On double-covering stationary points of a constrained Dirichlet energy

Jonathan Bevan [*]

Department of Mathematics, University of Surrey, Guildford, Surrey, GU2 7XH, UK

Abstract

The double-covering map $\mathbf{u}_{\mathrm{dc}} : \mathbb{R}^2 \to \mathbb{R}^2$ is given by

$$\mathbf{u}_{\mathrm{dc}}(\mathbf{x}) = \frac{1}{\sqrt{2}|\mathbf{x}|} \begin{pmatrix} x_2^2 - x_1^2 \\ 2x_1 x_2 \end{pmatrix}$$

in cartesian coordinates. This paper examines the conjecture that $\mathbf{u}_{\mathrm{dc}}$ is the global minimizer of the Dirichlet energy $I(\mathbf{u}) = \int_B |\nabla \mathbf{u}|^2 \, d\mathbf{x}$ among all $W^{1,2}$ mappings $\mathbf{u}$ of the unit ball $B \subset \mathbb{R}^2$ satisfying (i) $\mathbf{u} = \mathbf{u}_{\mathrm{dc}}$ on ∂B, and (ii) $\det \nabla \mathbf{u} = 1$ almost everywhere. Let the class of such admissible maps be $\mathcal{A}$. The chief innovation is to express $I(\mathbf{u})$ in terms of an auxiliary functional $G(\mathbf{u} - \mathbf{u}_{\mathrm{dc}})$, using which we show that $\mathbf{u}_{\mathrm{dc}}$ is a stationary point of I in $\mathcal{A}$, and that $\mathbf{u}_{\mathrm{dc}}$ is a global minimizer of the Dirichlet energy among members of $\mathcal{A}$ whose Fourier decomposition can be controlled in a way made precise in the paper. By constructing variations about $\mathbf{u}_{\mathrm{dc}}$ in $\mathcal{A}$ using ODE techniques, we also show that $\mathbf{u}_{\mathrm{dc}}$ is a local minimizer among variations whose tangent $\boldsymbol{\psi}$ to $\mathcal{A}$ at $\mathbf{u}_{\mathrm{dc}}$ obeys $G(\boldsymbol{\psi}^{\mathrm{o}}) > 0$, where $\boldsymbol{\psi}^{\mathrm{o}}$ is the odd part of $\boldsymbol{\psi}$. In addition, a Lagrange multiplier corresponding to the constraint $\det \nabla \mathbf{u} = 1$ is identified by an analysis which exploits the well-known Fefferman–Stein duality.

Résumé

Le double-revêtement $\mathbf{u}_{\mathrm{dc}} : \mathbb{R}^2 \to \mathbb{R}^2$ est donné par

$$\mathbf{u}_{\mathrm{dc}}(\mathbf{x}) = \frac{1}{\sqrt{2}|\mathbf{x}|} \begin{pmatrix} x_2^2 - x_1^2 \\ 2x_1 x_2 \end{pmatrix}$$

en coordonnées cartésiennes. Cet article examine la conjecture selon laquelle $\mathbf{u}_{\mathrm{dc}}$ est le minimiseur global de l'énergie de Dirichlet $I(\mathbf{u}) = \int_B |\nabla \mathbf{u}|^2 \, d\mathbf{x}$ pour les fonctions satisfaisant (i) $\mathbf{u} \in W^{1,2}(B)$, où B est la boule unité de $\mathbb{R}^2$, (ii) $\mathbf{u} = \mathbf{u}_{\mathrm{dc}}$ sur ∂B, et (iii) $\det \nabla \mathbf{u} = 1$ presque partout. Soit $\mathcal{A}$ la classe admissible de telles fonctions. La principale innovation est ici d'exprimer $I(\mathbf{u})$ sous forme d'une fonction auxiliaire $G(\mathbf{u} - \mathbf{u}_{\mathrm{dc}})$, avec laquelle nous montrons que $\mathbf{u}_{\mathrm{dc}}$ est un point stationnaire de I en $\mathcal{A}$, et que $\mathbf{u}_{\mathrm{dc}}$ est un minimiseur global de l'énergie de Dirichlet parmi les membres de $\mathcal{A}$ dont la décomposition de Fourier peut être contrôlée d'une manière détaillée dans l'article. En construisant des variations autour de $\mathbf{u}_{\mathrm{dc}}$ en $\mathcal{A}$ par des techniques variationnelles, nous montrons également que $\mathbf{u}_{\mathrm{dc}}$ est un minimiseur local parmi les variations dont la tangente $\boldsymbol{\psi}$ de $\mathbf{u}_{\mathrm{dc}}$ vers $\mathcal{A}$ obéissent à $G(\boldsymbol{\psi}^{\mathrm{o}}) > 0$, où $\boldsymbol{\psi}^{\mathrm{o}}$ est la partie impaire de $\boldsymbol{\psi}$. Additionnellement, un multiplicateur

[*] Tel.: +44 (0) 1483 682620.

E-mail address: j.bevan@surrey.ac.uk.

de Lagrange correspondant à la contrainte $\det \nabla \mathbf{u} = 1$ est identifié par une analyse qui exploite la dualité de Fefferman–Stein.

1. Introduction

We study the problem of minimizing the Dirichlet integral

$$I(\mathbf{u}) = \int_B |\nabla \mathbf{u}|^2 \, d\mathbf{x}$$

among mappings belonging to the class

$$\mathcal{A} = \left\{ \mathbf{u} \in W^{1,2}(B, \mathbb{R}^2) \colon \mathbf{u}(\mathbf{x}) = \mathbf{u}_{\mathrm{dc}}(\mathbf{x}) \text{ if } \mathbf{x} \in \partial B, \ \det \nabla \mathbf{u} = 1 \text{ a.e.} \right\}. \tag{1.1}$$

Here, B represents the unit ball in $\mathbb{R}^2$, a.e. refers to two-dimensional Lebesgue measure, and $\mathbf{u}_{\mathrm{dc}}$ refers to the double-covering map given by

$$\mathbf{u}_{\mathrm{dc}}(\mathbf{x}) = \frac{1}{\sqrt{2}|\mathbf{x}|} \begin{pmatrix} x_2^2 - x_1^2 \\ 2x_1 x_2 \end{pmatrix}$$

in cartesian coordinates $\mathbf{x} = (x_1, x_2)^T$. The mapping $\mathbf{u}_{\mathrm{dc}}$ takes the unit ball B to the ball centred at zero and with radius $\frac{1}{\sqrt{2}}$, covering the image twice as it does so.

First of all, it is a consequence of the standard theory of harmonic functions that, without the constraint $\det \nabla \mathbf{u} = 1$ a.e., the Dirichlet functional I has a smooth minimizer $\mathbf{v}$, say, which, when the boundary condition $\mathbf{v}(\mathbf{x}) = \mathbf{u}_{\mathrm{dc}}(\mathbf{x})$ for $\mathbf{x} \in \partial B$ is applied, is

$$\mathbf{v}(\mathbf{x}) = \frac{1}{\sqrt{2}} \begin{pmatrix} x_2^2 - x_1^2 \\ 2x_1 x_2 \end{pmatrix}.$$

Note that this map is radially symmetric in the sense that $|\mathbf{v}(\mathbf{x})| = |\mathbf{v}(\mathbf{x}')|$ whenever $|\mathbf{x}| = |\mathbf{x}'|$. Based on this, it is natural to suppose that the global minimizer $\mathbf{u}$, say, of I in $\mathcal{A}$, should it exist, will have a similar radial symmetry. Thus it is reasonable to expect that

$$\mathbf{u}(\mathbf{x}) = \rho(|\mathbf{x}|) \begin{pmatrix} x_2^2 - x_1^2 \\ 2x_1 x_2 \end{pmatrix}$$

for a suitable scalar function ρ. But the constraint $\det \nabla \mathbf{u} = 1$ a.e., together with the boundary condition $\mathbf{u}(\mathbf{x}) = \mathbf{u}_{\mathrm{dc}}(\mathbf{x})$ for $\mathbf{x} \in \partial B$, then implies that $\rho(|\mathbf{x}|) = \frac{|\mathbf{x}|}{\sqrt{2}}$. Hence, $\mathbf{u} = \mathbf{u}_{\mathrm{dc}}$, so that the global minimizer of I in $\mathcal{A}$ ought to be $\mathbf{u}_{\mathrm{dc}}$ if our intuition regarding its symmetry is correct. To an extent this intuition is accurate: among the results of this paper is the assertion that $\mathbf{u}_{\mathrm{dc}}$ globally minimizes the Dirichlet energy among all variations whose Fourier decomposition is suitably controlled. The main advance is to write

$$I(\mathbf{u}) = I(\mathbf{u}_{\mathrm{dc}}) + G(\mathbf{u} - \mathbf{u}_{\mathrm{dc}}),$$

where the auxiliary functional G is defined by

$$G(\boldsymbol{\varphi}) = \int_B |\nabla \boldsymbol{\varphi}|^2 + 3 \ln R \det \nabla \boldsymbol{\varphi} \, d\mathbf{x},$$

and then apply ideas of [5] to G. See Section 3 for details, and Theorem 3.1 in particular.

A similar analysis leads to the conclusion that $\mathbf{u}_{\mathrm{dc}}$ is a stationary point of I in the class $\mathcal{A}$, a necessary condition for it to be a local or global minimizer, as well as to the inference that $3 \ln R$ acts as a Lagrange multiplier for the Dirichlet energy subject to the constraint $\det \nabla \mathbf{u} = 1$ a.e. See Theorem 4.2 and Proposition 3.4, respectively, for details.

We note that the symmetry intuition also stems from recent striking results [17,18], which apply to deformations of annular domains and with more regular (affine) boundary conditions than those considered in this note. In [17], the authors identify the global minimizers of a quite general class of polyconvex stored-energy functionals (which includes the Dirichlet energy—see below), subject to the pointwise constraint det $\nabla \mathbf{u} = 1$ a.e. They show in particular that the minimizer is radially symmetric.

Our other main motivation originates in incompressible nonlinear elasticity. In the planar version of that theory, the stored energy $E(\mathbf{u})$ of a neo-Hookean, rubber-like material subject to a deformation $\mathbf{u} : B \to \mathbb{R}^2$ is expressed as

$$E(\mathbf{u}) = \int_B \Phi(\nabla \mathbf{u}, \det \nabla \mathbf{u}) \, d\mathbf{x},$$

where $\Phi : \mathbb{R}^{2 \times 2} \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is a convex function such that $\Phi(A, \delta) = +\infty$ whenever $\delta \neq 1$. A typical choice for Φ might be

$$\Phi(A, \delta) = |A|^2 + h(\delta)$$

for an appropriate function $h : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ (see [1,2], for example). When restricted to the class $\mathcal{A}$ defined in (1.1), the functional $E(\mathbf{u})$ differs from the Dirichlet energy only by a constant. The interpretation of $I(\mathbf{u})$ on $\mathcal{A}$ as a stored-energy functional is, however, somewhat artificial: the boundary condition is aphysical in the sense that it is the restriction of an essentially two-to-one mapping, which one would not expect to be able to impose on a realistic material. Nevertheless, the effect of the volume constraint and double-covering boundary conditions on the symmetry, or otherwise, of the global minimizer of stored-energy functions remains an interesting open question. This question is studied in [4,5] in the compressible case, that is, where the pointwise a.e. constraint det $\nabla \mathbf{u} = 1$ is replaced by det $\nabla \mathbf{u} > 0$ a.e.

The symmetry hypothesis above is not only dependent on the structure of the boundary condition but also on the integrand and function space setting of the problem. For example, the more general polyconvex stored-energies considered in [19,20], which model a rectangular bar under uniaxial tension subject to a volume-preserving constraint, are indeed minimized by the most symmetric deformation, provided, that is, one assumes competing functions are C^1. Without the latter assumption, by enlarging the function space to a subspace of *SBV*, and by including a surface energy term in the functional, [14] shows that the minimizer need not be the symmetric, homogeneous deformation identified as the minimizer in [19].

Remarkable examples of asymmetric minimizers are also given in [15]; in this case, and in common with ours, the functional minimized is polyconvex and the boundary mapping $\mathbf{u}_0$ in the planar case is not one–one. However, in the notation of [15], the Jacobian of each minimizer $\mathbf{u}_{\min}$, $\mathbf{v}_{\min}$ changes sign non-trivially, reflecting the change in topology associated with the boundary mapping $\mathbf{u}_0$. See also [16] for uniqueness results in the case that the boundary mapping is the identity.

The structure of the paper is as follows. After some preliminary details of notation and basic facts concerning the Dirichlet energy on the class $\mathcal{A}$, we identify and study in Section 3 the auxiliary functional G. There it is shown in Theorem 3.1 that $\mathbf{u}_{\mathrm{dc}}$ is a global minimizer of I among those maps in $\mathcal{A}$ such that $G((\mathbf{u} - \mathbf{u}_{\mathrm{dc}})^{(1)}) \geqslant 0$, where in general $\boldsymbol{\varphi}^{(1)}$ denotes the Fourier one-mode of the mapping $\boldsymbol{\varphi}$. We also establish that G is expressed in terms of a Lagrange multiplier related to the problem of minimizing I in the constrained class $\mathcal{A}$, and we give an example of a mapping $\boldsymbol{\phi}$ for which $G(\boldsymbol{\phi}) < 0$, showing that the positivity of G among all maps cannot be taken for granted. In Section 4.1, we show that $\mathbf{u}_{\mathrm{dc}}$ is a stationary point of I in $\mathcal{A}$, and we use ODE techniques to construct variations about $\mathbf{u}_{\mathrm{dc}}$ in $\mathcal{A}$ as flows. The latter is exploited in the remainder of the paper to prove that variations whose tangents to $\mathcal{A}$ at $\mathbf{u}_{\mathrm{dc}}$ satisfy $G(\boldsymbol{\psi}^o) > 0$ cannot lower the energy. Various open questions are posed in Section 5.

2. Notation and preliminaries

We denote the $m \times n$ real matrices by $\mathbb{R}^{m \times n}$, and unless stated otherwise we sum over repeated indices. We denote the identity matrix by $\mathbf{1}$, and throughout B is the unit ball in $\mathbb{R}^2$. A function $f : \mathbb{R}^{2 \times 2} \to \mathbb{R} \cup \{\infty\}$ is said to be polyconvex if there exists a convex function $\phi : \mathbb{R}^{2 \times 2} \times \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ such that

$$f(A) = \phi(A, \det A) \tag{2.1}$$

for all 2×2 real matrices A.

Other standard notation includes $\| \cdot \|_{k,p;\Omega}$ for the norm on the Sobolev space $W^{k,p}(\Omega)$, $\| \cdot \|_{p;\Omega}$ for the norm on $L^p(\Omega)$, and $\rightharpoonup$, $\overset{*}{\rightharpoonup}$ to represent weak and weak* convergence respectively in both of these spaces. Here, Ω is a domain in $\mathbb{R}^n$. As usual, we denote by $B(a, R)$ the ball in $\mathbb{R}^n$ centred at a with radius R. $\mathcal{H}^1(\Omega)$ represents the Hardy space dual to $BMO(\Omega)$, the space of functions of Bounded Mean Oscillation (see [11,6]). We use the notation $\mathrm{lip}_0(B, \mathbb{R}^2)$ for the space of Lipschitz continuous functions with compact support in the ball B. The odd and even parts of a function $\boldsymbol{\psi}$ are denoted respectively by $\boldsymbol{\psi}^{\mathrm{o}}(\mathbf{z}) := \frac{1}{2}(\boldsymbol{\psi}(\mathbf{z}) - \boldsymbol{\psi}(-\mathbf{z}))$ and $\boldsymbol{\psi}^{\mathrm{e}} := \frac{1}{2}(\boldsymbol{\psi}(\mathbf{z}) + \boldsymbol{\psi}(-\mathbf{z}))$; the decomposition $\boldsymbol{\psi}(\mathbf{z}) = \boldsymbol{\psi}^{\mathrm{o}}(\mathbf{z}) + \boldsymbol{\psi}^{\mathrm{e}}(\mathbf{z})$ is immediate.

The tensor product of two vectors $\mathbf{a} \in \mathbb{R}^m$ and $\mathbf{b} \in \mathbb{R}^n$ is written $\mathbf{a} \otimes \mathbf{b}$; it is the $m \times n$ matrix whose (i, j) entry is $a_i b_j$. The inner product of two matrices $X, Y \in \mathbb{R}^{m \times n}$ is $X \cdot Y = \mathrm{tr}(X^T Y)$. This obviously holds for vectors too. In plane polar coordinates (R, θ) the gradient of $\boldsymbol{\varphi} : B \to \mathbb{R}^2$ is

$$\nabla\boldsymbol{\varphi} = \boldsymbol{\varphi}_{,R} \otimes \mathbf{e}_R(\theta) + \frac{1}{R}\boldsymbol{\varphi}_{,\theta} \otimes \mathbf{e}_\theta(\theta),$$

where $\mathbf{e}_R(\theta) = (\cos\theta, \sin\theta)^T$ and $\mathbf{e}_\theta(\theta) = (-\sin\theta, \cos\theta)^T$. Throughout the paper we write $\boldsymbol{\varphi}_{,R} = \partial_R\boldsymbol{\varphi}$ and $\boldsymbol{\varphi}_{,\theta} = \partial_\theta\boldsymbol{\varphi}$. In this notation the formula

$$\det\nabla\boldsymbol{\varphi} = \frac{1}{R}J\boldsymbol{\varphi}_{,R} \cdot \boldsymbol{\varphi}_{,\theta}$$

holds, where J is the 2×2 matrix corresponding to a rotation of $\frac{\pi}{2}$ radians in the plane, i.e.,

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

The two most useful properties of J are that (i) $J^T = -J$, so that in particular $a \cdot Jb = -Ja \cdot b$ for any two $a, b \in \mathbb{R}^2$, and (ii) $\mathrm{cof}\, A = J^T A J$ for any 2×2 matrix A.

In the notation introduced above,

$$\mathbf{u}_{\mathrm{dc}}(\mathbf{x}) = \frac{R}{\sqrt{2}}\mathbf{e}_R(2\theta)$$

and

$$\nabla\mathbf{u}_{\mathrm{dc}}(\mathbf{x}) = \frac{1}{\sqrt{2}}\mathbf{e}_R(2\theta) \otimes \mathbf{e}_R(\theta) + \sqrt{2}\mathbf{e}_\theta(2\theta) \otimes \mathbf{e}_\theta(\theta).$$

In particular,

$$\det\nabla\mathbf{u}_{\mathrm{dc}} = 1$$

except at the point $\mathbf{x} = 0$.

2.1. The class $\mathcal{A}$ of admissible functions

As a first step, we note that $\mathcal{A}$ defined in (1.1) contains not only the map $\mathbf{u}_{\mathrm{dc}}$ itself, but also many more so-called twist maps based on $\mathbf{u}_{\mathrm{dc}}$. Let $g \in W^{1,2}((0, 1); \mathbb{R})$ satisfy $g(1) = 2k\pi$ for some integer k, and define

$$\mathbf{u}_g(\mathbf{x}) := \frac{R}{\sqrt{2}}\mathbf{e}_R(2\theta + g(R)),$$

where $R = |\mathbf{x}|$. It is then straightforward to check that $\mathbf{u}_g$ lies in the class $\mathcal{A}$ for each g as described above. In general, if $\boldsymbol{\Phi} : B \to B$ is a diffeomorphism with $\det\nabla\boldsymbol{\Phi} = 1$ a.e. and $\boldsymbol{\Phi}(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in \partial B$, then $\boldsymbol{\varphi} = \mathbf{u}_{\mathrm{dc}} \circ \boldsymbol{\Phi}$ lies in $\mathcal{A}$.

The existence of a minimizer of the Dirichlet energy I in the class $\mathcal{A}$ is now achieved using the direct method of the Calculus of Variations. The following result is included for completeness only.

Proposition 2.1. *There exists a minimizer of the Dirichlet energy in the class $\mathcal{A}$ of admissible functions.*

Proof. The class $\mathcal{A}$ contains $\mathbf{u}_{\mathrm{dc}}$ and the functional I is bounded below, so it follows that $\alpha := \inf_{\mathcal{A}} I$ exists. Let $(\mathbf{u}^{(j)})$ be a sequence in $\mathcal{A}$ such that $I(\mathbf{u}^{(j)}) \to \alpha$. Passing to a subsequence, and without relabelling, we see that $\mathbf{u}^{(j)} \rightharpoonup \mathbf{u}$ in $W^{1,2}(B, \mathbb{R}^2)$. The only potential difficulty lies in proving that $\mathbf{u}$ belongs to $\mathcal{A}$. But, by classical results (see e.g. [1] or [7, Theorem 8.20, part (i)]), $\det \nabla \mathbf{u}^{(j)} \rightharpoonup \det \nabla \mathbf{u}$ in $\mathcal{D}'(B)$, so that in particular $\det \nabla \mathbf{u} = 1$ a.e. The trace theorems for Sobolev functions further imply $\mathbf{u} = \mathbf{u}_{\mathrm{dc}}$ on ∂B. Hence $\mathbf{u} \in \mathcal{A}$, and it follows from the convexity of the integrand of the Dirichlet energy that $\mathbf{u}$ minimizes I in $\mathcal{A}$. $\quad\square$

The regularity of a general member of $\mathcal{A}$ is controlled by the boundary condition and the constraint. Ball remarks in [3] that no member of $\mathcal{A}$ is C^1; see [5] for a proof of this fact. In the absence of the doubling boundary condition, Evans and Gariepy examine in [10] the effect of the constraint $\det \nabla \mathbf{v} = 1$ a.e. on the possible regularity of planar Lipschitz maps $\mathbf{v} = (v^1, v^2)$ satisfying the additional nondegeneracy condition

$$\mathbf{v}_{,2}^1(\mathbf{x}) \geqslant c \quad \text{a.e. } \mathbf{x} \in B', \tag{2.2}$$

where $B' \Subset B$ and c is a positive constant. They show that if $\mathbf{v}$ minimizes a suitably quasiconvex energy, which includes the case of the Dirichlet energy, then $\nabla \mathbf{v}$ must be Hölder continuous on a dense subset of B'. Interesting results in this vein, although under different assumptions involving the dual pressure, have recently been obtained by Karakhanyan in [13]. In the degree two case considered in this paper, the regularity of a general stationary point of the energy would still seem to be open.

3. A global result for the double-covering map

In this section we study the Dirichlet energy $I(\mathbf{u})$ and make use of the constraint $\det \nabla \mathbf{u} = 1$ a.e. The analysis gives rise to an auxiliary functional $G(\mathbf{u} - \mathbf{u}_{\mathrm{dc}})$ which, via a Fourier decomposition and other arguments, can be used to control the sign of $I(\mathbf{u}) - I(\mathbf{u}_{\mathrm{dc}})$.

For any function $\mathbf{u}$ belonging to $W^{1,2}(B, \mathbb{R}^2)$, we define its (discrete) Fourier decomposition in plane polar coordinates by

$$\mathbf{u}(\mathbf{x}) = \sum_{j \geqslant 0} \mathbf{u}^{(j)}(\mathbf{x}), \tag{3.1}$$

where $\mathbf{u}^{(0)} = \frac{1}{2}\mathbf{A}_0(R)$ and $\mathbf{u}^{(j)} = \mathbf{A}_j(R) \cos j\theta + \mathbf{B}_j(R) \sin j\theta$ for all $j \geqslant 1$. In terms of such a decomposition, the main result of this section is the following.

Theorem 3.1. *Let $\mathbf{u} \in \mathcal{A}$ and define $\boldsymbol{\varphi} = \mathbf{u} - \mathbf{u}_{\mathrm{dc}}$. Let the functional $G : \mathcal{A} - \mathbf{u}_{\mathrm{dc}} \to \mathbb{R} \cup \{\pm\infty\}$ be given by*

$$G(\boldsymbol{\varphi}) = \int_B |\nabla \boldsymbol{\varphi}|^2 + 3 \ln R \det \nabla \boldsymbol{\varphi} \, d\mathbf{x}. \tag{3.2}$$

Then

(i) $I(\mathbf{u}) = I(\mathbf{u}_{\mathrm{dc}}) + G(\boldsymbol{\varphi})$;
(ii) *if $G(\boldsymbol{\varphi}^{(1)}) \geqslant 0$ then $G(\boldsymbol{\varphi}) \geqslant 0$.*

In particular, $\mathbf{u}_{\mathrm{dc}}$ is a global minimizer of I in the class

$$\mathcal{A}' = \big\{ \mathbf{u} \in \mathcal{A} \colon G\big((\mathbf{u} - \mathbf{u}_{\mathrm{dc}})^{(1)}\big) \geqslant 0 \big\}.$$

The theorem is proved in several steps, the first of which explains the appearance of $\ln R$ in the functional G. Later, in Proposition 3.4, we note that $\ln R$ serves as a Lagrange multiplier for this minimization problem. We also note that the logarithm plays a similarly important role in the analysis of the elastic stored-energy functionals considered in [5].

Lemma 3.1. *Let* $\boldsymbol{\varphi} \in \mathcal{A} - \mathcal{A}$. *Then*

(a)

$$\int_B \nabla \mathbf{u}_{\mathrm{dc}} \cdot \nabla \boldsymbol{\varphi} \, d\mathbf{x} = \frac{3}{\sqrt{2}} \int_B \boldsymbol{\varphi} \cdot \mathbf{e}_R(2\theta) \, dR \, d\theta. \tag{3.3}$$

(b) *Let* $\lambda \in W^{1,1}((0,1), \mathbb{R})$ *be such that* $R\lambda'(R)$ *is essentially bounded on the interval* $(0,1)$. *Then*

$$\int_B \lambda(R) \operatorname{cof} \nabla \mathbf{u}_{\mathrm{dc}} \cdot \nabla \boldsymbol{\varphi} \, d\mathbf{x} = - \int_B \sqrt{2} R\lambda'(R) \boldsymbol{\varphi} \cdot \mathbf{e}_R(2\theta) \, dR \, d\theta. \tag{3.4}$$

Proof. (a) The proof of (3.3) for maps $\boldsymbol{\varphi}$ in the class $C_c^1(B)$ is a straightforward integration by parts based on

$$\int_B \nabla \mathbf{u}_{\mathrm{dc}} \cdot \nabla \boldsymbol{\varphi} \, d\mathbf{x} = \int_0^{2\pi} \int_0^1 \left[\frac{1}{\sqrt{2}} R\boldsymbol{\varphi}_{,R} \cdot \mathbf{e}_R(2\theta) + \sqrt{2}\boldsymbol{\varphi}_{,\theta} \cdot \mathbf{e}_\theta(2\theta) \right] dR \, d\theta.$$

The identity holds for general $\boldsymbol{\varphi} \in \mathcal{A} - \mathcal{A}$ by an approximation argument using, among other well-known results, the fact that all maps in $\mathcal{A}$ are continuous. (See [21], or [12, Theorem 5.17].)

(b) The argument leading to (3.4) is similar. The only additional ingredient is the observation that the mapping $\mathbf{u}_{\mathrm{dc}}$ satisfies Piola's identity at all points except the origin, that is,

$$\operatorname{div} \operatorname{cof} \nabla \mathbf{u}_{\mathrm{dc}}(\mathbf{x}) = 0$$

for all non-zero $\mathbf{x} \in B$. Therefore, if $\boldsymbol{\varphi} \in C_c^1(B)$,

$$\int_B \lambda(R) \operatorname{cof} \nabla \mathbf{u}_{\mathrm{dc}} \cdot \nabla \boldsymbol{\varphi} \, d\mathbf{x} = - \int_B \boldsymbol{\varphi} \cdot \operatorname{cof} \nabla \mathbf{u}_{\mathrm{dc}} \nabla \lambda \, d\mathbf{x}.$$

Using

$$\operatorname{cof} \nabla \mathbf{u}_{\mathrm{dc}} = \sqrt{2}\mathbf{e}_R(2\theta) \otimes \mathbf{e}_R(\theta) + \frac{1}{\sqrt{2}}\mathbf{e}_\theta(2\theta) \otimes \mathbf{e}_\theta(\theta),$$

we see that

$$\operatorname{cof} \nabla \mathbf{u}_{\mathrm{dc}} \nabla \lambda = \sqrt{2}\lambda'(R)\mathbf{e}_R(2\theta),$$

from which (3.4) follows for $C_c^1(B)$ functions. Approximating a general $\boldsymbol{\varphi} \in \mathcal{A} - \mathcal{A}$ by a sequence of sufficiently smooth functions, as above, it follows that we can pass to the limit $k \to \infty$ in the equations

$$\int_B \lambda(R) \operatorname{cof} \nabla \mathbf{u}_{\mathrm{dc}} \cdot \nabla \boldsymbol{\varphi}^{(k)} \, d\mathbf{x} = - \int_B \sqrt{2} R\lambda'(R)\boldsymbol{\varphi}^{(k)} \cdot \mathbf{e}_R(2\theta) \, dR \, d\theta$$

provided $R\lambda'(R)$ is essentially bounded. $\square$

Now let $\mathbf{u} \in \mathcal{A}$. Note that because the constraint $\det \nabla \mathbf{u} = 1$ a.e. applies to both $\mathbf{u}$ and $\mathbf{u}_{\mathrm{dc}}$, it follows that

$$\det \nabla(\mathbf{u} - \mathbf{u}_{\mathrm{dc}}) + \operatorname{cof} \nabla \mathbf{u}_{\mathrm{dc}} \cdot \nabla(\mathbf{u} - \mathbf{u}_{\mathrm{dc}}) = 0 \tag{3.5}$$

almost everywhere in B. This simple observation is used to prove part (i) of Theorem 3.1 as follows:

Proposition 3.1. *Let* $\mathbf{u} \in \mathcal{A}$ *and define* $\boldsymbol{\varphi} = \mathbf{u} - \mathbf{u}_{\mathrm{dc}}$. *Then*

$$I(\mathbf{u}) = I(\mathbf{u}_{\mathrm{dc}}) + G(\boldsymbol{\varphi}).$$

Proof. It is straightforward to check that

$$I(\mathbf{u}) = I(\mathbf{u}_{\mathrm{dc}}) + \int_B |\nabla\boldsymbol{\varphi}|^2 \, d\mathbf{x} + 3\sqrt{2} \int_B \boldsymbol{\varphi} \cdot \mathbf{e}_R(2\theta) \, dR \, d\theta,$$

where (3.3) has been used to rewrite the term $\int_B \nabla\mathbf{u}_{\mathrm{dc}} \cdot \nabla\boldsymbol{\varphi} \, d\mathbf{x}$. By adding and subtracting a term $2\int_B \lambda(R) \operatorname{cof} \nabla\mathbf{u}_{\mathrm{dc}} \cdot \nabla\boldsymbol{\varphi}$, where λ satisfies the hypotheses of Lemma 3.1(b), and using (3.4) above, we see that

$$\begin{aligned}
I(\mathbf{u}) = I(\mathbf{u}_{\mathrm{dc}}) + I(\boldsymbol{\varphi}) + \sqrt{2} \int_B \big(3 - 2R\lambda'(R)\big)\boldsymbol{\varphi} \cdot \mathbf{e}_R(2\theta) \, dR \, d\theta \\
- \int_B 2\lambda(R) \operatorname{cof} \nabla\mathbf{u}_{\mathrm{dc}} \cdot \nabla\boldsymbol{\varphi} \, d\mathbf{x}.
\end{aligned} \tag{3.6}$$

The term involving $\boldsymbol{\varphi} \cdot \mathbf{e}_R(2\theta)$ vanishes provided $R\lambda'(R) = \frac{3}{2}$, to which the solution is clearly $\lambda = \frac{3}{2}\ln R + c$. Note that the hypotheses of Lemma 3.1(b) are indeed satisfied. Also, if λ_1 and λ_2 are two possible choices for λ then they differ by a constant, and hence

$$\int_B (\lambda_1 - \lambda_2) \operatorname{cof} \nabla\mathbf{u}_{\mathrm{dc}} \cdot \nabla\boldsymbol{\varphi} \, d\mathbf{x} = 0$$

by Piola's identity. Therefore the constant c can without loss of generality be set to zero. Finally, integrating (3.5) gives

$$\int_B 2\lambda(R) \det \nabla\boldsymbol{\varphi} \, d\mathbf{x} = -\int_B 2\lambda(R) \operatorname{cof} \nabla\mathbf{u}_{\mathrm{dc}} \cdot \nabla\boldsymbol{\varphi},$$

so that

$$I(\mathbf{u}) = I(\mathbf{u}_{\mathrm{dc}}) + \int_B |\nabla\boldsymbol{\varphi}|^2 + 3\ln R \det \nabla\boldsymbol{\varphi} \, d\mathbf{x}.$$

The right-hand side of this equation is $I(\mathbf{u}_{\mathrm{dc}}) + G(\boldsymbol{\varphi})$, which concludes the proof of the proposition. □

The technical condition on λ in Lemma 3.1(b) is used to approximate and manipulate the functional $\int_B \lambda(R) \det \nabla\boldsymbol{\varphi} \, d\mathbf{x}$. The condition and application can be summarised as follows:

Proposition 3.2. *Let $\lambda \in W^{1,1}((0,1), \mathbb{R})$ be such that $R\lambda'(R)$ is essentially bounded on the interval $(0,1)$. Then*

(i) $w(\mathbf{x}) := \lambda(R)$ *is of bounded mean oscillation, that is, $w \in BMO(B)$;*
(ii) *if $\boldsymbol{\varphi}_n \to \boldsymbol{\varphi}$ in $W^{1,2}(B, \mathbb{R}^2)$ then*

$$\int_B \lambda(R) \det \nabla\boldsymbol{\varphi}_n \, d\mathbf{x} \to \int_B \lambda(R) \det \nabla\boldsymbol{\varphi} \, d\mathbf{x} \tag{3.7}$$

as $n \to \infty$.

Proof. (i) Let

$$w_{\mathbf{x},t} = \fint_{B(\mathbf{x},t)} w(y) \, dy.$$

By a version of Poincaré's inequality (see [9, Section 4.5.2, Theorem 2]),

$$\left(\fint_{B(\mathbf{x},t)} |w - w_{\mathbf{x},t}|^2 \, dy \right)^{\frac{1}{2}} \leqslant C_1 t \fint_{B(\mathbf{x},t)} |\nabla w| \, dy.$$

Therefore,

$$\fint_{B(\mathbf{x},t)} |w - w_{\mathbf{x},t}| \, dy \leqslant C_1 t \fint_{B(\mathbf{x},t)} |\lambda'(R)| \, dy.$$

It can be checked that $\int_{B(\mathbf{x},t)} |\lambda'(R)| R \, dR \, d\theta \leqslant C_2 t$, so that

$$\sup_{\mathbf{x} \in B} \fint_{B(\mathbf{x},t)} |w - w_{\mathbf{x},t}| \, dy \leqslant C_1 C_2.$$

Hence w is a function of bounded mean oscillation.

(ii) Here we use the div-curl lemma of [6] to infer $\det \nabla \boldsymbol{\varphi}_n \to \det \nabla \boldsymbol{\varphi}$ in $\mathcal{H}^1$ from the assumed convergence $\boldsymbol{\varphi}_n \to \boldsymbol{\varphi}$ in $W^{1,2}$. Indeed, by using the identity

$$2(\det \nabla \boldsymbol{\varphi}_n - \det \nabla \boldsymbol{\varphi}) = \operatorname{cof} \nabla \boldsymbol{\varphi}_n \cdot \nabla \Phi_n + \nabla \boldsymbol{\varphi} \cdot \operatorname{cof} \nabla \Phi_n,$$

where $\Phi_n = \boldsymbol{\varphi}_n - \boldsymbol{\varphi}$, we see that $2(\det \nabla \boldsymbol{\varphi}_n - \det \nabla \boldsymbol{\varphi})$ is a sum of terms $a_n \cdot b_n$ with $\operatorname{div} a_n = 0$, $\operatorname{curl} b_n = 0$ and at least one of a_n, b_n converging strongly to zero in L^2. By the argument in [6, Lemma II.1],

$$\left\| \sup_{t>0} |h_t * (a_n \cdot b_n)| \right\|_1 \leqslant C \|a_n\|_2 \|b_n\|_2$$

for some constant $C > 0$ independent of t and n, where $h_t(y) = \frac{1}{t^2} h(\frac{y}{t})$ and $h : C_c^\infty(B, \mathbb{R}) \to \mathbb{R}$ is a standard mollifier. In particular, the latter inequality implies $a_n \cdot b_n \to 0$ strongly in $\mathcal{H}^1$. By part (i) above and the Fefferman–Stein [11] duality $(\mathcal{H}^1)^* = BMO$, the convergence (3.7) follows. $\quad\square$

The next two results are concerned with the analysis of the functional G using the Fourier decomposition (3.1). Some basic facts are presented in Lemma 3.2 below, using which we show in Proposition 3.3 that $G(\boldsymbol{\varphi})$ is nonnegative provided $\boldsymbol{\varphi}$ consists of Fourier two-modes or higher, i.e. when $\boldsymbol{\varphi}^{(0)} = \boldsymbol{\varphi}^{(1)} = 0$.

Lemma 3.2. *Let* $\mathbf{u} \in \mathcal{A}$ *and define* $\boldsymbol{\varphi} = \mathbf{u} - \mathbf{u}_{\mathrm{dc}}$. *Let* $\lambda \in W^{1,1}((0,1), \mathbb{R})$ *be such that* $R\lambda'(R)$ *is essentially bounded on the interval* $(0,1)$. *Then*:

(i) $\int_B |\nabla \boldsymbol{\varphi}|^2 \, d\mathbf{x} = \sum_{j \geqslant 0} \int_B |\nabla \boldsymbol{\varphi}^{(j)}|^2 \, d\mathbf{x}$;
(ii) $\det \nabla \boldsymbol{\varphi}^{(0)} = 0$;
(iii) $\int_B \lambda(R) \det \nabla \boldsymbol{\varphi} \, d\mathbf{x} = \sum_{j \geqslant 1} \int_B \lambda(R) \det \nabla \boldsymbol{\varphi}^{(j)} \, d\mathbf{x}$;
(iv) $\int_B \lambda(R) \det \nabla \boldsymbol{\varphi} \, d\mathbf{x} = \frac{1}{2} \int_B \lambda'(R) \boldsymbol{\varphi} \cdot J\boldsymbol{\varphi}_{,\theta} \, dR \, d\theta$.

Proof. Part (i) is standard and so its proof is omitted.

(ii) Note that $\boldsymbol{\varphi}^{(0)}$ is a function of R only, so that $\nabla \boldsymbol{\varphi}^{(0)} = \boldsymbol{\varphi}^{(0)}_{,R} \otimes \mathbf{e}_R(\theta)$. Part (ii) follows.

(iii) We exploit the orthogonality of the Fourier modes with respect to integration in θ. Write

$$\det \nabla \boldsymbol{\varphi} = \det \nabla \boldsymbol{\varphi}^{(0)} + \det \nabla (\boldsymbol{\varphi} - \boldsymbol{\varphi}^{(0)}) + \operatorname{cof}(\nabla \boldsymbol{\varphi} - \nabla \boldsymbol{\varphi}^{(0)}) \cdot \nabla \boldsymbol{\varphi}^{(0)},$$

multiply both sides by $\lambda(R)$ and integrate with respect to θ. Since $\nabla \boldsymbol{\varphi} - \nabla \boldsymbol{\varphi}^{(0)}$ contains no zero-modes while $\nabla \boldsymbol{\varphi}^{(0)}$ consists only of zero-modes, it follows that

$$\int_0^{2\pi} \lambda(R) \operatorname{cof} \nabla (\boldsymbol{\varphi} - \boldsymbol{\varphi}^{(0)}) \cdot \nabla \boldsymbol{\varphi}^{(0)} \, d\theta = 0.$$

Hence

$$\int_B \lambda(R) \det \nabla \boldsymbol{\varphi} \, d\mathbf{x} = \int_B \lambda(R) \det \nabla \boldsymbol{\varphi}^{(0)} + \int_B \lambda(R) \det \nabla (\boldsymbol{\varphi} - \boldsymbol{\varphi}^{(0)}) \, d\mathbf{x}.$$

Proceeding inductively, part (iii) follows. We have implicitly used part (ii) here to begin the summation from $j = 1$ rather than $j = 0$.

(iv) Part (iv) uses the expression $\det \nabla \boldsymbol{\varphi} = \frac{1}{R} J \boldsymbol{\varphi}_{,R} \cdot \boldsymbol{\varphi}_{,\theta}$ established in the introduction. Indeed, for smooth $\boldsymbol{\varphi}$

$$
\begin{aligned}
\int_B \lambda(R) \det \nabla \boldsymbol{\varphi} \, d\mathbf{x} &= \int_B \lambda(R) J \boldsymbol{\varphi}_{,R} \cdot \boldsymbol{\varphi}_{,\theta} \, dR \, d\theta \\
&= -\int_B \boldsymbol{\varphi} \cdot \lambda(R) J \boldsymbol{\varphi}_{,\theta R} \, d\theta \, dR \\
&= \int_B \left(\lambda(R) \boldsymbol{\varphi} \right)_{,R} \cdot J \boldsymbol{\varphi}_{,\theta} \\
&= \int_B \lambda'(R) \boldsymbol{\varphi} \cdot J \boldsymbol{\varphi}_{,\theta} - \lambda(R) J \boldsymbol{\varphi}_{,R} \cdot \boldsymbol{\varphi}_{,\theta} \, dR \, d\theta.
\end{aligned}
$$

Note that the rightmost term of the line above is $-\int_B \lambda(R) \det \nabla \boldsymbol{\varphi} \, d\mathbf{x}$. Rearranging gives (iv) in the case that $\boldsymbol{\varphi}$ is smooth. By Proposition 3.2, the left-hand side of the expression in (iv) can be approximated by smooth $\boldsymbol{\varphi}$ provided λ satisfies the hypotheses in the statement of the lemma.

It remains to check that

$$
\int_B \lambda'(R) \boldsymbol{\varphi}_n \cdot J \boldsymbol{\varphi}_{n,\theta} \, d\mathbf{x} \to \int_B \lambda'(R) \boldsymbol{\varphi} \cdot J \boldsymbol{\varphi}_{,\theta} \, d\mathbf{x}
$$

for any sequence $\boldsymbol{\varphi}_n \to \boldsymbol{\varphi}$ in $W^{1,2}$. By parts (ii) and (iii) above, it suffices to show this for functions $\boldsymbol{\varphi}$ in $W^{1,2}$ for which $\boldsymbol{\varphi}^{(0)} = 0$ a.e. Indeed, for any $\boldsymbol{\varphi} \in W^{1,2}$ we let

$$
P(\boldsymbol{\varphi}) = \int_B \lambda'(R) \boldsymbol{\varphi} \cdot \frac{J \boldsymbol{\varphi}_{,\theta}}{R} \, d\mathbf{x}
$$

and note that

$$
P(\boldsymbol{\varphi}) = P(\boldsymbol{\varphi} - \boldsymbol{\varphi}^{(0)}).
$$

(For details see parts of the calculation leading to part (iii) above which exploit the orthogonality of the various Fourier modes with respect to integration in the polar angle θ.) We also note that if $\boldsymbol{\varphi}_n \to \boldsymbol{\varphi}$ in $W^{1,2}$ then in particular $\boldsymbol{\varphi}_n - \boldsymbol{\varphi}_n^{(0)} \to \boldsymbol{\varphi} - \boldsymbol{\varphi}^{(0)}$ in that space. Therefore we may assume without loss of generality that $\boldsymbol{\varphi}^{(0)} = 0$ and that $\boldsymbol{\varphi}_n \to \boldsymbol{\varphi}$ in $W^{1,2}$ norm for a sequence $\boldsymbol{\varphi}_n$ with $\boldsymbol{\varphi}_n^{(0)} = 0$. Write

$$
P(\boldsymbol{\varphi}_n) - P(\boldsymbol{\varphi}) = -\int_B \lambda'(R) J (\boldsymbol{\varphi} + \boldsymbol{\varphi}_n) \cdot \left(\frac{\boldsymbol{\varphi}_{n,\theta} - \boldsymbol{\varphi}_{,\theta}}{R} \right) d\mathbf{x}.
$$

Using the hypothesis that $R\lambda'(R)$ is essentially bounded, it follows from the estimate

$$
\int_B \frac{|\boldsymbol{\varphi}|^2}{R^2} \, d\mathbf{x} \leqslant \int_B |\nabla \boldsymbol{\varphi}|^2 \, d\mathbf{x},
$$

which applies to functions such that $\boldsymbol{\varphi}^{(0)} = 0$ a.e., that

$$
\left| P(\boldsymbol{\varphi}_n) - P(\boldsymbol{\varphi}) \right| \leqslant C \left\| \nabla(\boldsymbol{\varphi} + \boldsymbol{\varphi}_n) \right\|_2 \left\| \nabla(\boldsymbol{\varphi} - \boldsymbol{\varphi}_n) \right\|_2.
$$

Hence $P(\boldsymbol{\varphi}_n) \to P(\boldsymbol{\varphi})$ as $n \to \infty$, concluding the proof of (iv) in the general case. $\quad\square$

The following result uses ideas from [5].

Proposition 3.3. $G(\boldsymbol{\varphi}) \geqslant \int_B |\boldsymbol{\varphi}_{,R}|^2 + \frac{1}{4} |\frac{\boldsymbol{\varphi}_{,\theta}}{R}|^2 \, d\mathbf{x}$ *if $\boldsymbol{\varphi}$ contains only Fourier 2-modes or higher.*

Proof. The term involving the determinant in $G(\boldsymbol{\varphi})$ is $\int_B 2\lambda(R) \det \nabla \boldsymbol{\varphi} \, d\mathbf{x}$, where $\lambda = \frac{3}{2} \ln R$. According to part (iv) of Lemma 3.2 (with $\lambda(R) = \frac{3}{2} \ln R$), we calculate

$$\int\limits_B 2\lambda \det \nabla \boldsymbol{\varphi} \, d\mathbf{x} = \int\limits_B \frac{3}{2R} \boldsymbol{\varphi} \cdot J \boldsymbol{\varphi}_{,\theta} \, dR \, d\theta.$$

Now $\boldsymbol{\varphi} \cdot J \boldsymbol{\varphi}_{,\theta} = 2\boldsymbol{\varphi}_2 \boldsymbol{\varphi}_{1,\theta}$, so that

$$\int\limits_B \frac{3}{2R} \boldsymbol{\varphi} \cdot J \boldsymbol{\varphi}_{,\theta} \, dR \, d\theta = 3 \int\limits_B \frac{\boldsymbol{\varphi}_1}{R} \frac{\boldsymbol{\varphi}_{2,\theta}}{R} \, d\mathbf{x}.$$

Therefore by the Cauchy–Schwarz inequality

$$3 \int\limits_B \ln R \det \nabla \boldsymbol{\varphi} \, d\mathbf{x} \geqslant -3 \left(\int\limits_B \left| \frac{\boldsymbol{\varphi}_1}{R} \right|^2 d\mathbf{x} \right)^{\frac{1}{2}} \left(\int\limits_B \left| \frac{\boldsymbol{\varphi}_{2,\theta}}{R} \right|^2 d\mathbf{x} \right)^{\frac{1}{2}}.$$

Because $\boldsymbol{\varphi}$ consists of Fourier two-modes or higher, it follows that

$$\int\limits_B \left| \frac{\boldsymbol{\varphi}_{1,\theta}}{R} \right|^2 d\mathbf{x} \geqslant 4 \int\limits_B \left| \frac{\boldsymbol{\varphi}_1}{R} \right|^2 d\mathbf{x},$$

so that in particular (and on applying standard inequalities)

$$3 \int\limits_B \ln R \det \nabla \boldsymbol{\varphi} \, d\mathbf{x} \geqslant -\frac{3}{2} \left(\int\limits_B \left| \frac{\boldsymbol{\varphi}_{1,\theta}}{R} \right|^2 d\mathbf{x} \right)^{\frac{1}{2}} \left(\int\limits_B \left| \frac{\boldsymbol{\varphi}_{2,\theta}}{R} \right|^2 d\mathbf{x} \right)^{\frac{1}{2}}$$

$$\geqslant -\frac{3}{4} \int\limits_B \left(\left| \frac{\boldsymbol{\varphi}_{1,\theta}}{R} \right|^2 + \left| \frac{\boldsymbol{\varphi}_{2,\theta}}{R} \right|^2 \right) d\mathbf{x}.$$

Hence

$$G(\boldsymbol{\varphi}) \geqslant \int\limits_B |\boldsymbol{\varphi}_{,R}|^2 + \frac{1}{4} \left| \frac{\boldsymbol{\varphi}_{,\theta}}{R} \right|^2 d\mathbf{x},$$

concluding the proof of the proposition. □

The preceding results can be combined as follows:

Proof of Theorem 3.1. Let $\mathbf{u} \in \mathcal{A}$ and define $\boldsymbol{\varphi} = \mathbf{u} - \mathbf{u}_{\text{dc}}$. By Proposition 3.1,

$$I(\mathbf{u}) = I(\mathbf{u}_{\text{dc}}) + G(\boldsymbol{\varphi}),$$

and by Lemma 3.2 parts (i) and (iii),

$$G(\boldsymbol{\varphi}) = I(\boldsymbol{\varphi}^{(0)}) + G(\boldsymbol{\varphi}^{(1)}) + \sum_{j \geqslant 2} G(\boldsymbol{\varphi}^{(j)} - \boldsymbol{\varphi}^{(0)} - \boldsymbol{\varphi}^{(1)}).$$

Applying Proposition 3.3 to $\boldsymbol{\varphi}^{(2+)} := \boldsymbol{\varphi} - \boldsymbol{\varphi}^{(0)} - \boldsymbol{\varphi}^{(1)}$, we see that

$$G(\boldsymbol{\varphi}^{(2+)}) \geqslant \int\limits_B |\boldsymbol{\varphi}_{,R}^{(2+)}|^2 + \frac{1}{4} \left| \frac{\boldsymbol{\varphi}_{,\theta}^{(2+)}}{R} \right|^2 d\mathbf{x}.$$

In accordance with the hypotheses of Theorem 3.1, we suppose that $\mathbf{u}$ satisfies $G((\mathbf{u} - \mathbf{u}_{\text{dc}})^{(1)}) \geqslant 0$, which in terms of $\boldsymbol{\varphi}$ means $G(\boldsymbol{\varphi}^{(1)}) \geqslant 0$. Since $I(\boldsymbol{\varphi}^{(0)})$ is clearly positive, we conclude that $G(\boldsymbol{\varphi}) \geqslant 0$, and hence that $I(\mathbf{u}) \geqslant I(\mathbf{u}_{\text{dc}})$. This concludes the proof of Theorem 3.1. □

Finally, we connect the auxiliary functional G with the classical Lagrange multiplier theory. Recall that in the case of a pointwise constraint such as $\det \nabla \mathbf{u} = 1$ a.e., we seek pairs of functions $(\mathbf{u}, \mu)$ such that $\mathbf{u}$ solves the Euler–Lagrange equation associated to the functional

$$I'(\mathbf{v}) := I(\mathbf{v}) + \int_B \mu(\mathbf{x})(\det \nabla v - 1)\, d\mathbf{x}. \tag{3.8}$$

In the case at hand, the weak form of the Euler–Lagrange equations is

$$\int_B \nabla \boldsymbol{\varphi} \cdot \left(2\nabla \mathbf{u} + \mu(\mathbf{x})\operatorname{cof}\nabla \mathbf{u}\right) d\mathbf{x} = 0 \quad \forall \boldsymbol{\varphi} \in C_c^1(B, \mathbb{R}^2). \tag{3.9}$$

When such an equation holds, μ is referred to as a Lagrange multiplier associated with $\mathbf{u}$ and the energy I, or, more generally, as a pressure. Given sufficient smoothness of both $\mathbf{u}$ and μ, stationarity in this broad sense (since there are no restrictions on the test functions $\boldsymbol{\varphi}$) conventionally implies that $\mathbf{u}$ is stationary with respect to variations in the restricted class $\mathcal{A}$. We state the following result in terms of the functional G defined in (3.2), which is natural when one looks at the form of the functional I' in (3.8) above.

Proposition 3.4. *The double-covering map* $\mathbf{u}_{\mathrm{dc}}$ *solves the Euler–Lagrange equation associated to the functional*

$$G(\mathbf{v}) = \int_B |\nabla \mathbf{v}|^2 + 3\ln R \det \nabla \mathbf{v}\, d\mathbf{x}.$$

In particular, the function $\mu(\mathbf{x}) = 3\ln R$ *is a Lagrange multiplier associated with* $\mathbf{u}_{\mathrm{dc}}$ *and the Dirichlet energy* I.

Proof. We are required to show that (3.9) holds with $\mathbf{u}_{\mathrm{dc}}$ in place of $\mathbf{u}$, $\mu(\mathbf{x}) = -3\ln R$, and all $\boldsymbol{\varphi} \in C_c^1(B, \mathbb{R}^2)$. That is,

$$2\int_B \nabla \mathbf{u}_{\mathrm{dc}} \cdot \nabla \boldsymbol{\varphi}\, d\mathbf{x} = \int_B -3\ln R \operatorname{cof}\nabla \mathbf{u}_{\mathrm{dc}} \cdot \nabla \boldsymbol{\varphi}\, d\mathbf{x}.$$

But Lemma 3.1 applies in particular to these $\boldsymbol{\varphi}$, so that the desired equation follows by choosing $\lambda(\mathbf{x}) = -\frac{3}{2}\ln R$ in (3.3) and (3.4). $\quad\square$

3.1. Fourier one-modes $\boldsymbol{\phi}$ such that $G(\boldsymbol{\phi}) < 0$

In view of Proposition 3.1, it is natural to ask whether $G(\boldsymbol{\varphi})$ is in general nonnegative: if this were so then $\mathbf{u}_{\mathrm{dc}}$ would automatically be a global minimizer of the Dirichlet energy in $\mathcal{A}$. The following result shows that G need not even be bounded below. However, the possibility that G is nonnegative on $\mathcal{A} - \mathbf{u}_{\mathrm{dc}}$ remains open: the function $\boldsymbol{\phi}$ such that $G(\boldsymbol{\phi}) < 0$ constructed below is not a member of this class.

Proposition 3.5. *Let G be the functional defined in (3.2). Then*

(i) *there exists $\boldsymbol{\phi}$ in $W_0^{1,2}(B, \mathbb{R}^2)$ such that $G(\boldsymbol{\phi}) < 0$;*
(ii) *any function $\boldsymbol{\phi}$ consisting purely of Fourier 1-modes and lying in $\mathcal{A} - \mathbf{u}_{\mathrm{dc}}$ satisfies $G(\boldsymbol{\phi}) \geqslant 0$.*

Remark 3.6. Functions such as $\boldsymbol{\phi}$ can be derived by supposing that

$$\boldsymbol{\phi} = \mathbf{A}_1(R)\cos\theta + \mathbf{B}_1(R)\sin\theta$$

and then examining conditions under which the inequalities in Proposition 3.3 are sharp. We omit this derivation in the proof below, preferring instead to give the required $\boldsymbol{\phi}$ in its most compact form.

Proof of Proposition 3.5. Let

$$\boldsymbol{\phi}(R,\theta) = f(R)\big(\mathbf{e}_R(\theta) - \mathbf{e}_\theta(\theta)\big),$$

and, for brevity, define $\mathbf{e}(\theta) = \mathbf{e}_R(\theta) - \mathbf{e}_\theta(\theta)$. Then

$$\nabla\boldsymbol{\phi} = f'(R)\mathbf{e}(\theta) \otimes \mathbf{e}_R(\theta) - \frac{f(R)}{R} J\mathbf{e}(\theta) \otimes \mathbf{e}_\theta(\theta),$$

and hence

$$G(\boldsymbol{\phi}) = 2\pi\left(\int\limits_0^1\left(R\big(f'(R)\big)^2 + \frac{(f(R))^2}{R^2}\right)dR + \int\limits_0^1 3(\ln R)f(R)f'(R)\,dR\right).$$

Therefore

$$G(\boldsymbol{\phi}) = 2\pi E(f), \tag{3.10}$$

where

$$E(f) := \int\limits_0^1 R\big(f'(R)\big)^2 - \frac{(f(R))^2}{2R}\,dR.$$

The result is proven if we can find f such that $\boldsymbol{\phi} \in W_0^{1,2}(B,\mathbb{R}^2)$ and $E(f) < 0$. To this end, let the parameters δ and ε satisfy $0 < \delta < \delta + \varepsilon < \frac{1}{2}$, and suppose $\sigma > 0$ is constant: its value will be chosen shortly. Define f^σ by

$$f^\sigma(R) = \begin{cases} R^\sigma & \text{if } 0 < R < \delta, \\ \delta^\sigma \frac{\delta+\varepsilon-R}{\varepsilon} & \text{if } \delta < R < \delta + \varepsilon, \\ 0 & \text{otherwise.} \end{cases}$$

It can be checked that

$$\frac{E(f^\sigma)}{\delta^\sigma} = \frac{\sigma}{2} - \frac{1}{4\sigma} + \frac{1}{4} + \frac{\delta+\varepsilon}{\delta} - \frac{(\delta+\varepsilon)^2}{2\varepsilon}\ln\left(\frac{\delta+\varepsilon}{\delta}\right).$$

Freezing δ and ε while allowing $\sigma \to 0+$ shows that $\lim_{\sigma\to 0} E(f^\sigma) = -\infty$. Hence $E(f^\sigma) < 0$ provided σ is sufficiently small. Finally, note that $\boldsymbol{\phi}$ obviously has compact support in B, and that $\boldsymbol{\phi} := f^\sigma(R)\mathbf{e}(\theta)$ lies in $W^{1,2}(B,\mathbb{R}^2)$ if and only if $\sigma > 0$. Hence $\boldsymbol{\phi} \in W_0^{1,2}(B,\mathbb{R}^2)$, and by (3.10), $G(\boldsymbol{\phi}) < 0$, concluding the proof of part (i).

Suppose for a contradiction that the statement in part (ii) of the proposition is false. Then there exists $\mathbf{u}$ in $\mathcal{A}$ such that $\boldsymbol{\phi} = \mathbf{u} - \mathbf{u}_{\mathrm{dc}}$ satisfies $G(\boldsymbol{\phi}) < 0$. Now (3.5) applies, giving

$$\det\nabla\boldsymbol{\phi} = -\operatorname{cof}\nabla\mathbf{u}_{\mathrm{dc}} \cdot \nabla\boldsymbol{\phi}.$$

Integrating this expression with respect to θ and using the fact that $\boldsymbol{\phi}$ consists purely of Fourier one-modes while $\operatorname{cof}\nabla\mathbf{u}_{\mathrm{dc}}$ consists purely of two-modes, it follows that

$$\int\limits_0^{2\pi}\det\nabla\boldsymbol{\phi}\,d\theta = 0.$$

But then $\int_B \lambda(R)\det\nabla\boldsymbol{\phi}\,d\mathbf{x} = 0$, so that $G(\boldsymbol{\phi}) = \int_B |\nabla\boldsymbol{\phi}|^2\,d\mathbf{x} \geqslant 0$, contradicting the assumption $G(\boldsymbol{\phi}) < 0$. □

Note that G is bounded below on the class $W_0^{1,2}(B,\mathbb{R})$ if and only if G is nonnegative on that class, as a standard argument shows. This simple method cannot, however, be used to establish the positivity of G on the smaller class $\mathcal{A} - \mathbf{u}_{\mathrm{dc}}$: for it to work we would require $k\boldsymbol{\varphi}$ to belong to $\mathcal{A} - \mathbf{u}_{\mathrm{dc}}$ whenever $k \in \mathbb{R}$ and $\boldsymbol{\varphi} \in \mathcal{A} - \mathbf{u}_{\mathrm{dc}}$, and such an assertion is false.

However, the following variant of the argument may be more useful: first note that in view of (3.5),

$$\int_B 2\lambda(R)\det\nabla\boldsymbol{\varphi}\,d\mathbf{x} \geqslant -3\big\|(\ln R)\operatorname{cof}\nabla\mathbf{u}_{\mathrm{dc}}\big\|_2\|\nabla\boldsymbol{\varphi}\|_2,$$

giving

$$G(\boldsymbol{\varphi}) \geqslant \|\nabla\boldsymbol{\varphi}\|_2\big(\|\nabla\boldsymbol{\varphi}\|_2 - C\big),$$

where $C = 3\|(\ln R)\operatorname{cof}\nabla\mathbf{u}_{\mathrm{dc}}\|_2$. In particular,

$$G(\boldsymbol{\varphi}) \geqslant -\frac{C^2}{4} \tag{3.11}$$

for all $\boldsymbol{\varphi} \in \mathcal{A} - \mathbf{u}_{\mathrm{dc}}$. Let k be a real parameter and suppose there are functions $\boldsymbol{\varphi}(\mathbf{x};k)$ such that $\boldsymbol{\varphi}(\cdot;k) \in \mathcal{A} - \mathbf{u}_{\mathrm{dc}}$ for all $k \geqslant 0$, $\boldsymbol{\varphi}(\mathbf{x},0) = \boldsymbol{\varphi}(\mathbf{x})$ for all $\mathbf{x}$, and

$$G\big(\boldsymbol{\varphi}(\cdot;k)\big) \leqslant g(k)G\big(\boldsymbol{\varphi}(\cdot,0)\big)$$

for some increasing function g. Then, by bounding the left-hand side of this inequality from below using (3.11) and letting $k \to \infty$, the conclusion $G(\boldsymbol{\varphi}) \geqslant 0$ could again be reached.

4. The double-covering map as a stationary point

In order that $\mathbf{u}_{\mathrm{dc}}$ minimizes the functional I in $\mathcal{A}$ it is necessary for I to be stationary with respect to sufficiently regular variations about $\mathbf{u}_{\mathrm{dc}}$, provided such variations exist. In particular, if the variations take the form of a one-parameter family $(\mathbf{u}(\mathbf{x},;\varepsilon))_{\varepsilon\in[0,\varepsilon_0)} \subset \mathcal{A}$ differentiable in ε and such that $\mathbf{u}(\cdot,0) = \mathbf{u}_{\mathrm{dc}}(\cdot)$, then it is natural to determine whether

$$\partial_\varepsilon|_{\varepsilon=0} I\big(\mathbf{u}(\cdot,\varepsilon)\big) = 0. \tag{4.1}$$

Notice that if $(\mathbf{u}(\mathbf{x},;\varepsilon))_{\varepsilon\in[0,\varepsilon_0)} \subset \mathcal{A}$ with the differentiability properties outlined above, then by differentiating the constraint $\det\nabla\mathbf{u}(\mathbf{x},\varepsilon) = 1$ at $\varepsilon = 0$ we obtain

$$\operatorname{cof}\nabla\mathbf{u}_{\mathrm{dc}}(\mathbf{x}) \cdot \nabla\mathbf{u}_{,\varepsilon}(\mathbf{x},0) = 0. \tag{4.2}$$

It turns out that this condition automatically implies the stationarity (4.1): see Theorem 4.1 below for details. It is a more delicate matter, however, to construct one-parameter variations about $\mathbf{u}_{\mathrm{dc}}$ in the class $\mathcal{A}$, chiefly because of the doubling nature of $\mathbf{u}_{\mathrm{dc}}$. Section 4.1 is devoted to this topic. The main result is that the even part of a variation with Lipschitz tangent $\boldsymbol{\psi}$, say, at $\mathbf{u}_{\mathrm{dc}}$ essentially arises as a flow, that is, as the solution to a certain vector ODE. It is not clear what role the odd part of the tangent $\boldsymbol{\psi}^{\mathrm{o}}$ plays in such one-parameter families, nor indeed whether such variations exist.

The next lemma effectively converts (4.3) into the stationarity condition (4.1).

Lemma 4.1. *Let $\boldsymbol{\psi} \in W_0^{1,2}(B,\mathbb{R}^2)$ satisfy*

$$\operatorname{cof}\nabla\mathbf{u}_{\mathrm{dc}}(\mathbf{x}) \cdot \nabla\boldsymbol{\psi}(\mathbf{x}) = 0 \tag{4.3}$$

for a.e. $\mathbf{x}$ in B. Then

$$\int_B \nabla\mathbf{u}_{\mathrm{dc}} \cdot \nabla\boldsymbol{\psi}\,d\mathbf{x} = 0. \tag{4.4}$$

Proof. An approximation argument shows that it is enough to prove (4.4) for smooth $\boldsymbol{\psi}$. Taking $\lambda(R) = \ln R$ in part (b) of Lemma 3.1, it follows that

$$\int_B \ln R\operatorname{cof}\nabla\mathbf{u}_{\mathrm{dc}} \cdot \nabla\boldsymbol{\psi}\,d\mathbf{x} = -\sqrt{2}\int_B \boldsymbol{\psi} \cdot \mathbf{e}_R(2\theta)\,dR\,d\theta.$$

The integral on the left is zero by (4.3), and that on the right is proportional to $\int_B \nabla\boldsymbol{\psi} \cdot \nabla\mathbf{u}_{\mathrm{dc}}\,d\mathbf{x}$ by part (a) of Lemma 3.1. Thus Eq. (4.4). □

We record the preceding discussion and results in the following:

Theorem 4.1. *Let* $(\mathbf{u}(\cdot\,; \varepsilon))_{\varepsilon \in [0,\varepsilon_0)} \subset \mathcal{A}$ *be a one-parameter family of maps differentiable in* ε *such that*

$$\mathbf{u}(\cdot, 0) = \mathbf{u}_{\mathrm{dc}}(\cdot),$$

$$\mathbf{u}_{,\varepsilon}(\cdot, 0) = \boldsymbol{\psi}(\cdot),$$

where $\boldsymbol{\psi} \in W_0^{1,2}(B, \mathbb{R}^2)$. *Then*

$$\mathrm{cof}\, \nabla \mathbf{u}_{\mathrm{dc}} \cdot \nabla \boldsymbol{\psi} = 0 \quad a.e., \tag{4.5}$$

and

$$\partial_\varepsilon |_{\varepsilon = 0} I\big(u(\cdot, \varepsilon)\big) = 0.$$

Proof. We have already noted that the constraint $\det \nabla \mathbf{u}(\mathbf{x}, \varepsilon) = 1$ a.e. $\mathbf{x} \in B$ together with the assumed differentiability with respect to ε implies (4.5). The result now follows by applying Lemma 4.1. □

Remark 4.1. We note that the hypothesis $\boldsymbol{\psi} \in W^{1,2}(B, \mathbb{R}^2)$ would in fact suffice. The reason is that $u(\cdot, \varepsilon)$ and $\mathbf{u}_{\mathrm{dc}}$ agree on ∂B for all $\varepsilon \in [0, \varepsilon_0)$, so that in particular $\boldsymbol{\psi}(\mathbf{x}) = \mathbf{u}_{,\varepsilon}(\mathbf{x}, 0) = 0$ if $\mathbf{x} \in \partial B$.

4.1. Variations via divergence-free flows

The tactic of this section is to generate variations belonging to $\mathcal{A}$ and satisfying (4.2) via 'divergence-free' flows. Proposition 4.3 details the construction in the case that the prescribed tangent $\boldsymbol{\psi}$ to $\mathcal{A}$ at $\mathbf{u}_{\mathrm{dc}}$ satisfies the necessary condition (4.2), that is,

$$\mathrm{cof}\, \nabla \mathbf{u}_{\mathrm{dc}} \cdot \nabla \boldsymbol{\psi} = 0.$$

We show in Lemma 4.2 below that this condition holds only if it holds separately for each of $\boldsymbol{\psi}^{\mathrm{o}}$ and $\boldsymbol{\psi}^{\mathrm{e}}$, where we have used the notation $\boldsymbol{\psi}^{\mathrm{o}}$, $\boldsymbol{\psi}^{\mathrm{e}}$ for the odd, respectively even, parts of $\boldsymbol{\psi}$, as defined in the introduction.

Inspired by the technique of Dacorogna and Marcellini, a useful summary of which can be found in [7], the equation

$$\mathrm{cof}\, \nabla \mathbf{u}_{\mathrm{dc}} \cdot \nabla \boldsymbol{\psi}^{\mathrm{e}} = 0$$

can be used to generate an even flow in $\mathcal{A}$ emanating from $\mathbf{u}_{\mathrm{dc}}$, provided $\boldsymbol{\psi}$ is Lipschitz. This technical restriction ensures that the relevant ODE theory applies. (See also [8] for related flow ideas but in a more straightforward topology.) Naturally, the even flow generated in this way is tangent at $\varepsilon = 0$ to the even part $\mathbf{u}^{\mathrm{e}}(\cdot, \varepsilon)$ of any sufficiently regular variation $(\mathbf{u}(\cdot, \varepsilon))_{\varepsilon \in [0,\varepsilon_0)}$ with $\mathbf{u}_{,\varepsilon}(\cdot, 0)$ proportional to $\boldsymbol{\psi}$. This is easily seen: a short calculation shows that $\mathbf{u}^{\mathrm{e}}_{,\varepsilon}(\cdot, 0) = \boldsymbol{\psi}^{\mathrm{e}}$, for example. We also remark that there is an abundance of such flows: see Corollary 4.1 for details.

Lemma 4.2. *Let* $\boldsymbol{\psi} \in \mathrm{lip}_0(B, \mathbb{R}^2)$ *and write* $\boldsymbol{\psi}(\mathbf{z}) = \boldsymbol{\psi}^{\mathrm{o}}(\mathbf{z}) + \boldsymbol{\psi}^{\mathrm{e}}(\mathbf{z})$, *where* $\boldsymbol{\psi}^{\mathrm{e}}(\mathbf{z}) = \boldsymbol{\psi}^{\mathrm{e}}(-\mathbf{z})$ *and* $\boldsymbol{\psi}^{\mathrm{o}}(\mathbf{z}) = -\boldsymbol{\psi}^{\mathrm{o}}(-\mathbf{z})$ *for all* $\mathbf{z} \in B$. *Suppose that*

$$\mathrm{cof}\, \nabla \mathbf{u}_{\mathrm{dc}}(\mathbf{z}) \cdot \nabla \boldsymbol{\psi}(\mathbf{z}) = 0 \quad a.e.\ \mathbf{z} \in B. \tag{4.6}$$

Then

$$\mathrm{cof}\, \nabla \mathbf{u}_{\mathrm{dc}}(\mathbf{z}) \cdot \nabla \boldsymbol{\psi}^{\mathrm{o}}(\mathbf{z}) = 0 \quad a.e.\ \mathbf{z} \in B.$$

Proof. Let B' be the set of $\mathbf{z}$ in B such that (4.6) holds for both $\mathbf{z}$ and $-\mathbf{z}$. Since (4.6) holds almost everywhere, $B \setminus B'$ is null. Let $\mathbf{z} \in B'$. Since $\nabla \boldsymbol{\psi}^{\mathrm{e}}(\mathbf{z}) = -\nabla \boldsymbol{\psi}^{\mathrm{e}}(-\mathbf{z})$, $\nabla \boldsymbol{\psi}^{\mathrm{o}}(\mathbf{z}) = \nabla \boldsymbol{\psi}^{\mathrm{o}}(-\mathbf{z})$, and $\nabla \mathbf{u}_{\mathrm{dc}}(-\mathbf{z}) = -\nabla \mathbf{u}_{\mathrm{dc}}(\mathbf{z})$, it follows from (4.6) evaluated at $-\mathbf{z}$ that

$$\big(\nabla \boldsymbol{\psi}^{\mathrm{e}}(\mathbf{z}) - \nabla \boldsymbol{\psi}^{\mathrm{o}}(\mathbf{z})\big) \cdot \mathrm{cof}\, \nabla \mathbf{u}_{\mathrm{dc}}(\mathbf{z}) = 0$$

for all $\mathbf{z}$ in B'. Evaluating (4.6) at $\mathbf{z} \in B'$ gives

$$\left(\nabla \boldsymbol{\psi}^{\mathrm{e}}(\mathbf{z}) + \nabla \boldsymbol{\psi}^{\mathrm{o}}(\mathbf{z}) \right) \cdot \mathrm{cof} \, \nabla \mathbf{u}_{\mathrm{dc}}(\mathbf{z}) = 0.$$

The proof is concluded by subtracting the former expression from the latter. $\qquad \square$

Remark 4.2. The same argument shows that if $\mathrm{cof} \, \nabla \mathbf{u}_{\mathrm{dc}} \cdot \nabla \boldsymbol{\psi} = F(\mathbf{z})$, where $F(\mathbf{z}) = F(-\mathbf{z})$ for all $\mathbf{z}$, then $\mathrm{cof} \, \nabla \mathbf{u}_{\mathrm{dc}}(\mathbf{z}) \cdot \nabla \boldsymbol{\psi}^{\mathrm{o}}(\mathbf{z}) = 0$ for non-zero $\mathbf{z}$.

A short calculation shows that for the $\boldsymbol{\psi}$ of Lemma 4.2

$$\mathrm{cof} \, \nabla \mathbf{u}_{\mathrm{dc}}(\mathbf{x}) \cdot \nabla \boldsymbol{\psi}(\mathbf{x}) = \frac{1}{R} \left(\sqrt{2} \big(R \boldsymbol{\psi}(\mathbf{x}) \cdot \mathbf{e}_R(2\theta) \big)_{,R} + \frac{1}{\sqrt{2}} \big(\boldsymbol{\psi}(\mathbf{x}) \cdot \mathbf{e}_\theta(2\theta) \big)_{,\theta} \right) \tag{4.7}$$

at any non-zero point $\mathbf{x} = R \mathbf{e}_R(\theta)$. This observation will be used in the result below, where the existence of one-parameter families of variations $(\mathbf{u}(\mathbf{x}, ; \varepsilon))_{\varepsilon \in [0, \varepsilon_0)}$ of $\mathbf{u}_{\mathrm{dc}}$ in $\mathcal{A}$ is related to solutions $\boldsymbol{\psi}$ of the equation $\mathrm{cof} \, \nabla \mathbf{u}_{\mathrm{dc}} \cdot \nabla \boldsymbol{\psi} = 0$.

Proposition 4.3. *Let* $\boldsymbol{\psi} \in \mathrm{lip}_0(B, \mathbb{R}^2)$ *satisfy*

$$\sqrt{2} \big(R \boldsymbol{\psi}(\mathbf{x}) \cdot \mathbf{e}_R(2\theta) \big)_{,R} + \frac{1}{\sqrt{2}} \big(\boldsymbol{\psi}(\mathbf{x}) \cdot \mathbf{e}_\theta(2\theta) \big)_{,\theta} = 0. \tag{4.8}$$

Then there exists a one-parameter family $(\mathbf{u}(\mathbf{x}, ; \varepsilon))_{\varepsilon \in [0, \varepsilon_0)}$ *such that*

(i) *each map* $\mathbf{u}(\mathbf{x}; \varepsilon)$ *is differentiable in* ε *for* $\varepsilon \in (0, \varepsilon_0)$ *provided* $\mathbf{x} \neq 0$;
(ii) $\mathbf{u}(\mathbf{x}, 0) = \mathbf{u}_{\mathrm{dc}}(\mathbf{x})$ *for each non-zero* $\mathbf{x} \in B$;
(iii) $\nabla \boldsymbol{\psi}(\mathbf{x}) = \nabla \mathbf{u}_{,\varepsilon}(\mathbf{x}, 0)$ *for each non-zero* $\mathbf{x} \in B$;
(iv) $\mathbf{u}(\cdot, \varepsilon) \in \mathcal{A}$ *for each* $\varepsilon \in [0, \varepsilon_0)$;
(v) $\mathbf{u}(\mathbf{x}, \varepsilon) = \mathbf{u}(-\mathbf{x}, \varepsilon)$ *for each* $\varepsilon \in [0, \varepsilon_0)$ *and each* $\mathbf{x} \in B$.

Conversely, any one-parameter family satisfying conditions (i)–(iv) *above generates a solution to* (4.8).

Proof. Let $\boldsymbol{\psi}$ satisfy (4.8). By Lemma 4.2 above, Eq. (4.8) implies that $\mathrm{cof} \, \nabla \mathbf{u}_{\mathrm{dc}}(\mathbf{x}) \cdot \nabla \boldsymbol{\psi}^{\mathrm{o}}(\mathbf{x}) = 0$ and $\mathrm{cof} \, \nabla \mathbf{u}_{\mathrm{dc}}(\mathbf{x}) \cdot \nabla \boldsymbol{\psi}^{\mathrm{e}}(\mathbf{x}) = 0$ for a.e. $\mathbf{x}$. Therefore we may assume without loss of generality that $\boldsymbol{\psi} = \boldsymbol{\psi}^{\mathrm{e}}$ in (4.8), so that $\boldsymbol{\psi}(\mathbf{x}) = \boldsymbol{\psi}(-\mathbf{x})$ for a.e. $\mathbf{x}$. Now define $\mathbf{v} \in \mathrm{lip}_0(B_{\frac{1}{\sqrt{2}}}, \mathbb{R}^2)$ by

$$\mathbf{v}\big(\mathbf{u}_{\mathrm{dc}}(\mathbf{x})\big) = \boldsymbol{\psi}(\mathbf{x}). \tag{4.9}$$

The function $\mathbf{v}$ is well-defined because if $\mathbf{u}_{\mathrm{dc}}(\mathbf{x}) = \mathbf{u}_{\mathrm{dc}}(\mathbf{y})$ then $\mathbf{x} = \pm \mathbf{y}$, and since $\boldsymbol{\psi}$ is by hypothesis an even function it follows that $\boldsymbol{\psi}(\mathbf{x}) = \boldsymbol{\psi}(\mathbf{y})$. Hence $\mathbf{v}(\mathbf{u}_{\mathrm{dc}}(\mathbf{x})) = \mathbf{v}(\mathbf{u}_{\mathrm{dc}}(\mathbf{y}))$. Thus

$$\mathbf{v}\big(\rho \mathbf{e}_R(\alpha)\big) = \boldsymbol{\psi}\left(\sqrt{2} \rho \mathbf{e}_R\left(\frac{\alpha}{2} \right) \right). \tag{4.10}$$

The argument above shows that this representation holds for $0 < \rho < \frac{1}{\sqrt{2}}$, $0 \leqslant \alpha < 2\pi$. In particular, since $\boldsymbol{\psi}$ has compact support in B, it follows that $\mathbf{v}(\frac{1}{\sqrt{2}} \mathbf{e}_R(\alpha)) = 0$ for all α.

Define $\mathbf{u}(\cdot, \varepsilon)$ as the solution of the ODE

$$\frac{d}{d\varepsilon} \mathbf{u}(\cdot, \varepsilon) = \mathbf{v}\big(\mathbf{u}(\cdot, \varepsilon)\big), \tag{4.11}$$

$$\mathbf{u}(\mathbf{x}, 0) = \mathbf{u}_{\mathrm{dc}}(\mathbf{x}) \tag{4.12}$$

for $\mathbf{x} \in B$. Since $\mathbf{v}$ is Lipschitz, solutions to (4.11) are in particular unique.

(i), (ii): Part (i) is immediate from standard ODE theory based on the assumption that $\mathbf{v}$ is Lipschitz. Part (ii) is included in the definition of the flow.

(iii): Taking the gradient of the first equation in (4.11) and setting $\varepsilon = 0$ yields

$$\nabla \mathbf{u}_{,\varepsilon}(\mathbf{x},0) = \nabla \mathbf{v}\big(\mathbf{u}_{\mathrm{dc}}(\mathbf{x})\big)\nabla \mathbf{u}_{\mathrm{dc}}(\mathbf{x}) \tag{4.13}$$

if $\mathbf{x} \neq 0$. Referring to the definition of $\mathbf{v}$ in (4.9), it is immediate that the right-hand side of (4.13) is $\nabla \boldsymbol{\psi}(\mathbf{x})$. Hence part (iii) of the proposition holds.

(iv): We must show that for each $\varepsilon \in [0,\varepsilon_0)$, $\mathbf{u}(\mathbf{x},\varepsilon) = \mathbf{u}_{\mathrm{dc}}(\mathbf{x})$ for $\mathbf{x} \in \partial B$ and $\det \nabla \mathbf{u}(\mathbf{x},\varepsilon) = 1$ for almost every $\mathbf{x}$ in B. To see the first of these conditions, note that $\mathbf{u}(\mathbf{x},\varepsilon) = \mathbf{u}_{\mathrm{dc}}(\mathbf{x})$ for $\mathbf{x} \in \partial B$ is a solution to both equations in (4.11). Here we use the fact derived above that $\mathbf{v}(\mathbf{u}_{\mathrm{dc}}(\mathbf{x})) = 0$ when $\mathbf{x} \in \partial B$. By the uniqueness of solutions to (4.11), it must be that $\mathbf{u}(\mathbf{x},\varepsilon) = \mathbf{u}_{\mathrm{dc}}(\mathbf{x})$ if $\mathbf{x} \in \partial B$ and $\varepsilon \in (0,\varepsilon_0)$.

The condition $\det \nabla \mathbf{u}(\mathbf{x},\varepsilon) = 1$ a.e. is a consequence of the following argument. Let $\mathbf{h}(\mathbf{x},\varepsilon) = \nabla \mathbf{u}(\mathbf{x},\varepsilon)$, and note that again taking the gradient of the first equation in (4.11) gives

$$\partial_\varepsilon \mathbf{h}(\mathbf{x},\varepsilon) = \nabla \mathbf{v}\big(\mathbf{u}(\mathbf{x},\varepsilon)\big)\mathbf{h}(\mathbf{x},\varepsilon)$$

for non-zero $\mathbf{x} \in B$ and $\varepsilon \in (0,\varepsilon_0)$, which equation is clearly of the form $\partial_\varepsilon \mathbf{h} = A\mathbf{h}$ with $A = \nabla \mathbf{v}(\mathbf{u}(\mathbf{x},\varepsilon))$. Hence, on applying the identity

$$\partial_\varepsilon \big(\det \mathbf{h}(\mathbf{x},\varepsilon)\big) = \mathrm{tr}\big(A(\mathbf{x},\varepsilon)\big)\det \mathbf{h}(\mathbf{x},\varepsilon),$$

we see that

$$\partial_\varepsilon \big(\det \mathbf{h}(\mathbf{x},\varepsilon)\big) = \mathrm{div}\,\mathbf{v}\big(\mathbf{u}(\mathbf{x},\varepsilon)\big)\det \mathbf{h}(\mathbf{x},\varepsilon).$$

We claim that (4.8) implies $\mathrm{div}\,\mathbf{v}(\mathbf{z}) = 0$ for non-zero $\mathbf{z}$, which when applied to the equation above implies that $\det \nabla \mathbf{u}(\mathbf{x},\varepsilon)$ is constant in ε. The initial condition $\mathbf{h}(\mathbf{x},0) = \det \nabla \mathbf{u}_{\mathrm{dc}}(\mathbf{x})$ further implies $\det \nabla \mathbf{u}(\mathbf{x},\varepsilon) = 1$ for non-zero $\mathbf{x} \in B$ and $\varepsilon \in (0,\varepsilon_0)$.

Now, on putting $\mathbf{z} = \rho \mathbf{e}_R(\alpha)$,

$$\mathrm{div}\,\mathbf{v}(\mathbf{z}) = \frac{1}{\rho}\big(\big(\rho \mathbf{v} \cdot \mathbf{e}_R(\alpha)\big)_{,\rho} + \big(\mathbf{v} \cdot \mathbf{e}_\theta(\alpha)\big)_{,\alpha}\big),$$

which, when rewritten in terms of $\boldsymbol{\psi}$, R and θ via (4.10), gives

$$\mathrm{div}\,\mathbf{v}(\mathbf{z}) = \frac{\sqrt{2}}{R}\bigg(\big(R\boldsymbol{\psi}(R,\theta) \cdot \mathbf{e}_R(2\theta)\big)_{,R} + \frac{1}{2}\big(\boldsymbol{\psi}(R,\theta) \cdot \mathbf{e}_\theta(2\theta)\big)_{,\theta}\bigg). \tag{4.14}$$

Comparing this expression with (4.8), we see that $\mathrm{div}\,\mathbf{v}(\mathbf{z}) = 0$ at non-zero $\mathbf{z}$. This completes the proof of part (iv) in the statement of the proposition.

(v): Let $\mathbf{w}(\mathbf{x}) = \mathbf{u}(-\mathbf{x},\varepsilon)$ and note that $\mathbf{w}$ solves the same ODE as $\mathbf{u}(\mathbf{x},\varepsilon)$ subject to the same initial condition. We again invoke the uniqueness of solutions to this class of ODEs to infer $\mathbf{w}(\mathbf{x}) = \mathbf{u}(\mathbf{x},\varepsilon)$ for $\mathbf{x} \in B$ and $\varepsilon \in (0,\varepsilon_0)$. Hence $\mathbf{u}(\cdot,\varepsilon)$ is even.

The opening remarks of Section 4 demonstrate that if the family of variations $(\mathbf{u}(\mathbf{x}, ;\varepsilon))_{\varepsilon \in [0,\varepsilon_0)}$ satisfies (i)–(iv), then in particular (4.2) holds a.e., that is,

$$\mathrm{cof}\,\nabla \mathbf{u}_{\mathrm{dc}} \cdot \nabla \mathbf{u}_{,\varepsilon}(\mathbf{x},0) = 0.$$

It follows from this equation and (4.7) that $\boldsymbol{\psi}(\mathbf{x}) := \mathbf{u}_{,\varepsilon}(\mathbf{x},0)$ solves (4.8) a.e. $\square$

The proof of Proposition 4.3 above shows that there are in fact many solutions to Eq. (4.8), as follows:

Corollary 4.1. *Let $\boldsymbol{\psi} \in \mathrm{lip}_0(B,\mathbb{R}^2)$. Then $\boldsymbol{\psi}$ is an even solution of*

$$\sqrt{2}\big(R\boldsymbol{\psi}(\mathbf{x}) \cdot \mathbf{e}_R(2\theta)\big)_{,R} + \frac{1}{\sqrt{2}}\big(\boldsymbol{\psi}(\mathbf{x}) \cdot \mathbf{e}_\theta(2\theta)\big)_{,\theta} = 0$$

if and only if

$$\boldsymbol{\psi}(\mathbf{z}) = \Psi\big(\mathbf{u}_{\mathrm{dc}}(\mathbf{z})\big) \quad \mathbf{z} \in B \setminus \{0\}, \tag{4.15}$$

$$\Psi(\mathbf{z}) = J\nabla \mathbf{w}(\mathbf{z}), \tag{4.16}$$

where $\mathbf{w}$ is an even function whose gradient $\nabla \mathbf{w} \in \mathrm{lip}_0(B_{\frac{1}{\sqrt{2}}},\mathbb{R}^2)$.

Proof. The calculation leading to (4.14) above shows that $\boldsymbol{\psi}$ is even and solves (4.8) if and only if Ψ defined by the first equation of (4.15) is even and divergence-free. It is well known that divergence-free functions are necessarily of the form $\Psi(\mathbf{z}) = J\nabla\mathbf{w}(\mathbf{z})$ for a suitable scalar-valued $\mathbf{w}$, so the second equation in (4.15) holds. Finally, Ψ is even if and only if $\nabla\mathbf{w}^{\mathrm{o}}(\mathbf{z}) = 0$ for non-zero $\mathbf{z}$. Therefore $\mathbf{w}^{\mathrm{o}}(\mathbf{z})$ is constant in $B \setminus \{0\}$, and hence $\mathbf{w}$ is even. $\quad\square$

Let $(\mathbf{u}(\mathbf{x}, ; \varepsilon))_{\varepsilon \in [0,\varepsilon_0)}$ be the one-parameter family constructed in Proposition 4.3 above, and let $\boldsymbol{\psi}(\mathbf{x}) = \mathbf{u}_{,\varepsilon}(\mathbf{x}, 0)$, so that in particular $\nabla\boldsymbol{\psi} = \nabla\mathbf{u}_{,\varepsilon}(\mathbf{x}, 0)$, and $\boldsymbol{\psi}$ satisfies

$$\operatorname{cof} \nabla\mathbf{u}_{\mathrm{dc}}(\mathbf{x}) \cdot \nabla\boldsymbol{\psi}(\mathbf{x}) = 0 \tag{4.17}$$

for a.e. $\mathbf{x}$ in B. Now, an easy calculation implies that

$$\partial_\varepsilon|_{\varepsilon=0} I\big(\mathbf{u}(\cdot, \varepsilon)\big) = \int_B \nabla\mathbf{u}_{\mathrm{dc}} \cdot \nabla\boldsymbol{\psi} \, d\mathbf{x}. \tag{4.18}$$

Lemma 4.2 and the argument leading to (4.18) together comprise a proof of the following stationarity result.

Theorem 4.2. *Let $\boldsymbol{\psi} \in \mathrm{lip}_0(B, \mathbb{R}^2)$ satisfy*

$$\sqrt{2}\big(R\boldsymbol{\psi}(\mathbf{x}) \cdot \mathbf{e}_R(2\theta)\big)_{,R} + \frac{1}{\sqrt{2}}\big(\boldsymbol{\psi}(\mathbf{x}) \cdot \mathbf{e}_\theta(2\theta)\big)_{,\theta} = 0 \quad a.e. \tag{4.19}$$

and let $(\mathbf{u}(\mathbf{x}, ; \varepsilon))_{\varepsilon \in [0,\varepsilon_0)}$ be the one-parameter family of variations about $\mathbf{u}_{\mathrm{dc}}$ in $\mathcal{A}$ associated with $\boldsymbol{\psi}^{\mathrm{e}}(\mathbf{x}) = \frac{1}{2}(\boldsymbol{\psi}(\mathbf{x}) + \boldsymbol{\psi}(-\mathbf{x}))$ by Proposition 4.3. Then

$$\partial_\varepsilon|_{\varepsilon=0} I\big(\mathbf{u}(\cdot, \varepsilon)\big) = 0.$$

4.2. Local minimality of $\mathbf{u}_{\mathrm{dc}}$: the role of the even variations

The stationarity result stated in Theorem 4.2 above clearly does not give any information on whether $\mathbf{u}_{\mathrm{dc}}$ is a local minimizer, a saddle point, or even a local maximizer of the Dirichlet energy in the class $\mathcal{A}$. In this section we focus on the role of even maps in determining how the energy $I(\mathbf{w})$ behaves when small variations $\mathbf{w}$ of $\mathbf{u}_{\mathrm{dc}}$ in $\mathcal{A}$ are made.

Specifically, we begin by combining Theorem 3.1 and Proposition 4.3 below to conclude that

$$I\big(\mathbf{u}(\cdot, \varepsilon)\big) \geqslant I(\mathbf{u}_{\mathrm{dc}})$$

for $\varepsilon \in [0,\varepsilon_0)$. Here, $(\mathbf{u}(\mathbf{x}, \varepsilon))_{\varepsilon \in (0,\varepsilon_0)}$ continues to denote the one-parameter family of variations of $\mathbf{u}_{\mathrm{dc}}$ in $\mathcal{A}$ constructed in Proposition 4.3. The key observation is that the condition

$$\mathbf{u}(\mathbf{x}, \varepsilon) = \mathbf{u}(-\mathbf{x}, \varepsilon)$$

for a.e. $\mathbf{x}$ in B provided by part (v) of Proposition 4.3 implies that $\mathbf{u}(\cdot, \varepsilon)$ has zero Fourier 1-mode. See Theorem 4.3 below for details.

This begs the question of whether the energy can be lowered by taking variations whose tangents at $\mathbf{u}_{\mathrm{dc}}$ are not assumed to be purely even. Such variations do not arise as flows, in contrast to those of Proposition 4.3: see below for details. We show in Proposition 4.4 below that the energy can be lowered provided there exists a tangent map $\boldsymbol{\psi}$ such that $G(\boldsymbol{\psi}^{\mathrm{o}}) < 0$, although at this time the question of whether such a map exists is open.

Theorem 4.3. *Let $\boldsymbol{\psi} \in \mathrm{lip}_0(B, \mathbb{R}^2)$ satisfy*

$$\sqrt{2}\big(R\boldsymbol{\psi}(\mathbf{x}) \cdot \mathbf{e}_R(2\theta)\big)_{,R} + \frac{1}{\sqrt{2}}\big(\boldsymbol{\psi}(\mathbf{x}) \cdot \mathbf{e}_\theta(2\theta)\big)_{,\theta} = 0,$$

and let $(\mathbf{u}(\mathbf{x}, ; \varepsilon))_{\varepsilon \in [0,\varepsilon_0)}$ be the one-parameter family of variations of $\mathbf{u}_{\mathrm{dc}}$ in $\mathcal{A}$ associated to $\boldsymbol{\psi}^{\mathrm{e}}(\mathbf{x}) = \frac{1}{2}(\boldsymbol{\psi}(\mathbf{x}) + \boldsymbol{\psi}(-\mathbf{x}))$ by Proposition 4.3. Then

$$I\big(\mathbf{u}(\cdot, \varepsilon)\big) \geqslant I(\mathbf{u}_{\mathrm{dc}}) \tag{4.20}$$

with equality if and only if $\mathbf{u}(\mathbf{x}, \varepsilon) = \mathbf{u}_{\mathrm{dc}}(\mathbf{x})$ for a.e. $\mathbf{x}$.

Proof. By Proposition 4.3 part (v), we see that $\mathbf{u}(\cdot, \varepsilon)$ is a family of even maps. Referring to the Fourier decomposition set out in (3.1), the Fourier 1-mode of $\mathbf{u}(\mathbf{x}, \varepsilon)$ is

$$\mathbf{u}^{(1)}(\mathbf{x}, \varepsilon) = \mathbf{A}_1(R) \cos \theta + \mathbf{B}_1(R) \sin \theta,$$

where

$$\mathbf{A}_1(R) = \frac{1}{\pi} \int_0^{2\pi} \mathbf{u}\big(R\mathbf{e}_R(\theta), \varepsilon\big) \cos \theta \, d\theta$$

and

$$\mathbf{B}_1(R) = \frac{1}{\pi} \int_0^{2\pi} \mathbf{u}\big(R\mathbf{e}_R(\theta), \varepsilon\big) \sin \theta \, d\theta.$$

Replacing θ with $\theta + \pi$ and using $\mathbf{u}(R\mathbf{e}_R(\theta)) = \mathbf{u}(-R\mathbf{e}_R(\theta))$ gives $\mathbf{A}_1(R) = -\mathbf{A}_1(R)$, and similarly for $\mathbf{B}_1(R)$. Hence $\mathbf{u}^{(1)}(\mathbf{x}, \varepsilon) = 0$ for a.e. $\mathbf{x}$ in B and all ε such that $0 \leqslant \varepsilon < \varepsilon_0$. In particular, taking $\boldsymbol{\varphi} = \mathbf{u}(\cdot, \varepsilon) - \mathbf{u}_{\mathrm{dc}}$ in Theorem 3.1, we see that $\boldsymbol{\varphi}^{(1)} = 0$, and hence by part (ii) of that result, $G(\boldsymbol{\varphi}) \geqslant 0$. Since by part (i) of Theorem 3.1 we have

$$I\big(\mathbf{u}(\cdot, \varepsilon)\big) = I(\mathbf{u}_{\mathrm{dc}}) + G(\boldsymbol{\varphi}),$$

inequality (4.20) follows. Finally, note that equality holds in (4.20) if and only if $G(\boldsymbol{\varphi}) = 0$, which, by Proposition 3.3, is true if and only if $\nabla \boldsymbol{\varphi} = 0$. But $\boldsymbol{\varphi} = 0$ on ∂B, so the latter holds iff $\boldsymbol{\varphi}$ is almost everywhere zero, whence the condition for equality stated above. $\quad \square$

Let $\mathbf{u}(\mathbf{x}, \varepsilon)$ be a variation of $\mathbf{u}_{\mathrm{dc}}$ in $\mathcal{A}$ such that $\mathbf{u}_{,\varepsilon}(\mathbf{x}, 0) = \boldsymbol{\psi}$. The constraint $\det \nabla \mathbf{u} = 1$ a.e. implies $\mathrm{cof}\, \nabla \mathbf{u}_{\mathrm{dc}} \cdot \nabla \boldsymbol{\psi} = 0$ a.e. By Lemma 4.2, the symmetry of $\mathbf{u}_{\mathrm{dc}}$ implies in particular that $\mathrm{cof}\, \nabla \mathbf{u}_{\mathrm{dc}} \cdot \nabla \boldsymbol{\psi}^{\mathrm{e}} = 0$ almost everywhere. Therefore, by Proposition 4.3, a flow in $\mathcal{A}$ with tangent $\boldsymbol{\psi}^{\mathrm{e}}$ to $\mathbf{u}_{\mathrm{dc}}$ can be generated. If we try to use a similar technique to generate a flow $\mathbf{w}(\cdot, \varepsilon)$, say, with tangent $\boldsymbol{\psi}^{\mathrm{o}}$ to $\mathbf{u}_{\mathrm{dc}}$, then we would require

$$\frac{d}{d\varepsilon} \mathbf{w}(\mathbf{x}, \varepsilon) = \mathbf{f}\big(\mathbf{w}(\mathbf{x}, \varepsilon), \varepsilon\big)$$

for a.e. $\mathbf{x}$, all sufficiently small ε and some $\mathbf{f}$, subject in addition to $\mathbf{w}_{,\varepsilon}(\mathbf{x}, 0) = \boldsymbol{\psi}^{\mathrm{o}}(\mathbf{x})$ a.e. and $\mathbf{w}(\mathbf{x}, 0) = \mathbf{u}_{\mathrm{dc}}(\mathbf{x})$ a.e. But by exchanging $\mathbf{x}$ and $-\mathbf{x}$ in the ODE defining $\mathbf{w}$ and setting $\varepsilon = 0$, it follows that $\boldsymbol{\psi}^{\mathrm{o}} = 0$. Hence the only odd flows (i.e., variations generated by ODEs in this way) are trivial.

Therefore we turn to the following technique. Let $\mathbf{w}(\cdot, \varepsilon)$ be a variation about $\mathbf{u}_{\mathrm{dc}}$ in $\mathcal{A}$ with tangent $\boldsymbol{\psi}$ at $\mathbf{u}_{\mathrm{dc}}$ such that $\mathbf{w}(\cdot, \varepsilon)$ is C^2 with respect to the parameter ε. We may suppose that

$$\mathbf{w}(\mathbf{x}, \varepsilon) = \mathbf{u}_{\mathrm{dc}}(\mathbf{x}) + \varepsilon \boldsymbol{\psi}(\mathbf{x}) + \varepsilon^2 \boldsymbol{\sigma}(\mathbf{x}) + o\big(\varepsilon^2\big) \quad \text{a.e.}$$

for a suitable function $\boldsymbol{\sigma}$, the expansion being understood in the asymptotic sense $\varepsilon \to 0+$. Consider the flow $\mathbf{u}(\cdot, \varepsilon)$ generated by Proposition 4.3 with initial tangent $\boldsymbol{\psi}^{\mathrm{e}}$. Then it can be written

$$\mathbf{u}(\mathbf{x}, \varepsilon) = \mathbf{u}_{\mathrm{dc}}(\mathbf{x}) + \varepsilon \boldsymbol{\psi}^{\mathrm{e}}(\mathbf{x}) + \varepsilon^2 \boldsymbol{\tau}(\mathbf{x}) + o\big(\varepsilon^2\big)$$

for an even function $\boldsymbol{\tau} : B \to \mathbb{R}^2$, and in the same asymptotic sense as above. In view of Theorem 4.3, we expect the energy to obey $I(\mathbf{u}(\cdot, \varepsilon)) \geqslant I(\mathbf{u}_{\mathrm{dc}})$, and so by expressing $I(\mathbf{w}(\cdot, \varepsilon))$ in terms of $I(\mathbf{u}(\cdot, \varepsilon))$ we can hope to conclude $I(\mathbf{w}(\cdot, \varepsilon)) \geqslant I(\mathbf{u}_{\mathrm{dc}})$. This is done in Proposition 4.4 below under the assumption that $G(\boldsymbol{\psi}^{\mathrm{o}}) > 0$, a condition reminiscent of that applied to $G((\mathbf{u} - \mathbf{u}_{\mathrm{dc}})^{(1)})$ in Theorem 3.1.

Proposition 4.4. *Let* $\mathbf{w}(\cdot, \varepsilon)$ *be a variation about* $\mathbf{u}_{\mathrm{dc}}$ *in* $\mathcal{A}$ *such that*

$$\mathbf{w}_{,\varepsilon}(\mathbf{x}, 0) = \boldsymbol{\psi}(\mathbf{x})$$

for some $\boldsymbol{\psi} \in \mathrm{lip}_0(B, \mathbb{R}^2)$. *Suppose further that*

$$\mathbf{w}(\mathbf{x}, \varepsilon) = \mathbf{u}_{\mathrm{dc}}(\mathbf{x}) + \varepsilon \boldsymbol{\psi}(\mathbf{x}) + \varepsilon^2 \boldsymbol{\sigma}(\mathbf{x}) + o\big(\varepsilon^2\big) \quad a.e. \tag{4.21}$$

for a function $\boldsymbol{\sigma}$ in $W^{1,2}(B, \mathbb{R}^2)$ and all sufficiently small ε. Let $\mathbf{u}(\cdot, \varepsilon)$ be the flow generated by the ODE

$$\frac{d}{d\varepsilon}\mathbf{u}(\mathbf{x}, \varepsilon) = \boldsymbol{\psi}^{\mathrm{e}}(\mathbf{x})$$

subject to $\mathbf{u}(\mathbf{x}, 0) = \mathbf{u}_{\mathrm{dc}}(\mathbf{x})$. Suppose that $\mathbf{u}(\cdot, \varepsilon)$ may be written

$$\mathbf{u}(\mathbf{x}, \varepsilon) = \mathbf{u}_{\mathrm{dc}}(\mathbf{x}) + \varepsilon\boldsymbol{\psi}^{\mathrm{e}}(\mathbf{x}) + \varepsilon^2\boldsymbol{\tau}(\mathbf{x}) + o(\varepsilon^2) \quad a.e. \tag{4.22}$$

for a function $\boldsymbol{\tau}$ in $W^{1,2}(B, \mathbb{R}^2)$ and all sufficiently small ε. Then

$$I(\mathbf{w}(\cdot, \varepsilon)) = I(\mathbf{u}(\cdot, \varepsilon)) + \varepsilon^2 G(\boldsymbol{\psi}^{\mathrm{o}}) + o(\varepsilon^2). \tag{4.23}$$

In particular,

(i) *if $G(\boldsymbol{\psi}^{\mathrm{o}}) > 0$, then $I(\mathbf{w}(\cdot, \varepsilon)) > I(\mathbf{u}_{\mathrm{dc}})$ for sufficiently small ε, and*
(ii) *if $\boldsymbol{\psi}$ is an odd function such that $G(\boldsymbol{\psi}^{\mathrm{o}}) < 0$, then $I(\mathbf{w}(\cdot, \varepsilon)) < I(\mathbf{u}_{\mathrm{dc}})$ for sufficiently small ε.*

Proof. In the following we use the shorthand $I(\mathbf{w}) = I(\mathbf{w}(\cdot, \varepsilon))$, and similarly for $I(\mathbf{u})$. Using the expression (4.21) above, we calculate

$$I(\mathbf{w}) = I(\mathbf{u}_{\mathrm{dc}}) + 2\varepsilon \int_B \nabla\boldsymbol{\psi} \cdot \nabla\mathbf{u}_{\mathrm{dc}}\, d\mathbf{x} + \varepsilon^2 \int_B |\nabla\boldsymbol{\psi}|^2 + 2\nabla\mathbf{u}_{\mathrm{dc}} \cdot \nabla\boldsymbol{\sigma}\, d\mathbf{x} + o(\varepsilon^2). \tag{4.24}$$

Similarly, using (4.22),

$$I(\mathbf{u}) = I(\mathbf{u}_{\mathrm{dc}}) + 2\varepsilon \int_B \nabla\boldsymbol{\psi}^{\mathrm{e}} \cdot \nabla\mathbf{u}_{\mathrm{dc}}\, d\mathbf{x} + \varepsilon^2 \int_B |\nabla\boldsymbol{\psi}^{\mathrm{e}}|^2 + 2\nabla\mathbf{u}_{\mathrm{dc}} \cdot \nabla\boldsymbol{\tau}\, d\mathbf{x} + o(\varepsilon^2). \tag{4.25}$$

Inserting (4.21) into the constraint $\det \nabla\mathbf{w} = 1$ a.e. and comparing terms in ε and ε^2 implies

$$\mathrm{cof}\, \nabla\mathbf{u}_{\mathrm{dc}} \cdot \nabla\boldsymbol{\psi} = 0 \quad a.e. \tag{4.26}$$

and

$$\mathrm{cof}\, \nabla\mathbf{u}_{\mathrm{dc}} \cdot \nabla\boldsymbol{\sigma} + \det \nabla\boldsymbol{\psi} = 0 \quad a.e., \tag{4.27}$$

respectively. Similarly, using (4.22) in place of (4.21),

$$\mathrm{cof}\, \nabla\mathbf{u}_{\mathrm{dc}} \cdot \nabla\boldsymbol{\psi}^{\mathrm{e}} = 0 \quad a.e. \tag{4.28}$$

and

$$\mathrm{cof}\, \nabla\mathbf{u}_{\mathrm{dc}} \cdot \nabla\boldsymbol{\tau} + \det \nabla\boldsymbol{\psi}^{\mathrm{e}} = 0 \quad a.e. \tag{4.29}$$

Now, since $\nabla\mathbf{u}_{\mathrm{dc}}$ is an odd function,

$$\int_B \nabla\mathbf{u}_{\mathrm{dc}} \cdot \nabla\boldsymbol{\psi}\, d\mathbf{x} = \int_B \nabla\mathbf{u}_{\mathrm{dc}} \cdot \nabla\boldsymbol{\psi}^{\mathrm{e}}\, d\mathbf{x}.$$

In particular, the expressions in (4.24) and (4.25) agree to order ε. In fact, by Theorem 4.3, the coefficients of ε in both (4.24) and (4.25) vanish.

Therefore it remains to consider the coefficient of ε^2 in (4.24). Since $\mathbf{w}$ and $\mathbf{u}$ are members of $\mathcal{A}$, the functions $\boldsymbol{\sigma}$ and $\boldsymbol{\tau}$ are continuous, and it follows by Lemma 3.1 that

$$2\int_B \nabla\mathbf{u}_{\mathrm{dc}} \cdot \nabla\boldsymbol{\sigma}\, d\mathbf{x} = -\int_B 3\ln R\, \mathrm{cof}\, \nabla\mathbf{u}_{\mathrm{dc}} \cdot \nabla\boldsymbol{\sigma}\, d\mathbf{x}$$

and

$$2 \int_B \nabla \mathbf{u}_{dc} \cdot \nabla \boldsymbol{\tau} \, d\mathbf{x} = - \int_B 3 \ln R \, \mathrm{cof} \, \nabla \mathbf{u}_{dc} \cdot \nabla \boldsymbol{\tau} \, d\mathbf{x}.$$

Inserting the decomposition $\nabla \boldsymbol{\psi} = \nabla \boldsymbol{\psi}^e + \nabla \boldsymbol{\psi}^o$ in (4.27) gives

$$\mathrm{cof} \, \nabla \mathbf{u}_{dc} \cdot \nabla \boldsymbol{\sigma} = - \det \nabla \boldsymbol{\psi}^e - \det \nabla \boldsymbol{\psi}^o - \mathrm{cof} \, \nabla \boldsymbol{\psi}^o \cdot \nabla \boldsymbol{\psi}^e \quad \text{a.e.,}$$

which, in view of (4.29), gives

$$\mathrm{cof} \, \nabla \mathbf{u}_{dc} \cdot \nabla \boldsymbol{\sigma} = \mathrm{cof} \, \nabla \mathbf{u}_{dc} \cdot \nabla \boldsymbol{\tau} - \det \nabla \boldsymbol{\psi}^o - \mathrm{cof} \, \nabla \boldsymbol{\psi}^o \cdot \nabla \boldsymbol{\psi}^e \quad \text{a.e.} \tag{4.30}$$

Now it is easy to verify that

$$\int_0^{2\pi} \mathrm{cof} \, \nabla \boldsymbol{\psi}^e \cdot \nabla \boldsymbol{\psi}^o \, d\theta = 0,$$

so that (4.30) gives, in particular,

$$\int_B \ln R \, \mathrm{cof} \, \nabla \mathbf{u}_{dc} \cdot \nabla \boldsymbol{\sigma} \, d\mathbf{x} = \int_B \ln R \, \mathrm{cof} \, \nabla \mathbf{u}_{dc} \cdot \nabla \boldsymbol{\tau} \, d\mathbf{x} - \int_B \ln R \, \det \nabla \boldsymbol{\psi}^o \, d\mathbf{x}.$$

The foregoing analysis shows that the coefficient of ε^2 in (4.24) satisfies

$$\begin{aligned}
\int_B |\nabla \boldsymbol{\psi}|^2 + 2\nabla \mathbf{u}_{dc} \cdot \nabla \boldsymbol{\sigma} \, d\mathbf{x} &= I(\boldsymbol{\psi}) - \int_B 3 \ln R \, \mathrm{cof} \, \nabla \mathbf{u}_{dc} \cdot \nabla \boldsymbol{\sigma} \, d\mathbf{x} \\
&= I(\boldsymbol{\psi}) - \int_B 3 \ln R \, \mathrm{cof} \, \nabla \mathbf{u}_{dc} \cdot \nabla \boldsymbol{\tau} \, d\mathbf{x} \\
&\quad + \int_B 3 \ln R \, \det \nabla \boldsymbol{\psi}^o \, d\mathbf{x} \\
&= \int_B \left| \nabla \boldsymbol{\psi}^e \right|^2 + 2\nabla \mathbf{u}_{dc} \cdot \nabla \boldsymbol{\tau} \, d\mathbf{x} \\
&\quad + \int_B \left| \nabla \boldsymbol{\psi}^o \right|^2 + 3 \ln R \, \det \nabla \boldsymbol{\psi}^o \, d\mathbf{x}.
\end{aligned}$$

We recognize the last line above as the sum of the coefficient of ε^2 in $I(\mathbf{u})$ and $G(\boldsymbol{\psi}^o)$, thereby proving (4.23).

To prove part (i) of the theorem let us suppose that $G(\boldsymbol{\psi}^o) > 0$. By (4.23),

$$I(\mathbf{w}) - I(\mathbf{u}_{dc}) = I(\mathbf{u}) - I(\mathbf{u}_{dc}) + \varepsilon^2 G(\boldsymbol{\psi}^o) + o(\varepsilon^2),$$

where, in view of Theorem 4.3, the first term is nonnegative. Combining this with the assumed positivity of $G(\boldsymbol{\psi}^o)$, it follows that $I(\mathbf{u}) > I(\mathbf{u}_{dc})$ for all sufficiently small ε, as claimed.

Finally, let $\boldsymbol{\psi}$ be an odd function such that $G(\boldsymbol{\psi}^o) < 0$. Then the flow associated to $\boldsymbol{\psi}^e$ is trivial, i.e., $\mathbf{u}(\cdot, \varepsilon) = \mathbf{u}_{dc}$, and (4.23) implies

$$I(\mathbf{w}) - I(\mathbf{u}_{dc}) = \varepsilon^2 G(\boldsymbol{\psi}^o) + o(\varepsilon^2)$$

for sufficiently small ε. Part (ii) of the theorem now follows. $\square$

Remark 4.5. Part (ii) of the theorem above only becomes useful once the existence of odd tangents $\boldsymbol{\psi}$ satisfying $G(\boldsymbol{\psi}) < 0$ is settled. This is effectively a linearized version of the search in Section 3.1 to find $\boldsymbol{\phi}$ such that $\mathbf{u}_{dc} + \boldsymbol{\phi} \in \mathcal{A}$ and $G(\boldsymbol{\phi}) < 0$.

5. Concluding remarks and open questions

Below are some observations on extensions, limitations and open questions related to the analysis contained in the preceding sections of the paper.

1. There may be variations about $\mathbf{u}_{dc}$ that are not one-parameter families, and much of the analysis in Section 4 will not apply in such cases. However, we note that the type of stationarity expressed in Theorem 4.1 is dependent on the permissible variations: less regular variations may not give rise to a meaningful notion of stationarity.
2. The functional G is clearly the key to determining whether or not $\mathbf{u}_{dc}$ is an energy minimizer in the class $\mathcal{A}$. When the boundary condition $\mathbf{u}_{dc}$ is replaced with the N-covering map

$$\mathbf{u}_{Nc}(R, \theta) = \frac{R}{\sqrt{N}} \mathbf{e}_R(N\theta),$$

where $N \geqslant 3$ is a positive integer, it is straightforward to compute a new auxiliary functional G_N with the same structure as G. Thus the analysis of G ought to generalize to G_N. For more general but topologically non-trivial boundary conditions, the associated Lagrange multiplier may well depend on the angular variable. The structure of the associated minimizer(s) in such cases is far from obvious.
3. G is polyconvex, but is it sequentially weakly lower semicontinuous in an appropriate function space? Traditional methods for dealing with such questions seem not to apply in this case.

References

[1] J.M. Ball, Convexity conditions and existence theorems in nonlinear elasticity, Arch. Ration. Mech. Anal. 63 (4) (1976/1977) 337–403.
[2] J.M. Ball, Discontinuous equilibrium solutions and cavitation in nonlinear elasticity, Philos. Trans. R. Soc. Lond. Ser. A 306 (1496) (1982) 557–611.
[3] J.M. Ball, Some open problems in elasticity, in: Geometry, Mechanics, and Dynamics, Springer, New York, 2002, pp. 3–59.
[4] P. Bauman, D. Phillips, N. Owen, Maximal smoothness of solutions to certain Euler–Lagrange equations from nonlinear elasticity, Proc. Roy. Soc. Edinburgh Sect. A 119 (3–4) (1991) 241–263.
[5] J. Bevan, X. Yan, Minimizers with topological singularities in two dimensional elasticity, ESAIM Control Optim. Calc. Var. 14 (1) (2008) 192–209.
[6] R. Coifman, P.-L. Lions, Y. Meyer, S. Semmes, Compensated compactness and Hardy spaces, J. Math. Pures Appl. (9) 72 (3) (1993) 247–286.
[7] B. Dacorogna, Direct Methods in the Calculus of Variations, second ed., Appl. Math. Sci., vol. 78, Springer, New York, 2008.
[8] G. Del Piero, R. Rizzoni, Weak local minimizers in finite elasticity, J. Elasticity 93 (2008) 203–244.
[9] L.C. Evans, R.F. Gariepy, Measure Theory and Fine Properties of Functions, Stud. Adv. Math., CRC Press, Boca Raton, FL, 1992.
[10] L.C. Evans, R.F. Gariepy, On the partial regularity of energy-minimizing, area-preserving maps, Calc. Var. Partial Differential Equations 9 (4) (1999) 357–372.
[11] C. Fefferman, E. Stein, H^p spaces of several variables, Acta Math. 129 (3–4) (1972) 137–193.
[12] I. Fonseca, W. Gangbo, Degree Theory in Analysis and Applications, Oxford Lecture Ser. Math. Appl., vol. 2, Oxford Sci. Publ., The Clarendon Press, Oxford University Press, New York, 1995.
[13] A.L. Karakhanyan, Sufficient conditions for regularity of area-preserving deformations, Manuscripta Math. 138 (2012) 463–476.
[14] C. Mora-Corral, Explicit energy-minimizers of incompressible elastic brittle bars under uniaxial extension, C. R. Math. Acad. Sci. Paris 348 (17–18) (2010) 1045–1048.
[15] E. Spadaro, Non-uniqueness of minimizers for strictly polyconvex functionals, Arch. Ration. Mech. Anal. 193 (3) (2009) 659–678.
[16] M. Shahrokhi-Dehkordi, A. Taheri, Quasiconvexity and uniqueness of stationary points on a space of measure preserving maps, J. Convex Anal. 17 (1) (2010) 69–79.
[17] J. Sivaloganathan, S. Spector, On the symmetry of energy-minimising deformations in nonlinear elasticity. I. Incompressible materials, Arch. Ration. Mech. Anal. 196 (2) (2010) 363–394.
[18] J. Sivaloganathan, S. Spector, On the symmetry of energy-minimising deformations in nonlinear elasticity. II. Compressible materials, Arch. Ration. Mech. Anal. 196 (2) (2010) 395–431.
[19] J. Sivaloganathan, S. Spector, On the global stability of two-dimensional, incompressible, elastic bars in uniaxial extension, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 466 (2116) (2010) 1167–1176.
[20] J. Sivaloganathan, S. Spector, On the stability of incompressible elastic cylinders in uniaxial tension, preprint, 2010.
[21] S. Vodopyanov, V. Goldstein, Quasiconformal mappings and spaces of functions with generalized first derivatives, Sib. Math. J. 17 (3) (1977) 515–531.