On the Optimal Design of Columns Against Buckling

Steven J. Cox

and

Michael L. Overton

June, 1990

TR90-13

ON THE OPTIMAL DESIGN OF COLUMNS AGAINST BUCKLING

Steven J. Cox† and Michael L. Overton‡

† Mathematical Sciences Department, Rice University, Houston, TX 77251. This author's research was supported by Defense Advanced Research Projects Agency Contract F49620-87-C-0065 and Air Force Office of Scientific Research Grant 89-0363.

‡ Courant Institute of Mathematical Sciences, New York University, New York, NY 10012. This author's research was supported by National Science Foundation Grant CCR-88-02408.

Abstract – We establish existence, derive necessary conditions, and construct and test an algorithm for the maximization of a column's Euler buckling load under a variety of boundary conditions over a general class of admissible designs. We prove that symmetric clamped–clamped columns possess a positive first eigenfunction and introduce a symmetric rearrangement that does not decrease the column's buckling load. Our necessary conditions, expressed in the language of Clarke's generalized gradient [10], subsume those proposed by Olhoff and Rasmussen [25], Masur [22], and Seiranian [32]. The work of [25], [22], and [32] sought to correct the necessary conditions of Tadjbakhsh and Keller [35] who had not foreseen the presence of a multiple least eigenvalue. This remedy has been hampered by Tadjbakhsh and Keller's miscalculation of the buckling loads of their clamped-clamped and clamped–hinged columns. We resolve this issue in the appendix.

In our numerical treatment of the associated finite dimensional optimization problem we build on the work of Overton [26] in devising an efficient means of extracting an ascent direction from the column's least eigenvalue. Owing to its possible multiplicity this is indeed a nonsmooth problem and again the ideas of Clarke [10] are exploited.

1. Introduction

We recall Todhunter's formulation [36, p. 66] of the following problem of Lagrange,

> "To find the curve which by its revolution about an axis in its plane determines the column of greatest efficiency."

For columns of unit length and volume, efficiency here denotes the structure's resistance to buckling under axial compression. When λ is the magnitude of the axial load and u the resulting transverse displacement we postulate the potential energy

$$\int_0^1 EI|u''|^2\,dx - \lambda \int_0^1 |u'|^2\,dx$$

with the two terms measuring bending and elongation respectively. Here I is the second moment of area of the column's cross section and E is its Young's modulus. For sufficiently small λ the minimum of this potential energy, over all admissible displacements, is zero. The (Euler) buckling load of the column is the greatest λ, call it λ_1, for which this minimum is zero. That is,

$$\lambda_1 \doteq \inf_{u \in V} \frac{\int_0^1 EI|u''|^2\,dx}{\int_0^1 |u'|^2\,dx} \tag{1.1}$$

where V is a closed subspace of H^2, the space of L^2 functions on the interval $(0,1)$ with first and second distributional derivatives in L^2, distinguished by the choice of boundary conditions. The choice that has generated the greatest interest is the clamped-clamped condition $u(0) = u'(0) = u(1) = u'(1) = 0$. With the corresponding V denoted by H_0^2 it is not difficult to show that the infimum in (1.1) is attained at some $u_1 \in H_0^2$. First order necessary conditions then require that u_1 satisfy

$$\int_0^1 EIu_1''v''\,dx = \lambda_1 \int_0^1 u_1'v'\,dx, \qquad \forall v \in H_0^2. \tag{1.2}$$

When I and E are smooth it follows from (1.2) that

$$-(EIu_1'')'' = \lambda_1 u_1'', \quad u_1(0) = u_1'(0) = u_1(1) = u_1'(1) = 0. \tag{1.3}$$

With this we recognize (1.1) as Rayleigh's principle for the least eigenvalue of (1.3) and u_1 as an associated first eigenfunction. For the problem of Lagrange the Young's modulus is assumed constant and, as the column is a solid of revolution, each cross section's second moment of area is simply a constant multiple of the square of its area, A, i.e., $I(x) = cA^2(x)$. Fixing our attention on columns of unit volume, we require

$$\int_0^1 A\,dx = 1. \tag{1.4}$$

We have reduced the problem of Lagrange to the search for that A which, subject to (1.4), maximizes the λ_1 of (1.1). This problem, with clamped-clamped boundary conditions, was first attacked in 1962 by Tadjbakhsh and Keller [35] in the continuation of work Keller [20] had begun at the suggestion of Clifford Truesdell. The work of [35] hinges on the necessary condition that the best A, and its corresponding eigenfunction u, satisfy

$$A^4|u''|^2 = A^3 \tag{1.5}$$

along the entire column. This was obtained on formally differentiating a second order analog (see equation (2.5)) of (1.3) with respect to A subject to the integral constraint. Upon reconciling (1.3) and (1.5), Tadjbakhsh and Keller arrived at the representation

$$\begin{aligned} A(x) &= \tfrac{4}{3}\sin^2\theta(x), \quad -\pi/2 \le \theta \le 3\pi/2, \\ \theta(x) &- \tfrac{1}{2}\sin 2\theta(x) + \pi/2 = 2\pi x, \quad 0 \le x \le 1. \end{aligned} \tag{1.6}$$

The most striking aspect of this claim is that it requires the cross sectional area to vanish at $1/4$ and $3/4$. This result should however come as no surprise, for implicit in (1.5) is the assumption that the optimal buckling load is simple, i.e., that the corresponding space of buckled configurations is one dimensional. This requires the optimal column to buckle in much the same way as the uniform column ($A \equiv 1$), the first eigenfunction of which is $U(x) = 1 - \cos(2\pi x)$. The fact that A vanishes at the inflection points of U agrees then with the heuristic (suggested by (1.1)) that the optimal column need be thick only in regions where it bends, i.e., where the magnitude of the linearized curvature $|u''|$ is large.

Tadjbakhsh and Keller claimed $16\pi^2/3$ as the buckling load of the resulting column. It was not until 1977 that Olhoff and Rasmussen [25], observing that (1.3) does not exclude multiple eigenvalues, noted that as the least eigenvalue does not vary smoothly with A at points where its multiplicity exceeds one, the formal differentiation in [35] would be hard to justify. As evidence that Tadjbakhsh and Keller had indeed taken the wrong course, Olhoff and Rasmussen claimed, on the basis of numerical work, 30.51 for the buckling load of the column constructed according to (1.6). Unfortunately, they neglected to describe the means by which this value was arrived at. Indeed the fact that A vanishes at $1/4$ and $3/4$ introduces computational difficulties. Although they did go on to suggest how $16\pi^2/3$ was incorrectly obtained, a number of workers have remained unconvinced, e.g., Myers and Spillers [24] and Barnes [4]. Upon fleshing out the relevant remarks of Olhoff and Rasmussen we shall see, in work relegated to an appendix, that the buckling load for the column proposed by Tadjbakhsh and Keller does not exceed $\pi^2/3$. These same arguments will serve to demonstrate that Tadjbakhsh and Keller's best clamped-hinged column also has a much lower buckling load than thought previously.

Having concluded that differentiating (1.3) would lead to less than optimal columns, Olhoff and Rasmussen presented a 'bimodal formulation' of the problem of Lagrange, i.e., one which would accommodate double eigenvalues. Their subsequent necessary condition paired the best A with two corresponding linearly independent eigenfunctions u, v and a scalar $t \in [0,1]$ so that

$$A\left(t|u''|^2 + (1-t)|v''|^2\right) = 1, \tag{1.7}$$

along the entire column. On implementing an algorithm that enforced this optimality condition, Olhoff and Rasmussen arrived at a column whose cross sectional area was positive throughout and which could withstand loads up to 52.3563. Their methods were however no more rigorous than those of Tadjbakhsh and Keller, and moreover, solely on the basis of claims, the latter still had the stronger column, for $52.3563 < 16\pi^2/3$. Those persuaded by Olhoff and Rasmussen's criticism of the work of Tadjbakhsh and Keller then set out to put (1.7) on a solid foundation. Actually, they joined the discussion of the more general problem: What conditions are necessary for a multiple eigenvalue to attain its extremum? The greatest advances on this question have come in finite dimensions and lie in the apparently little known work of Bratus and Seiranian [6]. These conditions, later discovered independently in a more general form by Overton [26], will be discussed in detail in §5. For now, we note that Bratus and Seiranian, upon applying their finite dimensional arguments to the problem of Lagrange arrived at the conclusion that the best A, must, with two corresponding orthogonal eigenfunctions u, v, satisfy

$$A(\delta_1|u''|^2 + \delta_2|v''|^2 + \delta_3 u''v'') = 1, \quad \text{where } \delta_1\delta_2 \geq \delta_3^2/4. \tag{1.8}$$

This condition was also proposed by Masur [22] who, like Seiranian [32], went on to represent the best A via a system of transcendental equations. Their approximate solutions to these systems are in good agreement, with respective buckling loads of 52.3564 and 52.3565, with that proposed by Olhoff and Rasmussen [25]. Note that (1.7) and (1.8), with the introduction of a second buckling mode, possess mechanisms which, at least in principle, rule out the possibility of columns with vanishing cross sectional area.

Our main contribution to the problem of Lagrange is essentially twofold. We employ the generalized gradient of Clarke in (i) a rigorous derivation of the necessary conditions (1.8) and (ii) the construction of an efficient algorithm to solve the associated finite dimensional optimization problem. Our initial focus on the clamped-clamped case will be extended in §5 to each of the boundary conditions considered by Tadjbakhsh and Keller.

In our discussion of the various optimality criteria something has been conspicuously lacking: the literature contains no proof of the existence of a best A for the problem of Lagrange. Before filling this gap we establish a number of preliminary results and look to a more general problem formulation.

2. The Optimal Design Problem

The moment I is more precisely the second moment of area of the cross section about a line through its centroid normal to the plane of buckling. That is, denoting the cross section by $\Omega(x)$ with centroid at the origin, if η is a unit normal to the plane of buckling then

$$I(x) = \int_{\Omega(x)} |\eta^T y|^2 \, dy. \tag{2.1}$$

When Ω is a circle, in fact when Ω is a regular polygon, this integral does not depend on η, and one finds that I varies as the square of the cross-sectional area, A. On considering so–called thin-walled columns we shall now see that I varies as an affine function of A. On the lateral surface of a cylinder with circular cross section of constant radius R we add a layer of variable thickness $\rho(x)$ with $\rho(x) \leq \varepsilon R$, $\varepsilon \ll 1$. Neglecting powers of ρ greater than one we find $I(x) = \pi R^3 \rho(x) + \pi R^4$ and $A(x) = 2\pi R\rho(x) + \pi R^2$. Taking $\tilde{A}(x) = A(x) - \pi R^2/2$ for our design variable we find $I(x) = (R^2/2)\tilde{A}(x)$. The effect of this choice on the integral constraint is trivial. Of greater interest is that $\tilde{A}$, by construction, must satisfy the pointwise bounds

$$\pi R^2/2 \leq \tilde{A}(x) \leq \pi R^2/2 + 2\varepsilon\pi R^2. \tag{2.2}$$

It is not difficult to continue this line of reasoning and collect a number of examples where I varies as some power of A. We proceed then to consider the case where $EI = \sigma^p$ for some $p > 0$. Compelled by our examination of the previous special cases we admit those σ in

$$ad = \{\sigma \in L^\infty : 0 < \alpha \leq \sigma(x) \leq \beta, \ \int_0^1 \sigma(x)\, dx = 1\}.$$

The weak formulation of the buckled column equation for $\sigma \in ad$ is

$$\int_0^1 \sigma^p u'' v'' \, dx = \lambda \int_0^1 u' v' \, dx, \qquad \forall v \in H_0^2. \tag{2.3}$$

As $\sigma \in L^\infty$ and $\alpha > 0$, (2.3) possesses the sequence of eigenvalues

$$0 < \lambda_1(\sigma) \leq \lambda_2(\sigma) \leq \cdots \uparrow \infty,$$

repeated according to their finite multiplicities and a corresponding sequence of eigenfunctions $\{u_k(\sigma)\}_{k=1}^\infty \subset H_0^2$, orthonormal in terms of the bilinear form associated with the right

side of (2.3). As $H_0^2(0,1) \subset C^1([0,1])$ we find $u_k \in C^1([0,1])$. Upon integrating by parts on the right side of (2.3) we find that $\sigma^p u_k''$ differs from $-\lambda_k(\sigma)u_k$ by an affine function of x. Hence, $\sigma^p u_k'' \in C^1([0,1])$, and, in fact

$$(\sigma^p u_k'')(x) = (\sigma^p u_k'')'(0)x + (\sigma^p u_k'')(0) - \lambda_k(\sigma)u_k(x), \tag{2.4}$$

We collect those eigenfunctions corresponding to $\lambda_1(\sigma)$ in

$$\mathcal{E}(\sigma) = \text{span}\,\{u_k(\sigma);\ \lambda_k(\sigma) = \lambda_1(\sigma)\},$$

a subspace of H_0^2 with dimension equal to the multiplicity of $\lambda_1(\sigma)$. Implicit in Olhoff and Rasmussen's bimodal formulation is the assumption that this multiplicity is at most two. Seiranian [32], has confirmed this through Kamke's analysis of the second order problem with nonseparated boundary conditions

$$w'' + \lambda\sigma^{-p}w = 0, \quad w(1) = w(0) + w'(0), \quad w'(1) = w'(0). \tag{2.5}$$

This is the strong version of (2.3) with $w = \sigma^p u''$ and was first considered in our context by Tadjbakhsh and Keller. Kamke, in [19, §4], proves that the multiplicity of each eigenvalue of (2.5) is no greater than two. Equation (2.4) however, suggests an approach that applies directly to the weak formulation.

If corresponding to $\lambda_k(\sigma)$ there existed three linearly independent eigenfunctions u_1, u_2, u_3 then one could choose scalars a, b, c not all zero such that $v = au_1 + bu_2 + cu_3$ satisfies, in addition to $v(0) = v'(0) = v(1) = v'(1) = 0$, the two conditions $(\sigma^p v'')'(0) = 0$ and $(\sigma^p v'')(0) = 0$. From (2.4) we conclude that v satisfies the homogeneous linear second order equation with zero initial conditions

$$\sigma^p(x)v''(x) + \lambda_k(\sigma)v(x) = 0, \quad v(0) = v'(0) = 0.$$

As the only solution to this equation is the identically zero function we have established

Lemma 2.1. *If $\sigma \in ad$ then the multiplicity of $\lambda_k(\sigma)$ is at most two.*

As the least eigenvalue of the uniform column is $4\pi^2$ we find, as a consequence of the monotonicity of the Rayleigh quotient, that

$$4\pi^2\alpha^p \le \lambda_1(\sigma) \le 4\pi^2\beta^p, \qquad \forall\,\sigma \in ad. \tag{2.6}$$

Corresponding to the least eigenvalue $\lambda_1(\sigma)$ one expects a positive eigenfunction. Indeed, this is the *only* type that Tadjbakhsh and Keller expected. To our knowledge however, there is no proof that a positive first eigenfunction need exist. We remark that on this point the oscillation theory of Kamke is insufficient, for it concludes only that eigenfunctions corresponding to the least nonzero eigenvalue of (2.5) possess either three *or* two zeros. This translates into either one *or* no zero(s) for eigenfunctions corresponding to $\lambda_1(\sigma)$. We now improve on this situation in the case where σ is even (about 1/2), i.e., $\sigma(x) = \sigma(1-x)$.

Theorem 2.2. *If $\sigma \in L^\infty$ is even and admits a positive lower bound then there exists a positive even eigenfunction corresponding to $\lambda_1(\sigma)$.*

Proof: We exploit the essential idea in inverse iteration, a popular technique for computing the least eigenvalue and eigenvector of a symmetric matrix. In our context this idea amounts to approximating the least eigenfunction by the solution of a related nonhomogeneous boundary value problem. Given $v_0 \in H_0^2$ we consider its expansion in the complete set of eigenfunctions $\{u_k(\sigma)\}$,

$$v_0(x) = \overline{v}(x) + \sum_{k=m+1}^{\infty} a_k u_k(x),$$

where $\lambda_m(\sigma) < \lambda_{m+1}(\sigma)$ and $\overline{v}$ is an eigenfunction corresponding to $\lambda_m(\sigma)$. From v_0 we construct the sequence $\{v_j\} \subset H_0^2$ according to

$$\int_0^1 \sigma^p v_j'' \phi''\, dx = \lambda_m(\sigma) \int_0^1 v_{j-1}' \phi'\, dx, \qquad \forall\, \phi \in H_0^2.$$

On expanding v_j in $\{u_k(\sigma)\}$ one finds

$$v_j(x) = \overline{v}(x) + \sum_{k=m+1}^{\infty} a_k \left(\frac{\lambda_m(\sigma)}{\lambda_k(\sigma)}\right)^j u_k(x).$$

As $\lambda_m(\sigma) < \lambda_k(\sigma)$ for all $k > m$ we find that v_j converges pointwise to $\overline{v}$ as $j \to \infty$. We shall now produce a v_0 whose corresponding $\overline{v}$ is even and positive. The earlier work of Kamke will then inform us that $m = 1$. Strictly speaking, Kamke requires that σ be continuous. Atkinson [1] demonstrates that σ need only be integrable.

Our choice for v_0 is the first eigenfunction of the uniform column, i.e., $1 - \cos(2\pi x)$, a positive even function with exactly two inflection points.

Lemma 2.3. *Let f be an even member of L^∞ with a positive lower bound and v be a positive, even member of H_0^2 with precisely two inflection points. If $u \in H_0^2$ satisfies*

$$\int_0^1 f u'' \phi''\, dx = \int_0^1 v' \phi'\, dx, \qquad \forall\, \phi \in H_0^2 \tag{2.7}$$

then u is positive, even, and possesses precisely two inflection points.

Proof of Lemma 2.3: Upon integrating by parts on the right of (2.7) we find that fu'' differs from v by an affine function. Dividing by f and integrating twice gives

$$u(x) = \int_0^x (x-y)(ay + b - v(y))g(y)\, dy \tag{2.8}$$

where $g = 1/f$ and a and b are determined by $u(1) = 0$ and $u'(1) = 0$, i.e., by

$$a \int_0^1 x g(x)\, dx + b \int_0^1 g(x)\, dx = \int_0^1 v(x) g(x)\, dx \tag{2.9}$$

$$a \int_0^1 x^2 g(x)\, dx + b \int_0^1 x g(x)\, dx = \int_0^1 x v(x) g(x)\, dx \tag{2.10}$$

That these equations uniquely determine a and b follows from Hölder's inequality

$$\left(\int_0^1 xg(x)\,dx\right)^2 < \int_0^1 g(x)\,dx \int_0^1 x^2 g(x)\,dx.$$

Our hypotheses in fact allow us to conclude that

$$a = 0, \qquad b = \frac{\int_0^1 v(x)g(x)\,dx}{\int_0^1 g(x)\,dx}.$$

This obviously satisfies (2.9). Regarding (2.10), recall that every even function satisfies $\int_0^1 \phi(x)\,dx = 2\int_0^1 x\phi(x)\,dx$. Consequently,

$$b = \frac{\int_0^1 v(x)g(x)\,dx}{\int_0^1 g(x)\,dx} = \frac{2\int_0^1 xv(x)g(x)\,dx}{2\int_0^1 xg(x)\,dx} = \frac{\int_0^1 xv(x)g(x)\,dx}{\int_0^1 xg(x)\,dx}$$

satisfies (2.10) as well. Labeling $s(x) = (b - v(x))g(x)$, equation (2.8), $u(1) = 0$, and $u'(1) = 0$ take the form

$$u(x) = \int_0^x (x-y)s(y)\,dy, \quad \int_0^1 s(y)\,dy = 0, \quad \int_0^1 ys(y) = 0.$$

With this and the fact that s is even we find

$$\begin{aligned} u(1-x) &= \int_0^{1-x} (1-x-y)s(y)\,dy \\ &= \int_0^1 (1-x-y)s(y)\,dy - \int_{1-x}^1 (1-x-y)s(y)\,dy \\ &= \int_0^x (x-y)s(y)\,dy = u(x). \end{aligned}$$

Regarding the convexity/concavity of u we observe that $f(x)u''(x) = b - v(x)$. That $b - v(x)$ has at least two zeros follows from $b > 0$, $v(0) = v(1) = 0$, and $0 < b < \|v\|_\infty$. For $b - v(x)$ to possess more than two zeros v must admit a local minimum, a condition that requires of v no less then four inflection points. These zeros, say x_0 and $1 - x_0$, are the inflection points of u. As u vanishes at 0 and is convex on $(0, x_0)$ it must be positive there, and, by symmetry, positive on $(1 - x_0, 1)$ as well. As u is positive at x_0 and $1 - x_0$ while concave between these points it must be positive on this interval as well. ∎

It now follows that $\{v_j\}$ is a sequence of positive even functions. The convergence of v_j to $\overline{v}$ being pointwise we conclude that $\overline{v}$ is itself a positive even function. From Kamke [19, §4] we recall that the mth eigenfunction possesses either m or $m - 1$ zeros. Hence $m = 1$, i.e., $\overline{v}$ is a positive even eigenfunction corresponding to $\lambda_1(\sigma)$. ∎

When σ is even and $\lambda_1(\sigma)$ is simple we now have, up to a scalar multiple, a unique positive even first eigenfunction, call it u_1. When $\lambda_1(\sigma)$ is double, in addition to u_1, we deduce from Kamke the existence of a first eigenfunction that changes sign, call it u_2. These two functions comprise a basis for $\mathcal{E}(\sigma)$ and, as u_2 changes sign, we see that au_1+bu_2 can not be even unless $b=0$. Hence, when σ is even, there exists, up to a scalar multiple, a unique positive even first eigenfunction.

Though Theorem 2.2 applies only to even functions we shall see in the next result that this suffices for our purposes. Note that Lemma 2.3 states that the operator $(d^2/dx^2(f\,d^2/dx^2))^{-1}(-d^2/dx^2)$ leaves a subcone of the positive H_0^2 functions invariant when f is even. This cone is however too 'thin' to allow one to invoke Krein-Rutman arguments. Regarding possible improvements of Lemma 2.3 we note that even the constant coefficient operator $(d^4/dx^4)^{-1}(-d^2/dx^2)$ does *not* leave the positive H_0^2 functions invariant. To see this we solve for b in (2.9-10) with $g=1$,

$$b = 4\int_0^1 v\,dx - 6\int_0^1 xv\,dx.$$

Taking for v any smooth positive function supported in $(2/3,1)$ produces $b<0$. As $u(0)=u'(0)=0$ and $u''(0)=b$ we conclude that u is not positive.

Theorem 2.4. *Given $\sigma\in ad$ there exists an even $\sigma_*\in ad$ for which $\lambda_1(\sigma)\le\lambda_1(\sigma_*)$.*

Proof: There is a very simple argument when $0<p\le 1$. Given a function ϕ on $(0,1)$ we denote its even part by $\phi_s(x)=\frac{1}{2}(\phi(x)+\phi(1-x))$. Consider the even function $\tilde\sigma\equiv((\sigma^p)_s)^{1/p}$ and its corresponding even first eigenfunction $\tilde u$. With the normalization $\|\tilde u'\|=1$ we find

$$\lambda_1(\sigma)\le\int_0^1\sigma^p|\tilde u''|^2\,dx=\int_0^1(\sigma^p)_s|\tilde u''|^2\,dx=\int_0^1\tilde\sigma^p|\tilde u''|^2\,dx=\lambda_1(\tilde\sigma). \tag{2.11}$$

As $t\mapsto t^{1/p}$ is convex we observe that

$$\tilde\sigma(x)=\left(\frac{\sigma^p(x)}{2}+\frac{\sigma^p(1-x)}{2}\right)^{1/p}\le\frac{\sigma(x)}{2}+\frac{\sigma(1-x)}{2}=\sigma_s(x). \tag{2.12}$$

Now $\sigma_s\in ad$ and (2.11-12) imply that $\lambda_1(\sigma)\le\lambda_1(\sigma_s)$. Our attempts to argue in a similar fashion for $p>1$ with $\tilde\sigma\equiv((\sigma^{-p})_s)^{-1/p}$ and (2.5) have been thwarted by the fact that $\lambda_1(\sigma)$ corresponds to the *third* eigenvalue of (2.5). What is needed is a rearrangement of σ that echoes the curvature of its corresponding first eigenfunction. To make this precise we first need the following extension of Lemma 2.3.

Recall that a function ϕ is said to be odd about the point $(x_0,\phi(x_0))$ on some interval containing x_0 when

$$\phi(x_0)-\phi(x_0-x)=\phi(x_0-x)-\phi(x_0)$$

for each x on the given interval. If, in addition to the original hypotheses of Lemma 2.3, we assume that f and v when restricted to $(0,1/2)$ are even about $1/4$ and odd about

$(1/4, v(1/4))$ respectively, we conclude that u, when restricted to $(0, 1/2)$, is odd about $(1/4, u(1/4))$.

To see this we recall that $\int_0^1 u''\,dx = 0$ and u'' is even about $1/2$, hence $\int_0^{1/2} u''\,dx = 0$. For the remainder of this paragraph all functions will be restricted to $(0, 1/2)$. Recall as well that $u'' = (b - v)/f$, the quotient of a function odd about $(1/4, b - v(1/4))$ and a function even about $1/4$. Hence u'' is odd about $(1/4, b - v(1/4))$. The condition that $\int_0^{1/2} u''\,dx = 0$ now forces $b = v(1/4)$. As u'' is now odd about $(1/4, 0)$ and $u(0) = u'(0) = 0$ we easily conclude that u is indeed odd about $(1/4, u(1/4))$.

If σ_* is now even about $1/2$ and even about $1/4$ when restricted to $(0, 1/2)$ then beginning the iteration of Theorem 2.2 with a v_0 that is even about $1/2$ and odd about $(1/4, v_0(1/4))$, e.g., $1 - \cos(2\pi x)$, will produce u_*, a positive eigenfunction corresponding to $\lambda_1(\sigma_*)$ that is even about $1/2$ and odd about $(1/4, u_*(1/4))$ on $(0, 1/2)$. We immediately note that σ_*^p and $|u_*''|^2$ are similarly ordered, i.e.,

$$(\sigma_*^p(x_1) - \sigma_*^p(x_2))(|u_*''(x_1)|^2 - |u_*''(x_2)|^2) \geq 0 \qquad \forall\, x_1, x_2 \in (0,1). \tag{2.13}$$

Given $\sigma \in ad$ we now define its rearrangement σ_*.

$$\ell_c = \{x \in (0,1) : \sigma(x) \geq c\}$$

$$\ell_c^* = \begin{cases} \{x \in \mathbf{R} : |x - 1/2| \leq 1/4|\ell_c|\} & \text{if } \ell_c \neq \emptyset \\ \emptyset & \text{otherwise} \end{cases}$$

$$\sigma_*(x) = \begin{cases} \sigma_*(1/2 - x), & \text{if } 0 \leq x \leq 1/4 \\ \sup\{c \in \mathbf{R} : x \in \ell_c^*\}, & \text{if } 1/4 \leq x \leq 3/4 \\ \sigma_*(1 - x), & \text{if } 3/4 \leq x \leq 1. \end{cases}$$

In essence, this distributes half of σ's mass in a symmetrically decreasing fashion about $1/2$ on $(1/4, 3/4)$, completing the rest via symmetry. By construction these two functions are equimeasurable, i.e.,

$$|\{x \in (0,1) : \sigma(x) \geq c\}| = |\{x \in (0,1) : \sigma_*(x) \geq c\}|, \qquad \forall\, c \in \mathbf{R},$$

and consequently, $\sigma_* \in ad$. We are now in position to apply the following result of Hardy, Littlewood and Pólya, see Pólya and Szegö [29, p. 153].

If f and f_1 are equimeasurable, g and g_1 are equimeasurable, $f \in L^q$, $g \in L^{q'}$, and f_1 and g_1 are similarly ordered, then

$$\int_0^1 fg\,dx \leq \int_0^1 f_1 g_1\,dx. \tag{2.14}$$

Given $\sigma \in ad$ we now rearrange it as above into σ_* and consider its corresponding $u_* \in \mathcal{E}(\sigma_*)$. Upon normalizing $\|u_*'\| = 1$ we find

$$\lambda_1(\sigma) \leq \int_0^1 \sigma^p |u_*''|^2\,dx \leq \int_0^1 \sigma_*^p |u_*''|^2\,dx = \lambda_1(\sigma_*).$$

The first inequality is a consequence of Rayleigh's principle, the second, of (2.13-14). ∎

The stage now set, we address, in the next two sections, existence and necessary conditions for the generalized problem of Lagrange

$$\sup_{\sigma \in ad} \lambda_1(\sigma). \tag{2.15}$$

3. Existence

We adopt the direct method of the calculus of variations and neglect to relabel subsequences. Denote by $\hat{\lambda}_1$ the value of (2.15) and by $\{\sigma_n\} \subset ad$ an associated maximizing sequence, i.e., $\lambda_1(\sigma_n) \uparrow \hat{\lambda}_1$. Without loss we may assume that each σ_n is even about 1/2. We abbreviate $\lambda_1(\sigma_n)$ to $\lambda_{1,n}$ and denote by $u_{1,n}$ a corresponding positive eigenfunction for which $\|u'_{1,n}\| = 1$ and $\int_0^1 \sigma_n^p |u''_{1,n}|^2 \, dx = \lambda_{1,n}$ where $\|\cdot\|$ denotes the L^2 norm. These normalizations, in light of (2.6), impose a uniform H^2 bound on the sequence $\{u_{1,n}\}$. As a result, there exists a subsequence with weak H^2 limit $\overline{u} \in H_0^2$. The imbedding of H^2 in H^1 being compact, we find $\|\overline{u}'\| = 1$, and so $\overline{u}$ is not identically zero. The natural question is whether $\hat{\lambda}_1$ and $\overline{u}$ are indeed an eigenvalue and eigenfunction for some column with corresponding $\overline{\sigma} \in ad$. If so, then $\overline{\sigma}$ is necessarily the desired optimal design. This question was first addressed by Senatorov [33] in the context of a second order problem. He discovered that one must consider weak convergence of the reciprocals of the coefficients of the highest order term. This observation continues to hold for fourth order problems, the details of which we now sketch.

Consider the weak formulation

$$\int_0^1 \sigma_n^p u''_{1,n} v'' \, dx = \lambda_{1,n} \int_0^1 u'_{1,n} v' \, dx \qquad \forall v \in H_0^2. \tag{3.1}$$

Our previous remarks reveal that the right hand side converges to $\hat{\lambda}_1 \int_0^1 \overline{u}' v' \, dx$ for each such v. Regarding the left side we define $\xi_n = \sigma_n^p u''_{1,n}$, and, as in (2.4), deduce from (3.1)

$$\xi_n(x) = a_n x + b_n - \lambda_{1,n} u_{1,n}(x), \tag{3.2}$$

for some $a_n, \, b_n \in \mathbf{R}$. As $\{\xi_n\}$ and $\{u_{1,n}\}$ are uniformly bounded in L^2 so to must be $\{a_n x + b_n\}$. From the simple inequality $(a_n^2 + b_n^2) \leq 16 \|a_n x + b_n\|^2$ we find $\{a_n\}$ and $\{b_n\}$ to be uniformly bounded, with the result that ξ_n converges strongly in L^2 to a function $\overline{\xi}$. The left side of (3.1) therefore converges to $\int_0^1 \overline{\xi} v'' \, dx$. It remains to characterize this $\overline{\xi}$. Recalling the pointwise bounds on the σ_n we may assume that σ_n^{-p} converges in the weak* topology of L^∞ to some function μ. Thus, $\xi_n \sigma_n^{-p}$ converges weakly in L^2 to $\mu\overline{\xi}$. But $\xi_n \sigma_n^{-p} = u''_{1,n}$, whose weak L^2 limit is $\overline{u}''$. Hence, $\overline{\xi} = \overline{u}'' \mu^{-1}$. Defining $\overline{\sigma} = \mu^{-1/p}$ we may pass to the limit in (3.1) and obtain

$$\int_0^1 \overline{\sigma}^p \overline{u}'' v'' \, dx = \hat{\lambda}_1 \int_0^1 \overline{u}' v' \, dx \qquad \forall v \in H_0^2.$$

As a result, $\hat{\lambda}_1 = \lambda_j(\overline{\sigma})$. Now $\overline{u}$, being the strong L^2 limit of a sequence of positive functions is itself positive, and, as a consequence, $j = 1$. We need only determine whether $\overline{\sigma} \in ad$. One may verify the pointwise bounds without trouble. With respect to the integral constraint we consider the convex function $f : \mathbf{R} \to \mathbf{R}$, $f(t) = t^{-1/p}$. The integral functional $\varphi \mapsto \int_0^1 f(\varphi(x))\,dx$ is then weak* lower semicontinuous on L^∞, see e.g., Dacorogna [12, Theorem 1.1]. As $1/\sigma_n^p$ converges weak* to $1/\overline{\sigma}^p$, this allows us to conclude that

$$\int_0^1 \overline{\sigma}\,dx = \int_0^1 f(1/\overline{\sigma}^p)\,dx \leq \liminf \int_0^1 f(1/\sigma_n^p)\,dx = \lim \int_0^1 \sigma_n\,dx = 1. \qquad (3.3)$$

If indeed equality does not hold in (3.3) then there exists a $\hat{\sigma} \in ad$ such that $\hat{\sigma}(x) \geq \overline{\sigma}(x)$ for almost every $x \in (0,1)$. From Rayleigh's principle we then easily deduce $\lambda_1(\hat{\sigma}) \geq \lambda_1(\overline{\sigma})$. Finally, from Theorem 2.4 we confirm that $\hat{\sigma} = \hat{\sigma}_s$. We have now proven

Theorem 3.1. *There exists an even $\hat{\sigma} \in ad$ for which $\lambda_1(\sigma) \leq \lambda_1(\hat{\sigma})$ for every $\sigma \in ad$.*

Our choice of ad was motivated by our interest in the "shape" of the strongest column. Theorem 3.1 however, may also be applied in the search for the "composition" of the strongest column. For example, consider the design problem where one must combine two materials in fixed proportion so as to maximize the buckling load of the resulting column. The set of admissible designs is then

$$ad_E = \{\alpha\chi(x) + \beta(1-\chi(x)) :\ \chi \text{ is the characteristic function of a subset of (0,1) with measure } \gamma\},$$

where α and β are the Youngs moduli of the respective materials with γ the volume fraction of the first. In this context, Theorem 3.1 states that λ_1 attains its maximum on the weak* closure of ad_E, i.e., on

$$ad_E^* = \{\alpha\theta(x) + \beta(1-\theta(x)) :\ 0 \leq \theta(x) \leq 1,\ \int_0^1 \theta\,dx = \gamma\}.$$

4. Necessary Conditions

We search now for a characterization of our optimal design, $\hat{\sigma}$. Typical of many such problems, two distinct approaches are possible. Taking advantage of the variational structure, the so–called direct approach attempts to swap the order of the limits in

$$\hat{\lambda}_1 = \lambda_1(\hat{\sigma}) = \sup_{\sigma \in ad} \inf_{u \in H_0^2} \frac{\int_0^1 \sigma^p |u''|^2\,dx}{\int_0^1 |u'|^2\,dx},$$

inferring necessary conditions from the resulting saddle point. The indirect approach strives to determine $\hat{\sigma}$ through knowledge of the tangents to the graph of $\sigma \mapsto \lambda_1(\sigma)$ and the normals to ad. Our implementation of these two approaches intersect in their reliance on (i) recent work of Auchmuty [2] on dual variational principles and (ii) a lopsided minimax principle.

Proposition 4.1. $\lambda_1^{-1}(\sigma) = \sup_{u \in H_0^2} \mathcal{A}(\sigma, u)$, $\qquad \mathcal{A}(\sigma, u) = \sqrt{2}\|u'\| - \frac{1}{2}\int_0^1 \sigma^p |u''|^2\, dx.$

$u \mapsto \mathcal{A}(\sigma, u)$ attains its maximum only on those $u \in \mathcal{E}(\sigma)$ for which $\|u'\| = \sqrt{2}\lambda_1^{-1}(\sigma)$.

Proof: In addition to being bounded above by $2\lambda_1^{-1}(\sigma)$, the map $u \mapsto \mathcal{A}(\sigma, u)$ is coercive and weakly upper semicontinuous on H_0^2 and therefore attains its maximum at some $\overline{u} \in H_0^2$. Necessarily, $\mathcal{D}_2\mathcal{A}(\sigma, \overline{u})$, the Gâteaux derivative of $u \mapsto \mathcal{A}(\sigma, u)$ at $\overline{u}$, must vanish. That is,

$$\int_0^1 \sigma^p \overline{u}'' v''\, dx = \sqrt{2}\|\overline{u}'\|^{-1} \int_0^1 \overline{u}' v'\, dx \qquad \forall\, v \in H_0^2.$$

As a result, $\overline{u}$ is an eigenfunction corresponding to the eigenvalue $\sqrt{2}\|\overline{u}'\|^{-1}$. As $\overline{u}$ maximizes $u \mapsto \mathcal{A}(\sigma, u)$, this must be the least eigenvalue, $\lambda_1(\sigma)$. ∎

Proposition 4.2. *For $F : X \times Y \to \mathbf{R}$ with $x \mapsto F(x, y)$ concave and upper semicontinuous and $y \mapsto F(x, y)$ convex and lower semicontinuous, if there exists a $y_o \in Y$ and $c_o \in \mathbf{R}$ such that $\{x \in X;\ F(x, y_o) \geq c_o\}$ is compact and $c_o < \inf_{y \in Y} \sup_{x \in X} F(x, y)$ then*

$$\sup_{x \in X} \inf_{y \in Y} F(x, y) = \inf_{y \in Y} \sup_{x \in X} F(x, y).$$

Proof: This is a weakening of Theorem 3.7, Chapter 2, in Barbu and Precupanu [3]. ∎

It is with the indirect approach that we shall meet with the greatest success. For prior attempts in this context see Haug and Rousselet [18] and Choi and Haug [7]. Our principal tool is the generalized gradient of Clarke [10].

For a real valued Lipschitz function F on a Banach space X we consider the generalized directional derivative of F at x in the direction v,

$$F^o(x; v) \equiv \limsup_{\substack{y \to x \\ t \downarrow 0}} \frac{F(y + tv) - F(y)}{t}.$$

Denoting the dual of X by X^* and $x^*(x)$ by $\langle x^*, x \rangle$ when $x^* \in X^*$ and $x \in X$, Clarke's generalized gradient of F at x is the nonempty, convex, weak* compact set

$$\partial F(x) \equiv \{\xi \in X^*;\ F^o(x; v) \geq \langle \xi, v \rangle\ \forall v \in X\}.$$

We demonstrate that $\sigma \mapsto \lambda_1^{-1}(\sigma)$ is Lipschitz on $\Sigma = \{\sigma \in L^\infty;\ \|\hat{\sigma} - \sigma\|_\infty < \alpha/2\}$. Choose $\sigma_1, \sigma_2 \in \Sigma$ such that $\lambda_1^{-1}(\sigma_1) > \lambda_1^{-1}(\sigma_2)$ and note that for $u_1 \in \operatorname{Argmax} \mathcal{A}(\sigma_1, \cdot)$, the set on which $u \mapsto \mathcal{A}(\sigma_1, u)$ attains its maximum,

$$\begin{aligned} |\lambda_1^{-1}(\sigma_1) - \lambda_1^{-1}(\sigma_2)| &\leq |\mathcal{A}(\sigma_1, u_1) - \mathcal{A}(\sigma_2, u_1)| \\ &\leq \int_0^1 |\sigma_1^p - \sigma_2^p| |u_1''|^2\, dx \\ &\leq \|u_1''\|^2 \|\sigma_1^p - \sigma_2^p\|_\infty \leq \alpha^{-2p} p |2\beta|^{p-1} \|\sigma_1 - \sigma_2\|_\infty. \end{aligned}$$

Without loss we assume that $\lambda_1(\hat{\sigma})$ is a double eigenvalue. Then $\mathcal{E}(\hat{\sigma})$ is two dimensional and $\operatorname{Argmax}\mathcal{A}(\hat{\sigma},\cdot)$ is the intersection of $\mathcal{E}(\hat{\sigma})$ with the sphere $\|u'\| = \sqrt{2}\hat{\lambda}_1^{-1}$. It will be convenient to choose a basis $\{\hat{u}_1, \hat{u}_2\}$ for $\mathcal{E}(\hat{\sigma})$ for which $\int_0^1 \hat{u}_i'\hat{u}_j'\,dx = 2\delta_{ij}\hat{\lambda}_1^{-2}$. For then,

$$\operatorname{Argmax}\mathcal{A}(\hat{\sigma},\cdot) = \{a\hat{u}_1 + b\hat{u}_2;\ a^2 + b^2 = 1\}. \tag{4.1}$$

Regarding the Gâteaux derivative of $\sigma \mapsto \mathcal{A}(\sigma, u)$ at $\hat{\sigma}$ in the direction η we have

$$\langle \mathcal{D}_1\mathcal{A}(\hat{\sigma}, u), \eta\rangle = -\tfrac{p}{2}\int_0^1 \eta\hat{\sigma}^{p-1}|u''|^2\,dx. \tag{4.2}$$

Denoting convex hull by 'co', the sense in which the gradient of a maximum is the maximum of the gradients is

Theorem 4.3. $\partial\lambda_1^{-1}(\hat{\sigma}) = \operatorname{co}\{-\frac{p}{2}\hat{\sigma}^{p-1}(a\hat{u}_1'' + b\hat{u}_2'')^2;\ a^2 + b^2 = 1\}$.

Proof: From (4.1) and (4.2) this set is precisely

$$\operatorname{co}\{\mathcal{D}_1\mathcal{A}(\hat{\sigma}, u);\ u \in \operatorname{Argmax}\mathcal{A}(\hat{\sigma},\cdot)\}. \tag{4.3}$$

Our claim does not fit neatly into Clarke's result [10, Theorem 2.8.2] on the generalized gradient of a pointwise maximum. The contortions involved in fitting our problem to Clarke's hypotheses are no less difficult, and far less instructive, than an independent proof.

Let us denote the set in (4.3) by Ξ. We show that $\Xi \subset \partial\lambda_1^{-1}(\hat{\sigma})$. For $\xi \in \Xi$ and $\eta \in L^\infty$

$$\begin{aligned}
\langle \xi, \eta\rangle &= \sum_{i=1}^n \mu_i \langle \mathcal{D}_1\mathcal{A}(\hat{\sigma}, u_i), \eta\rangle \\
&= \sum_{i=1}^n \mu_i \lim_{t\downarrow 0} \frac{\mathcal{A}(\hat{\sigma} + t\eta, u_i) - \mathcal{A}(\hat{\sigma}, u_i)}{t} \\
&\le \sum_{i=1}^n \mu_i \limsup_{t\downarrow 0} \frac{\lambda_1^{-1}(\hat{\sigma} + t\eta) - \lambda_1^{-1}(\hat{\sigma})}{t} \\
&\le (\lambda_1^{-1})^o(\hat{\sigma}; \eta),
\end{aligned}$$

hence $\xi \in \partial\lambda_1^{-1}(\hat{\sigma})$.

Regarding the reverse inclusion we define

$$g(\sigma;\eta) = \max_{\xi\in\Xi} \langle \xi, \eta\rangle, \qquad \sigma \in \Sigma,\ \eta \in L^\infty$$

and prove

$$(\lambda_1^{-1})^o(\sigma;\eta) \le g(\sigma;\eta).$$

Select $\sigma^n \to \sigma$ in L^∞ and $t_n \downarrow 0$ in $\mathbf{R}$ such that

$$q_n \equiv \frac{\lambda_1^{-1}(\sigma^n + t_n\eta) - \lambda_1^{-1}(\sigma^n)}{t_n}$$

converges to $(\lambda_1^{-1})^o(\sigma;\eta)$. Select $u_n \in \operatorname{Argmax} \mathcal{A}(\sigma^n + t_n\eta, \cdot)$ and note that

$$q_n \leq \frac{\mathcal{A}(\sigma^n + t_n\eta, u_n) - \mathcal{A}(\sigma^n, u_n)}{t_n}$$

with the right side equal to $\langle \mathcal{D}_1 \mathcal{A}(\sigma_n + \bar{t}_n\eta, u_n), \eta\rangle$, for some $\bar{t}_n \in (0, t_n)$, by the Mean Value Theorem. As $\sigma^n + t_n\eta \to \sigma$ in L^∞ and $u_n \in \operatorname{Argmax} \mathcal{A}(\sigma^n + t_n\eta, \cdot)$ we recall from our work in Theorem 3.1 that $u_n \to u \in \operatorname{Argmax} \mathcal{A}(\sigma, \cdot)$ in H^2 and $(\sigma^n + t_n\eta)^p u_n'' \to \sigma^p u''$ in L^2, and hence $u_n'' \to u''$ in L^2, i.e., $u_n \to u$ in H^2. Recalling (4.2) this establishes

$$\langle \mathcal{D}_1 \mathcal{A}(\sigma_n + \bar{t}_n\eta, u_n), \eta\rangle \to \langle \mathcal{D}_1 \mathcal{A}(\sigma, u), \eta\rangle$$

with $u \in \operatorname{Argmax} \mathcal{A}(\sigma, \cdot)$. As a result,

$$(\lambda_1^{-1})^o(\sigma;\eta) = \lim_{n\to\infty} q_n \leq \langle \mathcal{D}_1 \mathcal{A}(\sigma, u), \eta\rangle \leq g(\sigma;\eta).$$

If ζ is now an element of $\partial\lambda_1^{-1}(\hat{\sigma})$ then $g(\hat{\sigma};\eta) \geq \langle \zeta, \eta\rangle$ for each $\eta \in L^\infty$. Consequently,

$$0 = \min_{\eta\in L^\infty} \max_{\xi\in\Xi} \langle \xi - \zeta, \eta\rangle.$$

Noting that Ξ is closed and bounded in L^1 and finite dimensional (it lies in the span of $\{|\hat{u}_1''|^2, \hat{u}_1''\hat{u}_2'', |\hat{u}_2''|^2\}$) we find it compact in $(L^\infty)^*$. Invoking Proposition 4.2 yields a $\hat{\xi} \in \Xi$ for which

$$\langle \hat{\xi} - \zeta, \eta\rangle = 0 \qquad \forall \eta \in L^\infty.$$

It follows that $\zeta = \hat{\xi}$ and so $\partial\lambda_1^{-1}(\hat{\sigma}) \subset \Xi$. ∎

This proof, though identical in outline to Clarke [10, Theorem 2.8.2], has exploited additional properties of λ_1 and $\mathcal{A}$ to make up for the missing hypotheses. Observe that when $\lambda_1(\hat{\sigma})$ is simple the generalized gradient reduces to the singleton

$$\partial\lambda_1^{-1}(\hat{\sigma}) = \{\mathcal{D}_1 \mathcal{A}(\hat{\sigma}, \hat{u}_1)\}.$$

As zero is not a tangent direction to λ_1^{-1} at $\hat{\sigma}$, i.e., $0 \neq \partial\lambda_1^{-1}(\hat{\sigma})$, we are compelled to investigate the constraint set *ad*. Separating the equality from the inequality constraints brings

$$C \equiv \{\sigma \in L^\infty;\ \alpha \leq \sigma(x) \leq \beta\} \quad \text{and} \quad V(\sigma) \equiv \int_0^1 \sigma\, dx.$$

As $\hat{\sigma}$ minimizes $\sigma \mapsto \lambda_1^{-1}(\sigma)$ subject to $\sigma \in C$ and $V(\sigma) = 1$ we deduce from the Lagrange Multiplier Rule, [10, Theorem 6.1.1], that a nontrivial linear combination of elements in $\partial\lambda_1^{-1}(\hat{\sigma})$ and $\partial V(\hat{\sigma})$ is normal to C at $\hat{\sigma}$. In particular,

$$\big(\nu_1\partial\lambda_1^{-1}(\hat{\sigma}) + \nu_2\partial V(\hat{\sigma})\big) \cap N_C(\hat{\sigma}) \neq \emptyset,$$

where $\nu_1 \leq 0$, $\nu_1^2 + \nu_2^2 > 0$, and

$$N_C(\hat{\sigma}) = \{\zeta \in (L^\infty)^*;\ \int_0^1 (\hat{\sigma} - \sigma)\,\zeta(dx) \geq 0,\ \forall \sigma \in C\}$$

is the cone of normals to C at $\hat{\sigma}$. In light of our previous calculations, and the fact that $\partial V(\hat{\sigma}) = 1 \in L^1$, there exists a $\hat{\xi} \in \partial\lambda_1^{-1}(\hat{\sigma})$ for which

$$\int_0^1 (\hat{\sigma} - \sigma)(\nu_1\hat{\xi} + \nu_2)\,dx \geq 0, \qquad \forall\, \sigma \in C. \tag{4.4}$$

Observing that $\nu_1\hat{\xi} \geq 0$ we find that $\nu_2 \geq 0$ requires, through (4.4), that $\hat{\sigma} \equiv \beta$, an impossibility. Likewise, should $\nu_1 = 0$, (4.4) would require $\hat{\sigma} \equiv \alpha$ (since $\nu_2 < 0$). Taking $\ell^2 = \nu_2/\nu_1$ we arrive at

$$\int_0^1 (\hat{\sigma} - \sigma)(\hat{\xi} + \ell^2)\,dx \leq 0, \qquad \forall\, \sigma \in C.$$

The subsequent reduction to pointwise optimality conditions follows a well known course, see e.g., Cea and Malanowski [8]. In particular,

$$\hat{\sigma}(x) = \alpha \Rightarrow -\hat{\xi}(x) \leq \ell^2 \tag{4.5}$$

$$\alpha < \hat{\sigma}(x) < \beta \Rightarrow -\hat{\xi}(x) = \ell^2 \tag{4.6}$$

$$\hat{\sigma}(x) = \beta \Rightarrow -\hat{\xi}(x) \geq \ell^2 \tag{4.7}$$

for almost every $x \in (0,1)$. To appreciate this result we must recall that $\hat{\xi} \in \partial\lambda_1^{-1}(\hat{\sigma})$ means

$$-\hat{\xi}(x) = \tfrac{p}{2}\sum_{i=1}^n t_i\hat{\sigma}^{p-1}(x)(a_i\hat{u}_1''(x) + b_i\hat{u}_2''(x))^2, \quad \text{where}$$

$$t_i \geq 0,\ \sum_{i=1}^n t_i = 1,\ \text{and } a_i^2 + b_i^2 = 1.$$

On expanding this sum of squares, (4.6) becomes

$$\hat{\sigma}^{p-1}\left(\delta_1|\hat{u}_1''|^2 + \delta_2|\hat{u}_2''|^2 + \delta_3\hat{u}_1''\hat{u}_2''\right) = 1, \quad \text{where} \tag{4.8}$$

$$\delta_1 = \tfrac{p}{2}\sum_{i=1}^n t_i a_i^2/\ell^2, \quad \delta_2 = \tfrac{p}{2}\sum_{i=1}^n t_i b_i^2/\ell^2, \quad \delta_3 = p\sum_{i=1}^n t_i a_i b_i/\ell^2.$$

Observing that $\delta_1\delta_2$ indeed dominates $\delta_3^2/4$, we have recovered (1.8), the necessary condition of Bratus and Seiranian [6] and Masur [22]. With Seiranian [32] we remark that one may chose a basis for $\mathcal{E}(\hat{\sigma})$ that renders δ_3 zero. Ignoring any bound constraints, Masur [22] and Seiranian [32] found a σ and two orthogonal elements of $\mathcal{E}(\sigma)$ for which

(4.8) holds with $p = 2$. This appears to be the design obtained by Olhoff and Rasmussen [25] and, by all indications, the one preferred by our algorithm as well (see §7 Fig. 1). It appears likely that in this case the bound constraints are inactive due to the fact that where $\hat{\sigma}$ is less than one, $\hat{\sigma}^2$ is much less than one. As $\hat{\sigma}^2$ is the quantity that appears in the Rayleigh quotient we expect it to be as large as possible. This suggests that $\hat{\sigma}$ is bounded away from zero, independent of α. This lower bound with the integral constraint together support the conjecture that $\hat{\sigma}$ is in fact bounded above as well. Hence, when α and β are respectively chosen below and above these 'natural' bounds, condition (4.8) is free to stand on its own. Clearly these natural bounds must depend on p. In fact, we shall provide numerical evidence in §7 in favor of the argument that the natural lower (upper) bound is an increasing (decreasing) function of p for $p > 1$.

Unfortunately, it is not known whether (4.8) is a sufficient condition for optimality. The proof of sufficiency offered by Tadjbakhsh and Keller [35] is incorrect. They proceed as if $\lambda_1(\sigma)$ corresponds to the least eigenvalue of (2.5) and accordingly admit all functions that satisfy the boundary conditions as test functions in a Rayleigh principle argument. In fact, (2.5) possesses a double zero eigenvalue, hence only those functions that are orthogonal to the first two eigenfunctions can be admitted. We remark that Ramm's claim [30], that Tadjbakhsh and Keller mistakenly applied Hölders inequality in their sufficiency proof, is incorrect, though [35, §6 (25)] is only valid for $n < 0$.

Though (4.8) need not hold over the entire length of the column, we now show that where it does hold it requires that $\hat{\sigma}$ be smooth.

Theorem 4.4. *If $\alpha < \hat{\sigma}(x) < \beta$ for each $x \in (a, b) \subset (0, 1)$ then $\hat{\sigma} \in C^\infty(a, b)$.*

Proof: We observed in (2.4) that

$$\hat{\sigma}^p \hat{u}_i'' = l_i - \hat{\lambda}_1 \hat{u}_i, \tag{4.9}$$

where l_i is an affine function of x. Now multiply (4.8) by $\hat{\sigma}^{p+1}$,

$$\delta_1(\hat{\sigma}^p \hat{u}_1'')^2 + \delta_2(\hat{\sigma}^p \hat{u}_2'')^2 + \delta_3(\hat{\sigma}^p \hat{u}_1'')(\hat{\sigma}^p \hat{u}_2'') = \hat{\sigma}^{p+1}. \tag{4.10}$$

From (4.9) we find, on recalling $H_0^2 \subset C^1$, that each term on the left of (4.10) is C^1, and hence that $\hat{\sigma} \in C^1$. Writing (4.9) in the form

$$\hat{u}_i'' = \frac{l_i - \hat{\lambda}_1 \hat{u}_i}{\hat{\sigma}^p},$$

we conclude $\hat{u}_i'' \in C^1$, that is $\hat{u}_i \in C^3$. Repeating this exact argument leads to $\hat{\sigma} \in C^3$ and $\hat{u}_i \in C^5$. The result then follows from continued repetitions. ∎

Having succeeded in pursuing the indirect approach we now look to the possibility (and the implications) of exchanging the limits in the characterization

$$\hat{\lambda}_1^{-1} = \mathcal{A}(\hat{\sigma}, \hat{u}) = \inf_{\sigma \in ad} \sup_{u \in H_0^2} \mathcal{A}(\sigma, u).$$

Recalling Proposition 4.2, this will require convexity and lower semicontinuity of $\sigma \mapsto \mathcal{A}(\sigma, u)$, and concavity and upper semicontinuity of $u \mapsto \mathcal{A}(\sigma, u)$ as well as compactness of one of its upper level sets.

Remark 4.5. *We noted the weak H^2 upper semicontinuity of $u \mapsto \mathcal{A}(\sigma, u)$ in Proposition 4.1. As $\{u \in H_0^2;\ \mathcal{A}(\sigma, u) \geq c\}$ is bounded it is also weakly compact (independent of $c \in \mathbf{R}$ and $\sigma \in ad$). Convexity of $\sigma \mapsto \mathcal{A}(\sigma, u)$ follows on restricting $p \leq 1$.*

The two remaining properties will require more work. Note that $u \in \operatorname{Argmax} \mathcal{A}(\sigma, \cdot)$ implies $\mathcal{A}(\sigma, u) = \mathcal{A}(\sigma, -u) = \lambda_1^{-1}(\sigma)$ while $\mathcal{A}(\sigma, 0) = 0$. Hence, $u \mapsto \mathcal{A}(\sigma, u)$ is not concave on any set that contains $\operatorname{Argmax} \mathcal{A}(\sigma, \cdot)$. This suggests that we examine the half-spaces exterior to $\{u \in H_0^2;\ \|u'\| \leq \sqrt{2}\lambda_1^{-1}(\sigma)\}$. Unfortunately this ball, and hence its support planes, depend on σ. Consequently, if we expect these half-spaces to vary continuously with σ we must be careful in our choosing. This choice is greatly facilitated by the assumption that σ lies in ad_*, those functions in ad that are even about $1/2$. For in this case one may speak unambiguously of u_*, the positive even eigenfunction corresponding to $\lambda_1(\sigma_*)$. We normalize $\|u_*'\| = \sqrt{2}\lambda_1^{-1}(\sigma_*)$, and consider the associated half-space

$$\Pi_{\sigma_*} = \{v \in H_0^2;\ \lambda_1^2(\sigma_*) \int_0^1 u_*' v'\, dx > 2\}.$$

Proposition 4.6. *For $\sigma_* \in ad_*$, $u \mapsto \mathcal{A}(\sigma_*, u)$ is concave on Π_{σ_*}.*

Proof: The quadratic form associated with the second Gâteaux derivative of $u \mapsto \mathcal{A}(\sigma_*, u)$ at $\overline{u} \in \Pi_{\sigma_*}$ satisfies

$$\begin{aligned}
\langle \mathcal{D}_2^2 \mathcal{A}(\sigma_*, \overline{u}) v, v\rangle &= \sqrt{2}\|\overline{u}'\|^{-1} \int_0^1 |v'|^2\, dx - \int_0^1 \sigma_*^p |v''|^2\, dx - \sqrt{2}^{-1}\|\overline{u}'\|^{-3} \left(\int_0^1 \overline{u}' v'\, dx\right)^2 \\
&\leq (\sqrt{2}\|\overline{u}'\|^{-1} - \lambda_1(\sigma_*)) \int_0^1 |v'|^2\, dx \\
&\leq 0, \qquad \forall\, v \in H_0^2. \ \blacksquare
\end{aligned}$$

This suggests that we penalize $\mathcal{A}$ with the indicator function of Π_{σ_*},

$$\pi(\sigma_*, u) = \begin{cases} 0, & \text{if } u \in \Pi_{\sigma_*}; \\ \infty, & \text{otherwise.} \end{cases}$$

This not only guarantees concavity but also respects lower semicontinuity.

Proposition 4.7. *$\sigma_* \mapsto \mathcal{A}(\sigma_*, u) - \pi(\sigma_*, u)$ is lower semicontinuous for the strong L^∞ topology on ad_*.*

Proof: Now $\sigma_*^n \to \sigma_*$ in L^∞ clearly implies $\mathcal{A}(\sigma_*^n, u) \to \mathcal{A}(\sigma_*, u)$ for each $u \in H_0^2$. Regarding $\limsup \pi(\sigma_*^n, u) \leq \pi(\sigma_*, u)$, it suffices to show that

$$\pi(\sigma_*^n, u) \to 0, \qquad \forall\, u \in \Pi_{\sigma_*}.$$

From the proof of Theorem 3.1 it is clear that $\lambda_1(\sigma_*^n) \to \lambda_1(\sigma_*)$ and $u_*(\sigma_*^n) \to u_*(\sigma_*)$ in H_0^2. Hence

$$\lambda_1^2(\sigma_*^n) \int_0^1 u_*'(\sigma_*^n) u'\, dx \to \lambda_1^2(\sigma_*) \int_0^1 u_*(\sigma_*) u'\, dx > 2.$$

From this we conclude that u is eventually in each $\Pi_{\sigma_*^n}$, i.e., $\pi(\sigma_*^n, u) = 0$. $\blacksquare$

One may now modify Proposition 4.2 (see Cox and McLaughlin [11, §7]), and conclude

Theorem 4.8. *If $p \leq 1$ then $(\hat{\sigma}, u_*(\hat{\sigma}))$ is a saddle point for $\mathcal{A}$ over $ad_* \times H_0^2$. That is, denoting $u_*(\hat{\sigma})$ by $\hat{u}$,*

$$\mathcal{A}(\hat{\sigma}, u) \leq \mathcal{A}(\hat{\sigma}, \hat{u}) \leq \mathcal{A}(\sigma_*, \hat{u}) \qquad \forall (\sigma_*, u) \in ad_* \times H_0^2.$$

The latter inequality yields the following maximum principle

$$\int_0^1 \sigma_*^p |\hat{u}''|^2 \, dx \leq \int_0^1 \hat{\sigma}^p |\hat{u}''|^2 \, dx \qquad \forall \sigma_* \in ad_*.$$

The subsequent pointwise conditions call for an $\ell^2 > 0$ such that

$$\hat{\sigma}(x) = \alpha \Rightarrow \hat{\sigma}^{p-1}(x)|\hat{u}''(x)|^2 \leq \ell^2 \tag{4.11}$$

$$\alpha < \hat{\sigma}(x) < \beta \Rightarrow \hat{\sigma}^{p-1}(x)|\hat{u}''(x)|^2 = \ell^2 \tag{4.12}$$

$$\hat{\sigma}(x) = \beta \Rightarrow \hat{\sigma}^{p-1}(x)|\hat{u}''(x)|^2 \geq \ell^2 \tag{4.13}$$

for almost every $x \in (0,1)$. As $\sigma \mapsto \lambda_1^{-1}(\sigma)$ is convex when $p \leq 1$ these conditions are also sufficient. Hence, we see that where the direct method applies it gives more information. In particular, the necessary conditions (4.11-13) involve only a single buckling mode. Comparing these to the more general conditions in (4.5-7) suggests that $\lambda_1(\hat{\sigma})$ is indeed a simple eigenvalue when $p \leq 1$. We shall see numerical evidence of this in §7. The critical case, $p = 1$, where the optimal buckling load changes multiplicity, has received considerable attention. In this case, the right side of (4.11-13) is independent of $\hat{\sigma}$. In particular, a number of workers have claimed that

$$|\hat{u}''(x)| = \ell. \tag{4.14}$$

We remark however that in the absence of a second buckling mode the bound constraints must become active near the inflection points of $\hat{u}$, making (4.11) and (4.13) indeed necessary. Nonetheless, Seiranian [32], who deduced (4.14) from (1.8), proceeded to solve (4.14) in conjunction with (2.3), yielding

$$\hat{\sigma}(x) = \begin{cases} 3/2(1 - 16x^2), & \text{if } 0 \leq x \leq 1/4 \\ 3/2(16x - 16x^2 - 3), & \text{if } 1/4 \leq x \leq 3/4 \\ 3/2(32x - 16x^2 - 15), & \text{if } 3/4 \leq x \leq 1. \end{cases} \tag{4.15}$$

On evaluating the Rayleigh quotient with this $\hat{\sigma}$ and a specific C^1 test function Seiranian arrived at a buckling load of 48. This design, like that of Tadjbakhsh and Keller for $p = 2$, vanishes at $1/4$ and $3/4$. Unlike the design of Tadjbakhsh and Keller however, we are not able to show it to be suboptimal. We can only stress that lacking an existence proof for $\alpha = 0$, $p = 1$, there is no reason to believe that (4.14) is a necessary condition for optimality.

5. Other Boundary Conditions

Intent on a clean exposition we have to this point concentrated solely on the clamped-clamped boundary conditions $u(0) = u'(0) = u(1) = u'(1) = 0$. We now apply the work of the previous sections to the other standard sets of boundary conditions, in particular, hinged and free. A column is said to be free at a point when no conditions are prescribed, while it is hinged, or simply supported, when one requires its displacement to vanish there. As a matter of notation, the weak formulation of the buckled column equation will read

$$\int_0^1 \sigma^p u'' v'' \, dx = \mu \int_0^1 u' v' \, dx, \qquad \forall v \in V_{i,j} \tag{5.1}$$

where $V_{i,j}$ is a subspace of H^2, with i and j chosen from $\{0, 1, 2\}$ according to whether the respective end is either free, hinged, or clamped. For example,

$$V_{1,2} = \{u \in H^2;\ u(0) = 0,\ u(1) = u'(1) = 0\}$$

specifies the hinged-clamped column. We denote the least eigenvalue of (5.1) by $\mu_{i,j}(\sigma)$, and the corresponding space of eigenfunctions by $\mathcal{E}_{i,j}(\sigma)$. As before, $u \in \mathcal{E}_{i,j}(\sigma)$ implies that both u and $\sigma^p u''$ are elements of $C^1([0,1])$. In addition, such functions satisfy so called natural boundary conditions. In particular, if $i = 1$ then, in addition to $u(0) = 0$ we find

$$\sigma^p u''(0) = 0, \tag{5.2}$$

while if $i = 0$ we have, in addition to (5.2),

$$(\sigma^p u'')'(0) + \mu_{0,j}(\sigma) u'(0) = 0. \tag{5.3}$$

We shall consider only those $\mu_{i,j}(\sigma)$ for which $i + j \geq 2$, as otherwise $\mu_{i,j}(\sigma) = 0$. For comparison purposes we record these eigenvalues in the case of the uniform column.

$$\mu_{0,2}(1) = \pi^2/4, \quad \mu_{1,1}(1) = \pi^2, \quad \mu_{1,2}(1) \approx 2.046\pi^2, \quad \mu_{2,2}(1) = 4\pi^2. \tag{5.4}$$

Clearly, $\mu_{i,j}(1) = \mu_{j,i}(1)$. Analogous to (2.5), for $i + j \geq 2$, (5.4) gives the uniform bounds

$$\pi^2 \alpha^p / 4 \leq \mu_{i,j}(\sigma) \leq 4\pi^2 \beta^p, \qquad \forall \sigma \in ad. \tag{5.5}$$

As in §2 we address the multiplicity of $\mu_{i,j}(\sigma)$ and the presence of positive eigenfunctions.

Lemma 5.1. *For $\sigma \in ad$,*

(a) If $2 \leq i + j < 4$, then $\mu_{i,j}(\sigma)$ is simple and there exists a corresponding positive eigenfunction.

(b) $\mu_{0,2}(\sigma) < \mu_{1,2}(\sigma)$ and $\mu_{1,1}(\sigma) < \mu_{1,2}(\sigma) < \mu_{2,2}(\sigma)$.

Proof: (a) Seiranian noted for these boundary conditions that (5.1) is equivalent, except for the presence of a simple zero eigenvalue when $ij = 2$, to a second order problem with *separated* boundary conditions. It now follows from the oscillation theory of Stürm, see, e.g., Atkinson [1], that each $\mu_{i,j}(\sigma)$ is simple and that for $ij \neq 2$, $\sigma^p u''$ is of one sign for

each $u \in \mathcal{E}_{i,j}(\sigma)$. In case ij equals 0 or 1 this yields respectively a positive convex or concave element of $\mathcal{E}_{i,j}(\sigma)$. When $ij = 2$ we find that $\sigma^p u''$ vanishes exactly once on $(0,1)$ for each $u \in \mathcal{E}_{i,j}(\sigma)$. Here we find an eigenfunction that is convex on $(0, x_0)$ and concave on $(x_0, 1)$ for some x_0. As this function must vanish at 0 and 1 we conclude that it must be positive on $(0,1)$.

(b) As $V_{i+1,j} \subset V_{i,j}$ we find $\mu_{i,j}(\sigma) \leq \mu_{i+1,j}$. Should equality hold, we conclude $\mathcal{E}_{i+1,j}(\sigma) \subset \mathcal{E}_{i,j}(\sigma)$. As in (2.4), for $u \in \mathcal{E}_{i,j}(\sigma)$ we deduce from (5.1) that

$$(\sigma^p u'')(x) = ((\sigma^p u'')'(0) + \mu_{i,j}(\sigma)u'(0))\, x + \sigma^p u''(0) + \mu_{i,j}(\sigma)u(0) - \mu_{i,j}(\sigma)u(x). \quad (5.6)$$

If $\mu_{0,2}(\sigma) = \mu_{1,2}(\sigma)$ then for each $u \in \mathcal{E}_{1,2}(\sigma) \subset \mathcal{E}_{0,2}(\sigma)$ equation (5.6), in view of (5.3), reads

$$(\sigma^p u'')(x) = -\mu_{i,j} u(x). \quad (5.7)$$

On recalling that $u(1) = u'(1) = 0$ we see that u satisfies a linear homogeneous equation with zero terminal data, and hence, $u \equiv 0$.

If $\mu_{1,1}(\sigma) = \mu_{1,2}(\sigma)$ then for each $u \in \mathcal{E}_{1,2}(\sigma) \subset \mathcal{E}_{1,1}(\sigma)$ (5.6), in view of (5.2), reads

$$(\sigma^p u'')(x) = ((\sigma^p u'')'(0) + \mu_{i,j}(\sigma)u'(0))\, x - \mu_{i,j}(\sigma)u(x). \quad (5.8)$$

So $(\sigma^p u'')(1) = (\sigma^p u'')'(0) + \mu_{i,j}(\sigma)u'(0)$. But $(\sigma^p u'')(1) = 0$ so (5.8) reduces to (5.7) and again the clamped conditions at 1 imply that $u \equiv 0$.

If $\mu_{1,2}(\sigma) = \mu_{2,2}(\sigma)$ then for each $u \in \mathcal{E}_{2,2}(\sigma) \subset \mathcal{E}_{1,2}(\sigma)$ equation (5.6), in view of (5.2), reads

$$(\sigma^p u'')(x) = (\sigma^p u'')'(0)x - \mu_{i,j}(\sigma)u(x).$$

Hence $(\sigma^p u'')(1) = (\sigma^p u'')'(0)$, from which we conclude that u is either identically zero or not of one sign. This excludes the positive element of $\mathcal{E}_{1,2}(\sigma)$ established in part (a). ∎

Thanks to the presence of positive first eigenfunctions, the existence theory of §3 applies directly to the problem of Lagrange

$$\sup_{\sigma \in ad} \mu_{i,j}(\sigma), \qquad 2 \leq i + j \leq 4. \quad (5.9)$$

We note that only for symmetric boundary conditions, i.e., $i = j$, should one expect an even optimal design. As $\mu_{i,j}(\sigma)$ is simple when $i + j < 4$ we deduce from Theorem 4.3 and conditions (4.5-7) that

$$\hat{\sigma}_{i,j}(x) = \alpha \Rightarrow \hat{\sigma}_{i,j}^{p-1}(x)|\hat{u}''(x)|^2 \leq \ell^2 \quad (5.10)$$

$$\alpha < \hat{\sigma}_{i,j}(x) < \beta \Rightarrow \hat{\sigma}_{i,j}^{p-1}(x)|\hat{u}''(x)|^2 = \ell^2 \quad (5.11)$$

$$\hat{\sigma}_{i,j}(x) = \beta \Rightarrow \hat{\sigma}_{i,j}^{p-1}(x)|\hat{u}''(x)|^2 \geq \ell^2, \quad (5.12)$$

for almost every $x \in (0,1)$, where $\hat{u} \in \mathcal{E}_{i,j}(\hat{\sigma}_{i,j})$. As before, $\hat{\sigma}_{i,j}$ is smooth where (5.11) holds.

The right side of (5.11) is the sole necessary condition offered by Keller [20] and Tadjbakhsh and Keller [35]. We now investigate the extent to which their claim is valid.

Recall that their analysis of the clamped–clamped column erred in neglecting (a) double eigenvalues, and (b) bounds on σ. As the previous lemma precludes the former phenomenon we need only consider the latter. The observation to be made is that (5.10) and (5.12) are only needed near the zeros of $\hat{u}''$. As noted above, members of $\mathcal{E}_{1,1}(\sigma)$ and $\mathcal{E}_{2,0}(\sigma)$ have second derivatives of one sign. As such, in these cases, (5.11) stands on its own (with the minor adjustment that σ be allowed to vanish at 0 and/or 1). In addition, as the related second order problems are fully equivalent, i.e., there are no spurious eigenvalues, Tadjbakhsh and Keller's sufficiency proof is correct. In summary, Keller [20] has the correct necessary condition for the hinged–hinged column, Tadjbakhsh and Keller [35] have the correct necessary condition for the clamped–free column, and the proof of sufficiency in [35] holds for both. We now recall their analytical solutions to these problems.

Keller, in [20], with $p = 2$ and $i = j = 1$ reconciled (5.11) and (5.1) and found

$$\begin{aligned} \hat{\sigma}_{1,1}(x) &= \tfrac{4}{3}\sin^2\theta(x), \quad 0 \le \theta \le \pi, \\ \theta(x) - \tfrac{1}{2}\sin 2\theta(x) &= \pi x, \quad 0 \le x \le 1. \end{aligned} \tag{5.13}$$

We have observed that this is a shortened cycloid with parametrization

$$\begin{aligned} x(t) &= \tfrac{3}{4\pi}\big(\tfrac{2}{3}(t - \sin t)\big) \\ y(t) &= \tfrac{2}{3}(1 - \cos t) \end{aligned} \qquad 0 \le t \le 2\pi.$$

This column buckles under an axial load of $4\pi^2/3$. In [35], Tadjbakhsh and Keller with $p = 2$ and $i = 2,\ j = 0$ reconciled (5.11) and (5.1) and found

$$\begin{aligned} \hat{\sigma}_{2,0}(x) &= \tfrac{4}{3}\sin^2\theta(x), \quad -\pi/2 \le \theta \le 0, \\ \theta(x) - \tfrac{1}{2}\sin 2\theta(x) + \pi/2 &= \pi x/2, \quad 0 \le x \le 1, \end{aligned} \tag{5.14}$$

our parametrization being,

$$\begin{aligned} x(t) &= \tfrac{3}{2\pi}\big(\tfrac{2}{3}(t - \sin t)\big) + 1 \\ y(t) &= \tfrac{2}{3}(1 - \cos t) \end{aligned} \qquad -\pi \le t \le 0.$$

This column buckles under an axial load of $\pi^2/3$. Having argued in favor of the existing solutions to the clamped–free and hinged–hinged problems we now turn to the clamped–hinged problem.

We saw in Lemma 5.1 that the second derivative of each function in $\mathcal{E}_{2,1}(\sigma)$ must change sign. The effect of this is that (5.11) forces $\hat{\sigma}_{2,1}$ to vanish at an interior point. In particular, when Tadjbakhsh and Keller reconciled (5.1) and (5.11) they found

$$\begin{aligned} \hat{\sigma}_{2,1}(x) &= \frac{4\sin^2\theta(x)}{3\sin^2\theta(0)}, \quad \theta(0) \le \theta \le \pi, \\ \theta(x) - \tfrac{1}{2}\sin 2\theta(x) + \tfrac{1}{2}\sin 2\theta(0) - \theta(0) &= x(\pi + \tfrac{1}{2}\sin 2\theta(0) - \theta(0)), \quad 0 \le x \le 1, \\ \tfrac{1}{2}\sin 2\theta(0) - \theta(0) &= -\tfrac{2}{3}\sin^3\theta(0)\cos^{-1}\theta(0) - \pi. \end{aligned} \tag{5.15}$$

Taking $a = \frac{1}{2}\sin 2\theta(0) - \theta(0)$ note that this $\hat{\sigma}_{2,1}(x)$ vanishes at $a/(\pi + a)$. Tadjbakhsh and Keller assert that the column built according to (5.15) will not buckle under loads less than approximately 27.22 in magnitude. We show in the appendix that this column can not withstand loads exceeding $\pi^2/3$ – and so, in fact, is much weaker than the uniform column. In addition, as $\mu_{2,1}(\sigma)$ corresponds to the *second* eigenvalue of its associated second order problem, the sufficiency proof of [35] is invalid. Hence, (5.15) is not an optimal design. In summary, (5.11) can not stand alone in the clamped–hinged case, $\sigma(x) \geq \alpha$ is is indeed an active constraint, so rendering (5.10) absolutely necessary. We suspect that there exists no solution to (5.9) when $ij = 2$ and $\alpha = 0$.

6. The Finite Dimensional Problem

We discretize the interval $[0, h, 2h, \ldots, (N-1)h = 1]$ and approximate $V_{i,j}$ by the finite dimensional space $V_{i,j}^h$, the subspace of $V_{i,j}$ whose elements, when restricted to $[kh, (k+1)h]$, are cubic polynomials (see Strang and Fix [34]). As each member of $V_{i,j}^h$ is completely determined by the value of it and its derivative at each of the N mesh points, we identify $V_{i,j}^h$ with $\mathbf{R}^{2N-i-j}$. We next approximate ad with the class of piecewise constant functions

$$ad^h \equiv \{\sigma \in \mathbf{R}^{N-1} : \alpha \leq \sigma_k \leq \beta, \sum_{k=1}^{N-1} \sigma_k = N-1\}.$$

We have refrained from labeling elements of ad^h by σ^h to avoid confusion with powers of σ. In this context, the infinite dimensional eigenvalue problem of (5.1) is now approximated by

$$B_h(\sigma)q_h = \mu K_h q_h, \quad \sigma \in ad^h, \ q_h \in \mathbf{R}^{2N-i-j}. \tag{6.1}$$

$B_h(\sigma)$ and K_h, the so–called bending and stiffness matrices, are each real, $(2N-i-j) \times (2N-i-j)$, symmetric, positive definite, and banded with half bandwidth of four. Our interest is, of course, in $\mu_{i,j}^h(\sigma)$, the least eigenvalue of (6.1). For, as $h \to 0$, one finds e.g., in [34], that $\mu_{i,j}^h(\sigma) \to \mu_{i,j}(\sigma)$. The connection between the finite and infinite dimensional problems now understood, we concentrate solely on (6.1). It should cause no confusion if, in our presentation of the finite dimensional optimization problem, we suppress most dependence on h, i, and j. With this, (6.1) becomes

$$B(\sigma)q = \lambda K q, \quad \sigma \in ad^h, \ q \in \mathbf{R}^n, \tag{6.2}$$

and we denote its least eigenvalue by $\lambda_1(\sigma)$. Our finite dimensional problem of Lagrange is now

$$\max_{\sigma \in ad^h} \lambda_1(\sigma). \tag{6.3}$$

The care that was taken in differentiating $\sigma \mapsto \lambda_1(\sigma)$ in §4 must also be exercised here. The occurrence of multiple eigenvalues is still possible. Clarke [10, Proposition 2.8.8] specifies the generalized gradient of the largest eigenvalue of a symmetric positive definite matrix in terms of a convex hull. Though such a characterization may suffice for an analytical description, as in §4, for computational purposes we have found it more useful to specify

first order conditions in terms of, less well known, "dual matrices". We state the result in general terms. We shall need $\mathcal{S}^n$, the class of $n \times n$ real symmetric matrices, and the Frobenius matrix inner product, $\langle A, B \rangle = \operatorname{tr} A^T B$.

Theorem 6.1. *Let $B : \mathbf{R}^{N-1} \to \mathcal{S}^n$ be (Fréchet) differentiable with $B_k(\sigma) = \partial B(\sigma)/\partial \sigma_k$ and let K be a fixed symmetric positive semidefinite matrix of the same order n. Assume $\sigma \in ad^h$ is such that $\lambda_1(\sigma)$ has multiplicity t, with corresponding eigenvectors given by the columns of a matrix $Q_1 \in \mathbf{R}^{n \times t}$, normalized so that $Q_1^T K Q_1 = I$. Then a necessary condition for σ to solve (6.3) is that there exist a symmetric positive semidefinite matrix U of order t, with trace equal to one, and Lagrange multipliers ν and γ_k, $k = 1, \ldots, N-1$, such that*

$$\langle U, Q_1^T B_k(\sigma) Q_1 \rangle = \nu + \gamma_k, \quad \text{and} \tag{6.4}$$

$$\sigma_k = \alpha \Rightarrow \gamma_k \leq 0 \tag{6.5}$$

$$\alpha < \sigma_k < \beta \Rightarrow \gamma_k = 0 \tag{6.6}$$

$$\sigma_k = \beta \Rightarrow \gamma_k \geq 0 \tag{6.7}$$

for each k. Furthermore, this condition is also sufficient for optimality in the case that $\sigma \mapsto B(\sigma)$ is affine.

Proof: In the following we use the notation $U \geq 0$ to mean that a symmetric matrix U is positive semidefinite. Regarding $\lambda_1 : \mathcal{S}^n \to \mathbf{R}$, we invoke Rayleigh's principle in

$$\begin{aligned} \lambda_1 &= \min \{ \langle q, Bq \rangle ;\ q \in \mathbf{R}^n,\ \langle q, Kq \rangle = 1 \} \\ &= \min \{ \langle qq^T, B \rangle ;\ q \in \mathbf{R}^n,\ \langle q, Kq \rangle = 1 \} \end{aligned}$$

Let $Q \in \mathbf{R}^{n \times n}$ be any matrix satisfying

$$Q^T K Q = I. \tag{6.8}$$

It is easily shown that

$$\operatorname{co} \{ qq^T ;\ q \in \mathbf{R}^n,\ \langle q, Kq \rangle = 1 \} = \{ Q \hat{U} Q^T ;\ \hat{U} \in \mathcal{S}^n,\ \operatorname{tr} \hat{U} = 1,\ \hat{U} \geq 0 \},$$

by using the spectral decomposition of $\hat{U}$, which by assumption has nonnegative eigenvalues adding to one, to obtain the requisite convex combination showing that the second set is contained in the first. It follows that

$$\lambda_1 = \min \{ \langle Q \hat{U} Q^T, B \rangle ;\ \hat{U} \in \mathcal{S}^n,\ \operatorname{tr} \hat{U} = 1,\ \hat{U} \geq 0 \}. \tag{6.9}$$

Now take Q to be a matrix whose columns are eigenvectors of (6.2), normalized so that (6.8) holds. The first t columns of Q are the columns of Q_1 and

$$Q^T B Q = \operatorname{Diag}(\lambda_i),$$

where $\lambda_1 \leq \lambda_2 \leq \ldots$ are the eigenvalues of (6.2), repeated according to multiplicity. Therefore, the matrices achieving the minimization in (6.9) are those defined by

$$\hat{U} = Q \begin{pmatrix} U & 0 \\ 0 & 0 \end{pmatrix} Q^T = Q_1 U Q_1^T$$

where $U \in \mathcal{S}^t$, with $\operatorname{tr} U = 1$ and $U \geq 0$. Consequently, the generalized gradient of $B \mapsto -\lambda_1(B)$ is the set of such matrices $\hat{U}$ (see Rockafellar [31, pages 29 and 35] or Clarke [10, §2.8]), no convex hull operation being required since the set of such $\hat{U}$ is convex.

With $\lambda_1(\sigma) = (\lambda_1 \circ B)(\sigma)$, the desired necessary conditions now follow from (i) the chain rule for generalized gradients [10, Theorem 2.3.10], (ii) the standard Lagrange multiplier rule [10, Theorem 6.1.1], and (iii) properties of the inner product. In particular,

$$\partial(-\lambda_1(\sigma)) = \{v \in \mathbf{R}^{N-1};\ v_k = \langle U, Q_1^T B_k(\sigma) Q_1 \rangle,\ U \in \mathcal{S}^t,\ U \geq 0,\ \operatorname{tr} U = 1\}. \tag{6.10}$$

These necessary conditions are also sufficient in the case that $\sigma \mapsto B(\sigma)$ is affine because the composition of a concave function with an affine function is concave. ∎

Our attention to $\partial(-\lambda_1)$ in (6.10) and $\partial\lambda_1^{-1}$ in Theorem 4.3 rather than simply $\partial\lambda_1$ is merely an artifact of Clarke's concern with functions defined as pointwise maxima rather minima. Here, it was convenient to characterize $-\lambda_1$ as the maximum of a Rayleigh quotient where, in §4, we found it more advantageous to maximize a functional of Auchmuty, and hence to consider λ_1^{-1}.

A different proof of the unconstrained version of this theorem was given (when $K = I$) by Overton [26], following work of Fletcher [14]. The $n \times n$ matrix $\hat{U}$ is known as a "dual matrix" by analogy with "dual variables" (Lagrange multipliers) familiar from mathematical programming. The $t \times t$ matrix U may be called a "reduced dual matrix", but since it is the one we shall need as a computational tool we shall also refer to it as simply the dual matrix. The distinction between $\hat{U}$ and U is analogous to the notational question of whether inactive constraints in a nonlinear program should be assigned zero Lagrange multipliers.

In the case that $t = 1$ and the bound constraints are inactive, the necessary condition reduces to the requirement that the gradient of $\lambda_1(\sigma)$, whose elements are $q_1^T B_k(\sigma) q_1$, has the constant value ν. (Here q_1 is the only column in Q_1, and U is the scalar one.) In the case that $t = 2$, let

$$U = \begin{pmatrix} \delta_1 & \delta_3/2 \\ \delta_3/2 & \delta_2 \end{pmatrix}, \tag{6.11}$$

let the two columns of Q_1 be q_1 and q_2 and again assume that all bound constraints are inactive. The necessary condition then becomes

$$\delta_1 q_1^T B_k(\sigma) q_1 + \delta_2 q_1^T B_k(\sigma) q_2 + \delta_3 q_1^T B_k(\sigma) q_2 = \nu$$

together with the trace and positive semidefinite constraints on U. Without loss of generality, δ_1 and δ_2 may be taken to have nonnegative sign, and the normalizing trace condition

may be replaced by the assumption that $\nu = 1$. The positive semidefinite constraint is then simply

$$\delta_1 \delta_2 \geq \delta_3^2/4.$$

This is the same necessary condition given by Bratus and Seiranian [6] and Masur [22]. We note that the derivation given here not only applies for $t > 2$, but is much simpler than that given by [6] and [22] for the case $t = 2$. In a footnote, Masur conjectured that the positive semidefinite condition on U would also be the correct necessary condition for $t > 2$. Bratus [7] gave a discussion of necessary and sufficient conditions for general multiplicity t, but the given necessary condition concerns the necessary sign of the directional derivative of λ_1 for all feasible directions; the positive semidefinite condition on U was apparently not obtained.

Before discussing the algorithm that springs from Theorem 6.1 we investigate the extent to which it suggests a new tack on the infinite dimensional problem. Regarding the variational principle of (6.9) we consider $K^+(X)$, the space of positive compact linear operators on a real separable Hilbert space X. Each $T \in K^+(X)$ possesses a countable sequence of eigenvalues $\lambda_1(T) \geq \lambda_2(T) \geq \cdots \downarrow 0$ repeated according to multiplicity and a (possibly infinite) trace $\operatorname{tr} T = \sum_{i=1}^{\infty} \lambda_i$. In this context it is not difficult to show for symmetric $T \in K^+(X)$ that

$$\lambda_1(T) = \max \{\operatorname{tr} TU;\ U \in K^+(X),\ \operatorname{tr} U = 1\}.$$

Recall that $u \mapsto (\sigma^p u'')''$ and $u \mapsto -u''$ are positive symmetric isomorphisms of H_0^2 onto H^{-2} and H_0^1 onto H^{-1} respectively. We denote these maps by A and B, remark that $B^{1/2}$ is a positive isomorphism of H_0^1 onto L^2, and denote by I the compact imbedding of H_0^2 in H_0^1. With $\phi = B^{1/2} J u$, and * denoting adjoint, the buckled column equation receives the formulation

$$(1/\lambda)\phi = B^{1/2} I A^{-1} (B^{1/2} I)^* \phi.$$

By construction, $B^{1/2} I A^{-1} (B^{1/2} I)^*$ is a symmetric member of $K^+(L^2)$. Although we may now proceed to compute $\partial \lambda_1^{-1}$ as in the previous theorem, this representation suffers from its dependence on the unknown A^{-1} and $B^{1/2}$ in contrast to Theorem 4.3 that works directly with A and B.

We now turn to the question of how to solve the finite dimensional optimization problem (6.3). Few papers in the literature have attempted to do this in any way that respects the nonsmoothness of the objective function. Polak and Wardi [28] describe an algorithm for maximizing the least eigenvalue, of a variety of important structures, that accounts for the presence of multiple eigenvalues. They focus however on the clamped vibrating column, $(\sigma^p u'')'' = \lambda u$, $u \in H_0^2$, a problem which has long been known to admit only simple eigenvalues (see e.g., Leighton and Nehari [21, Lemma 4.1]). Our algorithm differs from that of [28] in our attention to the added structure of the generalized gradient of λ_1 as revealed in the theorem above. We use an algorithm specifically designed to exploit this structure, which is based on Overton [26], but modified to be far more efficient for moderate to large mesh size N.

Given $\sigma \in ad^h$ with $\lambda_1(\sigma)$ and $\lambda_2(\sigma)$ the two least eigenvalues of (6.2) we normalize the corresponding eigenvectors q_1 and q_2 so that $Q_1 = [q_1\ q_2]$ satisfies $Q_1^T K Q_1 = I$, the 2×2

identity matrix. These eigenvalues and eigenvectors are computed by subspace iteration with a block size of two, with the requisite linear systems solved directly in terms of the Cholesky factors of $B(\sigma)$ (see Bathe and Wilson [5] for details).

Define the approximate multiplicity t of λ_1 by $t = 2$ if

$$\lambda_2(\sigma) - \lambda_1(\sigma) \leq \tau\lambda_1(\sigma)$$

and $t = 1$ otherwise. Here τ is a tolerance which may be adjusted during the optimization process. A multiplicity higher than two is excluded (for sufficiently small h) by Lemmas 2.1 and 5.1. Now consider the following linear program (LP):

$$\max_{d \in R^N} d_N \tag{LP0}$$

subject to

$$Ed = e \tag{LP1}$$

$$Fd \leq f \tag{LP2}$$

$$d_k = 0, \quad k \in J \tag{LP3}$$

$$\alpha - \sigma_k \leq d_k \leq \beta - \sigma_k, \quad k = 1, \ldots N-1 \tag{LP4}$$

$$|d_k| \leq \rho, \quad k = 1, \ldots, N \tag{LP5}$$

where
(i) The first $N-1$ components of d represent proposed changes to σ, while the last component approximates the corresponding change in $\lambda_1(\sigma)$. Let us write $d = [\eta^T \omega]^T$, with $\eta \in \mathbf{R}^{N-1}$, $\omega \in \mathbf{R}$;
(ii) ρ is a scalar, whose purpose is to ensure $\|d\|$ is not too large;
(iii) J is an index set, which effectively removes the corresponding variables from the linear program;
(iv) The first row of the matrix E is $[1, \ldots, 1, 0]$, and the first element of the right-hand side vector e is 0. This ensures that the changes to σ respect the integral constraint;
(v) The second row of E is

$$[-q_1^T B_1(\sigma) q_1, \ldots, -q_1^T B_{N-1}(\sigma) q_1, 1]$$

and the corresponding element of e is 0. If $t = 2$, then E contains an additional two rows,

$$[-q_2^T B_1(\sigma) q_2, \ldots, -q_2^T B_{N-1}(\sigma) q_2, 1] \quad \text{and} \quad [-q_1^T B_1(\sigma) q_2, \ldots, -q_1^T B_{N-1}(\sigma) q_2, 0]$$

with corresponding elements of e set to $\lambda_2(\sigma) - \lambda_1(\sigma)$ and 0 respectively;
(vi) If $t = 1$, F contains the single row

$$[-q_2^T B_1(\sigma) q_2, \ldots, -q_2^T B_{N-1}(\sigma) q_2, 1]$$

and f is the scalar $\lambda_2(\sigma) - \lambda_1(\sigma)$. If $t = 2$, F and f are empty, i.e. (LP2) may be removed.

We note that the given rows may be computed very efficiently, since the derivative matrices $B_k(\sigma)$ are extremely sparse.

The justification for (v) and (vi) is as follows. If $t = 1$, the second row of E imposes a linearization of the nonlinear equation $\lambda_1(\sigma + \eta) = \lambda_1(\sigma) + \omega$, while the only row in F imposes a linearization of the inequality $\lambda_2(\sigma+\eta) \geq \lambda_1(\sigma)+\omega$. Since ω is being maximized, the solution to the linear program yields the steepest ascent direction for $\lambda_1(\sigma)$, projected to satisfy the integral and bound constraints, with steplength required to be short enough that the linearized value for $\lambda_2(\sigma + \eta)$ does not drop below that for $\lambda_1(\sigma + \eta)$, and that the various bounds are satisfied. If $t = 2$, the second through fourth rows of E give a linearization of the appropriate set of three nonlinear equations imposing the coalescence of $\lambda_1(\sigma+\eta)$ and $\lambda_2(\sigma+\eta)$, see [16]. The common linearized value, $\lambda_1(\sigma)+\omega$, is maximized, subject to the given constraints.

Theorem 6.2. *Suppose that $\tau = 0$ so that the multiplicity estimate t is exact, and suppose that $\rho > 0$ and J is the empty set. Then $d = 0$ is a (nonunique) solution to the linear program given above if and only if (6.4) holds for some $U \in \mathcal{S}^t$ with $\operatorname{tr} U = 1$, some $\nu \in \mathbf{R}$ and some $\gamma \in \mathbf{R}^{N-1}$ satisfying (6.5-7).*

Proof: By the usual Lagrange multiplier rule, the linear program admits the solution $d = 0$ if and only if there exist multipliers $\nu \in R$, $\delta \in \mathbf{R}^{t(t+1)/2}$ and $\gamma \in \mathbf{R}^{N-1}$ satisfying

$$E^T \begin{pmatrix} \nu \\ \delta \end{pmatrix} + \begin{pmatrix} \gamma \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

with γ subject to the standard sign condition. Setting $U = \delta_1 (= 1)$ if $t = 1$ and defining U by (6.11) if $t = 2$, we have (6.4-7) with $\operatorname{tr} U = 1$. The same argument holds in the reverse direction. ∎

Note two points: there is no positive semidefinite condition obtained on U, and the solution $d = 0$ is generally not at a vertex of the feasible region, so is not unique.

Our algorithm for solving (6.3) generates a sequence in ad^h. Each successive approximation is obtained from the previous one by consideration of a linear program (LP) of the form given above. We first define a simple version of the algorithm, but one which is too costly for practical use. In this version, we obtain $\sigma + \eta$, a candidate replacement for σ, by solving the LP for $d = [\eta^T \omega]^T$. If $\lambda_1(\sigma + \eta) > \lambda_1(\sigma)$, σ is replaced by $\sigma + \eta$ and the process repeated. Otherwise, σ remains unchanged, the trust region radius ρ is decreased by a factor of two, and the revised LP is considered. This kind of trust region approach can be made very effective by modifying the size of ρ according to how well the actual increase in $\lambda_1(\sigma)$ agrees with the linear prediction ω, as discussed in Fletcher [14] in the context of general nonlinear programming. As recommended by Fletcher, we double ρ if the ratio of actual to predicted increase exceeds 0.75 and halve ρ if the ratio is less than 0.25. The process is terminated when $\|d\| \leq \epsilon$, a convergence tolerance.

However, the expense of obtaining the optimal solution of each linear program is not justified. Although the "limit" LP defined by σ equal to a solution of (6.2-4) has an optimal solution which is not a vertex, generically, any LP solved during the successive approximation process can be expected to have a unique solution which must be at a vertex.

Since only a few constraints involve all the variables, most of the constraints defining a vertex are simple bounds, and most of these may be trust radius bounds of the form (LP5). Since the only purpose of the trust radius bounds is to avoid taking steps too large for the linearizations to be accurate, there is little to be gained by finding the exact set of active bounds. We therefore partially solve the LP as follows. We first attempt to obtain a feasible point for the LP by setting d to the least norm solution of (LP1,LP3), contracting this step if necessary to satisfy the various inequalities. This contraction effectively scales the right-hand side of the only possible inhomogeneous equality constraint in the LP, that corresponding to the third row of E in the case $t = 2$. The rationale here is that if the least norm step satisfying the equality constraints is not feasible, the underlying approximations are probably not good enough to justify the solution of the unmodified LP. We then start the LP solution process as in a projected gradient method, augmenting d by projected gradient steps with steplengths determined by the inequality constraints and bounds. The gradient being projected is that of the LP objective, i.e. the vector $[0, \ldots, 0, 1]$, while the constraints determining the projection are the equality constraints and any inequality constraint and bounds encountered during the process. This continues until either (a) a trust radius bound of the form (LP5) is encountered, or (b) the norm of the current projected gradient increment drops below the tolerance ϵ (unlikely to happen first). At this point the process of partially solving the LP is terminated. Anywhere from zero to many active bounds of the form (LP4) may be encountered by this process, as well as, possibly, the inequality (LP2) (in the case $t = 1$). By adding any active bounds encountered to the set J, we avoid having to process these bounds again during the (partial) solution of subsequent LP's. However, the signs of the associated bound multipliers must be checked after the partial LP solution and bounds with the wrong sign removed from J if necessary. The entire process is very efficient, requiring QR factorizations of matrices with only two to four columns, with rows removed corresponding to active bounds. For complete details of the process, see [27].

In the case $t = 2$, the LP partial solution process generates four multipliers corresponding to the rows of E, namely ν, δ_1, δ_2, and δ_3. If the corresponding dual matrix U, defined by (6.11), is not positive semidefinite, this is a clear indication that the multiplicity estimate t is incorrect, and so τ is reduced by a factor of ten. In principle, it might be necessary to use a more sophisticated technique to recover from a multiplicity estimate which is too large. For example, if the algorithm was started at a point where all the optimality conditions except $U \geq 0$ were satisfied, it would be necessary to split the multiple eigenvalue to obtain an ascent direction; this is explained further in [26] and [27]. However, this technique has not been required in our computational experiments for the Lagrange problem.

This completes our outline of the algorithm used to generate the numerical results given in the next section. For more algorithmic details, see [27]. We do not have any proof that the given algorithm will converge to a solution of (6.3), but given any approximate solution we may verify the required signs of γ and the eigenvalues of the dual matrix U, and compute the residual of the approximate equation (6.4). We have found the algorithm to be very effective in practice as the numerical results attest.

7. Computational Results

The algorithm outlined in the previous section has been implemented in Fortran 77 and tested extensively. Subroutines from the Linpack [13] and Eispack [17] libraries were used to (a) perform the QR factorizations required during the partial LP solution process (for matrices with at most four columns), (b) obtain the Cholesky factorizations of $B(\sigma)$ needed for subspace iteration (these matrices have only seven nonzero diagonals) and (c) solve the reduced generalized eigenvalue problems required for subspace iteration, producing eigenvectors which are orthogonal with respect to K (these matrices have order two). Parameters were set as follows: τ, the relative multiplicity tolerance, was initialized to 0.1; ρ, the trust radius, was initialized to 5.0; ϵ, the convergence tolerance, was set to .001. The initial σ was set to the constant one, corresponding to the uniform column. Runs were made for various values of N, the number of mesh points; p, the power of σ in the differential equation; α and β, the lower and upper bounds on σ, and the various homogeneous boundary conditions: clamped–clamped, clamped–hinged, clamped–free and hinged–hinged.

The algorithm was found to be very efficient, typically invoking subspace iteration, in the computation of the two least eigenvalues of (6.2), about 50 times prior to reaching the convergence tolerance. At the final iterate the residual of the approximate equation (6.4) was typically found to have norm about 10^-3. There was usually no difficulty in determining the correct final multiplicity t, with corresponding positive semidefinite dual matrix U. In the cases where the final multiplicity t was two the gap between λ_1 and λ_2 was typically reduced to 10^{-6}. The subspace iteration was itself very efficient, requiring only one or two steps on all but the first few steps of the optimization, reflecting the good separation of λ_2 from λ_3 and the availability of an excellent initial two-dimensional subspace, namely the span of the eigenvectors q_1 and q_2 computed at the previous optimization step. (The first subspace iteration was initialized using the first two columns of the identity matrix.) The initial σ in each case was that of the uniform column, $\sigma \equiv 1$. Symmetry was not imposed on the algorithm's subsequent choices of σ. A typical run for $N = 513$ required 1.5 hours on a Sparc station.

We begin our summary of the results with $p = 2$. Under the assumption that each transverse cross–section of the column is circular we recall that σ is proportional to the square of the cross–section's radius. Plotting both $\pm\sqrt{\sigma}$ then gives a lengthwise cross-section of the associated column. With this representation one may then view the corresponding buckling mode(s) simultaneously. Our figures, generated by Matlab [23], portray the column in the piecewise fashion produced by the algorithm of §6 while using dashed curves(s) to indicate the buckling mode(s). We remark that for those optimal designs with double buckling loads the corresponding buckling modes depend on our initial choice of subspace, in subspace iteration, and σ.

Figure 1 gives our strongest clamped–clamped column and its first two buckling modes. Here $p = 2$, $\alpha = 0$, $\beta = 10$, with a double buckling load of 52.3533. This value agrees to four figures with that obtained by Olhoff and Rasmussen [25], Masur [22], and Seiranian [32].

On increasing α or decreasing β these bound(s) will eventually become active. Figure 2 gives our strongest clamped–clamped column and its first two buckling modes when

$p = 2$, $\alpha = 0.25$, and $\beta = 10$. The least eigenvalue is still double, though now reduced to 52.3467.

As the uniform column has a simple first eigenvalue one expects that a sufficient increase in α would produce an optimal design with a simple first eigenvalue. Figure 3 gives our best strongest clamped–clamped column and its first buckling mode when $p = 2$, $\alpha = 0.5$, and $\beta = 10$. In this case the first two eigenvalues are 51.07086 and 62.3479.

Clearly there must exist (at least) one value of α between 1/4 and 1/2 at which the optimal buckling load switches multiplicity. Olhoff and Rasmussen [25] declare 0.28 to be the only such value. Our algorithm also indicates the presence of such a critical α in the vicinity of 0.28. We note that in addition to being able to approach 0.28 from above – proceeding until the gap between the least two eigenvalues vanishes, we have also approached from below, in this case proceeding until the least eigenvalue of the corresponding dual matrix vanishes.

Figure 4 gives our strongest clamped-hinged column and its first buckling mode when $p = 2$, $\alpha = 0.25$, and $\beta = 10$. The buckling load of 27.0762 is, as expected, simple. Although decreasing α increases the buckling load, our designs converge, as $\alpha \to 0$, to the Tadjbakhsh and Keller solution, (5.13). As shown in the appendix, this column buckles at $\pi^2/3$, and so can not possibly be optimal. The convergence of our algorithm to (5.13) only strengthens our belief that the problem, as stated by Tadjbakhsh and Keller, is without a solution. That is, $\sigma \mapsto \mu_{2,1}(\sigma)$ with $p = 2$, does *not* attain its maximum on *ad* when $\alpha = 0$.

Our numerical results also agree with Tadjbakhsh and Keller in the cases for which we have argued that they are correct. In particular, Figure 5 gives our strongest clamped–free column and first buckling mode when $p = 2$, $\alpha = 0$, $\beta = 10$. The buckling load, again simple, is 3.2897. Figure 6 gives our strongest hinged–hinged column and first buckling mode when $p = 2$, $\alpha = 0$, $\beta = 10$. Its simple buckling load is 13.1579.

We return to the clamped–clamped case and consider its dependence on p. Our analysis of (4.5-7) led us to believe that, for $p > 1$, the minimum (maximum) of the optimal design increases (decreases) with p. This is reinforced by Figure 7, whose lower (upper) curve traces the minimum (maximum) of the optimal design as a function of p. As the buckling load is double for each of these designs there must exist a curve, between the lower one and the curve that is constantly one, across which the optimal buckling load changes multiplicity. We have seen that $(2, 0.28)$ lies near such a curve. With respect to the range of p considered in Figure 7 we have found that both the optimal buckling load and the least eigenvalue of its corresponding dual matrix increase with p. Regarding the behavior as p tends to 1 from above we have found that the minimum of the optimal design tends to zero, and, though the optimal buckling load remains double, the least eigenvalue of the corresponding dual matrix tends to zero. Below $p = 1$ we found optimal designs with simple buckling loads regardless of our choice of α. Figure 8 gives our strongest clamped–clamped column when $p = 1$, $\alpha = 0$, and $\beta = 10$. The buckling load of 47.9898 and the design itself are very close to the analytical result of (4.15). Refining the mesh in neighborhoods of 1/4 and 3/4, and perhaps using piecewise linear elements for σ, would presumably bring us even closer to (4.15). We have not pursued this for two reasons. First,

we argued in §4 that in the absence of an existence proof one can not fully trust (4.15), and second, in both of the physical contexts that $p = 1$ has arisen there is an *a priori* strictly positive lower bound on the admissible σ. Regarding the latter, we present in Figure 9 our strongest clamped–clamped column when $p = 1$, $\alpha = 0.8$, and $\beta = 1.2$. Its simple buckling load is 43.4921.

Having addressed dependence on p and α at a particular level of discretization we now fix $p = 2$, $\alpha = 0$, $\beta = 10$ and with clamped–clamped boundary conditions demonstrate the convergence of several relevant parameters as N, the number of mesh points, becomes large. In particular, Table 1 lists $\hat{\mu}_{2,2}$ (the optimal buckling load), the least eigenvalue of the associated dual matrix U, and $\|\hat{\sigma}_N - \hat{\sigma}_{1025}\|_\infty$ (the greatest difference between the optimal design on a mesh of N points and the optimal design on a mesh of 1025 points) for values of N from 65 to 1025.

N	$\hat{\mu}_{2,2}$	min ev(U)	$\|\hat{\sigma}_N - \hat{\sigma}_{1025}\|_\infty$
65	52.14944	0.023859	0.1066
129	52.31027	0.043317	0.0415
257	52.33615	0.047034	0.0424
513	52.35332	0.046435	0.0059
1025	52.35548	0.046607	0.0000

We close our study with a glance at the numerical range of the buckling load over *ad*. To this point we have concentrated on its maximization, and, though we may compare this value to that of the associated uniform column, it would be of interest to weigh it against the minimum buckling load. Clearly, α must now be strictly positive, for one could produce a buckling load of zero by prescribing that σ vanish on some interval. Regarding the existence of a minimizer for $\sigma \mapsto \mu_{i,j}(\sigma)$ over *ad* we note that Theorem 3.1 is insufficient. Recall in (3.3) that the limit of the maximizing sequence integrated to less than one. This was not an obstacle, for adding mass could only increase the buckling load. As our goal now is to minimize this load it appears that one must either relax the cost functional or construct the so–called G–closure of *ad* to obtain a minimizing design. Instead of embarking on this we modified the strongest column algorithm to minimize instead of maximize $\sigma \mapsto \mu_{i,j}(\sigma)$.

The modification to the algorithm is very simple, namely changing the sign of (LP0) and requiring a decrease instead of an increase in the smallest eigenvalue. The modified algorithm generated plausible weakest designs in *ad*, and, though we lack a proof of optimality, we shall content ourselves with a discussion of these numerical results. In all cases (independent of p and α) the minimum buckling load was simple; this is to be expected since the minimization should tend to separate the least eigenvalue from the remainder of the spectrum. In addition, we find that the generated designs have their mass concentrated near the inflection points of their associated positive buckling mode. This too is to be expected, making the opposite argument to that made in §2.

Figure 10 gives our weakest clamped–clamped column with $p = 2$, $\alpha = 0.25$, and $\beta = 10$. This column buckles under a simple axial load of 2.5658. The buckling load of our strongest column in this class (see Fig. 2) is 52.3467.

Figure 11 gives our weakest clamped–clamped column with $p = 1$, $\alpha = 0.8$, and $\beta = 1.2$. This column buckles under the simple axial load of 33.5631. The buckling load of our strongest column in this class (see Fig. 9) is 43.4921.

8. Concluding Remarks

Having thoroughly discussed the work of Tadjbakhsh and Keller [35] we now take up two additional issues, first broached by Keller [20], that, though very important, have received little attention in the literature.

The first involves the optimal design of cylindrical columns. Here, given again a fixed amount of material to be distributed over a column of fixed length, we seek the shape of the cross–section that when used to generate a cylinder produces a column with the greatest buckling load. One is not allowed to "taper" the column as we have in past sections. Keller quickly reduced this problem to the search for that planar domain of fixed area with the greatest least second moment of area. Recall that the second moment of area of the domain Ω with centroid at the origin in the direction η (with $|\eta| = 1$) is

$$I(\Omega, \eta) \equiv \int_{\Omega} |\eta^T y|^2 \, dy.$$

Denoting the unit circle by S, Keller's problem takes the form

$$\sup_{|\Omega|=A} \inf_{\eta \in S} I(\Omega, \eta). \tag{8.1}$$

Keller noted the existence of Ω for which this value is infinite. To exclude such Ω he restricted himself to convex domains. Within this smaller class he then argued, without proof, that the equilateral triangle possesses the greatest least second moment of area.

This reduction to convex domains is too severe. It is not the lack of convexity that allows (8.1) to grow without bound but the possibility that Ω itself may be unbounded, though of finite area. To exclude this behavior one may simply bound $|\partial\Omega|$, the length of Ω's boundary. That this does indeed bound $I(\Omega, \eta)$ follows from the isoperimetric inequality

$$2I_p(\Omega)/\pi \leq (|\partial\Omega|/2\pi)^4$$

of Pólya and Szegö [29, §1.5], where $I_p(\Omega)$ is the polar moment of inertia. As the second moment of area will be independent of η for the best Ω, its value will be one half that of its polar moment of inertia. We must now consider,

$$\sup_{\substack{|\Omega|=A \\ |\partial\Omega| \leq L}} \inf_{\eta \in S} I(\Omega, \eta). \tag{8.2}$$

For fixed A the value of (8.2) will grow as one increases L, suggesting that for sufficiently large L one may produce a domain whose second moment of area exceeds that of the equilateral triangle of the same area. We shall accomplish this with $A = 1$ and $L = 16$. To produce large values in (8.2) one need only consider domains that are symmetric about the coordinate axes as well as the two diagonals and that stretch out towards infinity. The

symmetry will render $\eta \mapsto I(\Omega, \eta)$ constant, while the latter condition will ensure us that $I(\Omega, \eta)$ is large. The domain of Fig. 12 has an area of one, a boundary whose length does not exceed 16, and a second moment of area of $16/45 + 71/8^3 90$. This value is more than three times greater than $1/6\sqrt{3}$, the second moment of area of the equilateral triangle of the same area. Though we have not solved (8.2), this example demonstrates that (8.2) produces, through the designer's choice of L, columns with arbitrarily large buckling load.

The other issue that we wish to resurrect from Keller [20] is consideration of the nonlinear problem underlying (5.1). When $i = j = 1$ this problem is

$$-(EI\theta')' = \lambda \sin\theta, \quad \theta'(0) = \theta'(1) = 0, \tag{8.3}$$

where θ measures the angle between the column and a fixed axis in its plane of buckling. Equation (5.1) arises from linearizing $\sin\theta$ to θ, identifying $u' = \theta$, and differentiating (8.3). Keller shows that the least eigenvalue of (5.1) (with $i = j = 1$) is indeed the least bifurcation point of (8.3). A number of questions now suggest themselves. Assuming the least eigenvalue of (5.1) for the remaining boundary conditions to be the first bifurcation point of the associated nonlinear problem could the optimal designs have been found without recourse to the linearized problem, i.e., can our tools be extended to deal directly with the likes of (8.3)? More generally, can one find optimal designs for the quasilinear problems that come with nonlinear constituitive laws – for recall that (8.3) is essentially Euler's *elastica*, a column that can suffer neither compression nor shear?

Appendix

We have argued throughout that the necessary and sufficient conditions proposed by Tadjbakhsh and Keller are incorrect. This does not in itself however, invalidate their designs. Indeed, we argued that their solutions to the hinged–hinged and clamped–free problems are correct. This appendix serves to demonstrate that these are their only correct designs.

In particular, we show that Tadjbakhsh and Keller incorrectly calculated the buckling loads of their proposed solutions to the clamped-clamped and clamped–hinged problems. Recall their solution of the former,

$$\begin{aligned} A(x) &= \tfrac{4}{3}\sin^2\theta(x), \quad -\pi/2 \le \theta \le 3\pi/2, \\ \theta(x) &- \tfrac{1}{2}\sin 2\theta(x) + \pi/2 = 2\pi x, \quad 0 \le x \le 1. \end{aligned} \tag{A.1}$$

To compute the corresponding buckling load we minimize the Rayleigh quotient over all those elements in the closure of C_0^2 with respect to the norm $u \mapsto \int_0^1 A^2|u''|^2\,dx$. Denoting this space by $\mathcal{H}$, Tadjbakhsh and Keller's claim amounts to

$$\frac{16\pi^2}{3} = \inf_{u\in\mathcal{H}} \frac{\int_0^1 A^2|u''|^2\,dx}{\int_0^1 |u'|^2\,dx}.$$

We now produce an element of $\mathcal{H}$ for which the Rayleigh quotient is $\pi^2/3$. As Olhoff and Rasmussen [25] observed in their numerical work, the column constructed according to

(A.1) tends to deform (under axial compression) on $(0,1/4)$ and $(3/4,1)$, with the center of the column experiencing only a rigid motion. To make this precise we recall from [35] that this A, when restricted to $(0,1/4)$, is actually the optimal design of the clamped–free problem there, with a corresponding buckling load of $\pi^2/3$ (as the volume of this piece is also $1/4$) and first eigenfunction $u_{2,0}$. We normalize $u_{2,0}$ so that

$$\int_0^{1/4} |u'_{2,0}|^2\,dx = 1,$$

and consider

$$\tilde{u}(x) = \begin{cases} u_{2,0}(x), & \text{if } 0 \le x \le 1/4 \\ u_{2,0}(1/4), & \text{if } 1/4 \le x \le 3/4 \\ u_{2,0}(1-x), & \text{if } 3/4 \le x \le 1. \end{cases}$$

Evaluating the Rayleigh quotient of $\tilde{u}$ we find

$$\frac{\int_0^1 A^2|\tilde{u}''|^2\,dx}{\int_0^1 |\tilde{u}'|^2\,dx} = \frac{2\int_0^{1/4} A^2|u''_{2,0}|^2\,dx}{2\int_0^{1/4} |u'_{2,0}|^2\,dx} = \pi^2/3.$$

It remains to show that $\tilde{u} \in \mathcal{H}$. As $\tilde{u}$ is well behaved away from $1/4$ and $3/4$, where it is not even C^1, we must rely on the vanishing of A there. From (A.1) we find that

$$\theta(x) = O(|x-1/4|^{1/3}) \quad \text{and} \quad A(x) = O(|x-1/4|^{2/3}). \tag{A.2}$$

It will suffice to smooth the corners of $\tilde{u}$ with a quintic. In particular, we use

$$p_n(x) = ax + bx^2 + \sum_{i=3}^{5} c_i(n)x^i, \qquad c_i(n) = O(n^{i-1})$$

where $a = u'_{2,0}(1/4)$, $b = \frac{1}{2}u''_{2,0}(1/4)$, and the $c_i(n)$ are determined by $p_n(0) = p_n(1/n) = p'_n(1/n) = p''_n(1/n) = 0$. Defining

$$\varphi_n(1-x) = \varphi_n(x) = \begin{cases} \tilde{u}(x), & \text{if } 0 \le x \le 1/4 \\ \tilde{u}(1/4) + p_n(x-1/4), & \text{if } 1/4 \le x \le 1/4 + 1/n \\ \tilde{u}(1/4), & \text{if } 1/4 + 1/n \le x \le 1/2 \end{cases}$$

produces a sequence in $\mathcal{H}$. With (A.2) we now find

$$\begin{aligned} \int_0^1 A^2(\tilde{u}'' - \varphi_n'')^2\,dx &= 2\int_{1/4}^{1/4+1/n} A^2|\varphi_n''|^2\,dx \\ &\le cn^8 \int_{1/4}^{1/4+1/n} (x-1/4)^{4/3}(x-1/4)^6\,dx \\ &\le cn^{-1/3}, \end{aligned}$$

i.e., the φ_n converge in $\mathcal{H}$ to $\tilde{u}$ as $n \to \infty$. To recap, we have produced an admissible displacement with a Rayleigh quotient of $\pi^2/3$. By Rayleigh's principle, the buckling load of the clamped–clamped column with cross sectional area (A.1) can not exceed $\pi^2/3$.

Now recall Tadjbakhsh and Keller's solution to the clamped–hinged problem,

$$\hat{\sigma}_{2,1}(x) = \frac{4\sin^2\theta(x)}{3\sin^2\theta(0)}, \quad \theta(0) \le \theta \le \pi, \tag{A.3}$$

$$\theta(x) - \tfrac{1}{2}\sin 2\theta(x) + \tfrac{1}{2}\sin 2\theta(0) - \theta(0) = x(\pi + \tfrac{1}{2}\sin 2\theta(0) - \theta(0)), \quad 0 \le x \le 1,$$

$$\tfrac{1}{2}\sin 2\theta(0) - \theta(0) = -\tfrac{2}{3}\sin^3\theta(0)\cos^{-1}\theta(0) - \pi,$$

and the fact that it vanishes at $x_0 = y/(\pi + y)$ where $y = \frac{1}{2}\sin 2\theta(0) - \theta(0)$. Analogous to the above, this design is optimal for the clamped–free column on $(0, x_0)$. The volume of this piece being x_0 as well, we find that it buckles at $\pi^2/3$. Denoting by $u_{2,0}$ the corresponding clamped–free eigenfunction on $(0, x_0)$ we define,

$$\tilde{u}(x) = \begin{cases} u_{2,0}(x), & \text{if } 0 \le x \le x_0 \\ \frac{u_{2,0}(x_0)}{1-x_0}(1-x), & \text{if } x_0 \le x \le 1. \end{cases}$$

As above, on substituting $\tilde{u}$ into the Rayleigh quotient with the $\hat{\sigma}_{2,1}$ of (A.3) one finds $\pi^2/3$. To verify that $\tilde{u}$ is indeed an admissible displacement we argue exactly as before. The deciding factor above was the fact that $A^2(x)$ vanished at $1/4$ to an order strictly greater than one. Here we find

$$\theta(x) = O(|x - x_0|^{1/3}) \quad \text{and} \quad \hat{\sigma}_{2,1}(x) = O(|x - x_0|^{2/3}),$$

so $\tilde{u}$ is indeed the limit of a sequence of admissible displacements.

Acknowledgements

This work began during the first author's visit to the Courant Institute. The efforts of Bob Kohn and Cathleen Morawetz are here greatly appreciated. We thank Andre Cherkaev for introducing us to the work of Bratus and Seiranian.

References

[1] F. V. Atkinson, *Discrete and Continuous Boundary Problems*, Academic Press, 1964.

[2] G. Auchmuty, *Dual variational principles for eigenvalue problems*, in Nonlinear Functional Analysis and its Applications, F. Browder, ed., AMS, 1986, pp. 55–71.

[3] V. Barbu and T. Precupanu, *Convexity and Optimization in Banach Spaces*, Reidel, Boston, 1986.

[4] D. Barnes, *The shape of the strongest column is arbitrarily close to the shape of the weakest column*, Q. Applied Math., XLVI(1988), pp. 605–609.

[5] K. Bathe and E. Wilson, *Numerical Methods in Finite Element Analysis*, Prentice–Hall, Englewood Cliffs, NJ, 1976.

[6] A. Bratus and A. Seiranian, *Bimodal solutions in eigenvalue optimization problems*, Prikladnaia Matematika i Mekhanika, 47(1983), pp. 546–554. Applied Math. Mech., 47(1983), pp. 451–457.

[7] A. Bratus, *Multiple eigenvalues in problems of optimizing the spectral properties of systems with a finite number of degrees of freedom*, Zh. vychisl. Mat. mat. Fiz., 26(1986), pp. 645–654. USSR Comp. Math. and Math. Phys., 26(1986), pp. 1–7.

[8] J. Cea and K. Malanowski, *An example of a max-min problem in partial differential equations*, SIAM J. Control, 8(1970), pp. 305–316.

[9] K. K. Choi and E. J. Haug, E.J., *Optimization of structures with repeated eigenvalues*, in Optimal Design of Distributed Parameter Structures, Haug, E.J. and Cea, J., eds., Sijthoff–Noordhoff, Leyden, 1981, pp. 219–277.

[10] F. Clarke, *Optimization and Nonsmooth Analysis*, Centre de recherches mathematiques, Montreal, 1989.

[11] S. J. Cox and J. R. McLaughlin, *Extremal eigenvalue problems for composite membranes, II*, Applied Math. & Opt., to appear.

[12] B. Dacorogna, *Weak Continuity and Weak Lower Semicontinuity of Non-Linear Functionals*, Springer, Lecture Notes in Mathematics 922, New York, 1982.

[13] J. J. Dongarra et al., *LINPACK Users Guide*, SIAM, Philadelphia, 1978.

[14] R. Fletcher, *Semidefinite constraints in optimization*, SIAM J. Control and Optimization, 23(1985), 493-513.

[15] R. Fletcher, *Practical Methods of Optimization*, 2nd ed., John Wiley, New York, 1987.

[16] S. Friedland, J. Nocedal, and M. L. Overton, *The formulation and analysis of numerical methods for inverse eigenvalue problems*, SIAM J. Numer. Anal., 24(1987), pp. 634–667.

[17] B. S. Garbow et al., *Matrix Eigensystem Routines: EISPACK Guide Extension*, Lecture Notes in Computer Science 51, Springer-Verlag, New York, 1977.

[18] E. J. Haug and B. Rousselet, *Design sensitivity analysis in structural mechanics. II. Eigenvalue variations*, J. Struct. Mech., 8(1980), pp. 161–186.

[19] E. Kamke, *Neue Herleitung der Oszillationssätze für die linearen selbstadjungierten Randwetaufgaben zweiter Ordnung*, Math. Zeit., 44(1938), pp. 635–658.

[20] J. Keller, *The shape of the strongest column*, Arch. Rat. Mech. Anal., 5(1960), pp.275–285.

[21] W. Leighton and Z. Nehari, *On the oscillation of solutions of self–adjoint linear differential equations of the fourth order*, Trans. of AMS, 89(1958), pp. 325–378.

[22] E. Masur, *Optimal structural design under multiple eigenvalue constraints*, Int. J. Solids Struct., 20(1984), pp. 211–231.

[23] C. Moler et al., *Pro–Matlab Users Guide*, MathWorks, Sherborn, MA, 1987.

[24] M. Myers and W. Spillers, *A note on the strongest fixed-fixed column*, Q. Appl. Math., XLIV(1986), pp. 583–588.

[25] N. Olhoff and S. Rasmussen, *On single and bimodal optimum buckling loads of clamped columns*, Int. J. Solids Struct., 13(1977), pp. 605–614.

[26] M. L. Overton, *On minimizing the maximum eigenvalue of a symmetric matrix*, SIAM J. Matrix Anal. Appl., 9(1988), pp. 256–268.

[27] M. L. Overton, *Large-Scale Optimization of Eigenvalues*, NYU Computer Science Department technical report, no. 505, 1990.

[28] E. Polak and Y. Wardi, *A nondifferentiable optimization algorithm for structural problems with eigenvalue inequality constraints*, J. of Struct. Mech. 11(1983), pp. 561–577.

[29] G. Pólya and G. Szegö, *Isoperimetric Inequalities in Mathematical Physics*, Annals of Math. Studies, no. 27, Princeton, 1951.

[30] A. Ramm, *Queries*, Notices AMS, 29(1982), pp. 326–329.

[31] R. T. Rockafellar, *The Theory of Subgradients and its Applications to Problems of Optimization: Convex and Nonconvex Functions*, Series on Research and Education in Mathematics 1, Heldermann Verlag, Berlin, 1981.

[32] A. Seiranian, *On a problem of Lagrange*, Inzhenernyi Zh., Mekhanika Tverdogo Tela, 19(1984), pp. 101–111. 19(1984). Mechanics of Solids, 19(1984), pp. 100–111.

[33] P. Senatorov, *The stability of the eigenvalues and eigenfunctions of a Sturm-Liouville problem*, Differentsial'nye Uravneniia, 7(1971), pp. 1667–1671. Differential Equations, 7(1971), pp. 1266–1269.

[34] G. Strang and G.J. Fix, *An Analysis of the Finite Element Method*, Prentice–Hall, Englewood Cliffs, NJ, 1973.

[35] I. Tadjbakhsh and J. Keller, *Strongest columns and isoperimetric inequalities for eigenvalues*, J. Appl. Mech., 29(1962), pp. 159–164.

[36] I. Todhunter and K. Pearson, *A History of the Theory of Elasticity and of the Strength of Materials, I*, Cambridge, 1886.