

# A CALIBRATION METHOD FOR ESTIMATING CRITICAL CAVITATION LOADS FROM BELOW IN 3D NONLINEAR ELASTICITY

JONATHAN J. BEVAN AND JONATHAN H.B. DEANE

ABSTRACT. In this paper we give an explicit sufficient condition for the affine map  $u_\lambda(x) := \lambda x$  to be the global energy minimizer of a general class of elastic stored-energy functionals  $I(u) = \int_\Omega W(\nabla u) dx$  in three space dimensions, where  $W$  is a polyconvex function of  $3 \times 3$  matrices. The function space setting is such that cavitating (i.e., discontinuous) deformations are admissible. In the language of the calculus of variations, the condition ensures the quasiconvexity of  $I(\cdot)$  at  $\lambda \mathbf{1}$ , where  $\mathbf{1}$  is the  $3 \times 3$  identity matrix. Our approach relies on arguments involving null Lagrangians (in this case, affine combinations of the minors of  $3 \times 3$  matrices), on the previous work [4], and on a careful numerical treatment to make the calculation of certain constants tractable. We also derive a new condition, which seems to depend heavily on the smallest singular value  $\lambda_1(\nabla u)$  of a competing deformation  $u$ , that is necessary for the inequality  $I(u) < I(u_\lambda)$ , and which, in particular, does not exclude the possibility of cavitation.

## 1. INTRODUCTION

In this paper we consider an established model of elastic material that is capable of describing cavitation, that is, of admitting energy minimizers that are discontinuous. This phenomenon was first analysed in the setting of hyperelasticity by Ball in [2]; since then, a large and sophisticated literature has developed, including but not limited to [15, 12, 9, 13, 14, 7, 8], part of which focuses on finding boundary conditions which, when obeyed by all competing deformations, ensure that cavitation does *not* occur. It is to the latter body of work that we contribute by considering the case of purely bulk energy

$$I(u) = \int_\Omega W(\nabla u(x)) dx,$$

where  $u : \Omega \rightarrow \mathbb{R}^3$  represents a deformation of an elastic material occupying the domain  $\Omega$  in a reference configuration, and where  $W$  is a suitable stored-energy function. In the three dimensional setting, we give an explicit characterization of those affine boundary conditions of the form

$$u_\lambda(x) = \lambda x,$$

where  $\lambda > 0$  is a parameter, such that the quasiconvexity inequality

$$(1.1) \quad I(u) \geq I(u_\lambda)$$

holds among all suitable maps  $u$  agreeing with  $u_\lambda$  on  $\partial\Omega$ . It is by now well established that if  $\lambda$  is large enough,  $\lambda \geq \lambda_{\text{crit}}$ , say, then such an inequality cannot hold. Thus we probe  $\lambda_{\text{crit}}$  by

---

(J.J. Bevan) DEPARTMENT OF MATHEMATICS, UNIVERSITY OF SURREY, GUILDFORD, GU2 7XH, UNITED KINGDOM. (CORRESPONDING AUTHOR: t: +44(0)1483 682620.)

(J.H.B. Deane) DEPARTMENT OF MATHEMATICS, UNIVERSITY OF SURREY, GUILDFORD, GU2 7XH, UNITED KINGDOM.

*E-mail addresses:* j.bevan@surrey.ac.uk, j.deane@surrey.ac.uk.

finding  $\lambda_0$  such that (1.1) holds whenever  $\lambda \leq \lambda_0$ . This question has been addressed in [10] and, more recently, in [4]. In this paper we use a new approach, involving the addition of a suitable null Lagrangian (a method sometimes known as calibration), to deduce concrete lower bounds on  $\lambda_{\text{crit}}$  in the three dimensional case.

The analysis centres ostensibly on functions of the singular values of  $3 \times 3$  matrices. Let  $A$  be a  $3 \times 3$  matrix. Then the singular values of  $A$  are normally written as  $\lambda_j(A)$ , for  $j = 1, 2, 3$ , and their squares are the eigenvalues of  $A^T A$ . See [6, Chapter 13] or [5, Section 3.2] for useful introductions to singular values, as well as [1, 2, 12] for an illustration of their use in nonlinear elasticity. Singular values arise naturally in the stored-energy functions of isotropic elastic materials, and also in lower bounds which can be derived from them. Such was the case in [4], where, for  $2 < q < 3$  and for convex functions  $Z$  and  $h$ , a stored-energy function very similar to<sup>1</sup>

$$W(A) = |A|^q + Z(\text{cof } A) + h(\det A)$$

was shown to obey the inequality

$$(1.2) \quad I(u) - I(u_\lambda) \geq \int_{\Omega} \kappa |\nabla u - \nabla u_\lambda|^q + h'(\lambda^3) \Pi_{j=1}^3 (\lambda_i(\nabla u) - \lambda) + \lambda P(\nabla u) \, dx.$$

The function  $P$  is defined by

$$(1.3) \quad P(A) = \sum_{1 \leq i < j \leq 3} \lambda_i(A) \lambda_j(A) - \lambda \sum_{1 \leq i \leq 3} \lambda_i(A)$$

and the constant  $\kappa$  satisfies bounds defined in (3.3) below. By grouping the first two integrands in (1.2) together, it is possible to find conditions on  $\lambda$  such that  $\int_{\Omega} \kappa |\nabla u - \nabla u_\lambda|^q + h'(\lambda^3) \Pi_{j=1}^3 (\lambda_i(\nabla u) - \lambda) \, dx \geq 0$ . However, the corresponding inequality for  $P$ , namely

$$(1.4) \quad \int_{\Omega} P(\nabla u) \, dx \geq 0,$$

which, since  $P(\lambda \mathbf{1}) = 0$ , is equivalent to the quasiconvexity of  $P$  at  $\lambda \mathbf{1}$ , remains an open question. We show in this paper that  $P$  does satisfy a condition necessary for quasiconvexity at  $\lambda \mathbf{1}$  (see Proposition 2.3, part (a): rank-one convexity at  $\lambda \mathbf{1}$ ), but that the most tractable sufficient condition for (1.4) cannot hold (see Proposition 2.3, part (b): polyconvexity at  $\lambda \mathbf{1}$ ). Trying instead to find conditions under which each of

$$(1.5) \quad \begin{aligned} & \int_{\Omega} (\kappa/2) |\nabla u - \nabla u_\lambda|^q + h'(\lambda^3) \Pi_{j=1}^3 (\lambda_i(\nabla u) - \lambda) \, dx \geq 0 \\ & \int_{\Omega} (\kappa/2) |\nabla u - \nabla u_\lambda|^q + h'(\lambda^3) P(\nabla u) \, dx \geq 0 \end{aligned}$$

holds is closer to the right approach, although, for reasons connected with the curvature of  $P$  at  $\lambda \mathbf{1}$ , it does not seem to be possible to prove (1.5) without first modifying the integrand. This is what leads us to introduce the null Lagrangian

$$N(A) = \text{tr cof } A - \lambda \text{tr } A,$$

which has the property that  $\int_{\Omega} N(\nabla u) \, dx = 0$  for any admissible  $u$  and is such that there are conditions on  $\lambda$  under which

$$\int_{\Omega} (\kappa/2) |\nabla u - \nabla u_\lambda|^q + h'(\lambda^3) (P(\nabla u) - N(\nabla u)) \, dx \geq 0$$

---

<sup>1</sup>The original functional contained an ‘artificial’ quadratic term  $r|A|^2$ , with  $r$  large, to deal with the difficulties presented by the function  $P$ . This is no longer needed thanks to the calibration method we introduce.

for all admissible  $u$ . See Theorem 3.4 and (3.15) in particular. In fact,  $N$  is the unique null Lagrangian for which this method works: see Proposition 2.2. More generally, we remark that  $P$  and  $G := P - N$  possess properties that are both interesting in their own right and, at the same time, highly non-trivial to derive. (See Section 2.) A useful introduction to null Lagrangians can be found in [3].

The upper bound  $\lambda_0$  given in the right-hand side of (3.15) is investigated in Section 3 using a careful mixture of analysis and numerical techniques. The partnership between these approaches seems to be particularly fruitful when applied to  $G$  and to functions derived from it. Accordingly, we find an explicit constant  $\nu_1 \approx 0.4501$  such that if

$$0 \leq \lambda^{3-q} h'(\lambda^3) \leq \frac{\kappa}{2} (\sqrt{2})^{q-3} \nu_1^{2-q},$$

then  $I(u) \geq I(u_\lambda)$ . See Section 3, Subsection 3.1 and the appendices for details.

In Section 4, a careful analysis of the function

$$H(A) := \prod_{j=1}^3 (\lambda_j(\nabla u) - \lambda) + \lambda G(A)$$

yields, among other things, what we believe to be new necessary conditions for the inequality  $I(u) \leq I(u_\lambda)$ . A distinguished role seems to be played by the smallest singular value,  $\lambda_1(\nabla u)$ : see Proposition 4.6 in particular.

**1.1. Notation.** The inner product between two matrices  $A$  and  $B$  is given by  $A \cdot B = \text{tr } A^T B$ , and, as usual,  $\text{tr } A$  denotes the trace of  $A$ . For a function  $f : \mathbb{R}^{3 \times 3} \rightarrow \mathbb{R}$  and any  $3 \times 3$  matrix  $U$ , the shorthand

$$\begin{aligned} D_U f(A) &= \nabla f(A) \cdot U \\ D_U^2 f(A) &= \nabla^2 f(A)[U, U] \end{aligned}$$

will be used, where as usual  $\nabla^2 f(A)[U, U] = f_{,(ij)(kl)}(A) U_{ij} U_{kl}$  with the summation convention in force. When discussing polyconvexity, which is defined when it next features in the paper, we use the shorthand notation  $\mathbb{R}^{19}$  for the set  $\mathbb{R}^{3 \times 3} \times \mathbb{R}^{3 \times 3} \times \mathbb{R}$  containing the list of minors  $R(A) := (A, \text{cof } A, \det A)$  of any  $3 \times 3$  matrix. The set of  $3 \times 3$  square, orthogonal matrices is denoted by  $O(3)$ , and the subset of  $O(3)$  consisting of those matrices with determinant equal to 1 will be written  $SO(3)$ . For any two vectors  $a$  and  $n$  in  $\mathbb{R}^3$ , the notation  $a \otimes n$  will denote the matrix of rank one whose  $(i, j)$  entry is  $a_i n_j$ . Our notation for Sobolev spaces is standard.

## 2. CALIBRATION AND THE FUNCTION $G(A)$

In this section we give some properties of the function  $P$  and use them to derive the null Lagrangian  $N$  alluded to above. To start with, two technical results are required.

**Lemma 2.1.** *Let  $\lambda > 0$  and let  $U$  belong to  $\mathbb{R}^{3 \times 3}$ . Then the function  $P : \mathbb{R}^{3 \times 3} \rightarrow \mathbb{R}$  given by  $P(A) = \sum_{1 \leq i < j \leq 3} \lambda_i(A) \lambda_j(A) - \lambda \sum_{1 \leq i \leq 3} \lambda_i(A)$  admits the expansion*

$$(2.1) \quad P(\lambda \mathbf{1} + hU) = \lambda h \text{tr } U + h^2 \text{tr cof } U + o(h^2)$$

*as  $h \rightarrow 0$ . In particular, the first and second derivatives of  $P$  at  $\lambda \mathbf{1}$  exist and they satisfy  $D_U P(\lambda \mathbf{1}) = \lambda \text{tr } U$  and  $D_U^2 P(\lambda \mathbf{1}) = 2 \text{tr cof } U$  respectively.*

*Proof.* Let  $A = \lambda \mathbf{1} + hU$  and  $V_h = 2\lambda U_s + hU^T U$ , where  $U_s := (U + U^T)/2$  is the symmetric part of  $U$  and  $h$  is real. Thus  $A^T A = \lambda^2 \mathbf{1} + hV_h$ . By definition, each  $z = \lambda_i(A)$  is a root of the equation  $\det((\lambda^2 - z^2)\mathbf{1} + hV_h) = 0$ , i.e.  $z$  satisfies

$$0 = (\lambda^2 - z^2)^3 + h(\lambda^2 - z^2)^2 \operatorname{tr} V_h + h^2(\lambda^2 - z^2) \operatorname{tr} \operatorname{cof} V_h + h^3 \det V_h.$$

Let  $h \neq 0$ , define  $R(z, h) = \frac{\lambda^2 - z^2}{h}$  and note that

$$(2.2) \quad 0 = R(z, h)^3 + R(z, h)^2 \operatorname{tr} V_h + R(z, h) \operatorname{tr} \operatorname{cof} V_h + \det V_h.$$

Since the coefficients  $\operatorname{tr} V_h$ ,  $\operatorname{tr} \operatorname{cof} V_h$  and  $\det V_h$  vary continuously with  $h$ , it follows by standard arguments that the roots  $R_1(h)$ ,  $R_2(h)$  and  $R_3(h)$ , say, of (2.2), which are all real by construction, but not necessarily distinct, can be chosen such that they also vary continuously with  $h$  and satisfy  $R_j(h) \rightarrow R_j$  as  $h \rightarrow 0$ , where the  $R_j$  are roots of the polynomial equation

$$(2.3) \quad 0 = R^3 + R^2 \operatorname{tr} V_0 + R \operatorname{tr} \operatorname{cof} V_0 + \det V_0.$$

The foregoing argument implies that the two sets of roots  $\{R(\lambda_1(A), h), R(\lambda_2(A), h), R(\lambda_3(A), h)\}$  and  $\{R_1(h), R_2(h), R_3(h)\}$  must be equal. Hence for each  $i = 1, 2, 3$  and each  $h \neq 0$  there is  $j(i, h) \in \{1, 2, 3\}$  such that  $R(\lambda_i(A), h) = R_{j(i, h)}(h)$  and such that  $j(\cdot, h)$  is a permutation of  $\{1, 2, 3\}$ . For brevity, let us write  $j(i)$  in place of  $j(i, h)$ . We now have

$$(2.4) \quad \lambda^2 - \lambda_i^2(A) = h R_{j(i)}(h)$$

and, from (2.2),

$$(2.5) \quad \sum_{a=1}^3 R_a(h) = -\operatorname{tr} V_h$$

and

$$(2.6) \quad \sum_{1 \leq a < b \leq 3} R_a(h) R_b(h) = \operatorname{tr} \operatorname{cof} V_h.$$

Rewriting (2.4) gives

$$\lambda_i(A) = \lambda - \frac{h}{2\lambda} R_{j(i)}(h) - \frac{h^2}{8\lambda^3} (R_{j(i)}(h))^2 + o(h^2)$$

as  $h \rightarrow 0$  for  $i = 1, 2, 3$ . Substituting this into the expression for  $P(A)$  above implies that

$$P(A) = -\frac{h}{2} \sum_{a=1}^3 R_a(h) + \frac{h^2}{8\lambda^2} \left( 2 \sum_{1 \leq a < b \leq 3} R_a(h) R_b(h) - \sum_{a=1}^3 (R_a(h))^2 \right) + o(h^2).$$

Here the fact that  $\sum_{1 \leq i < k \leq 3} R_{j(i)}(h) + R_{j(k)}(h) = 2 \sum_{a=1}^3 R_a(h)$ , which is easily deduced from the fact that  $j(\cdot)$  is a bijection, proves useful. Next, the identity

$$\sum_{a=1}^3 R_a^2 = \left( \sum_{a=1}^3 R_a \right)^2 - 2 \sum_{1 \leq a < b \leq 3} R_a R_b$$

together with (2.5) and (2.6) yields

$$\begin{aligned} P(A) &= h\lambda \operatorname{tr} U + \frac{h^2}{2}|U|^2 + \frac{h^2}{8\lambda^2}(4\operatorname{tr} \operatorname{cof} V_h - (\operatorname{tr} V_h)^2) + o(h^2) \\ &= h\lambda \operatorname{tr} U + \frac{h^2}{2}|U|^2 + \frac{h^2}{2}(4\operatorname{tr} \operatorname{cof} U_s - (\operatorname{tr} U)^2) + o(h^2) \\ &= h\lambda \operatorname{tr} U + \frac{h^2}{2}((\operatorname{tr} U)^2 - \operatorname{tr}(U^2)) + o(h^2). \end{aligned}$$

In passing from the second to the final line above we have used the identity

$$\operatorname{tr} \operatorname{cof} U_s = \frac{(\operatorname{tr} U)^2}{2} - \frac{|U|^2}{4} - \frac{\operatorname{tr} U^2}{4},$$

and in the final line we recognise the coefficient of  $h^2$  as  $\operatorname{tr} \operatorname{cof} U$ , concluding the proof of (2.1). The final assertion of the lemma is now straightforward.  $\square$

We note that the precaution of writing  $R(\lambda_i(A), h) = R_{j(i,h)}(h)$ , where  $j(\cdot, h)$  is a suitable permutation of  $\{1, 2, 3\}$ , is not redundant: the quantities  $R(\lambda_i(A), h)$  really can ‘switch values’ within the set  $\{R_1(h), R_2(h), R_3(h)\}$ , as the following simple example shows. Suppose  $\lambda > 0$  and let  $U = e_1 \otimes e_1$ . Then, for  $h$  such that  $|h| < \lambda$ , the singular values of  $A := \lambda \mathbf{1} + hU$  are  $\lambda$ ,  $\lambda$  and  $\lambda + h$ , while the solutions  $R_i(h)$  to (2.2), which in this case is the equation  $R^2(R + 2\lambda + h) = 0$ , can be fixed as  $R_1(h) = 0$ ,  $R_2(h) = 0$  and  $R_3(h) = -2\lambda - h$ . Now, for  $h > 0$ , the ordered singular values of  $A$  are  $\lambda_1(A) = \lambda_2(A) = \lambda$  and  $\lambda_3(A) = \lambda + h$ , while for  $h < 0$ ,  $\lambda_1(A)$  and  $\lambda_3(A)$  exchange places, i.e.  $\lambda_1(A) = \lambda + h$ ,  $\lambda_2(A) = \lambda_3(A) = \lambda$ . Therefore the map  $j(\cdot, h)$  corresponds to the identity permutation on  $\{1, 2, 3\}$  when  $h > 0$ , whereas it corresponds to the transposition (13), in cycle notation, when  $h < 0$ . The switching comes about because we insist on working with ordered singular values.

For later use, we also note that the proof of Lemma 2.1 above shows that

$$(2.7) \quad P(\lambda + hU) = \lambda h \operatorname{tr} U + h^2 \operatorname{tr} \operatorname{cof} U + h^3 K(R_1, R_2, R_3) + o(h^3).$$

The term  $K$  is a symmetric polynomial in  $R_1$ ,  $R_2$  and  $R_3$ , and can therefore be expressed in terms of the coefficients of the polynomial (2.3). To see (2.7), expand (2.4) to third order in  $h$  and substitute into the expression for  $P(A)$ . The precise form of  $K$  does not matter, so we omit the proof.

**Proposition 2.2.** *Let  $C_1$  and  $C_2$  be fixed  $3 \times 3$  matrices, let  $C_3$  be a real number, and let*

$$N(A) = C_1 \cdot (A - \lambda \mathbf{1}) + C_2 \cdot (\operatorname{cof} A - \lambda^2 \mathbf{1}) + C_3 (\det A - \lambda^3)$$

for all  $A \in \mathbb{R}^{3 \times 3}$ . Then

$$(2.8) \quad D_U P(\lambda \mathbf{1}) = D_U N(\lambda \mathbf{1}) \quad \forall U \in \mathbb{R}^{3 \times 3}$$

if and only if  $C_1, C_2$  and  $C_3$  are related by the equation

$$(2.9) \quad (\lambda - \lambda^2 C_3 - \lambda \operatorname{tr} C_2) \mathbf{1} = C_1 - \lambda C_2^T.$$

Moreover,

$$(2.10) \quad D_U^2 P(\lambda \mathbf{1}) = D_U^2 N(\lambda \mathbf{1}) \quad \forall U \in \mathbb{R}^{3 \times 3}$$

if and only if  $C_2$  and  $C_3$  are related by the equation

$$(2.11) \quad C_2 = (1 - \lambda C_3) \mathbf{1}.$$

In particular, the unique quadratic null Lagrangian  $N$  satisfying both (2.9) and (2.11) is

$$(2.12) \quad N(A) = \operatorname{tr} \operatorname{cof} A - \lambda \operatorname{tr} A,$$

and it satisfies

$$(2.13) \quad \int_{\Omega} N(\nabla u) \, dx = 0$$

for all  $u$  belonging to  $W^{1,2}(\Omega, \mathbb{R}^3)$  such that  $u = u_{\lambda}$  on  $\partial\Omega$  (in the sense of trace).

*Proof.* Let  $A = \lambda \mathbf{1} + hU$  and note that

$$N(A) = [U \cdot C_1 + \lambda \operatorname{tr} U \operatorname{tr} C_2 - \lambda \operatorname{tr}(UC_2) + \lambda^2 C_3 \operatorname{tr} U] h + [C_2 \cdot \operatorname{cof} U + \lambda C_3 \operatorname{tr} \operatorname{cof} U] h^2 + C_3 h^3 \det U.$$

Comparing terms of order  $h$  in this expression with the expansion for  $P(A)$  given in (2.1), we see that  $D_U P(\lambda \mathbf{1}) = D_U N(\lambda \mathbf{1})$  for all  $U$  if and only if (2.9) holds. To prove the equivalence of (2.10) and (2.11), simply compare terms of order  $h^2$  to obtain

$$(2.14) \quad (C_2 + \lambda C_3 \mathbf{1}) \cdot \operatorname{cof} U = \operatorname{tr} \operatorname{cof} U$$

for all  $U$ , and then pick  $U$  such that  $\operatorname{cof} U = e_i \otimes e_j$ . It is then clear that (2.14) is equivalent to (2.11).

Finally, to prove that  $N(A) = \operatorname{tr} \operatorname{cof} A - \lambda \operatorname{tr} A$  is the unique, quadratic null Lagrangian satisfying (2.9) and (2.11) take  $C_3 = 0$  in (2.11) and (2.9). The former gives  $C_2 = \mathbf{1}$ , and the latter  $C_1 = -\lambda \mathbf{1}$ , which together imply (2.12). Equation (2.13) is a standard result about null Lagrangians; to see it without recourse to general theory, simply observe that, for sufficiently smooth  $\varphi$ ,  $N(\nabla \varphi)$  can be written as a divergence. The result then follows from the Green's theorem and an approximation argument. (The argument given in [6, Lemma 5.5 (ii)] serves as a useful template.)  $\square$

We remark that this establishes a simple pattern:  $N(A)$  can apparently be obtained from  $P$  by noting that if  $A = \operatorname{Diag}(\lambda_1, \lambda_2, \lambda_3)$  then  $P(A) = \operatorname{tr} \operatorname{cof} A - \lambda \operatorname{tr} A$ .

As was pointed out in the introduction, and originally conjectured in [4], it would be very useful if  $P$  were quasiconvex at the matrix  $\lambda \mathbf{1}$ . Our results in this direction are somewhat mixed. We find that  $P$  satisfies a condition necessary for quasiconvexity at  $\lambda \mathbf{1}$ , but that it does not satisfy a tractable, sufficient condition for quasiconvexity at  $\lambda \mathbf{1}$ . To be precise, (a)  $P$  is rank-one convex at  $\lambda \mathbf{1}$  but (b)  $P$  is not polyconvex at that point. These concepts are explained in more detail below. We note, incidentally, that  $P$  is not globally rank-one convex. The latter is relatively easy to see: one can immediately calculate that, for any rank-one matrix  $A = a \otimes n$ ,  $\lambda_1(A) = \lambda_2(A) = 0$  and  $\lambda_3(A) = |A|$ . In particular,  $P(A) = -\lambda |A|$ , which is a concave function of  $A$ . The foregoing discussion is summarised in the result below.

**Proposition 2.3.** *Let the function  $P$  be defined by (1.3). Then*

- (a)  $\lambda \mathbf{1}$  is a point of rank-one convexity of  $P$ , but
- (b)  $P$  is not polyconvex at  $\lambda \mathbf{1}$ .

*Proof. (a):* To show (a) we only need to verify that  $P(\lambda \mathbf{1} + a \otimes n)$  is convex as a function of the rank-one matrix  $a \otimes n$ . Let  $A = \lambda \mathbf{1} + a \otimes n$  and assume without loss of generality that  $n$  is a unit vector. The eigenvalues of  $A^T A$  do not depend on the particular basis with respect to which  $A$  is represented, so suppose to start with that the vector product  $a \wedge n \neq 0$  and let  $\{n, e_2, \frac{a \wedge n}{|a \wedge n|}\}$

be an orthonormal basis of  $\mathbb{R}^3$  in which  $a = a_1 n + a_2 e_2$  holds for some  $a_1$  and  $a_2$ . Then

$$A^T A = \begin{pmatrix} \lambda^2 + 2\lambda a_1 + |a|^2 & \lambda a_2 & 0 \\ \lambda a_2 & \lambda^2 & 0 \\ 0 & 0 & \lambda^2 \end{pmatrix}$$

with eigenvalues  $\lambda^2$  and  $\sigma_{\pm}^2$ , where

$$\sigma_{\pm}^2 = \lambda^2 + \frac{|a|^2}{2} + \lambda a_1 \pm |a| \sqrt{\frac{|a|^2}{4} + \lambda a_1 + \lambda^2}.$$

We can now calculate

$$\begin{aligned} P(A) &= \lambda \sigma_+ + \lambda \sigma_- + \sigma_+ \sigma_- - \lambda(\sigma_+ + \sigma_- + \lambda) \\ &= \sigma_+ \sigma_- - \lambda^2 \\ &= \lambda(|\lambda + a_1| - \lambda). \end{aligned}$$

This is clearly convex in  $a \otimes n$ , which proves part (a) of the lemma in the case that  $a \wedge n \neq 0$ . When  $a \wedge n = 0$  the same proof works with any orthonormal basis  $\{n, e_2, e_3\}$ .

**(b)** Assume for a contradiction that  $\lambda \mathbf{1}$  is a point of polyconvexity of  $P$ . This means that there is some point  $(C_1, C_2, C_3)$  in  $\mathbb{R}^{19}$  such that

$$(2.15) \quad P(A) \geq P(\lambda \mathbf{1}) + C_1 \cdot (A - \lambda \mathbf{1}) + C_2 \cdot (\text{cof } A - \lambda^2 \mathbf{1}) + C_3(\det A - \lambda^3)$$

for all  $A$  in  $\mathbb{R}^{3 \times 3}$ . Note that  $C_3$  has to be zero because  $P$  is at most quadratic. Next, take  $A = t e_1 \otimes e_1$  when  $(C_1)_{11} \geq 0$  and  $A = -t e_1 \otimes e_1$  otherwise, where  $t$  is a parameter to be chosen shortly. Recall that, when  $A$  is a rank-one matrix,  $P(A) = -\lambda|A|$ . This gives

$$-(\lambda + |(C_1)_{11}|)t \geq -\lambda \text{tr } C_1 - \lambda^2 \text{tr } C_2,$$

which is easily contradicted by taking  $t$  to be sufficiently large and positive. Therefore (2.15) is impossible, which concludes the proof.  $\square$

Next, with  $N$  as in Proposition 2.12, we define

$$(2.16) \quad G(A) := P(A) - N(A)$$

for all  $3 \times 3$  matrices  $A$ . We know by equation (2.8) in Proposition 2.2 that  $P$  and  $N$  are tangent at  $\lambda \mathbf{1}$ , so clearly  $D_U G(\lambda \mathbf{1}) = 0$  for all  $U$ . Moreover, by (2.11), we also have that  $D_U^2 G(\lambda \mathbf{1}) = 0$  for all  $U$ . Thus, by (2.7), the lowest order terms in the Taylor expansion of  $G$  about  $\lambda \mathbf{1}$  are cubic (or higher powers), and this, coupled with the assumption that the exponent  $q$  in the definition of  $W$  satisfies  $q < 3$ , is a key feature which enables us to find new lower bounds for  $\lambda_{\text{crit}}$ . The technique for doing so is described in the next section. We also record the following useful property of  $G$ , which flows directly from (2.13):

$$(2.17) \quad \int_{\Omega} P(\nabla u) \, dx = \int_{\Omega} G(\nabla u) \, dx$$

for all  $u$  belonging to  $W^{1,2}(\Omega, \mathbb{R}^3)$  such that  $u = u_{\lambda}$  on  $\partial\Omega$  (in the sense of trace).

3. NEW LOWER BOUNDS ON  $\lambda_{\text{CRIT}}$ 

Let the stored-energy function  $W : \mathbb{R}^{3 \times 3} \rightarrow [0, +\infty]$  be given by

$$(3.1) \quad W(A) = |A|^q + Z(\text{cof } A) + h(\det A)$$

where  $Z : \mathbb{R}^{3 \times 3} \rightarrow [0, +\infty)$  is convex and  $h : \mathbb{R} \rightarrow [0, +\infty]$  has the following properties:

- (H1)  $h$  is convex and  $C^1$  on  $(0, +\infty)$ ;
- (H2)  $\lim_{t \rightarrow 0^+} h(t) = +\infty$  and  $\liminf_{t \rightarrow +\infty} \frac{h(t)}{t} > 0$ ;
- (H3)  $h(t) = +\infty$  if  $t \leq 0$ .

The exponent  $q$  satisfies  $2 < q < 3$ . Let

$$I(u) = \int_{\Omega} W(\nabla u) \, dx$$

and define the class of admissible maps as

$$\mathcal{A}_{\lambda} = \{u \in W^{1,q}(\Omega; \mathbb{R}^3) : u = u_{\lambda} \text{ on } \partial\Omega, I(u) < +\infty\}.$$

The following argument is straightforward and can be found in [4, Section 3]. We include it here both for completeness and as a means of deriving the function  $P$  defined by (1.3). Applying [10, Lemma A.1] to  $A \mapsto |A|^q$  gives

$$(3.2) \quad |\nabla u|^q \geq |\lambda \mathbf{1}|^q + q|\lambda \mathbf{1}|^{q-2} \lambda \mathbf{1} \cdot (\nabla u - \lambda \mathbf{1}) + \kappa |\nabla u - \lambda \mathbf{1}|^q,$$

where

$$(3.3) \quad 2^{2-q} \leq \kappa \leq q2^{1-q}.$$

Therefore, by (3.2) and by appealing to the convexity of  $Z$  and  $h$ , we obtain

$$(3.4) \quad \begin{aligned} W(\nabla u) &\geq W(\nabla u_{\lambda}) + q|\lambda \mathbf{1}|^{q-2} \lambda \mathbf{1} \cdot (\nabla u - \lambda \mathbf{1}) + \kappa |\nabla u - \lambda \mathbf{1}|^q \\ &\quad + D_A Z(\text{cof } \lambda \mathbf{1}) \cdot (\text{cof } \nabla u - \text{cof } \lambda \mathbf{1}) \\ &\quad + h'(\lambda^3)(\det \nabla u - \lambda^3), \end{aligned}$$

for any  $u \in \mathcal{A}_{\lambda}$ . Integrating (3.4), and using the facts that both  $\nabla u$  and  $\text{cof } \nabla u$  are null Lagrangians in  $W^{1,q}(\Omega, \mathbb{R}^3)$  for  $q \geq 2$ , we obtain

$$(3.5) \quad \begin{aligned} I(u) - I(u_{\lambda}) &\geq \int_{\Omega} \kappa |\nabla u - \lambda \mathbf{1}|^q + h'(\lambda^3)(\lambda_1 \lambda_2 \lambda_3 - \lambda^3) \, dx \\ &= \int_{\Omega} \kappa |\nabla u - \lambda \mathbf{1}|^q + h'(\lambda^3) \hat{\lambda}_1 \hat{\lambda}_2 \hat{\lambda}_3 \, dx + \lambda h'(\lambda^3) \int_{\Omega} P(\nabla u) \, dx. \end{aligned}$$

In deriving this, it may help to recall the identity

$$(3.6) \quad \lambda_1 \lambda_2 \lambda_3 = \hat{\lambda}_1 \hat{\lambda}_2 \hat{\lambda}_3 + \lambda \sum_{1 \leq i < j \leq 3} \lambda_i \lambda_j - \lambda^2 \sum_{i=1}^3 \lambda_i + \lambda^3,$$

where the notation  $\lambda_i$  abbreviates  $\lambda_i(A)$  and, for each  $i$ ,  $\hat{\lambda}_i := \lambda_i - \lambda$ .

Continuing from (3.5), we split the first term into two equal parts and recall the property of  $G$  and  $P$  given in (2.17), thereby obtaining:

$$\begin{aligned} I(u) - I(u_\lambda) &\geq \int_{\Omega} (\kappa/2) |\nabla u - \lambda \mathbf{1}|^q + h'(\lambda^3) \hat{\lambda}_1 \hat{\lambda}_2 \hat{\lambda}_3 dx + \\ &\quad + \int_{\Omega} (\kappa/2) |\nabla u - \lambda \mathbf{1}|^q + \lambda h'(\lambda^3) G(\nabla u) dx \\ &\geq \int_{\Omega} (\kappa/2) |\Lambda(\nabla u) - \Lambda_0|^q + h'(\lambda^3) \hat{\lambda}_1 \hat{\lambda}_2 \hat{\lambda}_3 dx + \\ &\quad + \int_{\Omega} (\kappa/2) |\nabla u - \lambda \mathbf{1}|^q + \lambda h'(\lambda^3) G(\nabla u) dx. \end{aligned}$$

Here,  $\Lambda(A)$  is the 3-vector with entries  $\lambda_i(A)$  and  $\Lambda_0 = (\lambda, \lambda, \lambda)$ . We have used the well-known inequality  $|A - \lambda \mathbf{1}| \geq |\Lambda(A) - \Lambda_0|$ .

In keeping with the notation introduced in [4, Lemma 3.2], let

$$(3.7) \quad \mathcal{F}_1(\Lambda) = (\kappa/2) |\Lambda - \Lambda_0|^q + h'(\lambda^3) \hat{\lambda}_1 \hat{\lambda}_2 \hat{\lambda}_3,$$

and, in contrast to the approach of [4], let

$$\mathcal{F}_2(A) = (\kappa/2) |A - \lambda \mathbf{1}|^q + \lambda h'(\lambda^3) G(A).$$

In these terms we then have

$$(3.8) \quad I(u) - I(u_\lambda) \geq \int_{\Omega} \mathcal{F}_1(\Lambda) dx + \int_{\Omega} \mathcal{F}_2(\nabla u) dx.$$

The sign of the first integral can be controlled by appealing to the following result:

**Lemma 3.1.** ([4, Lemma 3.3]) *The function  $\mathcal{F}_1(\Lambda)$  defined in (3.7) is pointwise nonnegative on  $\mathbb{R}^{+++} := \{(x_1, x_2, x_3) : x_i > 0 \text{ for } i = 1, 2, 3\}$  provided  $h'(\lambda^3) > 0$  and*

$$(3.9) \quad \frac{(\kappa/2)}{h'(\lambda^3) \lambda^{3-q}} \geq (q-2)^{(q-2)/2} q^{-q/2}.$$

The pointwise nonnegativity of  $\mathcal{F}_2$ , on the other hand, relies primarily on the argument given in Lemma 3.2 below. In short, the idea is that  $\mathcal{F}_2(\nabla u)$  is dominated by  $|\nabla u - \lambda|^q$  for both small and large values of  $|\nabla u - \lambda|$  provided  $h'(\lambda^3)$  is itself not too large. Thus we generate a new upper bound on  $\lambda$  which must be imposed along with (3.9) in order to guarantee that  $I(u) \geq I(u_\lambda)$ .

**Lemma 3.2.** *With  $G$  as defined in (2.16) and for any positive constant  $c_0$ , let*

$$(3.10) \quad M_2(\lambda, c_0) = \sup \left\{ \frac{|G(A)|}{|A - \lambda \mathbf{1}|^2} : |A - \lambda \mathbf{1}| \geq c_0 \right\}$$

$$(3.11) \quad M_3(\lambda, c_0) = \sup \left\{ \frac{|G(A)|}{|A - \lambda \mathbf{1}|^3} : 0 < |A - \lambda \mathbf{1}| < c_0 \right\}.$$

*Then  $\mathcal{F}_2(A) \geq 0$  for all  $3 \times 3$  matrices  $A$  provided*

$$(3.12) \quad \lambda h'(\lambda^3) \leq \min_{c_0} \left\{ (\kappa/2) \max \left\{ \frac{c_0^{q-2}}{M_2(\lambda, c_0)}, \frac{c_0^{q-3}}{M_3(\lambda, c_0)} \right\} \right\}.$$

*Proof.* By (2.7), the quantity  $M_3(\lambda, c_0)$  is finite and, in view of the at most quadratic growth of  $G$ ,  $M_3(\lambda, c_0)$  is uniformly bounded as a function of  $c_0$ . By (2.11),  $M_2(\lambda, c_0)$  is finite for all  $c_0 \geq 0$ . Let  $A \neq \lambda \mathbf{1}$  and let  $c = |A - \lambda \mathbf{1}|$ . It is immediately clear that

$$\mathcal{F}_2(A) \geq \frac{\kappa}{2} |A - \lambda \mathbf{1}|^q - \lambda h'(\lambda^3) |G(A)|,$$

and we express the right-hand side in two ways:

$$(3.13) \quad \frac{\kappa}{2} |A - \lambda \mathbf{1}|^q - \lambda h'(\lambda^3) |G(A)| = \left( \frac{\kappa}{2} |A - \lambda \mathbf{1}|^{q-j} - \lambda h'(\lambda^3) \frac{|G(A)|}{|A - \lambda \mathbf{1}|^j} \right) |A - \lambda \mathbf{1}|^j,$$

where  $j$  is either 2 or 3. Now let  $w := 2\lambda h'(\lambda^3)/\kappa$  and suppose that (3.12) holds. Then, in particular,

$$(3.14) \quad w \leq \max\{c^{q-2}/M_2(\lambda, c), c^{q-3}/M_3(\lambda, c)\}.$$

If the maximum in (3.14) is given by  $c^{q-2}/M_2(\lambda, c)$  then  $wM_2(\lambda, c) \leq c^{q-2}$ , and hence  $\frac{\kappa}{2}c^{q-2} - \lambda h'(\lambda^3) \frac{|G(A)|}{c^{q-2}} \geq 0$ . Using (3.13) with  $j = 2$ , we see that  $\mathcal{F}_2(A) \geq 0$ . If the maximum in (3.14) is given by  $c^{q-3}/M_3(\lambda, c)$  then we can argue similarly, this time using (3.13) with  $j = 3$ , to conclude that  $\mathcal{F}_2(A) \geq 0$ .  $\square$

**Lemma 3.3.** *Let  $f_1(c_0) = \frac{c_0^{q-2}}{M_2(\lambda, c_0)}$ ,  $f_2(c_0) = \frac{c_0^{q-3}}{M_3(\lambda, c_0)}$  and define  $\xi(c_0) = \max\{f_1(c_0), f_2(c_0)\}$ . Then  $f_1$  is nondecreasing,  $f_2$  is nonincreasing, and*

$$\inf_{c_0 > 0} \xi(c_0) = M_2(\lambda, c^*)^{q-3} M_3(\lambda, c^*)^{2-q}$$

where  $c^*$  is the unique fixed point of the function  $c_0 \mapsto \frac{M_2(\lambda, c_0)}{M_3(\lambda, c_0)}$ .

*Proof.* Let  $M_2(c_0) = M_2(\lambda, c_0)$  and  $M_3(c_0) = M_3(\lambda, c_0)$  for brevity. It is clear from their definitions that  $M_2(c_0)$  and  $M_3(c_0)$  are nonincreasing and nondecreasing respectively. From this and the fact that  $2 < q < 3$ , it follows that  $f_1$  is nondecreasing and  $f_2$  is nonincreasing. Since  $f_1(c_0) \rightarrow +\infty$  as  $c_0 \rightarrow +\infty$  and  $f_2(c_0) \rightarrow +\infty$  as  $c_0 \rightarrow 0+$ , there is a unique point  $c^*$  such that  $f_1(c^*) = f_2(c^*)$ ,

$$\xi(c_0) = \begin{cases} f_2(c_0) & \text{if } c_0 \leq c^* \\ f_1(c_0) & \text{if } c_0 \geq c^* \end{cases}$$

and where, moreover,  $\xi(c^*)$  is the global minimum of  $\xi$  on  $\mathbb{R}^+$ . It is straightforward to see that the condition  $f_1(c^*) = f_2(c^*)$  is equivalent to the condition  $c^* = M_2(c^*)/M_3(c^*)$ , and that  $f_1(c^*) = M_2(c^*)^{q-3} M_3(c^*)^{2-q}$ .  $\square$

We are now in a position to state the main theorem of this section.

**Theorem 3.4.** *Let  $W$  be as in (3.1) and suppose that  $\lambda$  is chosen so that*

$$(3.15) \quad 0 \leq \lambda^{3-q} h'(\lambda^3) \leq \min\{(\kappa/2)(q-2)^{(2-q)/2} q^{q/2}, (\kappa/2)\lambda^{2-q} M_2(\lambda, c^*)^{q-3} M_3(\lambda, c^*)^{2-q}\},$$

where  $c^*$  is the unique fixed point of the function  $c_0 \mapsto M_2(\lambda, c_0)/M_3(\lambda, c_0)$ . Then  $I(u) \geq I(u_\lambda)$  for any map  $u \in H^1(\Omega, \mathbb{R}^3)$  whose boundary values agree with those of  $u_\lambda$  in the sense of trace. In particular, the largest possible value  $\lambda_0$  satisfying (3.15) is a lower bound for  $\lambda_{crit}$ . Moreover, if the rightmost inequality of (3.15) is strict then there is  $C = C(\Omega) > 0$  such that

$$(3.16) \quad I(u) - I(u_\lambda) \geq C \int_{\Omega} |\nabla u - \lambda \mathbf{1}|^q + |G(\nabla u)| \, dx.$$

for all admissible  $u$ .

*Proof.* According to (3.8),  $I(u) - I(u_\lambda)$  is bounded below by the sum of  $\int_\Omega \mathcal{F}_1(\nabla u) dx$  and  $\int_\Omega \mathcal{F}_2(\nabla u) dx$ . By inequality (3.15) and Lemma 3.1, the first of these integrals is nonnegative, while Lemmas 3.2 and 3.3 together imply that the second integral is nonnegative. Either way, it follows that  $I(u) \geq I(u_\lambda)$ , as claimed. It is then clear that  $\lambda_0$ , as defined above, is not larger than  $\lambda_{\text{crit}}$ .

Now suppose that the rightmost inequality in (3.15) is strict. The proof of [4, Theorem 3.6] shows that the inequality  $\lambda^{3-q} h'(\lambda^3) < \frac{\kappa}{2}(q-2)^{\frac{2-q}{2}} q^{\frac{q}{2}}$  implies, for some  $\epsilon > 0$ , that

$$\mathcal{F}_1(A) \geq \epsilon |\Lambda(A) - (\lambda, \lambda, \lambda)|^q$$

for all  $A$  in  $\mathbb{R}^{3 \times 3}$ . Here,  $\Lambda(A) = (\lambda_1(A), \lambda_2(A), \lambda_3(A))$  is the vector of singular values of the matrix  $A$ . The rigidity argument, with minor modifications, given in [4, Theorem 3.6] then shows that there is a constant,  $\beta(\Omega)$ , say, such that

$$\int_\Omega |\Lambda(\nabla u) - (\lambda, \lambda, \lambda)|^q dx \geq \beta(\Omega) \int_\Omega |\nabla u - \lambda \mathbf{1}|^q dx.$$

Hence

$$\int_\Omega \mathcal{F}_1(\nabla u) dx \geq \epsilon \beta(\Omega) \int_\Omega |\nabla u - \lambda \mathbf{1}|^q dx.$$

Finally, we deal with the term involving  $\mathcal{F}_2(A)$ . Fix  $A$  in  $\mathbb{R}^{3 \times 3}$  and let  $c_0 = |A - \lambda \mathbf{1}|$ . According to the proof of Lemma 3.2,

$$(3.17) \quad \mathcal{F}_2(A) \geq \frac{\kappa}{2} \left( |A - \lambda \mathbf{1}|^{q-j} - \frac{2\lambda h'(\lambda^3)}{\kappa} \frac{|G(A)|}{|A - \lambda \mathbf{1}|^j} \right) |A - \lambda \mathbf{1}|^j$$

for  $j = 2$  and  $3$ . Reusing the notation  $w = 2\lambda h'(\lambda^3)/\kappa$ , and still assuming that the rightmost inequality of (3.15) is strict, there is  $\epsilon' > 0$ , which is independent of  $c_0$ , such that

$$(3.18) \quad w + \epsilon' \leq \max\{c_0^{q-2}/M_2(\lambda, c_0), c_0^{q-3}/M_3(\lambda, c_0)\}.$$

By rewriting (3.17), we obtain

$$\mathcal{F}_2(A) \geq \frac{\kappa}{2} \left( c_0^{q-j} - (w + \epsilon') M_j(\lambda, c_0) \right) c_0^j + \frac{\kappa \epsilon'}{2} |G(A)|$$

for  $j = 2$  and  $3$ . Thanks to (3.18), the term in brackets is nonnegative, which leaves  $\mathcal{F}_2(A) \geq \kappa \epsilon' |G(A)|/2$ . Both terms in the right-hand side of inequality (3.16) are now accounted for.  $\square$

**Remark 3.5.** The goal of Theorem 3.4 is to give the largest possible bound on  $\lambda$  such that  $I(u) \geq I(u_\lambda)$ . A careful look at the proof of Lemma 3.2 shows that one could replace  $|G(A)|$  by  $G^-(A) = -\min\{G(A), 0\}$  and that the same conclusions would result, but with

$$M_2^-(\lambda, c_0) = \sup \left\{ \frac{G^-(A)}{|A - \lambda \mathbf{1}|^2} : |A - \lambda \mathbf{1}| \geq c_0 \right\}$$

$$M_3^-(\lambda, c_0) = \sup \left\{ \frac{G^-(A)}{|A - \lambda \mathbf{1}|^3} : 0 < |A - \lambda \mathbf{1}| < c_0 \right\}$$

in place of  $M_2(\lambda, c_0)$  and  $M_3(\lambda, c_0)$  respectively. Since  $M_j^-(\lambda, c_0) \leq M_j(\lambda, c_0)$  for  $j = 2$  and  $3$ , it follows that the upper bound involving  $c^*$  in (3.15) would not decrease. Numerical evidence suggests that it does, in fact, increase, and thus provides a better (lower) bound for  $\lambda_{\text{crit}}$ . See Remark 3.11 and Section A.3 for further details.

**3.1. Calculations leading to a concrete upper bound.** The upper bound (3.15) given in the statement of Theorem 3.4 contains two terms, one of which is explicitly given in terms of the exponent  $q$  and one which depends on the fixed point  $c^*$  of the function  $c_0 \mapsto M_2(\lambda, c_0)/M_3(\lambda, c_0)$ . While it does not seem to be possible to find  $c^*$  purely analytically, one can nevertheless make progress using a mixture of analysis and a careful numerical calculation, as we now describe.

**Proposition 3.6.** *With  $M_2(\lambda, c_0)$  as defined by (3.10),  $\lim_{c_0 \rightarrow \infty} M_2(\lambda, c_0) = \sqrt{2}$ .*

*Proof.* First, note that the limit in the statement exists because  $M_2(c)$  is nonincreasing and bounded below. Now

$$\frac{G(A)}{|A - \lambda \mathbf{1}|^2} = K(\hat{A}) + \frac{\lambda(2\text{tr } A - 3\lambda)}{|A - \lambda \mathbf{1}|^2} K(\hat{A}) - \frac{\lambda}{|A - \lambda \mathbf{1}|^2} \left( \sum_{i=1}^3 \lambda_i(A) - \text{tr } A \right)$$

where  $\hat{A} = A/|A|$  and

$$K(\hat{A}) = \sum_{1 \leq i < j \leq 3} \lambda_i(\hat{A})\lambda_j(\hat{A}) - \text{tr } \text{cof } \hat{A}.$$

Let  $c_0 > 0$ . Since  $K$  is bounded and the term  $\sum_{i=1}^3 \lambda_i(A) - \text{tr } A$  has at most linear growth in  $A$ , it is clear that there are constants  $\alpha_1$  and  $\alpha_2$ , depending only on  $\lambda$ , such that

$$(3.19) \quad |K(\hat{A})| + \frac{\alpha_1}{c_0} \leq \frac{|G(A)|}{|A - \lambda \mathbf{1}|^2} \leq |K(\hat{A})| + \frac{\alpha_2}{c_0}$$

whenever  $|A - \lambda \mathbf{1}| \geq c_0$ . We now prove that (a)  $\max\{K(\hat{A})\} = \sqrt{2}$ , where the maximum is taken over all unit matrices  $\hat{A}$ , and (b) that given any  $c > 0$  there is  $A(c)$  such that  $|A(c) - \lambda \mathbf{1}| \geq c$  and  $K(\widehat{A(c)}) = \sqrt{2}$ . The proposition then follows from this and (3.19).

To simplify the notation we replace  $\hat{A}$  by  $A$  in the following. By the polar factorization theorem (see [5, Theorem 3.2-2]), there is a positive semidefinite and symmetric matrix  $U$  and a matrix  $R \in O(3)$  such that  $A = RU$ . The eigenvalues of  $U$  are the singular values  $\lambda_i(A)$ ,  $i = 1, 2, 3$ , and one has

$$\begin{aligned} K(A) &= \sum_{1 \leq i, j \leq 3} \lambda_i(A)\lambda_j(A) - \text{tr}(\text{cof } R \text{ cof } U) \\ &= \sum_{1 \leq i, j \leq 3} \lambda_i(A)\lambda_j(A) - \text{tr}((\det R) R Q^T \text{ cof } D Q), \end{aligned}$$

where the well-known decomposition  $U = Q^T D Q$  has been used, with  $D = \text{diag}(\lambda_1(A), \lambda_2(A), \lambda_3(A))$  and  $Q$  in  $O(3)$ . Note that  $\det R = \pm 1$ . It follows that

$$K(A) = \sum_{1 \leq i, j \leq 3} \lambda_i(A)\lambda_j(A) - \text{tr}(P \text{diag}(\lambda_2\lambda_3, \lambda_1\lambda_3, \lambda_1\lambda_2)),$$

where  $P := Q \det R R Q^T$  belongs to  $SO(3)$ . The term involving the trace satisfies

$$\text{tr}(P \text{diag}(\lambda_2\lambda_3, \lambda_1\lambda_3, \lambda_1\lambda_2)) = P_{11}\lambda_2\lambda_3 + P_{22}\lambda_1\lambda_3 + P_{33}\lambda_1\lambda_2.$$

The term  $\sum_{1 \leq i, j \leq 3} \lambda_i(A)\lambda_j(A)$  is independent of  $P$ , so we can vary  $P$  in order to minimize the term involving the trace given above. Since  $P$  belongs to  $SO(3)$ , its rows and columns are orthogonal unit vectors. In particular,  $|P_{11}| \leq 1$ , so to make  $P_{11}\lambda_2\lambda_3$  minimal we should take  $P_{11} = -1$ . (Here we have used the ordering  $\lambda_3 \geq \lambda_2 \geq \lambda_1$ .) Hence row 1 of  $P$  is  $(-1, 0, 0)$ , forcing the second and third rows to take the form  $(0, P_{22}, P_{33})$  and  $(0, P_{32}, P_{33})$  respectively. In

particular,  $\det P = -(P_{22}P_{33} - P_{23}P_{32}) = 1$ . Without loss of generality, we may take  $P_{22} = \cos \alpha$ ,  $P_{32} = P_{23} = \sin \alpha$ ,  $P_{33} = -\cos \alpha$  for some  $\alpha$ , in which case

$$\begin{aligned} P_{11}\lambda_2\lambda_3 + P_{22}\lambda_1\lambda_3 + P_{33}\lambda_1\lambda_2 &\geq -\lambda_2\lambda_3 + \cos \alpha(\lambda_1\lambda_3 - \lambda_1\lambda_2) \\ &\geq -\lambda_2\lambda_3 - \lambda_1(\lambda_3 - \lambda_2) \end{aligned}$$

by choosing  $\alpha = \pi$ . Hence

$$K(A) \leq 2\lambda_3(\lambda_1 + \lambda_2)$$

where the matrix  $P$  yielding this upper bound is given by  $P = \text{diag}(-1, -1, 1)$ . Bearing in mind that  $|A|^2 = \sum_{i=1}^3 \lambda_i^2 = 1$ , it can be shown that  $2\lambda_3(\lambda_1 + \lambda_2) \leq \sqrt{2}$ , and that a maximizing choice of singular values is  $\lambda_1 = \lambda_2 = 1/2$ ,  $\lambda_3 = 1/\sqrt{2}$ . A suitable choice for a maximizing  $A$  would therefore be  $A_0 = P \text{diag}(1/2, 1/2, 1/\sqrt{2}) = \text{diag}(-1/2, -1/2, 1/\sqrt{2})$ , and one can check that  $K(A_0) = \sqrt{2}$ . To conclude the proof, it is enough to choose  $A(c) := rA_0$  for any  $r = r(c)$  large enough that  $|A(c) - \lambda \mathbf{1}| \geq c$ .  $\square$

Since  $D^2G(\lambda \mathbf{1}) = 0$ , it can easily be shown that  $M_2(\lambda, c_0)$  is constant as a function of  $c_0$  for all sufficiently small and positive  $c_0$ . We are therefore justified in writing  $M_2(\lambda, 0) := \lim_{c_0 \rightarrow 0} M_2(\lambda, c_0)$ , and in fact

$$M_2(\lambda, 0) = \sup \left\{ \frac{|G(A)|}{|A - \lambda \mathbf{1}|^2} : A \neq \lambda \mathbf{1} \right\}.$$

According to Proposition 3.6, we must have

$$(3.20) \quad M_2(\lambda, 0) \geq M_2(\lambda, c_0) \geq \sqrt{2} \text{ for all } c_0 > 0.$$

Using a ‘brute force’ approach, which we describe in the appendix, we find that  $\sqrt{2} - M_2(\lambda, 0) \approx 2.7 \times 10^{-9}$ , independently of  $\lambda$  in the range  $[1, 2]$ . Thus, in view of (3.20), it seems that  $M_2(\lambda, c_0) \equiv \sqrt{2}$  for all  $c_0$  and  $\lambda$ , and we record this as:

**Conjecture 3.7.** For all  $\lambda > 0$  and  $c_0 > 0$ ,  $M_2(\lambda, c_0) = \sqrt{2}$ .

In order to explain the method used to approximate  $M_2(\lambda, 0)$ , we introduce the notation

$$m_l(A, \lambda) := \frac{|G(A)|}{|A - \lambda \mathbf{1}|^l}$$

for  $l = 2$  and  $3$ , and we restrict attention to  $\lambda$  in the range  $[1, 2]$ . Choose a random starting matrix  $A_1$ , compute  $g_1 := \nabla m_2(A_1, \lambda)$  and find the scalar  $\sigma_1$ , say, which maximises  $m_2(A_1 + \sigma g_1)$ . Let the maximising value of  $\sigma$  be  $\sigma_1$ ; then we set  $A_2 = A_1 + \sigma_1 g_1$ . Proceeding iteratively, we compute a sequence of matrices  $A_i$ . It turns out that  $m_2(A_i, \lambda)$  tends to  $\sqrt{2}$  as  $i$  increases, regardless of  $\lambda$ .

Supposing that Conjecture 3.7 is right, we are now required to find  $M_2(\lambda, c^*)^{q-3} M_3(\lambda, c^*)^{q-2} = 2^{\frac{q-3}{2}} M_3(\lambda, c^*)$ , where  $c^*$  satisfies  $c^* = \sqrt{2}/M_3(\lambda, c^*)$ . We begin by recalling that  $c_0 \mapsto M_3(\lambda, c_0)$  is nondecreasing, and that, thanks to the at most quadratic growth of the function  $G(A)$ , there is  $c_1(\lambda)$  such that  $M_3(\lambda, c_0) = M_3(\lambda, c_1)$  for all  $c_0 \geq c_1$ . We are therefore justified in defining  $M_3(\lambda, \infty) := M_3(\lambda, c_1)$ . Appealing again to numerical techniques, which we describe in the appendix, we find that, for  $\lambda$  in the range  $[1, 2]$ , there is very good agreement between  $c_1(\lambda)$  and the expression  $\nu_3 \lambda + \nu_2$ , where  $\nu_2 \approx 1.764 \times 10^{-3}$  and  $\nu_3 \approx 1.842$ . Moreover, for the same range of  $\lambda$ , there is strong numerical evidence for the approximation  $M_3(\lambda, \infty) \approx \nu_1/\lambda$ , where  $\nu_1 \approx 0.4501$ . See Fig. 1.

This leads naturally to the following:

FIGURE 1. Left: plot of  $\nu_1/\lambda = 0.4501/\lambda$  (continuous curve) fitted to a series of values of  $M_3(\lambda, \infty)$  (filled circles). Right: plot of  $c_1(\lambda)$  versus  $\lambda$ . The error bars, which are almost too narrow to see, show the maximum and minimum values of  $c_1$  from among the fifteen top-ranking matrices. (See the appendix for details.)

**Conjecture 3.8.**  $M_3(\lambda, c_0) = M_3(\lambda, c_1) = \nu_1/\lambda$  for all  $c_0 \geq c_1(\lambda)$  and for  $\lambda$  in the range  $[1, 2]$ , where  $c_1(\lambda) = \nu_3\lambda + \nu_2$  and the values of  $\nu_1, \nu_2$  and  $\nu_3$  are given above.

Let us now suppose that Conjectures 3.7 and 3.8 are correct. Note that then the function  $p : c \mapsto M_2(\lambda, c)/M_3(\lambda, c) = \sqrt{2}\lambda/\nu_1$  is independent of  $c$  for all  $c \geq c_1$ . There are thus two possibilities for the fixed point  $c^*$  of  $p$ : either  $c^* = \sqrt{2}\lambda/\nu_1$  or  $c^* < c_1$ . Suppose for a contradiction that  $c^* < c_1$ . Then  $M_3(\lambda, c^*) \leq M_3(\lambda, c_1)$ , and so  $p(c^*) \geq p(c_1)$ . But  $p(c^*) = c^*$  and  $p(c_1) = \sqrt{2}\lambda/\nu_1$ , which gives  $c^* \geq \sqrt{2}\lambda/\nu_1$ . By hypothesis,  $c_1 > c^*$ , which when combined with the preceding inequality implies  $c_1 > \sqrt{2}\lambda/\nu_1$ . Applying Conjecture 3.8 and rearranging, we see that this is equivalent to  $\nu_1\nu_2 \geq (\sqrt{2} - \nu_3\nu_1)\lambda$ , which can only hold if  $\lambda \leq 1.357 \times 10^{-3}$ . But we supposed that  $\lambda \geq 1$ , which is a contradiction. In summary, we have shown the following result.

**Proposition 3.9.** *Let  $\lambda$  belong to the interval  $[1, 2]$  and suppose that Conjectures 3.7 and 3.8 are correct. Then the unique fixed point of the function  $p(c) := M_2(\lambda, c)/M_3(\lambda, c)$  is given by  $c^* = \sqrt{2}\lambda/\nu_1$ , and*

$$M_2(\lambda, c^*)^{q-3} M_3(\lambda, c^*)^{2-q} = \frac{1}{\sqrt{2}} \left( \frac{\sqrt{2}\lambda}{\nu_1} \right)^{q-2}.$$

Referring back to the upper bound given in (3.15), we now have the following:

**Corollary 3.10.** *Let the assumptions of Proposition 3.9 hold. Then  $I(u) \geq I(u_\lambda)$  provided*

$$(3.21) \quad 0 \leq \lambda^{3-q} h'(\lambda^3) \leq (\kappa/2)(\sqrt{2})^{q-3} \nu_1^{2-q}.$$

*Proof.* It is enough to show that

$$\min\{(q-2)^{(2-q)/2} q^{q/2}, \lambda^{2-q} M_2(\lambda, c^*)^{q-3} M_3(\lambda, c^*)^{2-q}\} = (\sqrt{2})^{q-3} \nu_1^{2-q}.$$

By Proposition 3.9, we clearly have  $\lambda^{2-q} M_2(\lambda, c^*)^{q-3} M_3(\lambda, c^*)^{2-q} = (\sqrt{2})^{q-3} \nu_1^{2-q}$ . Therefore it remains to show that  $(q-2)^{(2-q)/2} q^{q/2} > (\sqrt{2})^{q-3} \nu_1^{2-q}$  for  $2 < q < 3$ . Let  $z(q) = (\sqrt{2})^{q-3} \nu_1^{2-q}$

and  $y(q) = (q-2)^{(2-q)/2}q^{q/2}$ . Note that  $z$  is convex, while a short calculation reveals that  $y$  is concave on the interval  $(2, 3)$ . Therefore the inequality  $z(q) < y(q)$  will follow from the pair of inequalities  $z(2) < y(2+)$  and  $z(3) < y(3)$ , both of which are easy to check. This concludes the proof.  $\square$

**Remark 3.11.** Corollary 3.10 suggests that the bound is not the best possible: we would really expect both terms in the bound given in (3.15) to play a role. One way to achieve this might be to use the quantities  $M_j^-$  in place of  $M_j$  for  $j = 2, 3$ . Indeed, we find, again numerically, that there is very good agreement between  $M_3^-(\lambda, \infty)$  and  $\nu'_1/\lambda$  for  $\lambda$  in the range  $[1, 2]$ , where  $\nu'_1 = 0.1923$ . Interestingly, if we then replace  $M_3$  by  $M_3^-$  everywhere in the preceding calculations, we find that  $z(3) < y(3)$  no longer holds. In other words, *both* terms in the upper bound given by (3.15) appear to contribute. This observation comes with some caveats, however; see Section A.3 in the appendix.

#### 4. A CONDITION FOR THE INEQUALITY $I(u) \leq I(u_\lambda)$

The results in the previous sections provide conditions on  $\lambda$  under which the inequality  $I(u) \geq I(u_\lambda)$  holds for admissible maps. It is natural to ask what information results from supposing that  $I(u) \leq I(u_\lambda)$ , and, of our results, Theorem 3.4 is the first place to look. Now, if the rightmost inequality of (3.15) is strict, then  $u_\lambda$  sits in a ‘potential well’, as expressed by the estimate (3.16), which we recall here for the reader’s convenience:

$$I(u) - I(u_\lambda) \geq C \int_{\Omega} |\nabla u - \lambda 1|^q + |G(\nabla u)| dx.$$

Thus, in these circumstances,  $I(u) \leq I(u_\lambda)$  is impossible. If (3.15) holds with equality then a similar remark applies, but with the additional possibility of losing one or both terms in the right-hand side of (3.16). And when (3.15) fails, the preceding analysis tells us nothing about those  $u$  whose energy satisfies  $I(u) \leq I(u_\lambda)$ . Therefore a different approach is called for.

Consider the following simplified model, in which we set the function  $Z$  appearing in (3.1) to zero. Thus we let

$$(4.1) \quad W(A) = |A|^q + h(\det A) \quad \text{for } A \in \mathbb{R}^{3 \times 3},$$

where, as before,  $2 < q < 3$ . Let  $\lambda > 1$  and note that, by the convexity of  $h$ ,

$$(4.2) \quad \begin{aligned} I(u) - I(u_\lambda) &\geq \int_{\Omega} h'(\lambda^3)(\det \nabla u - \lambda^3) + |\nabla u|^q - |\nabla u_\lambda|^q dx \\ &\geq \int_{\Omega} h'(\lambda^3)(\det \nabla u - \lambda^3) + \kappa |\nabla u - \nabla u_\lambda|^q dx, \end{aligned}$$

where we have used (3.2) and (3.4) with  $Z = 0$ , and where the constant  $\kappa$  obeys the bounds specified in inequality (3.3). Using identity (3.6) and the definition (1.3) of  $P$ , write

$$(4.3) \quad \det \nabla u - \lambda^3 = \hat{\lambda}_1 \hat{\lambda}_2 \hat{\lambda}_3 + \lambda P(\nabla u),$$

where  $\hat{\lambda}_i := \lambda_i - \lambda$  and  $\lambda_i := \lambda_i(\nabla u)$ . Finally, recall that the function  $G$  defined in (2.16) satisfies

$$\int_{\Omega} G(\nabla u) dx = \int_{\Omega} P(\nabla u) dx$$

whenever  $u$  is admissible. Combining this with (4.2) and (4.3), we have

$$(4.4) \quad I(u) - I(u_\lambda) \geq \int_{\Omega} h'(\lambda^3)(\hat{\lambda}_1 \hat{\lambda}_2 \hat{\lambda}_3 + \lambda G(\nabla u)) + \kappa |\nabla u - \nabla u_\lambda|^q dx.$$

It will be useful to have a shorthand for the function with prefactor  $h'(\lambda^3)$  appearing in (4.4); accordingly, let

$$(4.5) \quad H(A) = \hat{\lambda}_1(A) \hat{\lambda}_2(A) \hat{\lambda}_3(A) + \lambda G(A).$$

We now give a series of results which allow us to find a lower bound on the function  $H$ .

**Lemma 4.1.** *Let  $A$  be a  $3 \times 3$  matrix such that  $\det A > 0$  and whose singular values obey  $\lambda_1(A) \leq \lambda_2(A) \leq \lambda_3(A)$ . Then there is  $R$  in  $SO(3)$  such that*

$$(4.6) \quad G(A) = (1 - R_{11})\alpha + (1 - R_{22})\beta + (1 - R_{33})\gamma,$$

where  $\alpha := \lambda_2 \lambda_3 - \lambda \lambda_1$ ,  $\beta := \lambda_1 \lambda_3 - \lambda \lambda_2$  and  $\gamma := \lambda_1 \lambda_2 - \lambda \lambda_3$  satisfy  $\alpha \geq \beta \geq \gamma$ , and  $\lambda_j = \lambda_j(A)$  for  $j = 1, 2, 3$ .

*Proof.* For brevity, write  $\lambda_j$  in place of  $\lambda_j(A)$  for  $j = 1, 2, 3$ . Note that  $P(A) = \alpha + \beta + \gamma$  by definition, and that, by polar decomposition, there are matrices  $Q_1$  and  $Q_2$ , belonging to  $O(3)$ , such that  $A = Q_2 Q_1^T D Q_1$  and where  $D$  is a diagonal matrix with entries  $\lambda_1, \lambda_2$  and  $\lambda_3$ . (See [5, Theorem 3.2-2].) Since  $\det A > 0$ ,  $Q_2$  must belong to  $SO(3)$ . We see that

$$\begin{aligned} N(A) &= \operatorname{tr} \operatorname{cof}(Q_2 Q_1^T D Q_1) - \lambda \operatorname{tr} Q_2 Q_1^T D Q_1 \\ &= \operatorname{tr}(R(\operatorname{cof} D - \lambda D)), \end{aligned}$$

where  $R := Q_1 Q_2 Q_1^T$  belongs to  $SO(3)$ . Hence  $N(A) = R_{11}\alpha + R_{22}\beta + R_{33}\gamma$ , and (4.6) follows. Finally, since  $\lambda_3(A) \geq \lambda_2(A) \geq \lambda_1(A)$ , the inequalities  $\alpha \geq \beta \geq \gamma$  must hold.  $\square$

Our aim is to minimize  $G(A)$  by allowing  $R$  to vary in  $SO(3)$ . To that end, consider the following.

**Lemma 4.2.** *Let  $R$  belong to  $SO(3)$  and suppose that it minimizes*

$$g(R) := (1 - R_{11})\alpha + (1 - R_{22})\beta + (1 - R_{33})\gamma.$$

Then

$$(4.7) \quad \alpha R_{12} - \beta R_{21} = 0$$

$$(4.8) \quad \alpha R_{13} - \gamma R_{31} = 0$$

$$(4.9) \quad \beta R_{23} - \gamma R_{32} = 0.$$

If none of  $\alpha, \beta, \gamma$  is zero then

$$(4.10) \quad \left( \left( \frac{\alpha}{\beta} \right)^2 - 1 \right) R_{12}^2 + \left( \left( \frac{\alpha}{\gamma} \right)^2 - 1 \right) R_{13}^2 = 0$$

$$(4.11) \quad \left( \left( \frac{\beta}{\gamma} \right)^2 - 1 \right) R_{23}^2 - \left( \left( \frac{\alpha}{\beta} \right)^2 - 1 \right) R_{12}^2 = 0.$$

Moreover, if exactly one of  $\alpha, \beta, \gamma$  is zero, then either  $R_{11}^2 = 1$  or  $R_{22}^2 = 1$  or  $R_{33}^2 = 1$ . The same is true if none of  $\alpha, \beta$  and  $\gamma$  is zero, provided at least one of  $\alpha > \beta$  and  $\beta > \gamma$  holds.

*Proof.* We first show that (4.7)-(4.9) hold. It is well known that the tangent space to  $SO(3)$  at  $R$  consists of those  $\rho$  in  $R^{3 \times 3}$  such that  $R^T \rho$  is antisymmetric. From this, it easily follows that there are real numbers  $a, b$  and  $c$  such that

$$\begin{aligned}\rho_{11} &= -bR_{13} - aR_{12} \\ \rho_{22} &= aR_{21} - cR_{23} \\ \rho_{33} &= bR_{31} + cR_{32}.\end{aligned}$$

Now suppose that  $R(\epsilon)$  is a smooth path of matrices belonging to  $SO(3)$ , and satisfying  $R(0) = R$  and  $\dot{R}(0) = \rho$ . We then have

$$\partial_{\epsilon}|_{\epsilon=0}g(R(\epsilon)) = a(R_{12}\alpha - R_{21}\beta) + b(R_{13}\alpha - R_{31}\gamma) + c(R_{23}\beta - R_{32}\gamma),$$

so that, by varying  $a, b$  and  $c$  independently, the stationarity conditions (4.7)-(4.9) follow. If  $\alpha, \beta$  and  $\gamma$  are all nonzero then equation (4.10) follows by solving (4.7) and (4.8) for  $R_{21}$  and  $R_{31}$  respectively and then using the fact that  $R_{12}^2 + R_{13}^2 = R_{21}^2 + R_{31}^2$ . Equation (4.11) follows from a similar argument.

To prove the last part of the statement, we consider cases as follows.

Case (i):  $\alpha = \mathbf{0} > \beta \geq \gamma$ .

Using (4.7) and (4.8), we see that  $R_{21} = R_{31} = 0$ . Hence, since the first column of  $R$  is a unit vector, we must have  $R_{11}^2 = 1$ .

Case (ii):  $\alpha > \mathbf{0} = \beta > \gamma$ .

Using (4.7) and (4.9), we see that  $R_{12} = 0$  and  $R_{32} = 0$ . Since the second column of  $R$  is a unit vector, we must have  $R_{22}^2 = 1$ .

Case (iii):  $\alpha \geq \beta > \gamma = \mathbf{0}$ .

Equations (4.8) and (4.9) imply respectively that  $R_{13} = 0$  and  $R_{23} = 0$ . Hence  $R_{33}^2 = 1$ .

Case (iv):  $\alpha > \beta > \gamma; \alpha \neq \mathbf{0}, \beta \neq \mathbf{0}, \gamma \neq \mathbf{0}$ . In this case, (4.10) implies that  $R_{12} = R_{13} = 0$ , and so  $R_{11}^2 = 1$ .

Case (v):  $\alpha = \beta > \gamma; \alpha \neq \mathbf{0}, \beta \neq \mathbf{0}, \gamma \neq \mathbf{0}$ . Now (4.10) implies that  $R_{13} = 0$  and (4.11) that  $R_{23} = 0$ . Hence  $R_{33}^2 = 1$ .

Case (vi):  $\alpha > \beta = \gamma; \alpha \neq \mathbf{0}, \beta \neq \mathbf{0}, \gamma \neq \mathbf{0}$ . Equations (4.10) and (4.11) imply that  $R_{12} = R_{13} = 0$ , and hence  $R_{11}^2 = 1$ .  $\square$

The next result will enable us to deal with the case  $\alpha = \beta = \gamma$ .

**Lemma 4.3.** *Let  $R$  belong to  $SO(3)$  and suppose that  $R$  is symmetric. Then  $3 \geq \text{tr } R \geq -1$ .*

*Proof.* If  $R$  is a symmetric, orthogonal matrix then  $R^2 = \mathbf{1}$ , from which it follows that any eigenvalue  $\mu_i$  of  $R$  must satisfy  $\mu_i^2 = 1$ . Moreover,  $\mu_1\mu_2\mu_3 = 1$ , from which it follows that  $3 \geq \text{tr } R = \mu_1 + \mu_2 + \mu_3 \geq -1$ .  $\square$

**Lemma 4.4.** *Let  $R$  belong to  $SO(3)$ . Then*

$$g(R) = (1 - R_{11})\alpha + (1 - R_{22})\beta + (1 - R_{33})\gamma$$

*satisfies*

$$(4.12) \quad g(R) \geq \min\{2(\beta + \gamma), 0\}.$$

*Proof.* If two or more of  $\alpha, \beta$  and  $\gamma$  are zero then the lower bound  $g(R) \geq 2(\beta + \gamma)$  is trivial. Thus, to minimize  $g$ , we may begin by supposing that the conditions of Lemma 4.2 apply, so that, in all cases except  $\alpha = \beta = \gamma$ , we have either that  $R_{11}^2 = 1$  or  $R_{22}^2 = 1$  or  $R_{33}^2 = 1$ . First

suppose that  $R_{11}^2 = 1$ . Then the diagonal elements of  $R$  are either of the form  $1, \cos \sigma, \cos \sigma$  for some  $\sigma$ , or else of the form  $-1, \cos \sigma, -\cos \sigma$ . In the former case,

$$g(R) = (1 - \cos \sigma)(\beta + \gamma).$$

If  $\beta + \gamma \geq 0$  then clearly  $g(R) \geq 0$ . If  $\beta + \gamma < 0$  then to minimize  $g$  we take  $\cos \sigma = -1$  and the claimed lower bound follows. If  $R_{11} = -1$ , then

$$g(R) = 2\alpha + \beta + \gamma + (\gamma - \beta) \cos \sigma,$$

which, since  $\gamma \leq \beta$ , implies that we should take  $\cos \sigma = 1$  in order to minimize  $g$ . Hence

$$g(R) \geq 2(\alpha + \gamma) \geq 2(\beta + \gamma).$$

If  $R_{33}^2 = 1$  or  $R_{22}^2 = 1$  then the argument needed is similar. Finally, let us suppose that  $\alpha = \beta = \gamma$ . Then

$$g(R) = (3 - \operatorname{tr} R)\gamma.$$

If  $\gamma < 0$  then  $g$  is minimized when  $\operatorname{tr} R = -1$ , according to Lemma 4.3. Hence, in this case,  $g(R) \geq 4\gamma = 2(\beta + \gamma)$ . Otherwise,  $g(R) \geq 0$  because  $\operatorname{tr} R \leq 3$  by Lemma 4.3. This completes the proof.  $\square$

**Proposition 4.5.** *Let  $\lambda > 0$ , let  $H$  be given by (4.5) and let  $\lambda_1 \leq \lambda_2 \leq \lambda_3$  be the singular values of  $A$ . Then*

$$H(A) \geq \begin{cases} (\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda) & \text{if } \lambda_1 \geq \lambda \\ (\lambda_1 - \lambda)(\lambda_2 + \lambda)(\lambda_3 + \lambda) & \text{if } \lambda_1 \leq \lambda. \end{cases}$$

*Proof.* First suppose that  $\lambda_1 \geq \lambda$ . Note that  $\beta + \gamma = (\lambda_1 - \lambda)(\lambda_2 + \lambda_3)$  is then nonnegative, and so, by inequality (4.12) in Lemma 4.4, we have  $G(A) = g(R) \geq 0$ . Hence  $H(A) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda) + \lambda G(A) \geq (\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda)$ .

Now suppose that  $\lambda_1 \leq \lambda$ . Then the lower bound in (4.12) is  $2(\beta + \gamma) \leq 0$ , and so

$$\begin{aligned} H(A) &\geq (\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda) + 2\lambda(\lambda_1 - \lambda)(\lambda_2 + \lambda_3) \\ &= (\lambda_1 - \lambda)(\lambda_2 + \lambda)(\lambda_3 + \lambda). \end{aligned}$$

$\square$

The main result of this subsection is the following.

**Proposition 4.6.** *Suppose  $h'(\lambda^3) \geq 0$  and let  $W$  be given by (4.1). Then any admissible map  $u \neq u_\lambda$  is such that:*

$$I(u) - I(u_\lambda) \geq \int_{\Omega} \kappa |\nabla u - \nabla u_\lambda|^q + h'(\lambda^3) H(\nabla u) \, dx.$$

*In particular, if  $\int_{\Omega} H(\nabla u) \, dx \geq 0$  then  $I(u) > I(u_\lambda)$ , while if  $I(u) \leq I(u_\lambda)$  then*

$$(4.13) \quad \int_{\{x \in \Omega: \lambda_1(\nabla u(x)) \geq \lambda\}} h'(\lambda^3)(\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda) + \kappa |\nabla u - \nabla u_\lambda|^q \, dx \leq \int_{\{x \in \Omega: \lambda_1(\nabla u(x)) \leq \lambda\}} h'(\lambda^3)(\lambda - \lambda_1)(\lambda + \lambda_2)(\lambda + \lambda_3) \, dx.$$

*Proof.* This follows from (4.4) and Proposition 4.5.  $\square$

We remark that the results of Section 3 imply that inequality (4.13) ought *not* to be possible for  $\lambda$  such that  $\lambda^{3-q}h'(\lambda^3)$  is sufficiently small (see (3.15)). It is not immediately obvious from (4.13) why this should be so; nor is it clear why such a prominent role is played by the smallest singular value  $\lambda_1(\nabla u)$ . This surely warrants further investigation.

## 5. ACKNOWLEDGEMENT

The authors would like to express their thanks to two unnamed referees for their careful reading of the paper and for the very constructive comments they made.

## APPENDIX

Two different algorithms have been used to compute the quantity  $M_3(\lambda, \infty) := \lim_{c \rightarrow \infty} M_3(\lambda, c)$ , as defined in Subsection 3.1. We recall the notation

$$m_l(A, \lambda) = \frac{|G(A)|}{|A - \lambda \mathbf{1}|^l}$$

for  $l = 2, 3$ . The algorithms are also brought to bear on the problem of calculating  $M_3^-(\lambda, \infty)$ , and we summarise the results below.

**A.1. Algorithm A: conjugate gradient.** This is a ‘brute force’ approach, which consists of

- Choosing a value of  $\lambda \in [1, 2]$ ;
- Generating matrices  $A$  with all elements  $a_{ij}$  being uniformly distributed random numbers over the interval  $[-\alpha, \alpha]$ , for given  $\alpha$ ;
- Using the Polak-Ribiere variant of the Fletcher-Reeves algorithm [11], one of the so-called conjugate gradient methods for maximisation of smooth functions, starting from each of these matrices, in order to find a candidate matrix for  $A_l^*(\lambda)$ , which is an approximate maximiser of  $m_3(\cdot, \lambda)$ .

We choose  $\alpha = 5$ , the justification for which is as follows. Five thousand matrices were generated and those which gave the 15 largest values of  $m_3(\cdot, \lambda)$  were saved as the computation proceeded. Of these, none had an element whose modulus exceeded 3.1, hence reassuring us that the choice  $\alpha = 5$  is ‘safe’ for  $\lambda$  in the range  $[1, 2]$ .

Since the algorithm is iterative, a stopping condition is required, and this is that

$$(A.1) \quad |m^{(i+1)} - m^{(i)}| \leq \frac{\epsilon}{2} \left( |m^{(i+1)}| + |m^{(i)}| \right),$$

where  $m^{(i)}$  is the value of  $m_3(A_i, \lambda)$  at the  $i$ -th iteration and  $\epsilon = 10^{-9}$ .

Various data are saved as the computation progresses, including the current maximising matrix, which is the latest approximation to  $A_l^*(\lambda)$ . For half the simulations, the initial random matrix is symmetric, and for the other half it is not — we do not know, *a priori*, whether  $A_l^*(\lambda)$  will be symmetric or not. The numerics strongly indicated that  $A_3^*(\lambda)$  is indeed symmetric, at least for  $\lambda$  in the range  $[1, 2]$ .

All computations were carried out using 40 significant figures. Algorithm A leads the approximation  $M_3(\lambda, \infty) \approx \nu_1^A/\lambda$ , where  $\nu_1^A = 0.4501$  —see Table 1; the algorithm also produces the approximation to  $c_1(\lambda)$  shown in Fig. 2 and summarised in Table 2 below.

$\lambda$	1.01	1.1	1.2	1.3	1.4	1.5
$M_3(\lambda, \infty)$	0.44566175	0.40919852	0.37509864	0.34624489	0.32151312	0.30007890
$\nu_1^A/\lambda$	0.44566173	0.40919850	0.37509862	0.34624488	0.32151311	0.30007890
$\lambda$	1.6		1.7	1.8	1.9	2.0
$M_3(\lambda, \infty)$	0.28132398		0.26477551	0.25006575	0.23690440	0.22505900
$\nu_1^A/\lambda$	0.28132397		0.26477550	0.25006575	0.23690439	0.22505917

TABLE 1. Computed values of  $M_3(\lambda, \infty)$  compared with the approximation  $\nu_1^A/\lambda$ , for various values of  $\lambda$ . Least squares was used to find  $\nu_1^A$ . The largest absolute deviation,  $1.72 \times 10^{-7}$ , occurs at  $\lambda = 2$ .

$\lambda$	1.01	1.1	1.2	1.3	1.4	1.5
$ A_3^* - \lambda \mathbf{1} $	1.86212	2.02791	2.21231	2.39673	2.58098	2.76509
$c_1(\lambda)$	1.86240	2.02820	2.21242	2.39664	2.58087	2.76509
$\lambda$	1.6		1.7	1.8	1.9	2.0
$ A_3^* - \lambda \mathbf{1} $	2.94943		3.13421	3.31866	3.50261	3.68432
$c_1(\lambda)$	2.94931		3.13353	3.31775	3.50197	3.68619

TABLE 2. Computed values of  $|A_3^* - \lambda \mathbf{1}|$  compared with the approximation  $\nu_2 + \nu_3 \lambda$ , for various values of  $\lambda$ .

**A.2. Algorithm B: pointwise supremum.** This is based on a different idea, although a Monte Carlo approach it is still at its heart. We start by fixing an interval for  $\lambda$ ,  $\Lambda = [\lambda_-, \lambda_+]$ , which is not necessarily  $[1, 2]$  — the computation time is, at one level, independent of the interval. We then define  $N_p$  equally-spaced points in  $\Lambda$ , these points being  $\lambda_i = \lambda_- + i\delta\lambda$  with  $\delta\lambda = (\lambda_+ - \lambda_-)/N_p$  and  $i = 0, \dots, N_p$ .

Next, as before, a large number,  $N$ , of random matrices  $A_i$  are generated. As can be seen from its definition,

$$m_l(A_i, \lambda) = \frac{|a_1 + a_2\lambda|}{(b_1 + b_2\lambda + b_3\lambda^2)^{l/2}}$$

where the coefficients  $a_1, \dots, b_3$  are functions of the elements of  $A_i$  that we compute numerically. We define  $f_{l,i}(\lambda) := m_l(A_i, \lambda)$ , and clearly, once the coefficients have been computed,  $f_{l,i}(\lambda)$  can easily be found for any  $\lambda$ .

We then compute

$$F_{l,j} = \sup \{f_{l,i}(\lambda_j), i = 1, \dots, N\}$$

for  $j = 0, \dots, N_p$ ;  $F_{l,j}$  is then a discrete approximation to  $M_l(\lambda_j, \infty)$ . The convergence to  $M_3(\lambda, \infty)$  is quite slow, but nonetheless, choosing  $N$  large enough gives reasonable agreement with results produced by Algorithm A, thereby providing an independent check. Compare Fig. 1 with Fig. 2.

**A.3. Calculating  $M_3^-(\lambda, \infty)$ .** Using Algorithm A, the methodology is the same as for  $M_3(\lambda, \infty)$ , with the same number of random matrices generated, whose elements have the same bounds. The investigations lead us to conjecture that

$$M_3^-(\lambda, \infty) \approx \nu_1^{A,-}/\lambda,$$

FIGURE 2. Left: approximation from Algorithm B to  $M_3(\lambda, \infty)$  and, for comparison,  $\nu_1^B/\lambda$ , where  $\nu_1^B = 0.4492$ . Right: the difference between the two curves in the left-hand figure.

FIGURE 3. Left: The result of using Algorithm A to estimate  $M_3^-(\lambda, \infty)$  versus  $\lambda$ , for  $\lambda = 1.0, 1.1, \dots, 2.0$ , circles, where the data has been obtained from the maximum of several computations. The curve shows  $\nu_1^{A,-}/\lambda$  for  $\nu_1^{A,-} = 0.1923$ . Right: Algorithm B used to estimate  $M_3^-(\lambda, \infty)$  versus  $\lambda$ , (continuous line), compared with the model  $\nu_1^{B,-}/\lambda$  with  $\nu_1^{B,-} = 0.1925$  (dashed line).

where  $\nu_1^{A,-} \approx 0.1923$ . See Figure 3.

Recall that in the case of  $M_3(\lambda, \infty)$  it was possible to compute and then model the quantity  $c_1(\lambda)$  accurately on the interval  $1 \leq \lambda \leq 2$  using an affine function of  $\lambda$ . The same cannot be said of the corresponding quantity  $c_1^-(\lambda)$ , and indeed this seems to behave somewhat erratically as a function of  $\lambda$ . Thus the analysis leading up to Proposition 3.9 does not apply, and hence the caveat regarding the substitution of  $M_3^-(\lambda, \infty) \approx \nu_1'/\lambda$  promised in Remark 3.11.

#### REFERENCES

- [1] J. M. Ball: Convexity conditions and existence theorems in nonlinear elasticity. *Arch. Rat. Mech. Anal.*, **63**, no. 4 (1977), 337–403.

- [2] J. M. Ball. Discontinuous equilibrium solutions and cavitation in nonlinear elasticity. *Phil. Trans. R. Soc. Lond. A*, **306**, 557-611 (1982).
- [3] J. M. Ball, J. C. Currie, P. J. Olver. Null Lagrangians, weak continuity, and variational problems of arbitrary order. *J. Funct. Anal.*, **41** (1981), no.2, 135-174.
- [4] J. Bevan, C. Zeppieri. A simple sufficient condition for the quasiconvexity of elastic stored-energy functions in spaces which allow for cavitation. *Calculus of Variations and Partial Differential Equations*, **55** (2), 1-25, 2015.
- [5] Philippe G. Ciarlet. *Mathematical Elasticity Volume I: Three Dimensional Elasticity*, Elsevier, 2004.
- [6] B. Dacorogna: *Direct methods in the calculus of variations*. Second edition. Applied Mathematical Sciences, **78**. Springer, New York, 2008.
- [7] D. Henao, C. Mora-Corral: Invertibility and weak continuity of the determinant for the modelling of cavitation and fracture in nonlinear elasticity. *Arch. Rat. Mech. Anal.*, **197** (2010), 619–655.
- [8] D. Henao, C. Mora-Corral: Fracture Surfaces and the Regularity of Inverses for BV Deformations. *Arch. Rat. Mech. Anal.*, **201** (2011), 575–629.
- [9] S. Müller, S. Spector. An existence theory for nonlinear elasticity that allows for cavitation. *Arch. Rat. Mech. Anal.* **131**(1995), 1-66.
- [10] S. Müller, J. Sivaloganathan, S. Spector: An isoperimetric estimate and  $W^{1,p}$ -quasiconvexity in nonlinear elasticity. *Calc. Var. Partial Differential Equations*, **8**, no. 2 (1999), 159–176.
- [11] W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery, *Numerical Recipes in C*, ISBN 0-521-43108-5, Cambridge University Press, Cambridge, UK (1992).
- [12] J. Sivaloganathan. Uniqueness of regular and singular equilibria for spherically symmetric problems of nonlinear elasticity. *Arch. Rat. Mech. Anal.*, **96** (1986), 97-136.
- [13] J. Sivaloganathan, S. Spector. Energy minimising properties of the radial cavitation solution in incompressible nonlinear elasticity. *J. Elasticity*, **93** (2008), no. 2, 177-187.
- [14] J. Sivaloganathan, S. Spector. Necessary conditions for a minimum at a radial cavitating singularity in nonlinear elasticity. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, **25**, no. 1 (2008), 201-213.
- [15] C. A. Stuart. Radially symmetric cavitation for hyperelastic materials. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, **2**, no. 1 (1985), 33-66.