integrals of the form

$$\int_{\Omega} v_i \phi(u) dx,$$

where  $\phi$  is an arbitrary continuous function on  $[\sigma_0, +\infty)$ . Indeed, the latter integrals are used to express the family of conserved quantities in [21, 22]. Roughly speaking, the functions  $(s - \sigma)_+$  ( $\sigma_0 \leq \sigma < +\infty$ ) are chosen here as a particularly useful basis for the space of all functions  $\phi(s)$ . This allows us to give the above physical interpretation to these integrals. Moreover, this also permits us to discretize the constraints in precisely the same way as the prototype problem  $(P_n)$  is formed from  $(P_\infty)$ . Recalling the definition (2.9), we replace the infinite family of constraints in (6.1) by

$$\begin{cases} \int_{\Omega} v_1 f_i(u) dx = \gamma_{1i}, \\ \int_{\Omega} v_2 f_i(u) dx = \gamma_{2i}, \end{cases} \quad (i = 1, \dots, n)$$

relative to a (fixed) partition  $0 \leq \sigma_0 < \sigma_1 < \dots < \sigma_{n-1} < \sigma_n = +\infty$ . The resulting discretized version of the minimization problem (6.1) is now accessible to analysis and computation. Its variational equations can be calculated formally to be

$$(6.3) \quad \begin{cases} -\Delta u = v_1 \Phi'_1(u) + v_2 \Phi'_2(u), \\ h'_1(v_1) = \Phi_1(u), \\ h'_2(v_2) = \Phi_2(u), \end{cases}$$



where, by virtue of the Lagrange multiplier rule,

$$(6.4) \quad \Phi_l(s) := \sum_{i=1}^n \lambda_{li} f_i(s) \quad (l = 1, 2)$$

for some multipliers  $\lambda_{li}$ . The observation that  $v_1$  and  $v_2$  enter into these equilibrium equations algebraically suggests immediately that they be eliminated. This yields an equation for  $u$  alone:

$$(6.5) \quad -\Delta u = (h'_1)^{-1}(\Phi_1(u))\Phi'_1(u) + (h'_2)^{-1}(\Phi_2(u))\Phi'_2(u),$$

which is a generalized form of the familiar Grad-Shafranov equation. In order to express this equation more plainly, it is useful to introduce the convex conjugate functions to  $h_l$  ( $l = 1, 2$ ) defined by  $h_l^*(z^*) = \sup_z z z^* - h_l(z)$ . Then (6.5) can be written simply as

$$(6.6) \quad -\Delta u = P'_1(u) + P'_2(u) \quad \text{with } P_l(s) := h_l^*(\Phi_l(s)).$$

Returning to the physical case (6.2), we find that

$$h_1^*(z^*) = \frac{1}{2}(z^*)^2, \quad h_2^*(z^*) = \left(\frac{\gamma-1}{\gamma} z^*\right)^{\gamma/(\gamma-1)},$$

and consequently we obtain the relations

$$P_1 = \frac{1}{2}v_1^2, \quad P_2 = v_2^\gamma.$$

It follows from these relations that the profile functions  $P'_1$  and  $P'_2$  in (6.6) have the usual interpretations as current and pressure profiles, respectively, in the Grad-Shafranov equation [3, Chapter 4].

Of course, the principal novelty inherent in (6.6) lies in the fact that the functions  $P'_l$  are not specified, but are instead determined along with the solution  $u$  through (6.4). We see therefore that (6.6) is a "generalized (or queer) differential equation" (GDE) in the sense of Grad et al. [12, 13]. However, our viewpoint differs from that proposed by Grad, who viewed a GDE as a highly implicit equation for  $u$  which combined the elliptic operation  $\Delta$  and differentiation with respect to the volume variable  $\alpha = \alpha(\sigma) = |\{u > \sigma\}|$ . Rather, we recognize (6.6) as merely the  $v_1, v_2$ -eliminated form of the standard variational equations for a solution triple  $(u, v_1, v_2)$  of the classical minimization problem (6.1). Thus, by treating that variational problem directly we obviate the need to introduce the notion of a GDE. On the other hand, the compelling reasons put forward by Grad and others for prescribing conserved quantities associated with the evolution equations governing ideal MHD in place of the current and pressure profiles in the Grad-Shafranov equation retain their validity; indeed, they provide the main justification for our multiconstrained variational approach.

The relevant physical problem also involves a free boundary, the interface between the (confined) plasma and the (surrounding) vacuum. This feature can be introduced (as in §5) by fixing  $\sigma_0 > 0$  in (6.1). Then the free boundary is  $\{u = \sigma_0\}$ , and  $\{u > \sigma_0\}$  and  $\{0 < u < \sigma_0\}$  are the plasma and vacuum regions,

respectively. The flux constant  $\sigma_0$  effectively specifies the (total) poloidal flux in the vacuum region. An additional constraint  $\int_{\Omega} v_1 dx = \gamma_{10}$  can be imposed to specify the (total) toroidal flux in the vacuum region. (This is necessary since the basis functions  $f_i(s)$  vanish for  $s \leq \sigma_0$ .) The mass density  $v_2$  will be positive in  $\{u > \sigma_0\}$  and vanish in  $\{0 < u < \sigma_0\}$  whenever the solution satisfies  $\lambda_{2i} \geq 0$  ( $i = 1, \dots, n$ ), and this condition may be expected to hold in general. The free boundary,  $\partial\{v_2 > 0\}$ , is therefore a flux surface on which pressure vanishes, as required by the physical interface conditions.

We conclude this discussion by noting some special cases of the above development. First, we consider the case of an incompressible fluid (or plasma), which we obtain by setting  $v_2 = 1$  (uniform density) in (6.1). Equivalently, we may achieve this limit case by taking the exponent  $\gamma$  to infinity; then we can verify that the compressible solutions tend to an incompressible solution in the limit. Second, we have the case of a purely poloidal magnetic field, which we get by putting  $v_1 = 0$ . This specialization of (6.1) results in a substantial simplification, since one family of constraints is dropped. Third, we combine the two preceding cases and we arrive at the prototype problem ( $P_{\infty}$ ) or its discretized form ( $P_n$ ), both posed in terms of  $u$  alone.

With this result in mind, we remark that it is possible to construct algorithms analogous to ( $A_n$ ) which iteratively solve the more general problem (6.1), at least under favorable circumstances. Moreover, a convergence theory similar to that given in §§3–5 for the prototype problem can be furnished for this general problem, although it is considerably more complicated owing to the lack of joint convexity of the constraint functionals in (6.1). Our sequel paper [8] addresses these problems and documents the results of implemented computations based on the more general algorithm.

#### BIBLIOGRAPHY

1. V. I. Arnold, *Variational principles for three-dimensional steady state flows of an ideal fluid*, Appl. Math. Mech. **29** (1965), 1002–1008.
2. ———, *Mathematical methods of classical mechanics*, Appendix 2, Springer-Verlag, New York, 1978.
3. G. Bateman, *MHD instabilities*, MIT Press, Cambridge, Mass., 1978.
4. T. B. Benjamin, *Impulse, flow force and variational principles*, IMA J. Appl. Math. **32** (1984), 3–68.
5. F. Bauer, O. Betancourt, and P. Garabedian, *A computational method in plasma physics*, Springer-Verlag, Berlin, Heidelberg, and New York, 1978.
6. R. Courant and D. Hilbert, *Methods of mathematical physics*, Vol. 1, Interscience, New York, 1937.
7. A. Eydeland and J. Spruck, *The inverse power method for semilinear elliptic problems*, Non-linear Diffusion Equations and Their Equilibrium States I, MSRI series, vol. 12, Springer-Verlag, 1988.
8. A. Eydeland, J. Spruck, and B. Turkington, *A computational method for multiconstrained variational problems in magnetohydrodynamics* (in preparation).
9. A. Friedman, *Variational principles and free-boundary problems*, Wiley-Interscience, New York, 1982.

10. D. Gilbarg and N. Trudinger, *Elliptic partial differential equations of second order*, Springer-Verlag, Berlin, Heidelberg, and New York, 1977.
11. R. Glowinski, *Numerical methods for nonlinear variational problems*, Springer-Verlag, New York, 1984.
12. H. Grad, *Magnetic confinement fusion energy research*, Proc. Sympos. Appl. Math., vol. 21, Amer. Math. Soc., Providence, R.I., 1979, pp. 3–40.
13. H. Grad, P. N. Hu, and D. C. Stevens, *Adiabatic evolution of plasma equilibria*, Proc. Nat. Acad. Sci. U.S.A. **72** (1975), 3789–3793.
14. A. D. Ioffe and V. M. Tihomirov, *Theory of extremal problems*, Elsevier North-Holland, New York, 1979.
15. M. D. Kruskal and R. M. Kulsrud, *Equilibrium of a magnetically confined plasma in a toroid*, Phys. Fluids **1** (1958), 265–274.
16. P. Laurence and E. Stredulinsky, *A new approach to queer differential equations*, Comm. Pure Appl. Math. **38** (1985), 333–355.
17. T. Ch. Mouschovias, *Nonhomologous contraction and equilibria of self-gravitating magnetic interstellar clouds embedded in an intercloud medium: Star formation I, Formulation of the problem and method of solution*, Astrophys. J. **206** (1976), 753–767.
18. J. Mossino and R. Temam, *Directional derivative of the increasing rearrangement mapping and application to a queer differential equation in plasma physics*, Duke Math. J. **41** (1981), 475–495.
19. R. T. Rockafellar, *Convex analysis*, Princeton Univ. Press, Princeton, N.J., 1970.
20. R. Temam, *Monotone rearrangement of a function and the Grad-Mercier equation of plasma physics*, Recent Methods in Nonlinear Analysis and Applications (E. De Giorgi, E. Magenes, and U. Mosco, eds.), Petagora Editrice, Bologna, 1979, pp. 83–98.
21. L. Woltjer, *Hydromagnetic equilibrium, III: axisymmetric incompressible media*, Astrophys. J. **130** (1959), 400–404.
22. ———, *Hydromagnetic equilibrium, IV: axisymmetric compressible media*, Astrophys. J. **130** (1959), 405–413.

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF MASSACHUSETTS AT AMHERST,  
AMHERST, MASSACHUSETTS 01003