Multi-Constrained Variational Problems of Nonlinear Eigenvalue Type : New Formulations and Algorithms

Alexander Eydeland, Joel Spruck and Bruce Turkington
University of Massachusetts at Amherst

1 Introduction

The present work was motivated by a reexamination of some fundamental constrained variational problems that arise in equilibrium theory in fluid dynamics and Magnetohydrodynamics. These problems lead formally to nonlinear eigenvalue problems as variational equations. Indeed, it is traditional that many fundamental problems of mathematical physics can be formulated as nonlinear eigenvalue problems of the form

$$-\Delta u = \Lambda(u),\ u \in H_0^1(\Omega)$$

The profile function $\Lambda(u)$ is usually expressed in the form $\Lambda(u) = \lambda f(u)$ where $f(u)$ is a known function and λ is a Lagrange multiplier.

However, these models do not always represent well the original physical problem. For in the original variational formulation, the constraints of the equilibrium formulation are derived from the conserved quantities associated to the evolution of the dynamical equations and thus the profile function is completely unknown and must be determined implicitly from the constraints.In MHD theory,such an objection was raised by Grad who introduced so called generalized or queer differential equation involving differentiations of the distribution function of u. These equations are purely formal unless u happens to be smooth with simple level sets. Nevertheless, Grad and his colleagues introduced numerical techniques to solve these equations.

As a prototypical problem of the type that that arises in many physical problems we pose the variational problem (P_∞)

$$(1.1) \qquad E(u) = frac12 \int_\Omega (|\nabla u|^2 \, dx \ \to \min \text{ over } u \in M_\infty(\bar{u}).$$

where for a given $\bar{u} \in H_0^1(\Omega), M_\infty(\bar{u})$ is the class of functions which are equimeasurable with $\bar{u}$. For simplicity we will assume that$\bar{u}$ is non-negative.

The existence of a weak solution $u \in H_0^1(\Omega)$ of (P_∞) is straightforward . On the other hand, the construction of an appropriate variational equation (that is a Lagrange multiplier rule to determine the profile function $\Lambda(u)$) satisfied by such a solution is unclear. Also unclear is the question of the optimal regularity of the solution. The main source of difficulty in investigating these questions lies in the nature of the family of constraints. Our approach to this difficulty is to formulate a clear and conceptually simple classical variational problem (P_n) which replaces the family of constraints in (P_∞) by a finite family of simpler constraints, and which approximates (P_∞) when n is large. Besides being more tractable analytically, the problem (P_n) is also amenable to a numerical method of solution, the development of which is a principal goal in this paper. One of the key points of our approach is the construction of a finite element or constraint space in which the basis functions have very special properties.

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

$$(1.2) \qquad F_i(u) := \int_\Omega f_i(u) dx \ \ (i = 1, \ldots, n) \ \text{ where}$$

$$(1.3) \qquad f_i(s) := \int_{\sigma_{i-1}}^{\sigma_i} (s - \sigma)_+ \, d\sigma = l/2 \ (s - \sigma_{i-1})_+^2 - l/2 \ (s - \sigma_i)_+^2$$

for $s \in$ (we write $s_+ = \max\{s, 0\}$ and $s_+^2 = (s_+)^2$.) Then we let

$$(1.4) \qquad M_n(\bar{u}) = \{u \in H_0^1(\Omega) \colon \ F_i(u) = F_i(\bar{u}) \text{ for all } i = 1, \ldots, n\}.$$

The problem (P_n) is the following multiconstrained minimization problem which may be viewed as a natural discretization of (P_∞)

$$(1.5) \qquad E(u) \to \min \text{over } u \in M_n(\bar{u}).$$

Equivalently, (P_n) may be expressed as

$$E(u) \to \min \text{ subject to } F_i(u) = \gamma_i \quad (i = 1, \ldots, n), \tag{1.6}$$

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 following result.

If $\bar{u}$ is given so that $\gamma_i > 0$ for every i, then there exists a non-negative minimizer $u \in C^{2,\beta}$ for (P_n) satisfying

$$-\Delta u = \sum_{i=1}^{n} \lambda_i f_i^{'}(u) \text{ for some } \lambda_i \in .$$

One point that should be emphasized is that our finite constraint problem yields exact solution of the relevant physical governing equations(for example, the ideal MHD equations) and is an approximation only in the sense that the full constraint space is not prescribed. Our approach is very different to that of Laurence and Stredulinisky who introduce an approximation to the Grad G.D.E. model that doesn't enjoy this property.

As we have mentioned earlier, one of the principal goals of this paper is to develop an efficient and justifiable numerical method to compute the solutions of problem (P_n). To this end, we introduce an iterative procedure which is designed to converge to the solutions of P_n. Our algorithm A_n may be succinctly defined as follows:

Given $u^0 \in M_n(\bar{u})$, u^{k+1} is the unique solution of the variational inequality

$$E_\tau(u) = E(u) + \frac{\tau}{2}\langle u, u\rangle \to \min \text{subject to} \tag{1.7}$$

$$F_i(u^k) + \langle F_i'(u^k),\ u - u^k\rangle \geq \gamma_i,\ i = 1, \ldots, n \tag{1.8}$$

which is defined in terms of a positive parameter τ and with the admissible functions u lying in the constraint space $M_n(\bar{u})$ defined by the n linearized inequality constraints.

For a discussion of the main ideas underlying our construction of algorithm A_n the reader is referred to Appendix2. We do want to emphasize here that the algorithm is designed to take maximal advantage of the convexity of the basis functions $f_i(s)$ even though the variational problem P_n is highly non-convex.

Perhaps the main result of the paper is the following

Given any initialization $u^0 \in M_n(\bar{u})$*, there exists a constant* $C = C(\Omega,\ E(u^0),\ \gamma_1^{-1} \dots \gamma_n^{-1}, (\Delta\sigma)^{-1})$ *so that if* $\tau > C$*, the sequence* u^k *generated by the algorithm* A_n *converges strongly in* $H_0^1(\Omega)$ *to the set* $\mathcal{S}$ *of critical points of the variational problem* P_n*, in the sense that* $dist(u^k, \mathcal{S}) \to 0$ *as* $k \to \infty$ *where* $dist(u, \mathcal{S}) = inf\,\{||u - v||_{H_0^1(\Omega)}\colon v \in \mathcal{S}\}$)

We now discuss the organization of the paper. In Sections 2-4, we restrict our discussion to a prototypical case of the class of variational problems that hold our attention throughout the paper. In Section 2 we formulate the prototype finite constraint problem (P_n), and establish its relation to the problem$(P_\infty$. . Sections 2 contains the proof of Theorem 1.1 and illustrates in a very clear and concrete way, some of the fundamental ideas of classical optimization and duality theory. In particular, Lemma 2.3 and its relation to the linear independence of the constraints is of independent interest. In Section 3 we describe the algorithm $(A_n)$for solving (P_n), which is a globally convergent iterative procedure. The basic convergence properties of this algorithm are derived and a more concrete description of the algorithm is given in terms of the so called dual variational problem. Section 4 contains the proof of the convergence of the algorithm to a solution of our original problem. The key step in the proof of the convergence is an apriori estimate (Lemma 4.3) which is of independent interest. In Section 5, we briefly indicate how to extend the development of Sections 2-4 to include a free boundary interface, to apply to distributions which change sign, and finally to apply to a more general class of variational problems. Section 6(Appendix 2)contains a brief description of the fundamental variational problems in MHD equilibrium theory which partially motivated our prototype problems. In future work we hope to elaborate on the method and ideas of the present paper to study the more general problems of Appendix 1. In Section 7(Appendix 1) we collect some remarks concerning (A_n) which help to clarify the construction of this algorithm.

The following conventional notations will be used in the sequel. Let $\Omega \subseteq {}^N (N = 1, 2, \dots)$ be a bounded open set whose boundary, $\partial\Omega$, is smooth enough (say $C^{1+\beta}$). We denote as usual Lebesgue measure on Ω by $dx = dx_1 \dots dx_N$, $x = (x_1, \dots, x_N) \in \Omega$, 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

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

The Sobolev space $H^m(\Omega)(m = 1,2,\ldots)$ is identified as 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 $\langle .,.\rangle$. The class of continuous functions on $\bar{\Omega} = \Omega \cup \partial\Omega$ is denoted by $C^0(\bar{\Omega})$, and $C^m(\bar{\Omega}) \subset C^0(\bar{\Omega})(m = 1,2,\ldots)$ denotes the class of m times continuously differentiable functions. Also, $C^{m,a}(\bar{\Omega})(0 < \alpha \leq 1)$ is the subclass of $C^m(\bar{\Omega})$ consisting of those functions whose derivatives of order $= m$ are Hölder continuous with exponent α; thus, $C^{0,1}(\bar{\Omega})$ denotes the class of Lipschitz continuous functions on $\bar{\Omega}$. All of these notations are adopted from the standard reference [**?**], where the reader is referred for further details, if necessary.

2 Formulation of problems (P_n) and (P_∞).

Let Ω be a domain in ^{n}. We begin by posing the problem (P_∞), which we adopt as a model for the type of variational problems presented in Appendix 1 (where the reader is referred 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, we let

$$(2.1) \qquad M_\infty(\bar{u}) = \{u \in H_0^1(\Omega) \colon \mid \{u > \sigma\} \mid = \mid \{\bar{u} > \sigma\} \mid \text{ for all } \sigma \in \}.$$

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

$$(2.2) \qquad u_*(\alpha) \colon= \inf\{\sigma \in \colon \mid \{u > \sigma\} \mid > \alpha\} \ (0 \leq \alpha \leq |\Omega|).$$

Then we see that

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

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

Lemma 2.1 *The following statements are equivalent:*

$$(2.4) \qquad \begin{array}{ll} (a) & u \in M_\infty(\bar{u}) \\ (b) & \int_\Omega \phi(u) dx = \int_\Omega \phi(\bar{u})\, dx \text{ for all } 0 \leq \phi \in C_0() \\ (c) & \int_\Omega (u-\sigma)_+ \, dx = \int_\Omega \bar{u} - \sigma)_+ \, dx \text{ for all } \sigma \in . \end{array}$$

Proof. That (a) implies (b) is standard; (b) is equivalent to (a) when functions of the form $\phi(s) = \{ {}^{0 \ s \leq \sigma}_{1 \ s > \sigma}$ are used, 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 σ. Precisely, the right derivative with respect to σ yields the desired result since

$$t^{-1}[(s-\sigma-t)_+ - (s-\sigma)_+] \uparrow \{ {}^{0 \ s \leq \sigma}_{1 \ \leq \sigma} \text{ as } \tau \downarrow 0\};$$

then, by virtue of monotone convergence, it follows that

$$\lim_{t\to 0^+} t^{-1}\{\int_\Omega (u-\sigma-t)_+ dx - \int_\Omega (u-\sigma)_+ dx\} = |\{u>\sigma\}|$$

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

The problem (P_∞) is a constrained minimization problem whose objective functional is

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

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 Ω, we pose the variational problem (P_∞)

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

We assume that $\bar{u}$ is nonnegative for the sake of simplicity in the exposition here, deferring the general case until §4. Equivalently, (P_∞) may be written explicitly in terms of its constraints as (P_∞)

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

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_∞) is straightforward (see Theorem 2.). On the other hand, the construction of an appropriate variational equation (such as 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 (optional) regularity of the solution. The main source of difficulty in investigating these questions is, of course, the nature of the family of constraints. With this in mind, we therefore proceed to formulate the problem (P_n) which replaces the family of constraints in (P_∞) by a finite family of simpler constraints, and which approximates (P_∞) when n is large. Besides being more tractable analytically, the problem (P_n) is also amenable to a numerical method of solution, the development of which is a principal goal in this paper. Thus, (P_n) may be viewed as the natural discretization of (P_∞).

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

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

$$f_i(s) := \int_{\sigma_{i-1}}^{\sigma_i} (s-\sigma)_+ \, d\sigma = l/2 \; (s-\sigma_{i-1})_+^2 - l/2 \; (s-\sigma_i)_+^2 \tag{2.9}$$

for $s \in$. (We write $s_+ = \max\{s, 0\}$ and simply $s_+^2 = (s_+)^2$ for the sake of brevity.) Then we let

$$M_n(\bar{u}) = \{u \in H_0^1(\Omega)\colon \; F_i(u) = F_i(\bar{u}) \text{ for all } i = 1, \ldots, n\}. \tag{2.10}$$

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 $\max_{1\le i\le n-1} (\sigma_i - \sigma_{i-1}) \to 0$ and $\sigma_{n-1} \to +\infty$ as $n \to +\infty$, then $M_\infty(\bar{u}) = \cap_n M_n(\bar{u})$, the class $M_n(\bar{u})$ being defined by the n-th partition in the sequence. (If $\bar{u} \in C^0(\bar{\Omega})$, say, then it is not necessary that $\sigma_{n-1} \to +\infty$, but rather it suffices that $\sigma_{n-2} \le \sup_\Omega \bar{u} < \sigma_{n-1}$.)

The problem (P_n) is the following multiconstrained minimization problem

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

Equivalently, (P_n) may be expressed as

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

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 $\bar{u}$ is given so that $\gamma_i > 0$ for every i, then there exists a minimizer u for (P_n) satisfying*

$$\begin{aligned} &(a) \quad u \in H_0^1(\Omega) \cap C^{2,\beta}(\bar{\Omega}), \\ &(b) \quad u \ge 0 \text{ in } \Omega, \\ &(c) \quad -\Delta u = \sum_{i=1}^n \lambda_i f_i'(u) \text{ for some } \lambda_i \in . \end{aligned} \tag{2.13}$$

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

$$E(u_j) = \inf \langle E(\tilde{u})\colon \tilde{u} \in M_n(\bar{u})\rangle.$$

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

$$E(u) \leq \lim_{j\to\infty} E(u_j) = \inf \langle E(\tilde{u})\colon \tilde{u} \in M_n(\bar{n})\rangle,$$

$$F_i(u) = \lim_{j\to\infty} F_i(u_j) = \gamma_i \text{ for each } i = 1, \ldots, n,$$

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 (2.9b) 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 (2.9a,c) must be nonnegative.)

The variational equation (2.9c) satisfied by u is the standard Lagrange multiplier rule. As is shown in the reference text [**?**], it suffices that the functionals $E, F_i\colon H_0^1(\Omega) \to$ be C^1 and that the gradients $F_1'(u), \ldots, F_n'(u)$ be linearly independent. It is immediate from their definitions that the functionals involved are Fréchet differentiable with

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

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

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

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

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

Lemma 2.3 *Let $f(s) \in C^1[0, \infty) \cap C_p^2(0, \infty)$ satisfy $f(0) = f'(0) = 0$, $f''(s) \geq 0$. Then for any $0 \leq u \in H_0^1(\Omega)$*

(2.16) $$\int_\Omega f(u) \; dx \; \leq c|\Omega|^{\frac{1}{2}} \; ||f''||_{L^2(\Omega)} \int_\Omega |\nabla u|^2 \; dx \quad \textit{if } N = 2$$

$$\int_\Omega f(u)\,dx \;\le c||f''||_{L^{\frac{N}{2}}(\Omega)} \int_\Omega |\nabla u|^2\,dx \quad \text{if } N\ge 2 \tag{2.17}$$

Proof. Suppose $n \ge 2 > p \ge \frac{2n}{n+4}$. Then by convexity of f,

$$\int_\Omega f(u)\,dx \;\le\; \int_\Omega uf'(u)\,dx \;\le\; ||u||_{L^{\frac{np}{np-n+p}}(\Omega)}\,||f'(u)||_{L^{\frac{np}{n-p}}(\Omega)} \tag{2.18}$$

and by the Sobolev inequality

$$\begin{aligned} ||f'(u)||_{L^{\frac{np}{n-p}}(\Omega)} &\le\; c\,||\nabla(f'(u)||_{L^p(\Omega)} \\ &\le\; c\,||f''(u)||_{L^{\frac{2p}{2-p}}(\Omega)}\,||\nabla u||_{L^2(\Omega)} \end{aligned} \tag{2.19}$$

If $N=2$ we take $p=1$ and obtain from (2.18) (2.19)

$$\begin{aligned} &\textstyle\int_\Omega f(u)\,dx \;\le\; c||u||_{L^2(\Omega)}\,||f''(u)||_{L^2(\Omega)}\,||\nabla u||_{L^2(\Omega)} \\ &\le\; c||f''(u)||_{L^2(\Omega)}\;\textstyle\int_\Omega |\nabla u|^2\,dx \end{aligned}$$

proving (2.16). If $N>2$, we choose $p=\frac{2N}{N+4}$ and obtain from (2.18) (2.19),

$$\begin{aligned} &\textstyle\int_\Omega f(u)\,dx \;\le\; c||u||_{L^{\frac{2N}{N-2}}(\Omega)}\,||f''(u)||_{L^{\frac{N}{2}}(\Omega)}\,||\nabla u||_{L^2(\Omega)} \\ &\le\; c||f''(u)||_{L^{\frac{N}{2}}(\Omega)}\;\textstyle\int_\Omega |\nabla u|^2\,dx \end{aligned}$$

□

Indeed, applying Lemma (2.3) to $f_i(u)$ establishes 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 $\gamma_1^{-1},\ldots,\gamma_n^{-1}$, $E(u)$ and $|\Omega|$).

Thus, if for some $c_i \in$ the identity $0=\sum_{i=1}^n\; c_i f_i'(u)$ holds a.e. in Ω , then by restricting their identity to the sets $\{\sigma_{i-1} < u < \sigma_i\}$ we find successively that $c_1=0, c_2=0,\ldots,c_n=0$. Hence, the required linear independence of the gradients of the constraint functionals is proved, and so (2.9c) follows for some Lagrange multipliers λ_i $(i=1,\ldots,n)$ uniquely determined by u.

Finally, the smoothness of u stated in (2.9a) can be derived from the standard regularity theory for elliptic partial differential equations. We introduce the Green function $g(x,x')$, $x,x'\in\Omega$, defined by $-\Delta_x g(x,x')=\gamma(x-x')$ for $x\in\Omega$, $\quad g(x,x')=0$ for $x\in\partial\Omega$, then we have

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

Since $||\Lambda(u)|| \leq (\max_{1\leq i\leq u}|\lambda_i|)\,||u||$ we obtain $\Lambda(u) \in L^2(\Omega)$, and hence $u \in C^0(\bar{\Omega})$, using elementary properties of g. Now differentiating (2.12) in x we find that $u \in C^1(\bar{\Omega})$. Thus, $\Lambda(u) \in C^{0,1}(\bar{\Omega})$, and this is optional since the Lipshitz function $\Lambda(s)$ is piecewise linear. The statement (2.9a) that $u \in C^{2,\beta}(\bar{\Omega})$, is therefore a consequence of the global Schauder estimates applied to the equation (2.9c). (It is assumed here that $\partial\Omega$ is of class $C^{2,\beta}$, at least.) This completes the proof of (2.9). □

The function $\Lambda(s)$ occurring in (2.12) 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:

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

where $\lambda_1 \ldots, \lambda_n$ are real constants (the multipliers in Theorem 1.2). Such a profile function is piecewise linear with slope $\Lambda'(s) = \lambda_i$ on each interval $\sigma_{i=1} < s < \sigma_i$. (We recall that the partition of $[0, +\infty)$ into intervals $[\sigma_{i-1},\ \sigma_i)$, $i = 1, \ldots, n$, is indexed so that $0 = \sigma_0 < \sigma_1 < \ldots < \sigma_{n-1} < \sigma_n = +\infty$.) Clearly, $\Lambda(s)$ is nondecreasing on $[0, +\infty)$ if and only if $\lambda_i \geq 0$ for every $i = 1, \ldots, n$. The monotone functions $f_i'(s)$ $(i = 1, \ldots, n)$ may be viewed as forming a basis for the space of profile functions - namely, those functions $\Lambda(s)$ which satisfy (i) $\Lambda \in C^0()$, (ii) $\Lambda(s) = 0$ for $s \leq 0$, (iii) $\Lambda(s)$ is linear (actually affine) on each interval $\sigma_{i=1} \leq s \leq \sigma_i$. This basis, each member of which is monotone, is related to the the standard basis consisting of finite-elements $\phi_i(s)$ for the partition (or grid) $\{\sigma_i\}_{i=0}^n$ by the formulas

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

where $\Delta\sigma_i = \sigma_i - \sigma_{i-1}$ denotes the increments. Then, 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 satisfied as the linear interpolant of its values Λ_i at the partition points (or grid nodes) σ_i.

The representation (2.13) 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 functions $f_i(s)$.

Furthermore, it is a consequence of the definition of $f_i(s)$ in (2.7) that

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

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

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

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 (2.15). In view of this fact, we may say that the problem (P_n) constitutes a generalization or extension of the classical Rayleigh principle [?] characterizing variationally the first eigenvalue-eigenfunction pair for the Laplacian operator $-\Delta$ on $H_0^1(\Omega)$. When $n = 1$, the problem (P_1)

$$\int_\Omega |\nabla u|^2 dx \to \min \ \text{ subject to } \int_\Omega u_+^2 = 2\gamma_1 \tag{2.24}$$

yields a solution pair $(u^{(1)}, \lambda_1^{(1)}) \in H_0^1(\Omega) \times$ to the usual linear eigenvalue problem

$$-\Delta u^{(1)} = \lambda_1^{(1)}) u^{(1)} \text{ in } \Omega, \ \ u^{(1)} = 0 \text{ on } \partial\Omega,$$

which, in addition, is nonnegative in Ω. When $n \geq 2$, the problem (P_n) results in a solution pair $(u^{(n)},\ \lambda^{(n)}) \in H_0^1(\Omega) \times {}^n$ of

$$-\Delta u^{(n)} = \Lambda^{(n)}(u^{(n)}) \text{ in } \Omega, u^{(n)} = 0 \text{ on } \partial\Omega,\ \ u^{(n)} \geq 0 \text{ in } \Omega, \tag{2.25}$$

a nonlinear eigenvalue problem for which the vector of eigenvalues $\lambda^{(n)} = (\lambda_1^{(n)}, \ldots, \lambda_n^{(n)})$ determines the (piecewise linear) profile function $\Lambda^{(n)}(s)$ according to (2.13). In going from the linear problem (P_1) to the general (and nonlinear) problem (P_n) the familiar $L^2(\Omega)$-normalization $||u||^2 = 2\gamma_1$ (=1, say) is replaced by the family of constraints $F_i(u) = \gamma_i (i = 1, \ldots, n)$, which in effect amounts to constraining n $L^2(\Omega)$-expressions involving the lower functions $(u-\sigma_i)_+$ of u; namely, the constraints for (P_n) can be written as

$$(2.26)\qquad ||(u-\sigma_{i-1})_+\,||^2 - ||\,(u-\sigma_i)_+\,||^2 = 2\gamma_i \quad (i=1,\ldots,n).$$

The family of constraints (2.17) is therefore seen to be a strengthening of the single constraint (2.15) in the usual Rayleigh principle (P_1).

The sense in which the solutions $u^{(n)}$ of (P_n) approximate solutions $u^{(\infty)}$ (say) of (P_∞) as $n \to +\infty$ can now be determined. As mentioned above, we suppose that a sequence of partitions is taken with $\max_{1\le i\le n-1}(\sigma_{i-1}-\sigma_i) \to 0$ and $\sigma_{n-1} \to +\infty$ as $n \to \infty$ so that we are assured that $M_\infty(\bar{u}) = \cap_n M_n(\bar{u})$. By virtue of the bound $E(u^{(n)}) \le E(\bar{u}) < +\infty$, 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 such a limit point of the solution sequence $u^{(n)}$ for the multi-constrained problems (P_n) then we claim that $u^{(\infty)}$ is a solution of the (infinitely-constrained) problem (P_∞). Indeed, any admissible function $\tilde{u} \in M_\infty(\bar{u})$ for (P_∞) is admissible for each problem (P_n), and hence, using the weak - H_0^1 convergence, $E(u^{(\infty)}) \le \liminf E(u^{(n)}) \le E(\tilde{u})$, where n is understood to be 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 $\sigma \in (0,+\infty)$

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

where $i = i(n,\sigma)$ is chosen so that $\sigma \in [\sigma_{i-1},\ \sigma_i)$ for each sufficiently large n. (We recall that $\Delta\sigma_i := \sigma_i - \sigma_{i-1}$.) Thus the claim that $u^{(\infty)}$ solves (P_∞) follows. This result can be stated in the form

$$(2.27)\qquad dist(u^{(n)}, S_\infty) \to 0 \text{ as } n \to +\infty,$$

where $S_\infty = S_\infty(\bar{u})$ is defined to be the set of solutions of (P_∞) for a given function $\bar{u}$; the distance function $dist(u,\mathcal{S})$ from a point $u \in L^2(\Omega)$ set $\mathcal{S} \subseteq L^2(\Omega)$ is defined by

$$(2.28)\qquad dist(u,\mathcal{S}) = \inf\{||u-v||:\ v \in \mathcal{S}\}.$$

That (2.18) is true is obvious from the above discussion. We remark that it is necessary to consider the *set* of solutions of (P_∞) since the uniqueness of its solutions has not been established and, in fact, may not hold in general. The statement (2.18) allows us to conclude, at least, that for large n the solution $u^{(n)}$ of (P_n) approximates *some* solution of (P_∞) in the $L^2(\Omega)$-norm. This conclusion justifies in part the discretization of the constraints for (P_∞) 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 $0 = \sigma_0 < \sigma_1 < \cdots < \sigma_{n-1} < \sigma_n < +\infty$, we consider the problem (P_n')

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

This problem is of variational inequality type [**?**], being an obstacle problem with no additional constraints. It can be shown that the analogue of Theorem 1.2 holds in the sense that there is a solution $u \in H_0^1(\Omega) \cap C^{1,1}(\bar{\Omega})$ satisfying the equation

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

(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 Construction of the algorithm for solving (P_n).

We now proceed to define an iterative procedure which is designed to converge to the solutions of (P_n). In this section we concentrate on describing the algorithm in its general form and establishing its basic convergence property. For a discussion of the main ideas underlying the construction of the algorithm the reader is referred to Appendix 2.

For a (fixed) positive constant τ, let E_τ be the following modified objective functional

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

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

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

We notice that in contrast to the class $M_n(\bar{u})$ 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 notations.) In terms of these definitions we construct a sequence of approximations $(u^k, \lambda^k) \in H_0^1(\Omega) \times {}^n$ 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: u^{k+1} is the (unique) solution of

$$E_\tau(u) \longrightarrow \min \ \text{ over } u \in L_n(u^k) \tag{3.3}$$

and μ^{k+1} is the corresponding n-vector of nonnegative multipliers (uniquely determined by u^{k+1}); then let

$$\lambda^{k+1} = \mu^{k+1} - \tau. \tag{3.4}$$

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

$$\text{(3.5)} \qquad \begin{cases} E'(u) = \sum_{j=1}^{n} \lambda_j F_j'(u) \text{ for some } \lambda_1, \ldots, \lambda_n \in \\ F_i(u) = \gamma_i \;\; (i = 1, \ldots, n); \end{cases}$$

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 development τ will be viewed as a (fixed) positive parameter.

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

$$\text{(3.6)} \qquad \begin{cases} E'(u^{k+1}) \,+ \tau u^{k+1} = \sum_{j=1}^{n} \lambda_j^{k+1} F_j'(u^k) \\ \lambda_i^{k+1} \left[F_i(u^k) + \langle F_i'(u^k),\, u^k \rangle - \gamma_i \right] = 0 \quad (i = 1, \ldots, n) \end{cases}$$

Indeed, these equations are precisely the Kuhn-Tucker conditions associated with the convex optimization problem (2.22) which defines (A_n). The reader is referred to [**?**] for a proof that (2.22) and (2.24) are equivalent provided that there exists some $\tilde{u} \in L_n(u^k)$ for which

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

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

$$\text{(3.7)} \quad F_i(u^k) \geq F_i(u^{k+1}) \,+\, \langle F_i'(u^{k-1}),\ u^k - u^{k-1} \rangle \geq \gamma_i \quad (k = 1, 2, \ldots).$$

In addition, $F_i(u^k) \leq \langle F_i'(u^k),\ u^k \rangle$ since $F_i(0) = 0$. thus, it is readily verified that the required condition is satisfied by $\tilde{u} = (1 + \epsilon) u^k$ with $\epsilon > 0$. It should be emphasized that the equations (2.24) characterize the iterate $(u^{k+1},\ \mu^{k+1}) \in L_n(u^k) \times [0, +\infty)^n$; and it is implicitly assumed in (2.24) that

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

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

for every $i = 1, \ldots, n$. The vector $\mu^{k+1} = (\mu_1^{k+1}, \ldots, \mu_u^{k+1})$ is called the Kuhn-Tucker vector corresponding to the solution u^{k+1} of (2.22).

The above construction of the algorithm (A_n) depends fundamentally on the convexity of the optimization problem (2.22); namely, it relies on the strict convexity of the objective functional E_τ 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 (2.22) follows immediately from these attributes. In turn, the uniqueness of the Kuhn-Tucker vector μ^{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 1.2, now using (2.25) with $\gamma_i > 0$.

We describe next a more explicit construction of the iterative procedure (A_n) which furnishes us with a concrete numerical implementation of the algorithm. Let G_τ denote the Green operator for the elliptic boundary value problem

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

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

$$a_{ij}(v)\colon = \langle f_i'(v),\ G_\tau f_j'(v)\rangle \tag{3.11}$$

$$c_i(v)\colon = \gamma_i - F_i(v) + \langle F_i'(v),\ v\rangle. \tag{3.12}$$

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

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

This defines a *positive-definite* quadratic form on ^{n} whenever $|\{x \in \Omega\colon\ \sigma_{i-1} < v(x) < \sigma_i\}| > 0$ for every $i = 1, \ldots, n$. 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 Ω. Arguing as in the proof of Theorem 1.2, we deduce that $\mu_1 = 0$, $\mu_2 = 0, \ldots, \mu_n = 0$ successively, as required.

We now claim that 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

$$u^{k+1} = \sum_{j=1}^{n} \mu_j^{k+1}\ w_j \tag{3.14}$$

where

$$w_j = G_\tau\, f_j'(u^k), \tag{3.15}$$

$$\begin{aligned} &\mu_j^{k+1} \geq 0 \ \text{ and } \ Q(\mu^{k+1};\ u^k) \\ &= \min\{Q(\mu;\, u^k)\colon \mu_i \geq 0,\, i = 1, \ldots, n\}. \end{aligned} \tag{3.16}$$

In other words, $(u^{k+1},\ \mu^{k+1})$ is constructed from u^k in a three-step process: (1) each $w_j (j = 1, \ldots, n)$ is found by solving (2.28) with $h = f_j'(u^k)$; (2) μ^{k+1} is taken to be the unique solution of the quadratic programming problem (2.34); (3) u^{k+1} is assembled from μ_j^{k+1} and $w_j (j = 1, \ldots, n)$ according to (2.32). 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 μ^{k+1} of (2.34) we have for each $i = 1, \ldots, n$

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

Also, combining (2.32) and (2.33) we have

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

It is clear that (2.36) is a restatement of the first equation in (2.24); in turn, it is evident from the definitions of a_{ij} and c_i that (2.35) is equivalent to the second equation in (2.24) together with the inequalities (2.26) and (2.27) which supplement (2.24). Thus the claimed equivalence of the two forms of the algorithm is verified.

It is interesting to note that the optimization problems (2.34) and (2.22) are dual problems in the sense of convex analysis [**?**]. Indeed, we can directly verify that

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

Then general theory informs us that the Kuhn-Tucker vector μ^{k+1} corresponding to the minimizer u^{k+1} for (2.22) is itself the maximizer for the dual problem:

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

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*

$$E_\tau(u^{k+1} - u^k) \le E_\tau(u^k - E_\tau(u^{k+1}) \tag{3.19}$$

$$E(u^{k+1}) \le E(u^0). \tag{3.20}$$

Proof. To prove (2.37) we use the identity

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

By virtue of (2.25) we have $u^k \in L_n(u^k)$. Therefore, $(1-t)u^{k+1} + tu^k \in L_n(u^k)$ for all $0 \le t \le 1$, as this class of functions is convex. Consequently, since u^{k+1} solves (2.22) we get

$$\begin{aligned} &E_\tau(u^{k+1}) \le E_t(u^{k+1}) + t(u^k - u^{k+1})) = \\ &E_\tau(u^{k+1}) + t\langle E'_\tau(u^{k+1}),\ u^k - u^{k+1}\rangle + 0(t^2) \end{aligned}$$

as $t \to 0+$. Hence, $\langle E_\tau(u^{k+1}),\ u^{k+1} - u^{k+1}\rangle \ge 0$, and so (2.37) follows from (2.39).

To prove (2.38) we note that

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

This yields the desired inequality since

$$E(u^{k+1}) = E_\tau(u^{k+1}) - \frac{\tau}{2}\,||u^{k+1}||^2 \le E_\tau(u^0) - \frac{\tau}{2}\,||u^0||^2 = E(u^0).$$

$\square$

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 iterates 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 $\mu = (\mu_1, \ldots, \mu_n) \in {}^n$ such that*

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

$$\mu_i\,[F_i(u) - \gamma_i] \;=\; 0 \quad (i = 1, \ldots, n) \tag{3.23}$$

$$\mu_i \geq 0, \;\; F_i(u) \geq \gamma_i \quad (i + 1, \ldots n). \tag{3.24}$$

Proof. We recognize that (2.40), (2.41), (2.42) are precisely the Kuhn-Tucker conditions for the convex minimization problem

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

and that μ is the Kuhn-Tucker vector corresponding to the minimizer u. Therefore, it suffices to show that u is indeed the solution of (2.43). Let $u^{k_p} \to u$ weakly in $H_0^1(\Omega)$, and hence strongly in $L^2(\Omega)$, as $p \to +\infty$. Given an arbitrary $\tilde{u} \in L_n(u)$ we consider the perturbation $\tilde{u} + \epsilon u$ with $\epsilon > 0$. Recalling the definition of $L_n(u)$, we find that for each $i = 1, \ldots, n$

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

This implies that for sufficiently large p

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

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, $\tilde{u} + \epsilon u \in L_n(u^{k_p})$ for sufficiently large p, and so we obtain

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

where, in the first inequality, we invoke the monotonicity property (2.37) and the lower semi-continuity of E_τ with respect to weak H_0^1 convergence. Since this inequality holds for arbitrarily small positive ϵ, we conclude that the limit point u solves (2.43), 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 2.3 and 2.4 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 (2.40), (2.41), (2.42). But (2.42) can be written equivalently as

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

in view of the basic identity (2.14) which implies that

$$\sum_{i=1}^{n} f_j'(s) = s_+ \quad (s \in). \tag{3.27}$$

(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) - namely, (2.32), (2.33), (2.34). 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 τ, sending λ_i into $\mu_i = \lambda_i + \tau$. Accordingly, a solution u of (2.40), (2.41), (2.42) 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 τ large enough ($\tau > \max\{0, -\lambda_1, \dots, -\lambda_n\}$) depending on the multipliers γ_i corresponding to *any* solution u of (2.44). The *a priori* estimate on $\max_{1 \leq i \leq n} |\lambda_i|$ needed to complete this argument is given in Lemma 3.2, and so 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:

Given $u^0 \in M_n(\bar{u}),\ u^{k+1}$ is the unique solution of the variational problem

$$
(4.1) \qquad \begin{array}{l} E_\tau(u) = \frac{1}{2} \int_\Omega (|\nabla u|^2 + \tau u^2) \to \min \\ u \in H_0^1(\Omega):\ F_i(u^k) + \langle F_i'(u^k),\ u - u^k \rangle \geq \gamma_i,\ i = 1, \ldots, n \end{array}
$$

which is defined in terms of a positive parameter τ and with linearized inequality constraints.

In this section we will show that τ can be chosen apriori so that the limit points of the Algorithm A_n are solutions of our original problem P_n. More precisely, we have

Theorem 4.1 *Given any initialization $u^0 \in M_n(\bar{u})$, there exists $C = C(\Omega,\ E(u^0),\ \gamma_1^{-1} \ldots \gamma_n^{-1}, (\Delta\sigma)^{-1})$ so that if $\tau > C$, the algorithm A_n given by (4.1) converges strongly in $H_0^1(\Omega)$ to the set $\mathcal{S}$ of critical points of our original problem P_n, in the sense*

$$
dist(u^k,\ \mathcal{S}) \to 0 \quad as\ h \to \infty
$$

$$
where \quad dist(u, \mathcal{S}) = \ inf\,\{||u - v||_{H_0^1(\Omega)}\colon v \in \mathcal{S}\})
$$

Theorem 4.1 says that any subsequence of $\{u^k\}$ has a subsequence which converges strongly in $H_0^1(\Omega)$ to an element of $\mathcal{S}$. In numerical experiments, the entire sequence $\{u^k\}$ converges. Whether or not this is the case is a deep question but we have the following result which indicates that generically there is only one limit point.

Proposition 4.2 *Let $S(u^0)$ be the set of limit points of Algorithm A_n for a given u^0. Then either*

(i) $S(u^0)$ contains exactly one point

or

(ii) $S(u^0)$ contains infinitely many points, none of which are isolated.

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

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

$$(4.2) \qquad \begin{cases} -\Delta u = \sum_{i=1}^n \lambda_i f_i'(u) \text{ in } \Omega \\ F_i(u) = \int_\Omega f_i(u) \geq \gamma_i > 0,\ i = 1, \ldots, n \end{cases}$$

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

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

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. As we indicated in the previous section, Lemmas 2.3 and 2.4 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 (4.2) with

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

Applying Lemma 4.3 we can conclude that $\lambda_i > -C$, or $\mu_i > \tau - C$. Hence if we apriori choose a $\tau > C$, we insure that the multipliers $\mu_i > 0$, $i = 1, \ldots n$. This implies, by Lemma 2.4 that $F_i(u) = \gamma_i$, that is, u is a critical point of our original 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, observe

$$\begin{aligned} ||u^{k+1} - u||^2_{H_0^1(\Omega)} &= \textstyle\int_\Omega \nabla u^{k+1} \cdot \nabla(u^{k+1} - u) - \int_\Omega \nabla u \cdot \nabla(u^{k+1} - u) \\ &= \textstyle\sum \mu_i^{k+1} \int_\Omega (u^{k+1} - u) f_i'(u^k) - \tau \int u^{k+1}(u^{k+1} - u) - \int_\Omega \nabla u \cdot (u^{k+1} - u) \\ &= \textstyle\sum \mu_i^{k+1} \int_\Omega (u^{k+1} - u)\, f_i'(u^k) + o(1) \text{ as } k \to \infty \end{aligned} \tag{4.3}$$

We claim that the μ_i^{k+1} are uniformly bounded. For u^{k+1} satisfies

$$-\Delta u^{k+1} + \tau\, u^{k+1} = \sum \mu_i^{k+1}\, f_i'(u^k)$$

and so

$$\begin{aligned} &\textstyle\sum \mu_i^{k+1} \int_\Omega u^k f_i'(u^k) = \int \nabla u^{k+1} \nabla u^k + \tau u^{k+1} u^k \\ &\leq E_\tau(u^{k+1}) + E_\tau(u^k \leq 2E_\tau(u^0) \end{aligned}$$

But, $\int_\Omega u^k f_i'(u^k) \geq \int_\Omega f_i(u^k) = F_i(u^k) \geq \gamma_i$, hence $\sum_i \mu_i^{k+1} \gamma_i \leq 2E_\tau(u^0)$. Since $\gamma_i > 0 \quad \forall_i$, the μ_i^{k+1} are uniformly bounded as claimed.

Now observe that for any $v \in H_0^1(\Omega) \cap L^\infty(\Omega)$, $v \geq 0$, we have

$$\sum_i \int_\Omega f_i'(v)^2\, dx \leq \lambda_l^{-1}(\Omega) \sum_i \int_\Omega f_i''(v)|\nabla v|^2 = \frac{1}{\lambda_1(\Omega)} \int_\Omega |\nabla v|^2\, dx$$

since $f_i''(s) = \chi\{\sigma_{i-1} < s < \sigma_i\}$. Hence,

$$\begin{aligned} \sum \mu_i^{k+1} \textstyle\int_\Omega (u^{k+1} - u) f_i'(u^k)\, dx & \\ &= \leq ||u^{k+1} - u|| (\textstyle\sum_i \int_\Omega (\mu_i^{k+1})^2 f_i'^2(u^k))^{\frac{1}{2}}\, dx \\ &\leq c||u^{k+1} - u|| \to 0 \text{ as } k \to \infty \end{aligned}$$

(4.4)

Therefore, (4.4) (4.5) show that $||u^{k+1} - u||_{H_0^1(\Omega)} \to 0$ as $k \to \infty$. The proof of Theorem 4.1 is complete. □

We can also give the

Proof of Proposition 4.2. Recall from Lemma 2.3 that

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

and consequently $||u^{k+1} - u^k||_{H_0^1(\Omega)} \to 0$ as $k \to \infty$. Using Theorem 4.1, we can repeat the argument of [], [Proposition 2]. For the readers convenience we will repeat the simple argument. Namely, it suffices to show that if $S(u^0)$ contains one isolated point u, then the entire sequence converges to u.

Since u is isolated we can find disjoint neighborhoods N_1 of u and N_2 containing all other elements of $S(u^0)$. Let $2\epsilon = dist(N_1, N_2)$. Since $||u^{k+1} - u^k||H_0^1(\Omega) \leq \epsilon$. We claim that there is an integer k so that for $k \geq k_1, u^k$ belongs either to N_1 or N_2. For otherwise, we can find a subsequence u^{k_i} lying in the complement of $N_1 \cup N_2$, contradicting Theorem 4.1. Since $u \in S(u^0)$ we can find a $u^k \in N_1$, with $k \geq \max(k_0, k_1)$. The $dist(u^{k+1}, N_2) \geq \epsilon$ and so $u^{k+1} \in N_1$. By induction $u^j \in N_1$ for $j \geq k$ and the proposition is proven. □

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

Proof of Lemma 4.3. We first estimate $|\lambda_1|, \ldots, |\lambda_{n-1}|$. Set $\Lambda(u) = \sum_{i=1}^n \lambda_i f_i'(u)$ and let $\Lambda_i = \Lambda(\sigma_i), i = 0, \ldots, n-1$. Then

(4.5) $$\begin{aligned} &\lambda_0 = 0 \text{ and } \lambda_i = \Lambda_{i-1} + \lambda_i \Delta\sigma_i \Lambda(s) = \Lambda_{i-1} + \lambda_i(s - \sigma_{i-1}) \text{ on } [\sigma_{i-1}, \sigma_i] \\ &= \Lambda_i + \lambda_i(s - \sigma_i) \end{aligned}$$

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

(4.6) $$|\Lambda_i| = \max |\Lambda_j|_{1\le j\le n-1}$$

Then for $s \in I =: [\sigma_i - \frac{1}{3}\Delta\sigma_i,\, \sigma_i - \frac{1}{6}\Delta\sigma_i]$,

(4.7) $$|\Lambda(s)| \ge \frac{1}{3}|\Lambda_i|$$

since $|\Lambda(s)| \ge |\lambda_i| \cdot \frac{1}{3}\Delta\sigma_i$ and $|\lambda_i| \le 2\frac{|\Lambda_i|}{\Delta\sigma_i}$.

Let $\phi(s)$ be a Lipschitz cut off function with $\phi(s) = 1$ for $s \in I, \phi(s) = 0$ for $s \le \bar{\sigma}_i = \frac{1}{2}(\sigma_{i-1} + \sigma_i)$ and $s \ge \sigma_i$ and $0 \le \phi(s) \le 1$. Using $\phi(u)$ as a test function in (4.21) we have the identity

(4.8) $$\int_\Omega \phi'(u)|\nabla u|^2\,dx = \int_\Omega \phi(u)\Lambda(u)\,dx$$

Since $\Lambda(s)$ has one sign on the support of $\phi(s)$, it follows from (4.9) that

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

Combining (4.10) and (4.8), we have

(4.10) $$|\Lambda_i| \le \frac{36E(u)}{(\Delta\sigma_i)|\{u \in I\}|}$$

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

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

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

Then

(4.11) $$||f_I'(u)||_L^{\frac{n}{n-1}\Omega} \ge |I|\,|\{u > \Gamma_{n-1}\}|^{\frac{n-1}{n}}$$

On the other hand, by the Sobolev inequality

(4.12) $$||f_I'(u)||_L^{\frac{n}{n-1}(\Omega)} \le c\int_\Omega f_I''(u)|\nabla u|\,dx \le c|\{u \in I\}|E(u)^{\frac{1}{2}}$$

and by the Poincaré inequality.

$$2F_n(u) = ||(u - \sigma_{n-1})_+||^2_{L^2(\Omega)} \le C_n|\{u > \sigma_{n-1}\}|^{\frac{2}{N}} E(u). \tag{4.13}$$

Combining (4.12) - (4.14) gives the required lower bound

$$|\{u \in I\}| \ge c_N |I|^2\, F_n(u)^{N-1}\, E(u)^{-N}. \tag{4.14}$$

Together with (4.11) this gives the estimate

$$|\Lambda_i| \le C_N E(u)^{N+1} F_n(u)^{1-N} (\Delta\sigma_i)^{-3} \le C_N E_0^{N+1} \gamma_n^{1-N} (\Delta\sigma)^{-3}. \tag{4.15}$$

To complete the proof we need only bound $|\lambda_n|$. Consider

$$-\Delta u = \Lambda_{n-1} + \lambda_n(u - \sigma_{n-1}) \text{ in } \{u > \sigma_{n-1}\} \tag{4.16}$$

Using $(u - \sigma_{n-1})_+$ as a test function in (4.17) yields the identity

$$\int_\Omega |\nabla(u - \sigma_{n-1})_+ + 2\lambda_n F_n(u) \tag{4.17}$$

Therefore,

$$\gamma_n|\lambda_n| \le E_0 + |\Lambda_{n-1}| \cdot (\frac{|\Omega|}{\lambda_1(\Omega)} E_0)^{\frac{l}{2}} \tag{4.18}$$

Recalling (4.16) this gives

$$|\lambda_n| \le C(\Omega, N_1 E_0)\gamma_n^{-2}(\Delta\sigma)^{-3} \tag{4.19}$$

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_∞) and (P_n). This is accomplished simply by informing the constraints (2.4c) only on the interval $\sigma_0 \leq \sigma < +\infty$ for some given $\sigma_0 > 0$. Thus, the problem (P_∞) becomes

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

Similarly, the problem (P_n) is defined with respect to a partition $0 < \sigma_0 < \sigma_1 < \ldots, < \sigma_{n-1} < \sigma_n = +\infty$, and the constraints for (P_n) are defined by (2.6), (2.7) as before. Now, however, $f_i(s) = 0$ for all $s \leq \sigma_0$ and $i = 1, \ldots, n$. the validity of Theorem 2.2 remains unaltered in this extension, and a solution u of (P_n) satisfies $-\Delta u = \Lambda(u)$ in Ω, $u = 0$ 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

$$(5.2)\qquad -\Delta u = \begin{cases} 0 & \text{in } \{0 < u < \sigma_0\} \\ \Lambda(u) & \text{in } \{u > \sigma_0\} \end{cases}$$

$$|\nabla u| \text{ is continuous across } \ \{u = \sigma_0\},$$

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

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

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, no essential changes in the above development are needed to include the free-boundary into (P_n).

Next, we consider an extension of (P_∞) and (P_n) (without free-boundaries) which allows u and $\bar{u}$ to change sign in Ω. The requirement that $\bar{u} \geq 0$ in Ω is now relaxed, and the constraints in (P_∞) 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 $0 = \sigma_0^+ < \sigma_1^+ < \ldots < \sigma_{n-1}^+ < \sigma_n^+ = +\infty$ and $0 = \sigma_0^- > \sigma_1^- > \ldots > \sigma_{m-1}^- > \sigma_m^- = -\infty$ for some n and m. The functionals $F_i^+ (i = 1, \ldots, n)$ and $F_j^- (j = 1, \ldots, m)$ 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

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

a piecewise linear function on $-\infty < s < +\infty$ which is increasing whenever $\lambda_1^+ \geq 0$ 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

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

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

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

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

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

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

$$\begin{cases} -\sum_{p,q} \frac{\partial}{\partial x_p}(a_{pq}(x)\frac{\partial u}{\partial x_q}) + a(x)u = b(x)\lambda(u) \;\text{ in }\; \Omega \\ u = 0 \quad \text{on} \;\; \partial\Omega, \end{cases}$$

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

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

6 Appendix 1. Variational Problems in Magneto Hydrodynamics.

Equilibrium problems in magneto hydrodynamics (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 [**?**, **?**], 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_∞) which encompasses the physical problems of interest as special cases. Let $\Omega \subseteq {}^2$ be the cross-section of a cylindrical domain $\Omega \times$, and let (x_1, x_2) denote the variable point in Ω and x_3 be the ignorable coordinate. Consider the minimization problem

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

where the admissible triple (u, v_1, v_2) belongs to $H^1_0(\Omega) \times L^{r_1}(\Omega) \times L^{r_2}(\Omega)$ 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 $h_\ell(0) = h_\ell'(v) = 0$ and $h_\ell(z) = O(|z|^{r_\ell})$ as $|z| \to \infty$ ($\ell = 1, 2$). As in (P_∞), the two infinite family 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 $\underline{B} = (B', B^2, B^3)$, which is independent of x_3, satisfies $\nabla \cdot B = 0$ in $\Omega \times$ and hence admits a representation $\underline{B} = (u_{x_2}, -u_{x_1}, v_1)$, 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 $\frac{1}{2}|\underline{B}|^2 = \frac{1}{2}|\nabla u|^2 + \frac{1}{2}v_1^2$. The mass density ρ 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

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

The interpretation of the constraints relies on differentiating them with respect to the parameter σ, 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 that all constraints involving integrals of the form

$$\int_\Omega v_\ell \phi(u) \, dx \ \text{ where } \phi \ \text{ is an arbitrary function.}$$

Indeed, the latter integrals are used to express the family of conserved quantities in [**?**, **?**]. Roughly speaking, the functions $(s - \sigma)_+$ ($\sigma_0 \le \sigma < +\infty$) are chosen here as a particularly useful basis for the space of all functions $\phi(s)$.

This allows us to give a more direct 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_∞). Recalling the definition (2.7), we replace the infinite family of constraints in (6.1) by

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

relative to a (fixed) partition $0 \leq \sigma_0 < \sigma_1 < \ldots < \sigma_{n-1} < \sigma_n = +\infty$. 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

$$\begin{cases} -\Delta u &= v_1 \Phi' 1(u) + v_2 \Phi_2'(u) \\ h_1'(v_1) &= \Phi_1(u) \\ h_2'(v_2) &= \Phi_2(u), \end{cases} \tag{6.3}$$

where, by virtue of the Lagrange multiplier rule,

$$\Phi_\ell(s) := \sum_{i=1}^{n} \lambda_{\ell i}\, f_i(s) \quad (\ell = 1, 2) \tag{6.4}$$

for some multiplier $\lambda_{\ell i}$. 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:

$$-\Delta u = (h_1'(\Phi_1(u))\Phi_1'(u) + (h_2')^{-1}(\Phi_2(u))\Phi_2'(u), \tag{6.5}$$

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 $h_\ell(\ell = 1, 2)$ defined by $h_\ell(z^*) = \sup_z zz^* - h_\ell(z)$. Then (6.5) can be written as simply

$$-\Delta u = P_1'(u) + P_2'(u) \ \text{ with } \ P_\ell(s) := h_\ell^* \left(\Phi_\ell(s)\right). \tag{6.6}$$

Returning to the physical case, we find that

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

and consequently we obtain the relations

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

It follows from these relations that the profile functions P_1' and P_2' in (6.6) have the usual interpretations or current and pressure profiles, respectively, in the Grad-Shafranov equation.

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.[**?**, **?**]. 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 Δ 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 multi-constrained 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 σ_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 $\lambda_{2_i} \geq 0 \;\; (i = 1, \dots, n)$, 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 conditons.

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 γ 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_∞) 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 last under favorable circumstances. However, a convergence theory similar to that given in §3-5 for the prototype problem is not yet available for this general problem, owing mainly to the lack of joint convexity of the constraint functionals in (6.1).

7 Appendix 2. Some Comments On Algorithm (A_n).

This appendix is included in order to indicate the main ideas underlying the algorithm (A_n) presented formally in the preceding section. As such the comments given here tend to emphasize the conceptual rather than logical development of general iterative procedure for solving the variational problem (P_n).

The globally convergent iterative procedure (A_n) can be considered as the product of a development consisting of three stages. At each of these stages a key ingredient in the algorithm is introduced. Therefore, it is useful when explaining the algorithm (A_n) to describe the form that it takes at each stage of its development. We now proceed to give such a description, stressing the ideas involved rather than the details of proof.

The first stage of (A_n) consists, in essence, of an iterative procedure which is based upon linearizing the nonlinear equality constraints occurring in the minimization problem (P_n). This first form of (A_n) is defined iteratively by the equations

$$-\Delta u^{k+1} = \sum_{j=1}^{n} \lambda_j^{k+1} f_j'(u^k) \text{ in } \Omega,\ u^{k+1} = 0 \text{ on } \partial\Omega \tag{7.1}$$

$$\int_\Omega [f_i(u^k) + f_i'(u^k)(u^{k+1} - u^k)]\, dx = \gamma_i\ (i = 1, \ldots, n). \tag{7.2}$$

For a fixed initialization $u^0 \in M_n(\bar{u})$, this iterative procedure defines a sequence $(u^{k+1}, \lambda^{k+1}) \in H^1 0(\Omega) \times {}^n$ which is intended to converge (in some suitable sense) to a solution of (P_n). Moreover, since the defining equations (7.1) and (7.2) are linear in (u^{k+1}, λ^{k+1}), the explicit construction of that iterative pair from u^k is a routine matter. First, we solve the linear elliptic problems

$$-\Delta w_j = f_j'(u^k) \text{ in } \Omega,\ w_j = 0 \text{ on } \partial\Omega\ (j = 1, \ldots, n); \tag{7.3}$$

second, we find $\lambda^{k+1} \in {}^n$ by solving the $n \times n$ system of linear algebraic equations

$$\sum a_{ij}(u^k)\lambda_j^{k+1} = c_i(u^k)\ (i = 1, \ldots, n), \tag{7.4}$$

where, as in (2.29) and (2.30) with $\tau = 0$,

$$a_{ij}(u^k) = \int_\Omega \nabla w_i \cdot \nabla w_j \, dx \;\; c_i(u^k) = \gamma_i - \int_\Omega [f_i(u^k) - f_i'(u^k)u^k] \, dx;$$

third, we set

$$u^{k+1} = \sum_{j=1}^{n} \lambda_j^{k+1} w_j. \tag{7.5}$$

This algorithm treats the constraints in (P_n) which the main source of difficulty in the problem, by replacing them at each iterative step with their linearization about the previous iterate. If the iterative sequence $(u^{k+1}, \, \lambda^{k+1})$ converges to $(u, \, \lambda) \in H_0^1(\Omega) \times {}^n$ (in norm, say), then u is a critical point of (P_n), or equivalently, $(u, \, \lambda)$ solves $(\mathcal{S})$, meaning that

$$\int_\Omega \nabla u \cdot \nabla v \, dx = \sum_{j=1}^{n} \lambda_j \int_\Omega f_j'(u) v \, dx \tag{7.6}$$

$$\int_\Omega f_i(u) \, dx = \gamma_i \;\; (i = 1, \ldots, n) \tag{7.7}$$

for all $v \in H_0^1(\Omega)$. Thus, the constraints (7.7) on solutions are retrieved from the linearized constraints (7.2) on iterates. Furthermore, it is noteworthy that the defining equations (7.1), (7.2) are precisely the variational equations for the convex minimization problem

$$\begin{array}{l} E(u) \to \min \;\; \text{subject to} \\ F_i(u^k) \, + \langle F_i'(u^k), \, u - u^k \rangle = \gamma_i \;\; (i = 1, \ldots, n); \end{array} \tag{7.8}$$

and it is obvious that u^{k+1} is the unique minimizer for (7.8) and λ^{k+1} is the uniquely determined multiplier vector for u^{k+1}.

Although the problem (P_n) is itself nonconvex having constraints that are nonlinear equalities, the convexity of its objective and constraint functionals implies that under certain conditions the sequence of iterates enjoys special monotonicity and convergence properties. One consequence of these convexity attributes is the inequality.

$$\int_\Omega f_i(u^{k+1}) \, dx \geq \int_\Omega [f_i(u^k) + f_i'(u^k - u^k)] \, dx = \gamma_i \tag{7.9}$$

which holds unconditionally for every k and each $i = 1, \ldots, n$. Hence, the iterate u^{k+1} may be said to satisfy the constraints (7.7) in a relaxed sense with inequalities replacing the equalities. From (7.9) combined with Lemma 4.2 it follows that the sets $\{x \in \Omega\colon \sigma_{i-1} < u(x) < \sigma_i\}$ have positive measure. Using the argument given in the proof of Theorem 2.2 the positive-definiteness of the $n \times n$ matrix $A = (a_{ij}(u^k))$ is therefore ensured, since for any $\xi \in {}^n$

$$\sum_{i,\,j=1}^{n} a_{ij}(u^k)\xi_i\xi_j = \int_\Omega |\nabla(\xi \cdot w|^2\, dx > 0$$

unless $\xi \cdot w = 0$ in Ω, which is equivalent to $\xi \cdot f'(u^k) = 0$ in Ω. Thus, the invertibility of A is derived from (7.9), and in turn the explicit construction of the iterate $(u^{k+1},\, \lambda^{k+1})$ is seen to be well-defined. Another consequence of convexity is the monotonicity of the values of the objective functional along the iterative sequence, but, unlike (7.9), this property requires that a further condition be satisfied–namely, $\lambda_i^{k+1} \geq 0$ for each $i = 1, \ldots, n$. A simple calculation making use of (7.1), (7.2) and (7.9) yields the result:

$$\begin{aligned} E(u^k) - E(u^{k+1}) - E(u^{k+1} - u^k) &= \textstyle\int_\Omega \nabla u^{k+1} \cdot \nabla(u^k - u^{k+1})\, dx \\ &= \textstyle\sum_{j=1}^n \lambda_j^{k+1} \int_\Omega f_j'(u^k)(u^k - u^{k+1})\, dx \\ &= \textstyle\sum_{j=1}^n \lambda_j^{k+1}\{\int_\Omega f_j(u^k)\, dx - \gamma_j\} \\ &\geq 0. \end{aligned}$$

Therefore,

$$E(u^{k+1}) \leq E(u^k) \text{ and } E(u^{k+1} - u^k) \to 0 \text{ as } k \to +\infty, \tag{7.10}$$

provided that every λ_i^{k+1} is nonnegative. Moreover, this condition also furnishes an upper bound on $|\lambda^{k+1}|$, as is evident from the estimate

$$\begin{aligned} E(u^{k+1}) + E(u^k) &\geq \textstyle\int_\Omega \nabla u^{k+1} \cdot \nabla u^k\, dx \\ &= \textstyle\sum_{j=1}^n \lambda_j^{k+1} \int_\Omega f_j'(u^k) u^k\, dx \\ &\geq \textstyle\sum_{j=1}^n \lambda_j^{k+1}\, \gamma_j. \end{aligned}$$

These properties lead directly to a (generalized) convergence result for the iterative sequence. Specifically, it follows that every subsequence of $(u^{k+1},\, \lambda^{k+1})$ has a further subsequence which converges to a solution $u,\, \lambda)$ of (7.6), (7.7), with the subsequence of u^{k+1} converging to u weakly in H_0^1 and strongly in L^2. In other words, the iterative procedure produces a sequence which converges (in the sense of minimum distance) to the *set* $\mathcal{S}$ of solutions whenever the condition $\lambda_j^{k+1} \geq 0$ can be enforced.

One would expect that the latter condition could be satisfied if the initialization u^0 of the above iterative procedure is taken close enough to a solution of (P_n) which itself satisfies $\lambda_i \geq \delta > 0 (i = 1, \ldots, n)$ for some δ. In addition, the monotonicity property (7.10) suggests that, at least typically, the iterates should converge to a local minimum of E rather then merely a solution of $\mathcal{S}$. Thus, the utility of this provisional form of the algorithm for computing solutions of (P_n) is demonstrated within certain limits. On the other hand, the condition that $\lambda_j^{k+1} \geq 0$ is quite stringent and cannot be expected necessarily to hold when either (i) u^0 is not close enough to a local minimizer or (ii) a solution u for which some λ_i are negative is sought.

The second stage of (A_n) is a modification of the above iterative procedure which weakens the above condition. Now the parameter $\tau > 0$ is introduced and (u^{k+1}, λ^{k+1}) is defined by the analogues of equations (7.1), (7.2) where $-\Delta$ is replaced by $-\Delta + \tau$ and λ_j^{k+1} is replaced by $\lambda_j^{k+1} + \tau$. With these substitutions made throughout the above algorithm, it can be verified that the analogues of the above mentioned monotonicity and convergence properties hold provided that $\lambda_i^{k+1} \geq -\tau$ for each $i = 1, \ldots, n$; with these modification it can be seen that E is replaced by $E_\tau = E + \tau(F_1 + \cdots + F_n)$, which is exactly the functionals defined in (2.20). This device for shifting the iterative multipliers by τ rests crucially on the identity (2.14), and hence on the structure of the family of constraints imposed in (P_n). Presumably, with this second form of (A_n) in hand one would expect to be able to treat the general case of (P_n) provided that the initialization of the iterative procedure is close enough to a solution. The choice of a suitable τ would then be dictated by the magnitude of the negative λ_i values corresponding to the solution.

The third stage for (A_n) is precisely the content of §3. The final form of the algorithm differs from the preceding forms in that the (linearized) constraints imposed at each iteration are now relaxed to inequalities. As is shown in Lemma 2.3 and 2.4, the monotonicity and (partial) convergence properties of this iterative procedure hold unconditionally, since $\mu_i^{k+1} \geq 0$ by necessity. Therefore, the algorithm (A_n) defines a sequence (u^{k+1}, λ^{k+1}) which converges in the sense explained above given as *arbitrary* initialization u^0. This means that (A_n) is a globally convergent iterative procedure, in contrast to its two provisional forms with equality constraints. In essence, the reason for this property of (A_n) lies in the fact that it takes full advantage of the convexity attributes of the objective and constraint functionals.

It is necessary however to show that the resulting solution u satisfies the constraints (7.7) as *equalities*. This is accomplished by choosing τ large enough depending on the *solutions* of $\mathcal{S}$ (not on the *iterates*) so that $\lambda_i + \tau > 0$ for each i, and hence that each constraint must be active.