ON MULTIWELL LIOUVILLE THEOREMS IN HIGHER DIMENSIONS

ROBERT L. JERRARD AND ANDREW LORENT

Abstract. We consider certain subsets of the space of $n \times n$ matrices of the form $K = \cup_{i=1}^{m} SO(n)A_i$, and we prove that for $p > 1, q \geq 1$ and for connected $\Omega' \subset\subset \Omega \subset \mathbb{R}^n$, there exists positive constant $a < 1$ depending on n, p, q, Ω, Ω' such that for $\varepsilon = \|\mathrm{dist}(Du, K)\|_{L^p(\Omega)}^p$ we have $\inf_{R \in K} \|Du - R\|_{L^p(\Omega')}^p \leq M\varepsilon^{1/p}$ provided u satisfies the inequality $\|D^2 u\|_{L^q(\Omega)}^q \leq a\varepsilon^{1-q}$. Our main result holds whenever $m = 2$, and also for *generic* $m \leq n$ in every dimension $n \geq 3$, as long as the wells $SO(n)A_1, \ldots, SO(n)A_m$ satisfy a certain connectivity condition. These conclusions are mostly known when $n = 2$, and they are new for $n \geq 3$.

February 6, 2008

1. Introduction

Rigidity theorems for mappings whose gradient lie in a subset of the conformal matrices date back to 1850, when Liouville [Lio 50] proved that given a domain $\Omega \subset \mathbb{R}^n$ and a function $u \in C^3\left(\Omega, \mathbb{R}^3\right)$ with the property that $Du\,(x) = \lambda\,(x)\, O\,(x)$ where $\lambda\,(x) \in \mathbb{R}_+$ and $O\,(n) \in SO\,(n)$ then u is either affine or a Mobius transformation. A corollary to Liouville's Theorem is that a C^3 function whose gradient belongs everywhere to $SO\,(n)$ is an affine mapping. A striking quantitative version of this fact was recently proved by Friesecke, James and Müller [Fr-Ja-Mu 02], who showed that for every bounded open connected Lipschitz domain $U \subset \mathbb{R}^n$, $n \geq 2$, and every $q > 1$, there exists a constant $C(U, q)$ such that, writing $K := SO(n)$,

$$\inf_{R \in K} \|Dv - R\|_{L^q(U)} \leq C\,(U, q)\, \|d\,(Dv, K)\,\|_{L^q(U)} \qquad \text{for every } v \in W^{1,q}(U; \mathbb{R}^n). \tag{1}$$

Here and below, $d(M, K)$ denotes the distance from a matrix $M \in \mathbb{R}^{n\times n}$ to a subset $K \subset \mathbb{R}^{n\times n}$, measured in the Euclidean norm. This result strengthens earlier work of a series of authors, including John [Jo 61],[Jo 61], Reshetnyak [Re 67], and Kohn [Ko 82], and it has had a number of important applications. For example, it is a main tool used to provide a relatively complete analysis of the gamma limit of thin elastic structures, [Fr-Ja-Mu 02], [Fr-Ja-Mu 06].

A number of works have extended the above result (1) to cover various larger classes of matrices than $SO\,(n)$. Faraco and Zhong proved the corresponding result with $K = \Pi SO\,(n)$ where $\Pi \subset \mathbb{R}_+ \backslash \{0\}$ is a compact set, [Fa-Zh 05]. Chaudhuri and Müller [Cha-Mu 03] and later De Lellis and Szekelyhidi [De-Se 06] considered a set of the form $K = SO\,(n)\, A \cup SO\,(n)\, B$ where A and B are *strongly incompatible* in the sense of Matos [Mat 92].

If we consider two compatible wells $K = SO\,(n)\, A \cup SO\,(n)\, B$, i.e. wells for which there exists matrices $X \in SO\,(n)\, A$, $Y \in SO\,(n)\, B$ with $\mathrm{rank}\,(X - Y) = 1$, then the example of a piecewise affine function u such that $\mathrm{Image}(Du) = \{X, Y\}$ shows that no exact analog of (1) can hold. In this paper we show, however, that a sort of 2-well theorem can hold provided one has suitable control over second derivatives; indeed, this remains true for collections of $m \geq 3$ wells $K = \cup_{i=1}^{m} SO(n)A_i$ satisfying certain algebraic conditions. As we will recall in greater detail below, most of our main conclusions are known in 2 dimensions, however all are new in $\mathbb{R}^n, n \geq 3$. The main result of this paper is

Theorem 1. *Let $p, q \geq 1$, let $A_1, A_2, \ldots A_m \in \mathbb{R}^{n\times n}$ be matrices of non-zero determinant, and let $K = \bigcup_{i=1}^{m} SO\,(n)\, A_i$. Suppose that $m = 2$, or that for each $i \in \{1, \ldots, m\}$, there exists*

$v_i \in S^{n-1}$ such that either

$$|A_i v_i| > |A_j v_i| \text{ for all } j \neq i \tag{2}$$

or

$$|v_i^T A_i^{-1}| > |v_j^T A_j^{-1}| \text{ for all } j \neq i. \tag{3}$$

Then for any bounded, open, $\Omega \subset \mathbb{R}^n$ and connected $\Omega' \subset\subset \Omega$ there exists positive constants $a < 1$ and $M < \infty$, depending on K, Ω, Ω', p, q, such that for any $u \in W^{1,p} \cap W^{2,q}(\Omega; \mathbb{R}^n)$ that satisfies

$$\frac{1}{\varsigma} \int_\Omega d^p\left(Du, K\right) + \varsigma^q \left|D^2 u\right|^q dx \leq a \tag{4}$$

for some $\varsigma \in (0,1]$, there exists $i \in \{1, \ldots m\}$ such that

$$\int_{\Omega'} d^p\left(Du, SO\left(n\right) A_i\right) dx \leq M \varsigma^{1/p}, \tag{5}$$

and if $p > 1$ there exists $R \in SO\left(n\right) A_i$ such that

$$\int_{\Omega'} |Du - R|^p \, dx \leq M \varsigma^{1/p}. \tag{6}$$

The theorem is interesting in when $0 < \varsigma \ll a$. The result as stated follows easily from the case when Ω is the unit ball in $\mathbb{R}^n$ and Ω' is some small subball, so we will mostly focus on this situation. The conclusions of the theorem are generally not true if $\Omega' = \Omega$, as long as compatible wells are allowed; this is easily seen by taking u to be a suitable mollification of a piecewise affine function whose gradient assumes exactly two values. An example in [Co-Sc 06b], Remark 6.1, shows that the scaling in (5), (6) is sharp.

Remark 1. We suspect that the theorem remains true whenever $m = 3, n \geq 2$, and we verify in Section 7 that for $m = n \geq 3$, the hypotheses of the theorem are *generically* satisfied as long as the n wells have the property that they cannot be partitioned into two disjoint subfamilies of wells with no rank-1 connections between them.

However, for $m = 4$ and any $n \geq 2$, one can find examples of matrices $A_1, \ldots, A_4$ such that the conclusions of the theorem fail for $K = \cup_{i=1}^4 SO(n) A_i$. To construct an example for $\Omega' \subset\subset \Omega \subset \mathbb{R}^2$, we start with a equilateral triangle $T \subset \Omega'$ of diameter ℓ, and we partition T into three congruent subtriangles S_1, S_2, S_3. Let $S_4 = \Omega \backslash T$. We can then find a piecewise affine function u_0 and matrices $A_1, \ldots A_4 \in M^{2\times 2}$ such that $Du_0 = A_i$ a.e. on S_i, for $i = 1, \ldots, 4$. Let $u_\varsigma = u_0 * \phi_\epsilon$ where $\phi_\epsilon := \epsilon^{-n} \phi\left(\frac{z}{\epsilon}\right)$ and ϕ is a standard mollifier on $\mathbb{R}^2$.

One can fix $\ell \lessapprox a$ such that u_ς satisfies (4) for every $\varsigma \ll l$. However, as $\varsigma \to 0$, $\int_{\Omega'} d^p(u_\varsigma, SO(n) A_i) dx \gtrapprox c\ell^2$ for every i, so the conclusions of the theorem do not hold.

The first 2-well Liouville Theorem was due to the second author [Lor 05], who established essentially the above result in the case when $m = n = 2$ and $p = q = 1$, for matrices $A, B \in \mathbb{R}^{2\times 2}$ with $\det A = \det B$, with suboptimal scaling in (6), and under the assumption that u is Lipschitz and invertible, with Lipschitz inverse. This was greatly improved by Conti and Schweizer, [Co-Sc 06a], who proved Theorem 1 for $q = 1$, still for $m = n = 2$. In particular [Co-Sc 06a] established this case of the theorem with the optimal scaling as in (5), (6), and without either the assumption of invertibility or any conditions on the two wells. A different proof of Theorem 1 for $n = m = 2$, valid for general $p, q \geq 1$, was given in [Lor pr 06]. This argument is conceptually simple, and the proof clarifies some technical issues in [Co-Sc 06a], but it yields suboptimal scaling in (6) and requires the assumption $\det A = \det B$.

1.1. **Ingredients in the proof.** As mentioned above, we work mostly on $\Omega = B_1 \subset \mathbb{R}^n$. Straightforward arguments from previous work, recalled in Section 2, allow us easily to find some $i_* \in \{1, \ldots, m\}$ and a large set $U_0 \subset B_1$ with small perimeter in B_1, such that $d(Du, SO(n)A_{i_*}) = d(Du, K)$ in U_0. We always assume for concreteness that $i_* = 1$. Our first goal is to find many pairs of points $(x, y) \subset U_0 \times U_0$ such that

$$|u(x) - u(y)| = |A_1(x - y)| + O(\varsigma^{1/p}). \tag{7}$$

Further easy and well-known arguments, also recalled in Section 2, allow us to find many pairs of points for which the inequality $|u(x) - u(y)| \leq |A_1(x - y)| + O(\varsigma^{1/p})$ holds. Following previous work, we wish to prove the opposite inequality by applying the same argument to u^{-1}. In general u is not invertible, but in fact it is only necessary to prove that there are many line segments along which u can be inverted. One of the important contributions of [Co-Sc 06a] was to introduce arguments, using tools from degree theory, to support this contention. Their local invertibility arguments, however, rest on the Sobolev embedding $W^{2,1} \hookrightarrow H^1$ (in ways that are not made completely explicit), and so do not apply to $\mathbb{R}^n$ for $n \geq 3$.

To address this difficulty we prove a new Lipschitz truncation result, showing that one can find a Lipschitz function w such that the set $\{x \in B_1 : w(x) \neq u(x)\}$ is not only small, but also can be contained in a set of small perimeter. The new point is the perimeter estimate, which follows from the control over second derivatives of u supplied by (4). The specific facts we need about this Lipschitz approximation are proved in Section 3. They are deduced from a general truncation result that we prove in Section 8. Using the Lipschitz approximation and some elements from earlier work of various authors, we find in Section 4 a large subset $\mathcal{W}$ of $u(B_1)$ on which an inverse map is well-defined and Lipschitz, with its gradient near $A_1^{-1}SO(n)$ and, crucially, with control over the perimeter of $\mathcal{W}$. This allows us in Section 5 to complete the proof that (7) holds for a large set of pairs of points.

The proof of Theorem 1 is given in Section 6. We first consider the case when the majority phase, represented by A_1, satisfies (2). Then we can bound $d(\cdot, SO(n)A_1)$ by $d(\cdot, K)$ + a null lagrangian, and it directly follows, via integration by parts, that

$$\sum_{k=0}^{n} \int_{[x_k, x]} d(Du, SO(n)A_1) dH^1 \leq C \sum_{k=0}^{n} \int_{[x_k, x]} d(Du, K) dH^1 + \text{ boundary terms} \tag{8}$$

where $x_0, \ldots, x_n$ are the vertices of a long, thin simplex with long axis roughly parallel to v_i, x is a point near the barycenter, and $[x_k, x]$ denotes the line segment joining x_k and x. The boundary terms have the form $C \sum_{k=0}^{n} |u(x_k) - l_R(x_k)|$, where l_R is an affine map with $Dl_R = R \in SO(n)A_1$. The inequality (8) recasts and extends ideas developed in [Co-Sc 06a] for $n = 2$. We present the short proof of (8) in the next subsection.

If the majority phase A_1 satisfies (2), then by using (7) and a linear algebra lemma proved in Section 8, we can find a vertices $x_0, \ldots, x_n$ and an affine l_R map such that the boundary terms in (8) are less than $C\varsigma^{1/p}$. The proof of Theorem 1 in this case is essentially completed by integrating (8) over points x near the barycenter.

When the majority phase A_1 satisfies (3), the idea of the proof is to apply to u^{-1} the argument already used to prove the theorem under assumption (2). The fact that u need not be invertible again causes technical difficulties. Thus, we work with the Lipschitz approximant w found earlier, and we use a lemma, proved in Section 8, which asserts roughly speaking that *almost every* line segment passing through a large convex subset of $w(B_1)$ can be realized as the image via w of a Lipschitz path in B_1. Although the restriction of w to these Lipschitz paths is not injective in general, this lemma provides a good enough proxy for invertibility to allow us to complete the proof of the theorem under the hypothesis (3). The null lagrangian calculation that leads to (8) is a bit harder to implement in the inverse direction, and in its place we use an argument more directly related to a proof given in [Co-Sc 06a] when $n = 2$.

Finally, it is easy to see that when $m = 2$, each well must satisfy at least one of (2), (3).

The condition (3) does not appear in any previous work, so that our result yields new information when $m \geq 3$, even in 2 dimensions. In particular, in 2 dimensions [Co-Sc 06a] essentially proves the theorem if every A_i satisfies either (2) or the condition that

$$\text{for each } j \neq i, \det A_j > \det A_i > 0. \tag{9}$$

Only the case of $m = 2$ wells is discussed in [Co-Sc 06a], but the argument works almost without change for $m \geq 2$ under the assumptions discussed here. The proof given under condition (9) is intrinsically 2-dimensional and so is not available here in the generality we consider here.

1.2. **Proof of** (8)**.** As discussed above, a crucial point in the proof of Theorem 1 in the case when hypothesis (2) holds is that if $\{x_0, \ldots, x_n\}$ are the vertices of a suitable simplex (where "suitability" is related to the algebraic condition (2)) then one can bound $\int d(Du, SO(n)A_1)$ by $\int d(Du, K)$ + *boundary terms* along certain lines. We illustrate how this works in the simplest possible case, that of a 2-well Liouville Theorem in $\mathbb{R}^1$. For this, suppose that $K = \{a_1, a_2\}$ for $a_2 < a_1 \in \mathbb{R}$, and consider $u : (-1, 1) \to \mathbb{R}$. Since $a_1 > a_2$, we can find constants c_1, c_2 such that

$$|s - a_1| = d(s, a_1) \leq c_1 d(s, K) + c_2(a_1 - s) \qquad \text{for all } s \in \mathbb{R}. \tag{10}$$

We substitute $s = u'$ in (10) and integrate. If we let l_1 be an affine function with $l_1' = a_1$, then $a_1 - u' = (l_1 - u)'$, and we find that

$$\int_{-1}^{1} d(u', a_1) \leq c_1 \int_{-1}^{1} d(u', K) + c_2 \left(|l_1(1) - u(1)| + |l_1(-1) - u(-1)|\right). \tag{11}$$

The next lemma, which is not used until Section 6, is essentially the same argument, but now for m wells in $\mathbb{R}^n$. Note that if $i = 1$ satisfies (2), then condition (12) below is fulfilled if $\{x_0, \ldots, x_n\}$ are the vertices of a long thin simplex roughly parallel to v_1.

Lemma 1. *Assume that* $\{A_1, \ldots, A_m\}$ *are* $n \times n$ *matrices and let* $K = \bigcup_i SO(n)A_i$. *Let* $x_0, \ldots, x_n \in B_1 \subset \mathbb{R}^n$ *be vertices of a simplex with the property that*

$$\left| A_1 \frac{x - x_i}{|x - x_i|} \right| > (1 + \alpha) \left| A_j \frac{x - x_i}{|x - x_i|} \right| \qquad \textit{for all } j \in \{2, \ldots, m\} \textit{ and } i \in \{0, \ldots, n\} \tag{12}$$

for some x in the interior of the simplex $\operatorname{conv}\{x_0, \ldots, x_n\}$. *Then there exists a constant C such that*

$$\sum_{i=0}^{n} \int_{[x_i, x]} d(Du, SO(n)A_1)\, dH^1 \;\leq\; C \sum_{i=0}^{n} \int_{[x_i, x]} d(Du, K)\, dH^1 \; + C \sum_{i=0}^{n} |u(x_i) - l_R(x_i)| \tag{13}$$

for every smooth $u : B_1 \to \mathbb{R}^n$ *and every affine map* l_R *with* $Dl_R = R \in SO(n)A_1$.

Moreover, if we write $x = \sum_{i=0}^{n} \lambda_i x_i$ *with* $\sum \lambda_i = 1$ *and* $\lambda_i > 0$ *for all i, and if* $\lambda_i |x - x_i| \geq \alpha' > 0$ *for all i, then the constant C in* (13) *are uniformly bounded by constants depending only on* $\{A_i\}$, α, α'.

This lemma is inspired by an argument from [Co-Sc 06a]. In the context of the 1-dimensional toy problem discussed above, the idea in [Co-Sc 06a] would be to use information about $\int_{-1}^{1} d(u', K)$ and the boundary behavior of u at ± 1 to bound $L^1(\{x \in (-1, 1) : d(u', a_2) < d(u', a_1)\})$. We use arguments of this sort in Lemma 9, when considering the hypothesis (3). In fact, either argument —integration by parts or direct estimates of the size of a bad set — could be used to prove both halves of Theorem 1.

Proof Lemma 1. Step 1. Fix $\{x_0, \ldots, x_n\}$ and $x = \sum_{i=0}^{n} \lambda_i x_i$ satisfying (12), where $0 < \lambda_i$ for all i, and $\sum \lambda_i = 1$. Also fix an affine map l_R with $Dl_R = R \in SO(n)A_1$.

For $i = 0, \ldots, n$, let us write $\tau_i := \frac{x - x_i}{|x - x_i|}$ and $v_i := \lambda_i R(x - x_i) = \lambda_i |x - x_i| R\tau_i$. Note that $\sum v_i = R(\sum \lambda_i (x - x_i)) = 0$.

We first claim that (12) implies that there exist $c_1, c_2 > 0$ such that

$$d(M, SO(n)A_1) \ \leq c_1 d(M, K) + c_2 v_i^T (R - M)\tau_i \tag{14}$$

for every $n \times n$ matrix M and every $i \in \{0, \ldots, n\}$. Here v_i, τ_i are column vectors, and v_i^T denotes the transpose of v_i. Inequality (14) is the analog of (10) from the 1-dimensional case. To prove (14), we write $\tilde\lambda_i = \lambda_i |x - x_i|$ for simplicity, and we note that since $R \in O(n)A_1$,

$$v_i^T R\tau_i = \tilde\lambda_i |R\tau_i|^2 = \tilde\lambda_i |A_1\tau_i|^2 > \tilde\lambda_i (1+\alpha)|A_1\tau_i| \ |A_j\tau_i|$$

for $j \geq 2$, using (12). Similarly $v_i^T M\tau_i \leq \tilde\lambda_i |A_1\tau_i| \ |M\tau_i|$. In particular, if $M \in SO(n)A_j$ for some $j \geq 2$, then $|M\tau_i| = |A_j\tau_i|$, and so

$$v_i^T (R - M)\tau_i \geq \tilde\lambda_j \alpha |A_1\tau_i| \ |A_j\tau_i| \geq c > 0 \qquad \text{for} \quad M \in \cup_{j=2}^m SO(n)A_j.$$

Also, if $M \in \cup_{j\geq 2} SO(n)A_j$, then $d(M, SO(n)A_1) \leq C(\{A_1, \ldots, A_m\})$, It follows that we can fix positive constants c_2 so large and δ so small that

$$d(M, SO(n)A_1) \ \leq c_2 v_i^T (R - M)\tau_i$$

say for all M such that $d(M, \cup_{j=2}^m SO(n)A_j) \leq \delta$. Then by choosing c_1 large enough, we can arrange that

$$d(M, SO(n)A_1) - c_2 v_i^T (R - M)\tau_i \ \leq c_1 d(M, K)$$

whenever $d(M, \cup_{j=2}^m SO(n)A_j) \geq \delta$. Then (14) follows.

Step 2. Now we substitute $M = Du$ in (14), so that $R - M$ becomes $R - Du = D(l_R - u)$. We then integrate to find that

$$\sum_{i=0}^n \int_{[x_i, x]} d(Du, SO(n)A_1)\, dH^1 \leq \ \sum_{i=0}^n \int_{[x_i, x]} \left[c_1 d(Du, K) + c_2 v_i^T (R - Du)\tau_i \right] \ dH^1 \tag{15}$$

Since $R - Du = D(l_R - u)$ and τ_i is tangent to $[x_i, x]$, we can integrate by parts to find that

$$\begin{aligned} \sum_{i=0}^n \int_{[x_i, x]} v_i^T (R - Du)\tau_i \ dH^1 &= \sum_i v_i^T \left[(l_R - u)(x) - (l_R - u)(x_i) \right] \\ &= -\sum_i v_i^T (l_R - u)(x_i) \end{aligned} \tag{16}$$

since $\sum_i v_i^T (l_R - u)(x) = \left(\sum_i v_i\right)^T (l_R - u)(x) = 0$. Now (13) follows by combining (15), (16).

The statement about dependence of the constants in (13) on various other parameters follows from inspection of the above argument. $\square$

2. Preliminaries

In this section we introduce some notation and reformulate some arguments from [Lor 05] that provide the starting point for our analysis.

2.1. **Some notation.** Given matrices $A_1, \ldots, A_m$, we always write $K = \cup_{i=1}^m SO(n)A_i$.

We write $B_r(x)$ for the *open* ball in $\mathbb{R}^n$ of radius r, centered at x. We write B_r as an abbreviation for $B_r(0)$. Define $[x, y]$ to denote the line segment joining x and y. If S is a subset of $\mathbb{R}^n$, then $\mathbb{1}_S$ always denotes the characteristic function of S, so that $\mathbb{1}_S(x) = 1$ if $x \in S$ and 0 otherwise.

We will write $\sigma = \sigma(K)$ to denote a fixed small number depending only on the given matrices $A_1, \ldots, A_m$. We select $\sigma \leq 1$ to satisfy

$$\sigma < \frac{1}{4} dist(SO(n)A_i, SO(n)A_j) \qquad \text{for all } i \neq j, \tag{17}$$

$$d(M, K) > \sigma \qquad \text{for any matrix } M \text{ such that } \det M < \sigma. \tag{18}$$

Note that (17) implies that $d(Du, K) = d(Du, SO(n)A_i)$ whenever $d(Du, SO(n)A_i) < 2\sigma$.

All constants throughout, including generic constants C that appear in many estimates, as well as named constants such as κ_0 in Proposition 1 for example, may depend on the collection K of wells, the dimension n, and the powers p, q appearing in assumptions (4), for example. but are independent of the parameters ς, a.

We often (though not always) use latin letters to refer to the reference configuration B_1 and greek letters to refer to the image $u(B_1)$. Thus points in B_1 will be denoted x, y, z, whereas points in the image will be denoted ξ, η, ζ. In addition we will write β_ρ to denote an ellipsoid in the image with length-scale ρ; in fact β_ρ will be defined as $\beta_\rho = l_R(B_\rho)$, where l_R is a particular affine map we find that is close to u, see Section 3.

2.2. **Finding a majority phase.**

Lemma 2. *Let $K = \cup_{i=1}^m SO(n)\, A_i$. Let $u : B_1 \subset \mathbb{R}^n \to \mathbb{R}^n$ be a smooth function such that*

$$\frac{1}{\varsigma} \int_{B_1} \left(d^p (Du, K) + \varsigma^q \left| D^2 u \right|^q \right) dx \leq a. \tag{19}$$

Then we can find $i \in \{1, \ldots, m\}$ an open set $U_0 \subset B_1$ with smooth boundary such that

$$\mathrm{Per}_{B_1} (U_0) < Ca \text{ and } L^n (B_1 \backslash U_0) < Ca^{\frac{n}{n-1}}, \tag{20}$$

and

$$d(Du, SO(n)A_i) = d(Du, K) < \sigma \qquad \text{for all } x \in U_0. \tag{21}$$

We take U_0 to be smooth because it is convenient later to identify $\mathrm{Per}_{B_1} (U_0)$ with $H^{n-1}(B_1 \cap \partial U_0)$.

Proof of Lemma 2. Let q^* be the Holder conjugate of q, and let $s := 1 + \frac{p}{q^*}$ and $J(x) := d^{s} (Du(x), K)$. If $q^* = \infty$ we use the convention that $d^{\frac{p}{q^*}} (\cdot, K) := 1$. We have by Young's inequality

$$\int_{B_1} |DJ|\, dx \leq C \int_{B_1} d^{\frac{p}{q^*}} (Du, K) \left| D^2 u \right| dx \;\leq\; \frac{C}{\varsigma} \int_{B_1} d^p (Du, K) + \varsigma^q \left| D^2 u \right|^q dx \overset{(19)}{\leq} Ca. \tag{22}$$

Then by the coarea formula, we can find $\alpha \in \left((\frac{\sigma}{2})^s, \sigma^s \right)$ with $Per_{B_1}(\{x \in B_1 : J(x) < \alpha\}) \leq Ca$. Note that

$$\bigcup_{i=1}^{n} \{x \in B_1 : d^s (Du, SO(n)\, A_i) < \alpha\} = \{x \in B_1 : J(x) < \alpha\}.$$

Since the sets on the left-hand side above are disjoint by the choice (17) of σ, it follows that

$$Per_{B_1}(\{x \in B_1 : d^s (Du, SO(n)\, A_i) < \alpha\}) \;\leq\; Per_{B_1}(\{x \in B_1 : J(x) < \alpha\}) \;\leq\; Ca$$

for every i. So by the relative isoperimetric inequality we have

$$\begin{aligned} &\min \{L^n (\{x \in B_1 : d^s (Du, SO(n)\, A_i) > \alpha\}), L^n (\{x \in B_1 : d^s (Du, SO(n)\, A_i) < \alpha\})\} \\ &\qquad < Ca^{\frac{n}{n-1}} \end{aligned}$$

for every i. Since $\int d^p(Du, K) < a\varsigma$, it cannot be the case that $\{d^s (Du, SO(n)\, A_i) < \alpha\}$ has small measure for every i, and since a is small, there can be at most one i such that $L^n (\{x \in B_1 : d^s (Du, SO(n)\, A_i) < \alpha\}) > 1 - Ca^{\frac{n}{n-1}}$. We define

$$U_0 := \{x \in B_1 : d^s (Du, SO(n)\, A_i) < \alpha\} \text{ for this choice of } i.$$

Since J is a C^1 function by Sard's Theorem the image under J of the critical points of J have zero L^1 measure, so we can assume we choose α so that the level set $J^{-1}(\alpha)$ does not intersect the set of critical points of J. Then ∂U_0 is smooth, as required.

□

Upon relabeling, we may assume that $i = 1$ in Lemma 2, so that U_0 satisfies

$$U_0 \subset \{x \in B_1 : d\left(Du\left(x\right), SO\left(n\right) A_1\right) < \sigma\}. \tag{23}$$

It would of course be possible to perform a change of variables that sets A_1 equal to the identity matrix. We will mostly remain in the original coordinates, so that one can see explicitly where A_1 appears in our arguments.

2.3. **Non-stretching pairs.** We next show that can find many pairs of points that are not stretched by u (relative to the affine maps with gradient in $SO(n)A_1$). The argument we give is somewhat more complicated than necessary for the present lemma, but it will be needed again in Section 5.

Lemma 3. *Assume $u : B_1 \subset \mathbb{R}^n \to \mathbb{R}^n$ is a smooth function that satisfies (19) and that A_1 is the majority phase as in (23). Then there exists $\mathcal{G}_1 \subset B_1 \times B_1$ such that*

$$L^{2n}((B_1 \times B_1) \setminus \mathcal{G}_1) \leq Ca^{1/p} \tag{24}$$

and letting $\epsilon = \varsigma^{\frac{1}{p}}$,

$$\textit{if } (x, y) \in \mathcal{G}_1, \textit{ then } |u\left(y\right) - u\left(x\right)| \leq |A_1(y - x)| + C\epsilon. \tag{25}$$

Proof. We define

$$\mathcal{G}_1 := \{(x, y) \in B_1 \times B_1 \ : \ [x, y] \subset U_0, \quad \int_{[x,y]} d(Du, SO(n)A_1) \ dH^1 \ \leq \epsilon\}. \tag{26}$$

Note from (19) we have

$$\int_{B_1} d\left(Du, K\right) \leq C \left(\int_{B_1} d^p\left(Du, K\right)\right)^{\frac{1}{p}} \leq Ca^{\frac{1}{p}}\epsilon. \tag{27}$$

Step 1 . To prove (25), we fix $(x, y) \in \mathcal{G}_1$, and we write $\tau := \frac{y-x}{|y-x|}$. Note that if M is any $n \times n$ matrix, then, $|M\tau| \leq |A_1\tau| + Cd(M, SO(n)A_1)$. Thus

$$\begin{aligned} |u(y) - u(x)| &= \left| \int_{[x,y]} Du(z) \ \tau \ dH^1 z \right| \\ &\leq \int_{[x,y]} \left[|A_1\tau| + C\, d(Du, SO(n)A_1) \right] dH^1 \ \leq \ |A_1(x - y)| + C\epsilon \end{aligned}$$

for $(x, y) \in \mathcal{G}_1$.

Step 2. We next prove (24). If $(x, y) \in (B_1 \times B_1) \setminus \mathcal{G}_1$, then at least one of the following must hold: either x or y fails to belong to U_0, that is

$$(x, y) \in [B_1 \setminus U_0] \times B_1 \text{ or } (x, y) \in B_1 \times [B_1 \setminus U_0]; \tag{28}$$

or the segment $[x, y]$ meets $\partial U_0 \cap B_1$, that is

$$[x, y] \cap (\partial U_0 \cap B_1) \ \neq \ \emptyset; \tag{29}$$

or

$$\int_{[x,y]} 1\!\!1_{U_0} \ d(Du, SO(n)A_1) \ dH^1 \ > \epsilon. \tag{30}$$

We saw in Lemma 2 that $L^n(B_1 \setminus U_0) \leq Ca^{\frac{n}{n-1}}$, so clearly (28) holds on a set of measure at most $Ca^{\frac{n}{n-1}} \leq Ca$. And Lemma 5 (proved at the end of this section) shows that

$$L^{2n}(\{(x, y) \ : \ \ (29) \text{ holds } \}) \leq CH^{n-1}(\partial U_0 \cap B_1).$$

However, in Lemma 2 we showed that $H^{n-1}(\partial U_0 \cap B_1) \leq Ca$. Finally, Lemma 4 (proved immediately below) implies that

$$L^{2n}(\{(x,y) \in B_1 \times B_1 \ : \ (30) \text{ holds } \}) \leq \frac{C}{\epsilon} \int_{B_1} 1\!\!1_{U_0} \ d(Du, SO(n)A_1) \overset{(21)}{\leq} \frac{C}{\epsilon} \int_{B_1} d(Du, K) \overset{(27)}{\leq} Ca^{1/p}.$$

Together, these estimates imply (24).

□

The first of the lemmas used above is

Lemma 4. *Suppose that $f : B_1 \to \mathbb{R}$ is nonnegative and integrable. Then for any constant $b > 0$,*

$$L^{2n}\left(\{(x,y) \in B_1 \times B_1 : \int_{[x,y]} f \ dH^1 > b\}\right) \leq \frac{C}{b} \int_{B_1} f.$$

Proof. We extend f by 0 on the complement of B_1. Then by a change of variables, we find that

$$\begin{aligned}\int_{B_1} \int_{B_1} \int_{[x,y]} f \ dH^1 \ dy \ dx &= \int_{B_1} \int_{B_1} \int_0^1 f(x + s(y-x))|y-x| \ ds \ dy \ dx \\ &\leq \int_{B_1} \int_{B_2} \int_0^1 f(x+sp)|p| \ ds \ dp \ dx.\end{aligned}$$

But $\int_{B_1} \int_{B_2} \int_0^1 f(x+sp)|p| \ ds \ dp \ dx \leq \|f\|_{L^1(B_1)} \int_{B_2} \int_0^1 |p| \ ds \ dp = C\|f\|_{L^1(B_1)}$ by Fubini's Theorem, so the lemma follows by Chebyshev's inequality. □

The other lemma we used is

Lemma 5. *Suppose that Ω is a bounded, open convex subset of $\mathbb{R}^n$. Then there exists a constant $C = C(\Omega)$ such that for any set $S \subset \mathbb{R}^n$,*

$$L^{2n}\left(\{(x,y) \in \Omega \times \Omega : [x,y] \cap S \neq \emptyset\}\right) \leq CH^{n-1}_{\infty}(S) \tag{31}$$

where $H^{n-1}_{\infty}(S) := \inf\{\sum_i \gamma_{n-1} s_i^{n-1} : \ S \subset \bigcup_i B_{s_i}(x_i)\}$.

The constant γ_{n-1} appearing in the definition of H^{n-1}_{∞} is the same normalization factor appearing in the definition of Hausdorff measure, so that $H^{n-1}_{\infty}(S) \leq H^{n-1}(S)$ for every S.

Proof of Lemma 5 . Without loss of generality we assume $0 \in \Omega$. For any $S \subset \mathbb{R}^n$, we will write

$$\varphi(S) := L^{2n}\left(\{(x,y) \in \Omega \times \Omega : [x,y] \cap S \neq \emptyset\}\right).$$

We first claim that $\varphi(B_r(p)) \leq Cr^{n-1}$ for any $p \in \mathbb{R}^n$ and $r > 0$. To prove this, note that by Fubini's Theorem,

$$\begin{aligned}\varphi(B_r(p)) &= \int_{y \in \Omega} L^n(\{x \in \Omega \ : [x,y] \cap B_r(p) \neq \emptyset\} \ ) \ dy \\ &\leq \int_{|z| \leq \mathrm{diam}(\Omega)} L^n(\{x \in \Omega \ : [x, x+z] \cap B_r(p) \neq \emptyset\} \ ) \ dz.\end{aligned}$$

And for every fixed $z \neq 0$, if $[x, x+z] \cap B_r(p) \neq \emptyset$, then x belongs to the cylinder of radius r, with axis parallel to z, that contains $B_r(p)$. The intersection of such a cylinder with Ω has L^n measure bounded by Cr^{n-1}. Thus $L^n(\{x \in \Omega \ : [x, x+z] \cap B_r(p) \neq \emptyset\} \ ) \leq Cr^{n-1}$ for every $z \neq 0$. The claim follows.

Now given $S \subset \mathbb{R}^n$, let $\{B_{r_i}(p_i)\}$ be a collection of balls such that

$$S \subset \bigcup_i B_{r_i}(p_i) \qquad \text{and} \qquad \sum_i \gamma_{n-1} r_i^{n-1} \leq 2H_\infty^{n-1}(S).$$

Then any segment $[x, y]$ that intersects S also intersects some ball $B_{r_i}(p_i)$, so we deduce that

$$\varphi(S) \leq \sum_i \varphi(B_{r_i}(p_i)) \leq C \sum_i r_i^{n-1} \leq C H_\infty^{n-1}(S).$$

□

3. Lipschitz approximation

In this section we find a Lipschitz function w that agrees with u on the complement of a small set E and that is close to affine if a is small. Such arguments are standard. The main new ingredient here, which is crucial for our later arguments, is that we use information about the second derivatives of u to control the perimeter of the set $E = \{x : u(x) \neq w(x)\}$, or more precisely, of a set that contains E.

Proposition 1. *Suppose the smooth function $u : B_1 \subset \mathbb{R}^n \to \mathbb{R}^n$ satisfies (19), and assume as in (23) that $SO(n)A_1$ is the majority phase.*

Then there is a Lipschitz function $w : B_1 \to \mathbb{R}^n$ with $\|Dw\|_{L^\infty(B_1)} \leq C(K)$ and an open set $U_1 \subset B_1$ with countably piecewise smooth boundary such that $u = w$ in U_1, and letting $\epsilon = \varsigma^{\frac{1}{p}}$ the following hold:

(i) $\|d\,(Dw, K)\,\|_{L^1(B_1)} \leq C\epsilon.$

(ii) $d\,(Dw\,(x)\,, SO\,(n)\,A_1) = d\,(Dw\,(x)\,, K)$ *for every* $x \in U_1$.

(iii) $L^n\,(B_1 \backslash U_1) \leq ca^{\frac{n}{n-1}}$, *and* $Per_{B_1}\,(U_1) \leq ca.$

(iv) *There exists $R \in SO\,(n)\,A_1$ and an affine map l_R with $Dl_R = R$ such that*

$$\|w - l_R\|_{L^\infty(B_1)} \leq ca^{\frac{1}{n+1}}.$$

(v) *There exists $\kappa_0 > 0$ such that for $\rho_0 := 1 - \kappa_0 a^{\frac{1}{n+1}}$ and $\beta_{\rho_0} := l_R(B_{\rho_0})$ we have*

  (a) $H^{n-1}\,(w\,(\partial U_1) \cap \beta_{\rho_0}) \leq ca,$

  (b) $\deg(w, B_1, \xi) = 1$ *for all* $\xi \in \beta_{\rho_0}$.

The ellipsoid $\beta_\rho := l_R(B_\rho)$ should be thought of as the counterpart in the image $u(B_1)$ of the ball B_ρ in the reference configuration.

Proof. We will apply a general truncation result, Lemma 11, which is proved in an Appendix. Toward this end, will write $f(Du) = d^s(Du, K)$, where $s = 1 + \frac{p}{q^*}$ and q^* is the Holder conjugate of q. We first claim that

$$\|f(Du)\|_{W^{1,1}(B_1)} \leq ca. \tag{32}$$

We have already proved in (22) that $\|D(f(Du))\|_{L^1(B_1)} \leq a$, so we only need to estimate $\|f(Du)\|_{L^1}$. To do this, let us temporarily write $g(x) = d(Du\,(x)\,, K)$, so that $f(Du) = g^s$. Then (assuming ϵ is small enough) (32) follows from (19), (22), and the inequality

$$\|g^s\|_{L^1(B_1)} \leq C\left(\|g\|_{L^1(B_1)}^s + \|D(g^s)\|_{L^1(B_1)}\right) \tag{33}$$

since the terms on the right-hand side are just $\|d(Du, K)\|_{L^1}^s$ and $\|D(f(Du))\|_{L^1}$.

To prove (33), we use Holder's inequality and the Sobolev embedding theorem to deduce that

$$\begin{aligned}\|g^s\|_{L^1(B_1)} &= \|g\|^s_{L^s(B_1)} \leq C\|g\|^{s\theta}_{L^1(B_1)} \; \|g\|^{s(1-\theta)}_{L^{ns/(n-1)}(B_1)} \qquad \text{for some } \theta \in (0,1)\\ &= C\|g\|^{s\theta}_{L^1(B_1)} \; \|g^s\|^{(1-\theta)}_{L^{n/n-1}(B_1)} \\ &\leq C\|g\|^{s\theta}_{L^1(B_1)} \left(\|g^s\|_{L^1(B_1)} + \|D(g^s)\|_{L^1(B_1)}\right)^{(1-\theta)}.\end{aligned}$$

Then Young's inequality $ab \leq \theta a^{1/\theta} + (1-\theta)b^{1/(1-\theta)}$ implies that

$$\|g^s\|_{L^1(B_1)} \leq C\theta\|g\|^s_{L^1(B_1)} + (1-\theta)(\|g^s\|_{L^1(B_1)} + \|D(g^s)\|_{L^1(B_1)})$$

which proves (33).

We now fix $\lambda = \lambda(K)$ large enough that $d(Du, K) \geq \frac{1}{2}|Du|$ whenever $|Du| \geq \lambda$, and we apply Lemma 11 to u with this choice of λ, and with $q = 1$ and f as above, so that $f(Du) \geq |Du| - c$. This produces a Lipschitz function $w : B_1 \to \mathbb{R}^n$ with $\|Dw\|_{L^\infty(B_1)} \leq C\lambda = C$. From conclusion *(ii)* of Lemma 11,

$$\|Du - Dw\|_{L^1(B_1)} \leq \frac{C}{\lambda}\int_{\{x\in B_1:|Du(x)|>\lambda\}} |Du| \overset{(19)}{\leq} C\epsilon. \tag{34}$$

Since $d(Dw, K) \leq d(Du, K) + |Du - Dw|$, it follows that $\int_{B_1} d(Dw, K) \leq C\epsilon$. Thus we have proved *(i)*.

Now let U_0 be the set constructed in Lemma 2. Recall that $d(Du, SO(n)A_i) = d(Du, K)$ in U_0, so that

$$\begin{aligned}\int_{B_1} d\left(Dw, SO\left(n\right)A_i\right) &\leq \int_{U_0} d\left(Du, K\right) + |Du - Dw|\,dx + cL^n\left(B_1\backslash U_0\right)\\ &\overset{(19),(20),(34)}{\leq} Ca^{\frac{n}{n-1}}.\end{aligned}$$

So by the one-well L^1 Liouville Theorem[1] ([Co-Sc 06a] Proposition 2.6) there exists $R \in SO\left(n\right)A_i$ such that $\int_{B_1} |Dw - R| \leq Ca$. And by Poincaré there exists an affine function l_R with $Dl_R = R$ such that $\int_{B_1} |w - l_R| \leq Ca$. By an interpolation inequality, Theorem 5.9 [Ad-Fo 00], this gives

$$\|w - l_R\|_{L^\infty(B_1)} \leq C\|w - l_R\|^{\frac{1}{n+1}}_{L^1(B_1)}\|w - l_R\|^{\frac{n}{n+1}}_{W^{1,\infty}(B_1)} \leq Ca^{\frac{1}{n+1}}(\|w - l_R\|_{L^\infty(B_1)} + C)^{\frac{n}{n+1}},$$

and this is easily seen to imply *(iv)*.

Next, Lemma 11 also asserts that there exists an open set $E' \subset B_1$ with smooth boundary, such that $E := \{x \in B_1 : u(x) \neq w(x)\} \subset E'$, and

$$L^n(E')^{\frac{n-1}{n}} + Per_{B_1}(E') \leq \|f(Du)\|_{W^{1,1}} \overset{(32)}{\leq} Ca.$$

We define $U_1 := U_0 \setminus E'$. Then conclusion *(ii)* is immediate and conclusion *(iii)* follows directly from the above estimates of E' and corresponding properties of U_0 from Lemma 2.

We now fix a constant $\kappa_0 > 0$ such that, if we define $\rho_0 := 1 - \kappa_0 a^{\frac{1}{n+1}}$, then

$$\beta_{\rho_0} = l_R(B_{\rho_0}) \subset w(B_1) \setminus w(\partial B_1). \tag{35}$$

This is possible due to conclusion *(iv)*. It follows that

$$H^{n-1}(w(\partial U_1) \cap \beta_{\rho_0}) \leq H^{n-1}\left(w\left(\partial U_1 \cap B_1\right)\right) \leq Ca$$

[1] Strictly speaking, we should change variables in a way that transforms $SO(n)A_i$ into $SO(n)$, apply the L^1 Theorem, and then change variables back; this is justified, since the theorem we are citing is valid on any Lipschitz domain.

using conclusion *(iii),* together with the fact that w is Lipschitz, so we have shown *(v), (a).* Finally, for $t \in [0,1]$ and $x \in B_1$, define $w_t(x) = tw(x) + (1-t)l_R(x)$. It follows from conclusion *(iv)* that, taking κ_0 larger if necessary, $w_t(\partial B_1) \cap \beta_{\rho_0} = \emptyset$ for every $t \in [0,1]$. Thus the homotopy invariance of degree implies that for $\xi \in \beta_{\rho_0}$,

$$\deg(w, B_1, \xi) = \deg(w_1, B_1, \xi) = \deg(w_0, B_1, \xi) = \deg(l_R, B_1, \xi) = 1.$$

This completes the proof of *(v).* □

4. Partial invertibility of u

The main result of this section is the following

Proposition 2. *Suppose $u : B_1 \subset \mathbb{R}^n \to \mathbb{R}^n$ is a smooth function satisfying* (19)*, with the wells labeled as in* (23) *so that A_1 is the majority phase, and let $U_1 \subset B_1$ be the set found in Proposition 1. Then there exists a smooth open set $\mathcal{W}_1 \subset \mathbb{R}^n$ and a C^1 function $v : \mathcal{W}_1 \to U_1$ such that $\|Dv\|_{L^\infty(\mathcal{W}_1)} \leq C$,*

$$u(v(\xi)) = \xi \text{ for all } \xi \in \mathcal{W}_1, \tag{36}$$

and for $\epsilon = \varsigma^{\frac{1}{p}}$

$$\int_{\mathcal{W}_1} d(Dv, A_1^{-1}SO(n))\ dx \leq\ C\epsilon \tag{37}$$

Moreover, there is an affine map l_R with $Dl_R = R \in SO(n)A_1$ and a constant κ_1 such that for $\rho_1 := 1 - \kappa_1 a^{\frac{1}{n+1}}$ and $\beta_{\rho_1} := l_R(B_{\rho_1})$, the following hold:

$$\mathcal{W}_1 \subset \beta_{\rho_1} \quad Per_{\beta_{\rho_1}} \mathcal{W}_1 \leq Ca, \quad L^n(\beta_{\rho_1} \setminus \mathcal{W}_1) \leq Ca^{\frac{n}{n-1}} \tag{38}$$

Finally, there exists a constant $k_1 > \kappa_1$ such that for $r_1 = 1 - k_1 a^{\frac{1}{n+1}}$,

$$L^n(B_{r_1} \setminus v(\mathcal{W}_1)) \leq Ca^{\frac{n}{n-1}}. \tag{39}$$

We start by proving a lemma in which we find a set $D \subset B_1$ of small measure, and a radius r_0 close to 1, such that the Lipschitz approximation w found earlier is one-to-one on $B_{r_0} \setminus D$. We do not however have any information about the perimeter of D. The properties (41), (42) of D that we record in the statement of the Lemma are consequences of these facts that will be useful in our later arguments. The proof is follows arguments from [Co-Sc 06a].

Lemma 6. *Let β_{ρ_0} be as defined in Proposition 1 (v). Suppose $w : B_1 \to \mathbb{R}^n$ is a Lipschitz function that satisfies (i), (iv) and (v) (b) of Proposition 1. Let $V := \{x \in B_1 : det\,(Dw\,(x)) \leq \sigma\}$ and define*

$$D := \{x \in B_1 : w\,(x) \in w\,(V) \cap \beta_{\rho_0}\}\,. \tag{40}$$

Then $L^n\,(D) < C\epsilon$. In addition there exists a constant $k_0 > 0$ such that for $r_0 := 1 - k_0 a^{\frac{1}{n+1}}$, function w is one-to-one on $B_{r_0} \backslash D$ and for any set $S \subset B_{r_0}$,

$$\deg(w, S, \xi) = 1 \quad \text{ for } \xi \in w(S \setminus D) = w(S) \setminus w(V); \tag{41}$$

and

$$S \backslash D = w^{-1}\,(w\,(S))\,\backslash D. \tag{42}$$

This lemma will be used not only in the proof of Proposition 2, but also in the proof of Lemma 9.

Proof. Recall in Proposition 1 we defined $\rho_0 := 1 - \kappa_0 a^{\frac{1}{n+1}}$, where the constant κ_0 was chosen so that $\deg(w, B_1, \xi) = 1$ for all $\xi \in \beta_{\rho_0} = l_R(B_{\rho_0})$.

We choose k_0 so large that $w(B_{r_0}) \subset \beta_{\rho_0}$. This is possible by Proposition 1, *(iv).*

Step 1. We first verify (41) and (42). If $S \subset B_{r_0}$, then the choice of r_0 and the fact that w satisfies *(v) (b)* of Proposition 1 imply that $\deg(w, B_1, \xi) = 1$ for any $\xi \in w(S)$. Thus for $\xi \in w(S) \backslash w(V)$ we have

$$\text{Card}\left(w^{-1}(\xi)\right) = \sum_{x \in w^{-1}(\xi)} \text{sgn}\left(\det\left(Dw(x)\right)\right) = 1,$$

Now if $x \in w^{-1}(w(S)) \backslash D$, then $w(x) \in w(S) \backslash w(V)$. Hence $w(x)$ has a unique preimage, which necessarily belongs to S. Thus

$$w^{-1}(w(S)) \backslash D \subset S \backslash D.$$

The opposite inclusion is obvious, so we have proved (42). Similarly, if $\xi \in w(S \setminus D)$ then $w^{-1}(\xi)$ consists of one point, say x, which evidently belongs to $S \setminus D \subset S \setminus V$. This implies that $\det Dw(x) > \sigma$. Thus $\text{sgn}(\det(Dw(x))) = 1$, and (41) follows.

Step 2. We next show that $L^n(D) \leq C\epsilon$.

Note that from the choice (18) of σ and since w satisfies *(i)* of Proposition 1 we know that $L^n(V) \leq C\epsilon$.

Recall the change of variables degree formula (see [Fo-Ga 95] for example)

$$\int_{\mathbb{R}^n} \psi(\xi)\ deg(w, A, \xi)\, d\xi = \int_A \psi(w(x)) \det(Dw(x))\, dx \tag{43}$$

for open $A \subset B_1$ and $\psi \in L^\infty(\mathbb{R}^n)$. We define $\psi(\xi) = \mathbb{1}_{w(V) \cap \beta_{\rho_0}}(\xi)$, so that the definition (40) of D implies that $\psi(w(x)) = \mathbb{1}_D(x)$. Then (43) yields

$$\begin{aligned} \int_{w(V) \cap \beta_{\rho_0}} deg(w, B_1 \backslash V, \xi)\, d\xi &= \int_{D \cap (B_1 \backslash V)} \det(Dw(x))\, dx \\ &\geq \sigma\ L^n(D \backslash V). \end{aligned} \tag{44}$$

Recall from Proposition 1 (v) that $deg(w, B_1, \xi) = 1$ for all $\xi \in \beta_{\rho_0}$. Thus

$$\begin{aligned} deg(w, B_1 \backslash V, \xi) &\leq 1 + |deg(w, V, \xi)| \\ &\leq 1 + \text{Card}\left(w^{-1}(\xi) \cap V\right). \end{aligned}$$

Note that $\int_{w(V)} \text{Card}\left(w^{-1}(\xi) \cap V\right) d\xi = \int_V |\det(Dw(x))|\, dx \leq \|Dw\|^n_{L^\infty} L^n(V) \leq C\epsilon$. Similarly, $\int_{w(V)} d\xi = L^n(w(V)) \leq \|Dw\|^n_\infty L^n(V) \leq C\epsilon$. Thus

$$\begin{aligned} \int_{w(V) \cap \beta_{\rho_0}} deg(w, B_1 \backslash V, \xi)\, d\xi &\leq \int_{w(V) \cap \beta_{\rho_0}} 1 + \text{Card}\left(w^{-1}(\xi) \cap V\right) d\xi \\ &\leq C\epsilon. \end{aligned}$$

Putting this together with (44) we have $L^n(D \backslash V) \leq C\epsilon$. Since we know $L^n(D \cap V) \leq C\epsilon$ this establishes $L^n(D) \leq C\epsilon$. □

Lemma 7. *Let $w : B_1 \to \mathbb{R}^n$ be a Lipschitz function satisfying the conclusions of Proposition 1. If $S \subset U_1 \cap B_{r_0}$, for r_0 as defined in Lemma 6, then*

$$\left|L^n(w(S)) - \det A_1 L^n(S)\right| \leq \epsilon. \tag{45}$$

Proof. For $S \subset U_1 \cap B_{r_0}$ it follows from Proposition 1 *(i), (ii)* that $\int_S d(Dw, SO(n)A_1) \leq C\epsilon$. Then using the fact (Proposition 1 *(v)*) that $\deg(w, B_1, \xi) = 1$ for every $\xi \in w(S) \subset \beta_{\rho_0}$, we compute

$$\begin{aligned}
L^n(w(S)) &= \int_{\mathbb{R}^n} \mathbb{1}_{w(S)}(\xi)\ \deg(w, B_1, \xi)\ d\xi \\
&\overset{(43)}{=} \int_{w^{-1}(w(S))} \det Dw(x)\ dx \\
&= \int_{w^{-1}(w(S))\setminus D} \det Dw(x)\ dx + \int_{w^{-1}(w(S))\cap D} \det Dw(x)\ dx
\end{aligned}$$

where D was defined in the previous lemma, in which we also proved that $L^n(D) \le C\epsilon$. To estimate the second term, note that $|\int_{w^{-1}(w(S))\cap D} \det Dw(x)\ dx| \le CL^n(D) \le C\epsilon$. And in view of (42),

$$\begin{aligned}
\int_{w^{-1}(w(S))\setminus D} \det Dw(x)\ dx &= \int_{S\setminus D} \det Dw(x)\ dx \\
&= \det A_1[L^n(S) - L^n(S\cap D)] + \int_{S\setminus D} (\det Dw(x) - \det A_1)\ dx.
\end{aligned}$$

Since Dw is Lipschitz, $\int_S |\det Dw - \det A_1| \le C\int_S d(Dw, SO(n)A_1) \le C\epsilon$. So combining the above inequalities and using again the fact that $L^n(D) \le C\epsilon$, we obtain (45). □

Proof of Proposition 2. Throughout the proof we will use notation introduced in Proposition 1. Recall also that $r_0 = 1 - k_0 a^{\frac{1}{n+1}}$ was fixed in Lemma 6. We fix $\rho_1 = 1 - \kappa_1 a^{\frac{1}{n+1}}$ by choosing $\kappa_1 > k_0$ large enough that $l_R(B_{\rho_1}) = \beta_{\rho_1} \subset w(B_{r_0}) \setminus w(\partial B_{r_0})$. This is possible due to Proposition 1 *(iv)*.

Next, we define

$$\mathcal{W}_1 := \{\xi \in \beta_{\rho_1}\ :\ \deg(w, U_1 \cap B_{r_0}, \xi) = 1\}. \tag{46}$$

Step 1. First we establish some properties of $\mathcal{W}_1$. General facts about degree imply that $\deg(w, U_1\cap B_{r_0}, \cdot)$ is locally constant in $\mathbb{R}^n \setminus w(\partial(U_1 \cap B_{r_0}))$. Since $w(\partial(U_1\cap B_{r_0}))$ is closed, it follows that $\mathcal{W}_1$ is open. In addition, we deduce that

$$\partial\mathcal{W}_1 \cap \beta_{\rho_1} \subset w(B_{r_0} \cap \partial U_1).$$

Since w is Lipschitz, it follows from conclusion *(v)* of Proposition 1 that

$$Per_{\beta_{\rho_1}}(\mathcal{W}_1) \le H^{n-1}(w(B_{r_0}\cap\partial U_1))\ \le CH^{n-1}(B_{r_0}\cap \partial U_1) \le Ca. \tag{47}$$

Next, recall that Lemma 6 implies that $\deg(w, U_1\cap B_{r_0}, \xi) = 1$ for every $\xi \in w((U_1\cap B_{r_0})\setminus D)$, see (41), where $D \subset B_1$ is defined in (40) and has the property that $L^n(D) \le C\epsilon$. Again using Proposition 1 *(iv)*, we know that $w(B_{1/2}) \subset \beta_{\rho_1}$ if a is small enough. For such a, it follows that $w((U_1 \cap B_{1/2})\setminus D) \subset \mathcal{W}_1$, and Lemma 7 with Proposition 1 *(i)*, *(ii)* implies that

$$L^n(\mathcal{W}_1) \ge L^n(w((U_1\cap B_{\frac{1}{2}})\setminus D)\,) \ge\ \det A_1\ L^n((U_1\cap B_{\frac{1}{2}})\setminus D) - C\epsilon.$$

Then Proposition 1 *(iii)* and the fact that $L^n(D) \le C\epsilon$ yield $L^n(\mathcal{W}_1) \ge \det A_1 L^n(B_{1/2}) - Ca^{\frac{n}{n-1}} - C\epsilon$.

On the other hand, we know from (47) and the relative isoperimetric inequality that

$$\min\{L^n(\mathcal{W}_1), L^n(\beta_{\rho_1}\setminus\mathcal{W}_1)\} \le Ca^{\frac{n}{n-1}}$$

Combining these facts, we conclude that $L^n(\beta_{\rho_1}\setminus\mathcal{W}_1) \le Ca^{\frac{n}{n-1}}$.

Thus we have verified that $\mathcal{W}_1$ has all the properties asserted in (38).

Step 2. For $\xi \in \mathcal{W}_1$, recalling that $\det Dw > 0$ in U_1, we deduce from the definition of $\mathcal{W}_1$ that

$$1 = \deg(w, U_1 \cap B_{r_0}, \xi) = \sum_{\{x \in U_1 \cap B_{r_0} : w(x) = \xi\}} \operatorname{sgn}\left(\det\left(Dw\left(x\right)\right)\right)$$
$$= \operatorname{Card}(\{x \in U_1 \cap B_{r_0} : w(x) = \xi\}).$$

It follows that it makes sense to define $v : \mathcal{W}_1 \to \mathbb{R}^n$ by stipulating that

$$v(\xi) = x \iff x \in U_1 \cap B_{r_0} \text{ and } w(x) = \xi. \tag{48}$$

Since $u = w$ in U_1, we deduce that $u(v(\xi)) = \xi$ for all $\xi \in \mathcal{W}_1$, as required.

We next verify that v is C^1. To do this, fix any $\xi \in \mathcal{W}_1$, and let $x = v(\xi)$. Since w is smooth with $\det Dw \neq 0$ in U_1, the inverse function theorem implies that there is a neighborhood N_ξ of ξ and a C^1 map $\tilde{v} : N_\xi \to B_1$ such that $\tilde{v}(\xi) = x$ and $w(\tilde{v}(\eta)) = \eta$ for all η in N_ξ. Since U_1 is open, we may assume (by shrinking N_ξ if necessary) that $\tilde{v}(\eta) \in U_1 \cap B_{r_0}$ for all $\eta \in N_\xi$. Then it is clear that $\tilde{v} = v$ in N_ξ, and therefore that v is C^1.

Step 3. Since v is C^1, it follows that $Dv(\xi) = Dw(v(\xi))^{-1}$. Thus by the change of variables $\xi = w(x)$ (which is straightforward, since v is a bijection onto its image) we find that

$$\int_{\mathcal{W}_1} d(Dv(\xi), A_1^{-1} SO(n))\, d\xi = \int_{v(\mathcal{W}_1)} d(Dw(x)^{-1}, A^{-1} SO(n)) \det Dw(x)\, dx.$$

Because from (48) we know $v\left(\mathcal{W}_1\right) \subset U_1 \cap B_{r_0}$ and $d(Dw(x), SO(n)A_1) \leq \sigma$ for $x \in U_1 \subset U_0$, so there is a constant C such that $\det Dw(x) \leq C$ and

$$d(Dw(x)^{-1}, A_1^{-1} SO(n)) \leq C\ d(Dw(x), SO(n)A_1) = C\ d(Dw(x), K)$$

for all $x \in U_1$. We conclude that

$$\int_{\mathcal{W}_1} d(Dv(\xi), A_1^{-1} SO(n))\, d\xi \leq C \int_{B_1} d(Du, K) \overset{(19)}{\leq} C\epsilon.$$

Step 4. Finally, we fix k_1 such that, if we define $r_1 = 1 - k_1 a^{\frac{1}{n+1}}$, then $w(B_{r_1}) \subset \beta_{\rho_1}$. This is possible as usual due to Proposition 1 *(iv)*. The definitions imply that $r_1 < r_0$.

Then the definition (48) of v implies that $B_{r_1} \setminus v(\mathcal{W}_1) = (B_{r_1} \setminus U_1) \cup S$, where

$$S := \{x \in B_{r_1} \cap U_1 \ : \ w(x) \notin \mathcal{W}_1\}.$$

Then we can use Proposition 1 *(iv)* and Lemma 7 to estimate

$$L^n(B_{r_1} \setminus v(\mathcal{W}_1)) \leq L^n(B_{r_1} \setminus U_1) + L^n(S)$$
$$\leq C a^{\frac{n}{n-1}} + \frac{L^n(w(S))}{\det A_1} + C\epsilon.$$

And $w(S) \subset \beta_{\rho_1} \setminus \mathcal{W}_1$, so (38) implies that $L^n(w(S)) \leq C a^{\frac{n}{n-1}}$. This proves (39) and completes the proof of the Proposition. $\square$

5. Non-shrinking pairs

In this section we prove

Proposition 3. *Suppose that $u : B_1 \subset \mathbb{R}^n \to \mathbb{R}^n$ is a smooth function satisfying (19) and assume as in (23) that $SO(n)A_1$ is the majority phase. Then there exists $\mathcal{G} \subset B_{r_1} \times B_{r_1}$ (where $r_1 = 1 - k_1 a^{\frac{1}{n+1}}$ was fixed in Proposition 2 and satisfies* (39)*) such that for $\epsilon = \varsigma^{\frac{1}{p}}$,*

$$\textit{if } (x,y) \in \mathcal{G}, \textit{ then } \ \left|\left|u\left(y\right) - u\left(x\right)\right| \ - \ \left|A_1(x-y)\right|\right| \leq C\epsilon \tag{49}$$

and $L^{2n}((B_{r_1} \times B_{r_1}) \setminus \mathcal{G}) \leq C a^{\frac{1}{p}}$.

Following the proof we give a couple of corollaries that will be used in later sections. Throughout the proof we will use notation from the statement of Proposition 2.

Proof of Proposition 3. It suffices to find a set $\mathcal{G}_2 \subset B_{r_1} \times B_{r_1}$ such that

$$L^{2n}((B_{r_1} \times B_{r_1}) \setminus \mathcal{G}_2) \leq Ca^{\frac{1}{p}} \tag{50}$$

and

$$\text{if } (x,y) \in \mathcal{G}_2, \text{ then } \ |u(y) - u(x)| \geq \ |A_1(x-y)| - C\epsilon \tag{51}$$

since then the conclusions of the lemma follow if we define $\mathcal{G} := \mathcal{G}_1 \cap \mathcal{G}_2$, where $\mathcal{G}_1$ was constructed in Lemma 3.

Step 1 . We define

$$\Gamma := \{(\xi,\eta) \in \beta_{\rho_1} \times \beta_{\rho_1} \ : \ [\xi,\eta] \subset \mathcal{W}_1, \int_{[\xi,\eta]} d(Dv, A_1^{-1}SO(n)) \ dH^1 \ \leq \epsilon\} \tag{52}$$

and

$$\mathcal{G}_2 := \{(v(\xi), v(\eta)) \ : \ (\xi,\eta) \in \Gamma\}. \tag{53}$$

We claim that

$$L^{2n}((\beta_{\rho_1} \times \beta_{\rho_1}) \setminus \Gamma) \leq Ca^{\frac{1}{p}}. \tag{54}$$

and

$$\text{if } (\xi,\eta) \in \Gamma, \text{ then } |A_1(v(\xi) - v(\eta))| \leq |\xi - \eta| + C\epsilon \tag{55}$$

In fact, (54) follows from *exactly* the same argument used in the proof of Lemma 3 to establish (24). That proof relied on the facts that $L^n(B_1 \setminus U_0) \leq Ca^{\frac{n}{n-1}}$, $H^{n-1}(\partial U_0 \cap B_1) \leq Ca$, and $\int_{U_0} d(Du, SO(n)A_1) \leq C\epsilon$. In Proposition 2 we proved that the same estimates hold with B_1, U_0 and u replaced by $\beta_{\rho_1}, \mathcal{W}_1$ and v, respectively, so the earlier arguments can be repeated word for word. Next fix $(\xi,\eta) \in \Gamma$, and write $\tau := \frac{\eta-\xi}{|\eta-\xi|}$. Then

$$\begin{aligned} |A_1(v(\eta) - v(\xi))| &= \left| \int_{[\xi,\eta]} A_1 Dv \ \tau \ dH^1 \right| \\ &\leq \int_{[\xi,\eta]} \left[1 + C\, d(Dv, A_1^{-1}SO(n))\right] dH^1 \ \overset{(52)}{\leq} \ |\eta - \xi| + C\epsilon \end{aligned}$$

for $(\xi,\eta) \in \mathcal{G}_1$, proving (55).

Step 2. Observe that (51) is an immediate consequence of (55) and the definition (53) of $\mathcal{G}_2$ and equation (36) of Proposition 2. To verify (50), note that

$$(B_{r_1} \times B_{r_1}) \setminus \mathcal{G}_2 \quad \subset \quad [(B_{r_1} \times B_{r_1}) \setminus (v(\mathcal{W}_1) \times v(\mathcal{W}_1))] \ \cup \ \mathcal{S}$$

for $\mathcal{S} := \{(x,y) \in v(\mathcal{W}_1) \times v(\mathcal{W}_1) \ : \ (u(x), u(y)) \notin \Gamma\}$. We deduce from (39) that

$$L^{2n}(\ (B_{r_1} \times B_{r_1}) \setminus (v(\mathcal{W}_1) \times v(\mathcal{W}_1)) \ ) \leq Ca^{\frac{n}{n-1}}.$$

To estimate the measure of $\mathcal{S}$, we use a change of variables (which is straightforward, since v is a C^1 diffeomorphism onto its image) to compute

$$\begin{aligned} L^{2n}(\mathcal{S}) &= \int_{v(\mathcal{W}_1) \times v(\mathcal{W}_1)} \mathbb{1}_{(u(x),u(y)) \notin \Gamma} \ dx \ dy \\ &= \int_{\mathcal{W}_1 \times \mathcal{W}_1} \mathbb{1}_{(\xi,\eta) \notin \Gamma} \ \det Dv(\xi) \ \det Dv(\eta) \ d\xi \ d\eta \\ &\leq \ J^2 \ L^{2n}((\mathcal{W}_1 \times \mathcal{W}_1) \setminus \Gamma) + \int_{\mathcal{W}_1 \times \mathcal{W}_1} |J^2 - \det Dv(\xi) \ \det Dv(\eta)| \ d\xi \ d\eta \end{aligned}$$

for $J := \det A_1^{-1}$. The integral on the right-hand side is bounded above by

$$\int_{\mathcal{W}_1\times\mathcal{W}_1} J|J-\det Dv(\xi)| + |\det Dv(\xi)||J- \det Dv(\eta)|\, d\xi\, d\eta \overset{(37)}{\leq} C\epsilon$$

since $|J-\det Dv(\xi)| \leq Cd(Dv(\xi), A^{-1}SO(n))$ in $\mathcal{W}_1$ (because v is Lipschitz). And Step 1 (recall definition (46)) implies that

$$L^{2n}((\mathcal{W}_1\times\mathcal{W}_1)\setminus\Gamma) \overset{(46)}{\leq} L^{2n}((\beta_{\rho_1}\times\beta_{\rho_1})\setminus\Gamma) \overset{(54)}{\leq} Ca^{1/p}.$$

Combining the above estimates, we arrive at (50). □

Our first Corollary is

Corollary 1. *Assume the hypotheses of Proposition 3. Then for any $C_1>0$, the set of points $x\in B_{r_1}$ such that*

$$L^n(\{y\in B_{r_1}\ :\ (x,y)\notin\mathcal{G}\}) \geq C_1a^{1/p} \tag{56}$$

has measure at most $\frac{C}{C_1}$, where $\mathcal{G}$ is as found in Proposition 3.

Proof. This follows from Fubini's Theorem, Chebyshev's inequality, and the conclusion of Proposition 3, ie the fact that $L^{2n}(B_{r_1}\times B_{r_1}\setminus\mathcal{G}) \leq Ca^{1/p}$. □

Corollary 2. *Assume the hypotheses of Proposition 3.*

Suppose also that $B_\delta(y_0),\dots,B_\delta(y_n)$ are pairwise disjoint balls contained in B_{r_1}, and that $\mathcal{H}\subset B_1$ is a measurable set such that $L^n(B_1\setminus\mathcal{H})\leq\frac{1}{2}L^n(B_\delta)$.

Then for every sufficiently large $C_1>0$, there exists $a_0>0$ such that if $a\leq a_0$, then there are points $x_k\in B_\delta(y_k)\cap\mathcal{H}$ for $k=0,\dots,n$ such that

$$|\ |u\,(x_k)-u\,(x_l)\,|\ -\ |A_1(x_k-x_l)|\ | \leq C\epsilon \quad \text{for all } k\neq l \tag{57}$$

(for the same C as in (49)), and

$$L^n(\{y\in B_{r_1}\ :\ (x_k,y)\notin\mathcal{G}\}) \leq C_1a^{1/p} \quad \text{for all } k \tag{58}$$

We prove in Lemma 13, in an appendix, that if (57) holds, and if $\{x_0,\dots,x_n\}$ are the vertices of a nondegenerate simplex, then there exists an affine function l_R with $Dl_R = R\in SO(n)A_1$ such that $|u(x_i)-l_R(x_i)|\leq C\epsilon$ for every i.

Proof. Let us say that a point $x\in B_{r_1}$ is *good* if it does *not* satisfy (56), for some value $C_1>0$ to be selected below. Thus, for a fixed good point x, all $y\in B_{r_1}$ away from an exceptional set (depending on x) of measure at most $C_1a^{1/p}$ satisfy $|\ |u(x)-u(y)|-|A_1(x-y)|\ |\leq C\epsilon$.

Let us set $\mathcal{H}' := \{x\in\mathcal{H}\ :\ x \text{ is good}\}$. We fix C_1 in the definition of "good" so large that $L^n(B_{r_1}\setminus\mathcal{H}')\leq\frac{3}{4}L^n(B_\delta)$; it follows from the hypothesis on $\mathcal{H}$ and from Corollary 1 that this is possible. We will show that if a is small enough, then there exist x_k, $k=0,\dots,n$, such that

$$x_k\in B_\delta(y_k)\cap\mathcal{H}' \text{ and } (x_k,x_l)\in\mathcal{G} \text{ for every } k\neq l. \tag{59}$$

In view of the definitions, this will prove the corollary.

We fix x_0 be any point in $B_\delta(y_0)\cap\mathcal{H}'$. Now assume by induction that we have found $x_0,\dots,x_{k-1}$ satisfying (59), for some $k\leq n$. Since $x_0,\dots,x_{k-1}$ belong to $\mathcal{H}'$ and hence are good points, it follows that

$$L^n\left(\bigcup_{l=0}^{k-1}\{x\in B_\delta(y_k)\cap\mathcal{H}'\ :\ (x_l,x)\notin\mathcal{G}\}\right) \leq kC_1a^{1/p}.$$

Thus

$$\begin{aligned}
&L^n(\{x \in B_\delta(y_k) \cap \mathcal{H}' \ : \ (x_l, x) \in \mathcal{G} \text{ for } l = 0, \ldots, k-1\}) \\
&\quad \geq L^n(B_\delta(y_k) \cap \mathcal{H}') - kC_1 a^{1/p} \quad \geq \quad \frac{1}{4} L^n(B_\delta) - nC_1 a^{1/p}.
\end{aligned}$$

In particular, the above set is nonempty for every $k \leq n$ if $a \leq a_0 = \left[\frac{1}{8nC_1} L^n(B_\delta)\right]^p$. Then we can pick x_k to be any point in the above nonempty set, and this eventually yields a collection satisfying (59). □

6. Proof of Theorem 1

In this section we present the proof of Theorem 1. Most of the work of the proof is carried out in two lemmas, in which we consider the case $\Omega = B_1$ and $\Omega' = B_\delta$ for some small $\delta > 0$. We start by assuming these two lemmas hold, and we use them to complete the proof of the theorem. The proofs of the lemmas follow.

Proof of Theorem 1. Step 1. If $m = 2$, we claim that each well must satisfy either (2) or (3). To see this, suppose that $SO(n)A_1$ and $SO(n)A_2$ are two distinct wells such that (2) does not hold for $i = 1$. We may assume by a polar decomposition that A_1, A_2 are both symmetric. The assumption that (2) fails for $i = 1$ says that $|A_2 v|^2 \geq |A_1 v|^2$ for all v. In particular, since this holds for v of the form $v = A_2^{-1} w$, it follows that

$$|w|^2 \geq |A_1 A_2^{-1} w|^2 = w^T (A_2^{-1} A_1^2 A_2^{-1}) w \quad \text{for all } w, \tag{60}$$

or in other words that $A_2^{-1} A_1^2 A_2^{-1} \leq Id$. By taking inverses we find that $A_2 A_1^{-2} A_2 \geq Id$, or equivalently $|w^T A_2 A_1^{-1}|^2 \geq |w|^2$ for all w. This in turn implies that $|v^T A_1^{-1}|^2 \geq |v^T A_2^{-1}|^2$ for all v. To prove that (3) holds, we must show that strict inequality holds for some v, which however is clear, since otherwise equality would hold in (60), which would imply that $A_1 A_2^{-1} \in SO(n)$, and this is impossible since the two wells were assumed to be distinct.

Thus it suffices to show that the lemma holds if each well satisfies (2) or (3).

Step 2. Now fix a connected set $\Omega' \subset\subset \Omega$, and fix $r < \text{dist}(\Omega', \partial\Omega)$. For δ to be specified below, we fix points $x_1, \ldots, x_N \in \Omega'$ such that $\Omega' \subset \cup_{k=1}^n B_{\delta r}(x_k)$. For each k,

$$\frac{1}{\varsigma} \int_{B_r(x_k)} d^p\left(Du, K\right) + \frac{\left|D^2 u\right|^q}{\varsigma^q} dx \leq \frac{1}{\varsigma} \int_\Omega d^p\left(Du, K\right) + \frac{\left|D^2 u\right|^q}{\varsigma^q} dx \leq a \tag{61}$$

and so we can apply a suitable scaled version of Lemma 2 to find some $i = i(k)$ and a set $U_0^k \subset B_r(x_k)$ satisfying (20), (21) hold (with B_1 and A_i replaced by $B_r(x_k)$ and $A_{i(k)}$, and with the constants now depending on r, which however has been fixed.) These conclusions imply that $i(k) = i(k')$ for any k, k' such that $L^n(B_r(x_k) \cap B_r(x_k')) \geq C a^{\frac{n}{n-1}}$ for a suitable constant C. Thus by taking a small enough we deduce that $i(k)$ is in fact independent of k, so that every ball $B_r(x_k)$ has the same majority phase. We relabel the wells as usual so that A_1 represents this majority phase.

Step 3. By assumption A_1 satisfies (2) or (3). In the former case, it follows by continuity that A_1 satisfies the hypothesis (62) of Lemma 8 (proved below) for some $\alpha > 0$, and similarly, if (3) holds, then hypothesis (69) of Lemma 9 is valid for some $\alpha > 0$. We now require that $\delta \leq \delta_i$, $i = 1, 2$, where δ_1, δ_2 appear in the conclusions of Lemmas 8 and 9 respectively. Then in view of (61), if a is small enough then we can apply Lemma 8 or 9 (scaled to a ball of radius r) on each $B_r(x_k)$ to conclude that

$$\int_{\Omega'} d\left(Du, SO\left(n\right) A_1\right) \leq C \sum_{k=1}^N \int_{B_{\delta r}(x_k)} d\left(Du, SO\left(n\right) A_1\right) \leq C \varsigma^{1/p}.$$

Note also that $d^p(Du, SO(n)A_1) \leq C[d(Du, SO(n)A_1) + d^p(Du, K)]$. Thus the above inequalities immediately imply that $\int_{\Omega'} d^p\left(Du, SO\left(n\right)A_1\right) \leq C\varsigma^{1/p}$. Finally, by applying Theorem 3.1 of [Fr-Ja-Mu 02] (see (1)) we conclude that if $p > 1$, then

$$\inf_{R\in K} \|Du - R\|^p_{L^p(\Omega')} \leq \varsigma^{1/p}.$$

□

We now give the proofs of the two lemmas used above. The first uses Lemma 1, which is proved in the introduction.

Lemma 8. *Let $\{A_1, \ldots, A_m\}$ be a set of $n \times n$ matrices, and let $K = \bigcup_i SO(n)A_i$. Let $u : B_1 \to \mathbb{R}^n$ be a smooth function that satisfies (19) and assume the matrices have been labeled so that A_1 is the majority phase, i.e. the set U_0 we obtain from Lemma 2 satisfies (23).*

Suppose A_1 has the property that there exists $v_1 \in S^{n-1}$ and $\alpha > 0$ such that

$$|A_1 v| > (1+\alpha)\,|A_j v| \ \text{ for all } j \geq 2 \text{ and all } v \text{ such that } |v\cdot v_1| > (1-\alpha)|v|. \tag{62}$$

Then there exist constants $a_0, \delta_1 > 0$ such that if $a \leq a_0$ in (19)*, then*

$$\int_{B_{\delta_1}} d\left(Du, SO\left(n\right)A_1\right) dx \leq C\epsilon. \tag{63}$$

Proof. Step 1. We first find points $\{x_0, \ldots, x_n\}$ such that the hypotheses of Lemma 1 are satisfied, together with some other conditions that we will need below.

Fix $0 < \delta \leq \frac{1}{8}$, $\alpha > 0$, and $y_0, \ldots, y_n$ such that $|y_k| = \frac{1}{2}$ for all k, and if $x_k \in B_\delta(y_k)$ for $k = 0, \ldots, n$ and $x \in B_\delta(0)$, then

$$|\tau_k \cdot v_1| \geq 1 - \alpha \quad \text{ for } \tau_k := \frac{x - x_k}{|x - x_k|} \tag{64}$$

and

$$x = \sum_{k=0}^n \lambda_k x_k \text{ with } \sum_{k=0}^n \lambda_k = 1 \quad \text{and} \quad \lambda_k > \delta \text{ for all } k. \tag{65}$$

For example, we can take $y_0 = \frac{1}{2}v_1$, and $y_k = -s_1 v_1 + s_2 z_k$, where s_1, s_2 are constants such that $s_1^2 + s_2^2 = \frac{1}{4}$, and $\{z_1, \ldots, z_n\}$ are the vertices of a regular $n-1$-dimensional simplex sitting on the unit sphere in $v_1^\perp$. If s_2 is sufficiently small, then $|\frac{y_k}{|y_k|} + v_1| < \alpha$, and the above conditions hold if δ is sufficiently small.

We will write

$$\mathcal{H} := \{x \in B_1 \ : \ \int_{B_1} d(Du(z), K)|x - z|^{1-n}\ dz\ \leq C_2\epsilon\}.$$

for a constant C_2 to be determined below. Fubini's Theorem implies that

$$\int_{B_1}\int_{B_1} d(Du(z), K)|x-z|^{1-n}\ dz\ dx\ \leq C\int_{B_1} d(Du(z), K)\ dz\ \overset{(27)}{\leq}\ C\epsilon$$

so we deduce from Chebyshev's inequality that $L^n(B_1 \setminus \mathcal{H}) \leq \frac{C}{C_2}$. We now fix C_2 large enough that $L^n(B_1 \setminus \mathcal{H}) \leq \frac{1}{2}L^n(B_\delta)$. Then Corollary 2 implies that if a is smaller than an appropriate constant a_0, we can find points $x_k \in B_\delta(y_k) \cap \mathcal{H}$ for $k = 0, \ldots, n$ such that

$$|\ |u\left(x_k\right) - u\left(x_l\right)|\ -\ |A_1(x_k - x_l)|\ | \leq C\epsilon \quad \text{ for all } k \neq l.$$

This is exactly the hypothesis of Lemma 13 (proved in Section 8.3). The conclusion of this lemma is that there exists an affine map l_R with $Dl_R = R \in O(n)A_1$ such that

$$|u(x_k) - l_R(x_k)| < C\epsilon \text{ for } k = 0, \ldots, n. \tag{66}$$

Step 2. It follows from (64), and (62) that for $\{x_0, \ldots\ x_n\}$ as found above, the hypotheses (12) of Lemma 1 are satisfied for every $x \in B_\delta$. It follows from the lemma and (66) that

$$\sum_{k=0}^{n} \int_{[x_k,x]} d(Du, SO(n)A_1)\, dH^1 \ \leq\ C \sum_{k=0}^{n} \int_{[x_k,x]} d(Du, K)\, dH^1\ + C\epsilon \tag{67}$$

with a fixed constant C valid for all $x \in B_\delta$.

Step 3. To complete the proof we will integrate the above inequality over $x \in B_\delta$. Both sides of the resulting estimate contain terms of the form $\int_{B_\delta} F_k(x)\ dx$, for F_k of the form $F_k(x) = \int_{[x_k,x]} f(y) dH^1$. Note that by Lemma 14

$$\begin{aligned}\int_{B_\delta} F_k(x)\ dx &= \int_{\theta \in S^{n-1}} \int_0^\infty F_k(x_k + r\theta) r^{n-1} \mathbb{1}_{x_k + r\theta \in B_\delta}\ dr dH^{n-1}\theta \\ &= \int_{\theta \in S^{n-1}} \int_0^\infty \int_0^r f(x_k + s\theta) r^{n-1} \mathbb{1}_{x_0 + r\theta \in B_\delta}\ ds\ dr dH^{n-1}\theta.\end{aligned}$$

We apply Fubini's Theorem, integrate in the r variable, and undo the transformation to polar coordinates, to obtain

$$\int_{B_\delta} F_k(x)\ dx = \int f(x) G_k(x) |x - x_k|^{1-n}\ dx, \qquad \text{where} \quad G_k(x_k + s\theta) := \int_s^\infty r^{n-1} \mathbb{1}_{x_k + r\theta \in B_\delta} dr. \tag{68}$$

Now we integrate both sides of the inequality (67). Since G_k is clearly bounded,

$$\sum_k \int_{B_\delta} (\int_{[x_k,x]} d(Du, K)\, dH^1) dx \leq \sum_k C \int_{B_1} d(Du, K) |x - x_k|^{1-n}\ dx\ \leq\ C\epsilon$$

where we have used the fact that $x_k \in \mathcal{H}$ for $k = 0, \ldots, n$. It is also easy to check that $G_k(x)\,|x - x_k|^{1-n} \geq C^{-1}$ in B_δ, which implies that

$$\sum_k \int_{B_\delta} d(Du, SO(n)A_1) \leq C \sum_k \int_{B_\delta} (\int_{[x_k,x]} d(Du, SO(n)A_1)\, dH^1) dx.$$

By combining these with (67) and defining $\delta_1 := \delta/2$, we complete the proof of the lemma. $\square$

The proof of Theorem 1 will be completed by the following lemma. As mentioned in the introduction, the idea is roughly to apply to u^{-1} an argument like that used in the above lemma. Because u is not invertible, we work with the Lipschitz approximation w found earlier, and we use a lemma (proved in Section 8) that more or less allows us to find a Lipschitz path in B_1 in the inverse image of a.e. line segment in β_{ρ_1}.

Lemma 9. *Let $\{A_1, \ldots, A_m\}$ be a set of $n \times n$ matrices, and let $K = \bigcup_i SO(n)A_i$. Let $u : B_1 \to \mathbb{R}^n$ be a smooth function that satisfies (19) and assume the matrices have been labeled so that A_1 is the majority phase, i.e. the set U_0 we obtain from Lemma 2 satisfies (23).*

Suppose A_1 has the property that there exists $v \in S^{n-1}$ and $\alpha \in (0,1)$ such that

$$\left|v^T A_1^{-1}\right| > (1-\alpha)^{-1} \left|v^T A_j^{-1}\right| \ \textit{for all } j \geq 2 \textit{ and all } v \textit{ such that } |v \cdot v_1| > (1-\alpha)|v|. \tag{69}$$

Then there exist constants $a_0, \delta_2 > 0$ such that if $a \leq a_0$ in (19)*, then*

$$\int_{B_{\delta_2}} d\left(Du, SO\left(n\right) A_1\right) dx \leq C\epsilon. \tag{70}$$

Proof. By defining $\widetilde{A_i} := A_1^{-T} A_i$ and $\tilde{v} = A_1 v_1$ we find hypothesis (69) is preserved for the wells $SO\left(n\right) \widetilde{A_1} \cup \ldots SO\left(n\right) \widetilde{A_m}$ and so without loss of generality we can assume $A_1 = Id$.

Let w denote the Lipschitz approximation of u found in Proposition 1. Recall that in Lemma 6 we found a set $D \subset B_1$, with $L^n(D) \leq C\epsilon$, and such that w is one-to-one and $\det Dw(x) > \sigma$ in $B_{r_0} \setminus D$.

Note also that there exists a constant C such that

$$\text{if } x \notin D, \text{ then } d(Dw(x)^{-1}, A_j^{-1}SO(n)) \leq Cd(Dw(x), SO(n)A_j) \text{ for every } j. \tag{71}$$

This is clear, because the fact that w is Lipschitz implies that $\{Dw(x) : x \notin D\}$ is contained in the compact set $\{M : \det M > \sigma, \|M\| \leq C(K)\}$.

Define also

$$\mathcal{Y} := \{x \in B_1 \setminus D \;:\; d(Dw(x), SO(n)) > d(Dw(x), K)\}.$$

It suffices to prove that

$$L^n(\mathcal{Y} \cap B_{\delta_2}) \leq C\epsilon \tag{72}$$

for a suitable δ_2; this readily implies (70).

Step 1: We first claim that for L^{2n} a.e. $(x, y) \in (B_{r_0} \setminus D) \times (B_{r_0} \setminus D)$ such that $\nu := \frac{y-x}{|y-x|}$ satisfies $|\nu \cdot v_1| > 1 - \alpha$, we have the estimate

$$|x - y| \leq |w(x) - w(y)| \;+\; \int_{[w(x),w(y)]} \Theta \; dH^1 \;\; - c\alpha H^1([w(x), w(y)] \cap w(\mathcal{Y})) \tag{73}$$

where Θ is a nonnegative function, independent of x and y and given explicitly below, such that

$$\int_{\beta_{\rho_0}} \Theta(\xi) \; d\xi \;\; \leq C\epsilon, \tag{74}$$

where recall ball β_{ρ_0} is the large ball in the image we obtain from Proposition 1 *(v)*.

To prove this, we use Lemma 12, which implies that for a.e. $(x, y) \in (B_{r_0} \setminus D) \times (B_{r_0} \setminus D)$ there is an injective Lipschitz path $g : [0, 1] \to \mathbb{R}^n$, such that $g(0) = x, g(1) = y$, and $w(g(t)) \in [w(x), w(y)]$ for all $t \in [0, 1]$. We will write $\gamma(t) := w(g(t))$, so that $Dw(g(t))g'(t) = \gamma'(t)$. Then

$$|y - x| = \nu^T(y - x) = \nu^T \int_0^1 g'(t)dt = \int_0^1 \nu^T Dw(g(t))^{-1}\gamma'(t)dt.$$

Let us temporarily write $M(t)$, or simply M, for $Dw(g(t))$. Then applying the area formula to the right-hand side above (since the image of γ is the segment $[w(x), w(y)]$), we deduce that

$$|x - y| \;=\; \int_{[w(x),w(y)]} \sum_{\{t\in[0,1]:\gamma(t)=\xi\}} \nu^T M(t)^{-1} \frac{\gamma'(t)}{|\gamma'(t)|} \; dH^1(\xi).$$

It follows that

$$|x - y| = |w(x) - w(y)| + \int_{[w(x),w(y)]} \Theta_0(\xi) \; dH^1(\xi) \tag{75}$$

for

$$\Theta_0(\xi) = \left(\sum_{\{t\in[0,1]:\gamma(t)=\xi\}} \nu^T M(t)^{-1} \frac{\gamma'(t)}{|\gamma'(t)|} \right) - 1. \tag{76}$$

If $g(t) \in \mathcal{Y}$, then there exists some $j \geq 2$ such that $d(M(t), K) = d(M(t), SO(n)A_j)$. Recalling that $A_1 = Id$, we infer from (69) that $|\nu^T A_j|^{-1} \leq 1 - \alpha$, so that

$$\begin{aligned}
\nu^T M\,(t)^{-1} \frac{\gamma'}{|\gamma'|} \quad &\leq \quad \left|\nu^T M\,(t)^{-1}\right| \\
&\leq \quad |\nu^T A_j^{-1}| + d\left(M\,(t)^{-1}, A_j^{-1} SO\,(n)\right) \\
&\overset{(71)}{\leq} \quad 1 - \alpha + Cd(M\,(t)\,, K) \qquad\qquad \text{if } g(t) \in \mathcal{Y}.
\end{aligned}$$

Similarly,

$$\nu^T M\,(t)^{-1} \frac{\gamma'}{|\gamma'|} \leq 1 + d(M\,(t)^{-1}\,, SO(n)) \leq 1 + Cd(M\,(t)\,, K) \qquad\qquad \text{if } g(t) \notin (\mathcal{Y} \cup D).$$

For $g(t) \in D$, we claim that the fact that $\|M(t)\| \le C(K)$ implies that

$$\nu^T M\,(t)^{-1}\,\gamma' \le C|\det M(t)|^{-1}|\gamma'|$$

To see this, we recall the polar decomposition $M\,(t) = QS$, where $Q \in O(n)$ and $S = \sqrt{M^T M}$ is symmetric and nonnegative. Then $|\nu^T M\,(t)^{-1}\,\gamma'| \le |\nu^T S^{-1}|\,|\gamma'| \le C(\max\{\lambda_i^{-1}\})|\gamma'|$, where $\{\lambda_i\}$ are the eigenvalues of S. The fact that M is bounded implies that $\lambda_i^{-1} \ge C$ for all i, and it follows that $\max_i\{\lambda_i^{-1}\} \le C \det S^{-1} = C|\det M|^{-1}$. This proves the claim.

Since w is one-to-one on $B_{r_0} \setminus D$, the above estimates of $\nu^T M\,(t)^{-1}\,\frac{\gamma'}{|\gamma'|}$ imply that

$$\Theta_0(\zeta)\ \le -\alpha \mathbb{1}_{\zeta\in w(\mathcal{Y})} + \Theta(\zeta)$$

where

$$\Theta(\zeta) := \mathbb{1}_{\zeta\notin w(V)} d(Dw(w^{-1}(\zeta)), K) + c\mathbb{1}_{\zeta\in w(V)} \sum_{w(z)=\zeta} |\det Dw(z)|^{-1}. \tag{77}$$

We now see that (73) follows from the above with (76), (75). To verify (74), note that by a change of variables, Proposition 1 (v)(b) (recall definition (40))

$$\int_{\beta_{\rho_0}} \mathbb{1}_{\zeta\notin w(V)} d(Dw(w^{-1}(\zeta)), K)\,d\zeta = \int \mathbb{1}_{z\notin D} d(Dw(z), K)\,|\det Dw(z)|dz\ \le C\int d(Dw, K)\ \le C\epsilon.$$

And similarly,

$$\int_{\beta_{\rho_0}} \mathbb{1}_{\zeta\in w(V)} (\sum_{y\in w^{-1}(\zeta)} |\det Dw(y)|^{-1})\ d\zeta\ =\ \int \mathbb{1}_{z\in D}\ dz = L^n(D) \le C\epsilon.$$

Step 2. By arguing exactly as in Step 1 of the proof of Lemma 8 we can find points $x_0, \dots, x_n \in B_{1/2} \setminus D$, a number $0 < \delta < 1/8$, and an affine map l_Q with $Dl_Q = Q \in SO(n)$, such that $|x_k| \ge 3/8$ for all k, and all of the following hold. First, if $x \in B_{2\delta}$ then

$$|\tau_k \cdot v_1| \ge 1 - \alpha \quad \text{ for } \tau_k = \frac{x - x_k}{|x - x_k|},\quad k = 0, \dots, n, \quad \text{ and}$$

$$x = \sum_{k=0}^n \lambda_k x_k \text{ with } \sum_{k=0}^n \lambda_k = 1 \quad \text{and} \quad \lambda_k > \delta \text{ for all } k.$$

Second, $|w(x_k) - l_Q(x_k)| \le C\epsilon$. Third, (73) holds for $(x, y) = (x_k, x)$, for L^n a.e. $x \in B_{2\delta} \setminus D$. And finally,

$$\int_{\beta_{\rho_0}} \Theta(\zeta)|\xi_k - \zeta|^{1-n}\ d\zeta \le C\epsilon \qquad \text{for } \xi_k := w(x_k), k = 0, \dots, n. \tag{78}$$

Step 3. We have defined $\xi_k := w(x_k)$ for $k = 0, \dots, n$. We claim that for $\xi \in w(B_{2\delta} \setminus D)$

$$\sum_{k=0}^n \int_{[\xi_k,\xi]} \mathbb{1}_{w(\mathcal{Y})}\,dH^1 \le C\sum_{k=0}^n \int_{[\xi_k,\xi]} \Theta\ dH^1\ +\ C\epsilon. \tag{79}$$

We will write $e_k(\xi) := \int_{[\xi_k,\xi]} \Theta\ dH^1$. Let $x = w^{-1}(\xi)$, since w is injective on $B_\delta \backslash D$, point x is well defined. We first use (73) to see that

$$\begin{aligned}
H^1([\xi_k, w(x)] \cap w(\mathcal{Y})\ ) &\le |w(x) - \xi_k| - |x - x_k| + e_k(\xi)\\
&\le |l_Q(x) - w(x)| + |l_Q(x - x_k)| + |l_Q(x_k) - \xi_k| - |x - x_k| + e_k(\xi)\\
&= |l_Q(x) - w(x)| + C\epsilon + e_k(\xi).
\end{aligned}$$

To estimate $|l_Q(x) - w(x)|$ we argue as follows. Since l_Q is an isometry,

$$|l_Q(x) - w(x_k)| \le |l_Q(x - x_k)| + |w(x_k) - l_Q(x_k)| \le |x_k - x| + C\epsilon. \tag{80}$$

So we use (73) again to find that

$$|l_Q(x) - w(x) - (w(x_k) - w(x))| \overset{(80)}{\leq} |x_k - x| + C\epsilon \overset{(73)}{\leq} |w(x) - w(x_k)| + e_k(\xi) + C\epsilon.$$

It follows that

$$\frac{w(x) - w(x_k)}{|w(x) - w(x_k)|} \cdot (l_Q(x) - w(x)) + |w(x_k) - w(x)| \;\leq\; |w(x) - w(x_k)| + e_k(\xi) + C\epsilon$$

and hence that $\frac{w(x)-w(x_k)}{|w(x)-w(x_k)|} \cdot (l_Q(x) - w(x)) \leq e_k(\xi) + C\epsilon$ for $k = 0, \ldots, n$. Recall from Proposition 1 *(iv)* that there exists an affine map l_R such that

$$Dl_R = R \in SO(n) \text{ and } \|w - l_R\|_{L^\infty} \leq Ca^{\frac{1}{n+1}}. \tag{81}$$

It follows that the convex hull of $\{\frac{w(x)-w(x_k)}{|w(x)-w(x_k)|}\}_{i=0}^n$ contains a ball B_b of radius b bounded away from zero, if a is small enough. Thus Lemma 15 implies that $|l_Q(x) - w(x)| \leq C\sum_{k=0}^n e_k(\xi) + C\epsilon$, and we have proved (79).

Step 4. We use the notation $\Delta = w(D)$ and we write $\beta_\delta := l_R(B_\delta)$. Note that β_δ is just a ball of radius δ (although not necessarily centered at the origin), since we are assuming that $A_1 = Id$. Note that from (81) and (40) we have $\beta_\delta \backslash \Delta \subset w\left(B_{2\delta} \backslash D\right)$.

We next integrate (79) over $\xi \in \beta_\delta \setminus \Delta$. Both sides of the resulting inequality contain terms of the form $\int_{\beta_\delta \backslash \Delta} F_k(\xi)\; d\xi$, for F_k of the form $F_k(\xi) = \int_{[\xi_k,\xi]} f(\eta) dH^1$. Arguing exactly as in the proof of Step 3 of Lemma 8 we find that

$$\int_{\beta_\delta\backslash\Delta} F_k(\xi)\; d\xi = \int f(\xi) G_k(\xi) |\xi - \xi_k|^{1-n}\; d\xi, \quad \text{ where } G_k(\xi_k + s\theta) := \int_s^\infty r^{n-1} \mathbb{1}_{\xi_k + r\theta \in\; \beta_\delta\backslash\Delta} dr. \tag{82}$$

Note in particular that G_k is bounded. We also claim that

$$\mathcal{B}_k := \{\xi \in \beta_{\delta/2} \;:\; G_k(\xi) \leq c_0\} \text{ is such that } L^n\left(\mathcal{B}_k\right) \leq C_0\epsilon \tag{83}$$

for suitable constants c_0, C_0. To see this, fix $\xi \in \beta_{\delta/2}$, and write ξ in the form $\xi = l_R(x)$ with $|x| < \delta/2$. Then for any k, since by definition $\xi_k = w(x_k)$,

$$|\xi - \xi_k| = |l_R(x) - l_R(x_k) + l_R(x_k) - w(x_k)| \overset{(81)}{\geq} |x - x_k| - Ca^{\frac{1}{n+1}}.$$

It follows that if a is small enough, then $|\xi - \xi_k| \geq 1/8$, say, for $\xi \in \beta_{\delta/2}$. Thus for $\theta \in S^{n-1}$ and $s > \frac{1}{8}$ such that $\xi_k = \xi + s\theta$

$$\begin{aligned} G_k\left(\xi_k + s\theta\right) &\geq s^{n-1}\int_s^\infty \mathbb{1}_{\xi_k + r\theta \in \beta_\delta\backslash\Delta} dr \geq 8^{1-n} L^1\left(\{r \in [s,\infty] : \xi_k + r\theta \in \beta_\delta \setminus \Delta\}\right) \\ &\geq 8^{1-n}\left[L^1\left(\{r \in [s,\infty] : \xi_k + r\theta \in \beta_\delta\}\right) - L^1\left(\{r \in [s,\infty] : \xi_k + r\theta \in \Delta \cap \beta_\delta\}\right)\right]. \end{aligned}$$

Any ray starting at a point in $\beta_{\delta/2}$ must travel a distance at least $\delta/2$ before leaving β_δ, so the first term on the right-hand side above is greater than $\delta/2$ for $\xi_k + s\theta \in \beta_{\delta/2}$. Thus

$$G_k\left(\xi_k + s\theta\right) \;\;\geq\;\; 8^{1-n}\left(\delta/2 - L^1\left(\{r \in [s,\infty] : \xi_k + r\theta \in \Delta \cap \beta_\delta\}\right)\right) \tag{84}$$

for $\xi_k + s\theta \in \beta_{\delta/2}$. Let $\Psi := \left\{\theta : G_k\left(\xi_k + s\theta\right) \leq 8^{1-n}\frac{\delta}{4} \quad \text{for some } \xi_k + s\theta \in \beta_{\delta/2}\right\}$. Note that from (84)

$$L^1\left(\{r \in [s,\infty] : \xi_k + r\theta \in \Delta \cap \beta_\delta\}\right) \geq 8^{1-n}\frac{\delta}{4} \text{ for all } \theta \in \Psi. \tag{85}$$

We take $c_0 = 8^{1-n}\frac{\delta}{4}$, and we use the notation $l_\theta^z = \{z + \lambda\theta : \lambda > 0\}$. ¿From the definition (83) of $\mathcal{B}_k$, we compute (using Lemma 14)

$$\begin{aligned} L^n(\mathcal{B}_k) &\leq \int_{\beta_{\delta/2}} \mathbb{1}_{\mathcal{B}_k}(\xi)\,|\xi - \xi_k|^{1-n}\,d\xi \leq \int_{\theta\in\Psi}\int_{l_\theta^{\xi_k}\cap\beta_{\delta/2}} \mathbb{1}_{\mathcal{B}_k}(\xi)\,dH^1\xi dH^{n-1}\theta \\ &\leq H^{n-1}(\Psi) \\ &\overset{(85)}{\leq} \frac{1}{c_0}\int_{\theta\in\Psi}\int_{l_\theta^{\xi_k}} \mathbb{1}_{\beta_\delta\cap\Delta}(\xi)\,dH^1\xi dH^{n-1}\theta \\ &= \frac{1}{c_0}\int \mathbb{1}_{\Delta\cap\beta_\delta}(\xi)\,|\xi-\xi_k|^{1-n}\,d\xi \leq CL^n(\Delta) \leq C\epsilon. \end{aligned}$$

Thus (83) is established.

Step 5. Defining $f(\eta) = \mathbb{1}_{w(\mathcal{Y})}(\eta)$ and $F_k(\xi) = \int_{[\xi_k,\xi]} f(\eta)\,dH^1\eta$ recall the definition of Θ from (77), let $H_k(\xi) := \int_{[\xi,\xi_k]} \Theta(\eta)\,dH^1\eta$ for $\xi \in w(B_{2\delta}\backslash D)$, recall also that $\beta_\delta\backslash\Delta \subset w(B_{2\delta}\backslash D)$

$$\begin{aligned} C\epsilon + C\sum_{k=0}^n \int_{\beta_\delta\backslash\Delta} H_k(\xi)\,d\xi &\overset{(79)}{\geq} \sum_{k=0}^n \int_{\beta_\delta\backslash\Delta} F_k(\xi)\,d\xi \\ &\overset{(82)}{\geq} \sum_{k=0}^n \int_{\beta_\delta} f(\xi)\,G_k(\xi)\,|\xi-\xi_k|^{1-n}\,d\xi \\ &\overset{(83)}{\geq} c_0\int_{\beta_{\delta/2}} f(\xi)\,d\xi - C\epsilon. \end{aligned} \tag{86}$$

Note that $\sum_{k=0}^n \int_{\beta_\delta\backslash\Delta} H_k(\xi)\,d\xi \leq C\sum_{k=0}^n \int_{\beta_{\rho_0}} \Theta(\xi)\,|\xi-\xi_k|^{1-n}\,d\xi \overset{(78)}{\leq} C\epsilon$, so putting this together with (86) we have

$$L^n(\beta_{\delta/2}\cap w(\mathcal{Y})) = \int_{\beta_{\delta/2}} f(\xi)\,d\xi \overset{(86)}{\leq} C\epsilon. \tag{87}$$

We remark also that (81) implies that $w\left(B_{\delta/4}\cap\mathcal{Y}\right) \subset \beta_{\delta/2}\cap w(\mathcal{Y})$ if a is sufficiently small, so that $L^n\left(w\left(B_{\delta/4}\cap\mathcal{Y}\right)\right) \leq C\epsilon$. Next recall that $\det Dw \geq \sigma$ in $\mathcal{Y}$, since $\mathcal{Y}\cap D = \emptyset$. Now as w is injective on $B_{\frac{\delta}{4}}\backslash D$, the area formula implies that

$$C\epsilon \overset{(87)}{\geq} L^n\left(w\left(B_{\delta/4}\cap\mathcal{Y}\right)\right) = \left|\int_{B_{\delta/4}\cap\mathcal{Y}} \det(Dw(x))\,dx\right| \geq \sigma L^n(B_{\delta/4}\cap\mathcal{Y}),$$

which is (72). $\square$

6.1. **Sharp L^∞ control on a large subset.** The methods used above yield the following result, which is valid for m wells in $\mathbb{R}^n$ without any conditions on the wells.

Proposition 4. *Let $L = \cup_{i=1}^m SO(n)A_i$. Suppose $u : B_1 \to \mathbb{R}^n$ satisfies (19), and assume as in* (23) *that A_1 is the majority phase. Let $r_1 = 1 - k_1 a^{\frac{1}{n+1}}$ be the constant found in Proposition 2 and let $\epsilon = \varsigma^{\frac{1}{p}}$.*

Then there exists $\mathcal{O} \subset B_{r_1}$ and some $R \in SO(n)A_1$ where $L^n(B_{r_1}\backslash\mathcal{O}) \leq Ca^{1/p}$, and

$$\|u - l_Q\|_{L^\infty(\mathcal{O})} \leq C\epsilon. \tag{88}$$

Proof. It suffices to prove the Proposition for all $a < a_0$, for some fixed $a_0 > 0$.

Let $\{y_0, \ldots, y_n\}$ be the vertices of a regular simplex centered at 0 with $|y_k| = \frac{1}{2}\ \forall k$. Then using Corollary 2 (which is valid if a_0 is taken to be small enough) we find points $x_k \in B_{1/8}(y_k)$ such that

$$|\,|u(x_k) - u(x_l)| - |A_1(x_k - x_l)|\,| \leq C\epsilon \quad \text{for all } k \neq l \tag{89}$$

and

$$L^n\left(\{y \in B_{r_1} \ \ (x_k, y) \notin \mathcal{G}\} \right) \ \leq Ca^{1/p}, \tag{90}$$

where $\mathcal{G}$ is the set found in Proposition 3.

Now it follows from (89) and Lemma 13 (see section 8.3) that there exists an affine map l_R with $R \in O(n)A_1$ such that

$$|u(x_k) - l_R(x_k)| \leq C\epsilon \quad \text{ for } k = 0, \ldots, n$$

Let

$$\mathcal{O} := \bigcap_{k=0}^{n} \{y \in B_{r_1} \ : \ (x_k, y) \in \mathcal{G}\}$$

It is clear from (90) that $L^n(B_{r_1} \setminus \mathcal{O}) \leq Ca^{1/p}$. Next, note that the definition of $\mathcal{G}$ implies that if $y \in \mathcal{O}$, then

$$\left| |u(y) - u(x_k)| \ - \ |A_1(y - x_k)| \right| \leq C\epsilon \qquad\qquad \text{for } k = 0, \ldots, n. \tag{91}$$

Then it follows directly from the final conclusion of Lemma 13 that $|u(y) - l_R(y)| \leq C\epsilon$. This proves (88).

To complete the proof of Proposition 4 we only need to note that by Proposition 1, *(iv)* there exists some $R \in SO(n)A_1$ and affine function l_R with $Dl_R = R$ such that $\|w - l_R\|_{L^\infty(\mathcal{O})} \leq ca^{\frac{1}{n+1}}$. So in particular

$$\begin{aligned} \|l_O - l_R\|_{L^\infty(\mathcal{O})} &\leq \|w - l_O\|_{L^\infty(\mathcal{O})} + \|w - l_R\|_{L^\infty(\mathcal{O})} \\ &\leq a^{\frac{1}{n+1}}. \end{aligned}$$

Thus O and R must belong to the same connected component of $O(n)A_1$. Therefore $O \in SO(n)A_1$.

□

7. Totally rank-1 connected wells

Recall that we have shown that an m-well Liouville Theorem holds for $K = \cup_{i=1}^m SO(n)A_i$ satisfying the condition

$$\forall \quad i \in \{1, 2, \ldots m\}, \ \exists \ v_i \in S^{n-1} \text{ such that } |A_i v_i| > |A_j v_i| \text{ for all } j \neq i. \tag{92}$$

Given $K = \cup_{i=1}^n SO(n)A_i$, we form a graph $\mathcal{G}_K$ with vertices $v_1, v_2, \ldots v_n$ where $\mathcal{G}_K$ has edge (v_i, v_j) if and only if there exists $A \in SO(n)A_i$ and $B \in SO(n)A_j$ with $\operatorname{rank}(A - B) = 1$. We say that K is *totally rank*-1 *connected* if $\mathcal{G}_K$ forms a connected graph. We will prove that, loosely speaking, (92) is satisfied for most totally rank-1 connected collections of n wells in $\mathbb{R}^n$.

We say that a well $SO(n)A$ is *positive* if $\det A > 0$. We will restrict our attention to positive wells, since we are interested in orientation-preserving maps with nonvanishing determinant. The map $SO(n)A \mapsto A^T A$ defines a bijection between the set of positive wells and the set of positive definite symmetric matrices. This map is clearly well-defined, since $\tilde{A}^T\tilde{A} = A^T A$ for any $\tilde{A} \in SO(n)A$, and it is invertible, since $SO(n)A = SO(n)\sqrt{A^T A}$ when $SO(n)A$ is positive.

It is often convenient to describe properties of wells $SO(n)A$ in terms of the associated positive definite matrices $A^T A$. An instance of this is the following well-known

Lemma 10. *Two positive wells $SO(n)A$ and $SO(n)B$ are rank-1 connected if and only if there exist column vectors p, q, at least one of which is nonzero, such that $p \cdot q = 0$ and*

$$A^T A - B^T B = pp^T - qq^T. \tag{93}$$

The matrix $pp^T - qq^T$ is uniquely determined by the two wells, so that the wells determine the vectors p, q up to multiplication by -1. The degenerate cases $p = 0$, $q = 0$ are not excluded. (Clearly if $p = q = 0$ then $SO(n)A_1 = SO(n)A_2$.) We present a proof of Lemma 10 at the end of this section, since we have not been able to find a good reference.

The main result of this section is the following

Theorem 2. *Let $SO(n)A_i$ be positive wells for $i = 1, \dots, n$, and assume that there exists a set $\mathcal{C}$ of the form $\mathcal{C} = \{(i_k, j_k)\}_{k=1}^{n-1}$ such that $SO(n)A_{i_k}$ is rank-1 connected to $SO(n)A_{j_k}$ for every k. Assume moreover that*

$$\forall\ i, j \in \{1, \dots, n\}, \quad \exists w_0 = i, w_1, \dots, w_\ell = j \text{ such that } (w_k, w_{k+1}) \in \mathcal{C} \text{ or } (w_{k+1}, w_k) \in \mathcal{C}. \tag{94}$$

For $k = 1, \dots, n-1$, let p_k, q_k be the vectors characterized (up to a sign) by the conditions

$$p_k \cdot q_k = 0 \text{ and } A_{i_k}^T A_{i_k} - A_{j_k}^T A_{j_k} = p_k p_k^T - q_k q_k^T, \tag{95}$$

and assume that

$$\{p_1, \dots, p_{n-1}, q_1, \dots, q_{n-1}\} \text{ contains no linearly dependent subset of } n \text{ elements.} \tag{96}$$

Then $K = \cup_{i=1}^n SO(n)A_i$ satisfies condition (92).

Remark 2. The assumptions about $\mathcal{C}$ imply that K is totally rank-1 connected, since

$$\{\{i_k, j_k\} : k = 1, 2, \dots n-1\}$$

are the edges of the graph $\mathcal{G}_K$ that characterizes the rank-1 connectivity of K and assumption (94) implies that $\mathcal{G}_K$ is connected.

Conversely, whenever K is totally rank-1 connected, we can find a set $\mathcal{C}$ satisfying the above conditions. Indeed, by an elementary result in graph theory, every connected graph with n vertices has a connected subgraph with the same vertices and only $(n-1)$ edges (this subgraph is know as a *spanning tree*). Thus given any totally rank-1 connected K, we can select a spanning tree and use it to define $\mathcal{C}$, by listing the edges in some arbitrary order from 1 to $n-1$, and then orienting each edge by imposing an order on the associated vertices (i.e. replacing the unordered pair $\{i_k, j_k\}$ by the ordered pair (i_k, j_k) for example.)

Remark 3. We claim that Theorem 2 shows that (92) is satisfied in

$$\mathcal{R} := \{K = \cup_{i=1}^n SO(n)A_i \ : K \text{ is totally rank-1 connected}, SO(n)A_i \text{ positive for all } i\}$$

except on a closed set of measure zero. To see this, it suffices to argue that for every set $\mathcal{C} = \{(i_k, j_k)\}_{k=1}^{n-1}$ satisfying (94), corresponding to a possible way of connecting the different wells, the hypotheses of Theorem 2 are satisfied in

$$\begin{aligned}\mathcal{R}_{\mathcal{C}} := \{\cup_{i=1}^n SO(n)A_i : SO(n)A_i \text{ rank-1 connected to } SO(n)A_j \text{ if } (i,j) \in \mathcal{C},\\ SO(n)A_i \text{ positive for all } i\}\end{aligned}$$

away from a closed set of measure zero. For simplicity we consider $\mathcal{C}_0 = \{(1,2), \dots, (n-1,n)\}$, corresponding to collections of wells such that $SO(n)A_i$ is rank-1 connected to $SO(n)A_{i+1}$ for $i = 1, \dots, n-1$. (The argument is nearly identical for any other $\mathcal{C}$.) Consider the set

$$\mathcal{S} := \{(S, (p_1, q_1), \dots, (p_{n-1}, q_{n-1})) : S \in M^{n\times n} \text{ is symmetric}, p_i, q_i \in \mathbb{R}^n, p_i \cdot q_i = 0 \text{ for all } i\}.$$

Given $(S, (p_i, q_i)\,) \in \mathcal{S}$, we define symmetric matrices $S_1, \dots, S_n$ by

$$S_1 := S, \qquad S_{i+1} := S_i + p_i p_i^T - q_i q_i^T.$$

Let $\mathcal{S}_+ := \{(S, (p_i, q_i)) \in \mathcal{S} : S_i$ as defined above is positive definite for all $i\}$. Lemma 10 implies that for $(S, (p_i, q_i)) \in \mathcal{S}_+$, the collection $K = \cup_{i=1}^n SO(n)\sqrt{S_i}$ belongs to $\mathcal{R}_{\mathcal{C}_0}$, and also that every $K \in \mathcal{R}_{\mathcal{C}_0}$ arises in this fashion. Thus $\mathcal{R}_{\mathcal{C}_0}$ can be parameterized by points in $\mathcal{S}_+$, which is an open subset of $\mathcal{S}$.

The point is that one can easily check that (96) fails only on a union of hypersurfaces in $\mathcal{S}$, which is a closed set in $\mathcal{S}$ of $H^{\dim \mathcal{S}}$ measure 0.

Proof of Theorem 2. Let $\mathcal{H}_K$ be the graph with vertices $v_1, v_2, \dots v_n$ where (v_i, v_j) is an edge of $\mathcal{H}_K$ if and only (i,j) or (j,i) belong to $\mathcal{C}$ [2]. As noted in Remark 2, $\mathcal{H}_k$ is connected graph. We say an arbitrary connected graph is a *tree* if and only if it contains no *loops*, by this we mean it contains no non-trivial sequences of edges $(v_{i_1}, v_{i_2}), (v_{i_2}, v_{i_3}), \dots (v_{i_{n-1}}, v_{i_n})$ with $v_{i_1} = v_{i_n}$. It is well known (and easy to prove by induction) that every tree with n vertices has $(n-1)$ edges. Also well known is that any connected graph contains a subgraph with the same vertices that turns out to be a tree. From these two facts we can conclude $\mathcal{H}_K$ is itself a tree since it already has a minimal possible number of edges.

Step 1. We first claim that it suffices to show that for every $(\sigma_1, \dots, \sigma_{n-1}) \in \{\pm 1\}^{n-1}$, we can find a vector v such that

$$\sigma_k(|A_{i_k} v|^2 - |A_{j_k} v|^2) > 0 \quad \text{for every } k = 1, m \dots, n-1. \tag{97}$$

We fix $i \in \{1, \dots, n\}$. Let $\mathcal{D}_1 := \{j \in \{1, 2, \dots n\} : (i,j) \in \mathcal{C} \text{ or } (j,i) \in \mathcal{C}\}$. For any $j \in \mathcal{D}_1$ we chose the sign of σ_k is the obvious way if $(i,j) = (i_k, j_k) \in \mathcal{C}$ then chose $\sigma_k = 1$ and if $(i,j) = (j_k, i_k) \in \mathcal{C}$ then $\sigma_k = -1$. Then (97) gives us that $|A_i v|^2 > |A_j v|^2$ for all $j \in \mathcal{D}_1$.

Let $\mathcal{D}_2 := \{l \in \{1, 2, \dots n\} : \text{For some } p_l \in \mathcal{D}_1, (l, p_l) \in \mathcal{C} \text{ or } (p_l, l) \in \mathcal{C}\}$. For if $l \in \mathcal{D}_2$, if $(l, p_l) = (i_k, j_k) \in \mathcal{C}$ chose $\sigma_k = 1$ and if $(p_l, l) = (j_k, i_k) \in \mathcal{C}$ chose $\sigma_k = -1$. We then have $|A_i v|^2 > |A_{p_l} v|^2 > |A_l v|^2$ for any $l \in \mathcal{D}_2$. And we can continue inductively defining $\mathcal{D}_3, \mathcal{D}_4, \dots$ choosing signs such that (97) implies (92). We will never have $i \in \mathcal{D}_m$ (for any m) because the graph $\mathcal{H}_K$ is a *tree* (and recall its (oriented) edges are given by $\mathcal{C}$), for the same reason if $m_2 > m_1$ then $\mathcal{D}_{m_2} \cap \mathcal{D}_{m_1} = \emptyset$ and the chain of inequalities we build will be consistent. The geometric picture is that we start from a vertex on the graph $\mathcal{H}_K$ and expand outwards one edge at a time choosing signs σ_k one at the time.

Step 2. We now fix an arbitrary $(\sigma_1, \dots, \sigma_{n-1}) \in \{\pm 1\}^{n-1}$, and we show that the system of inequalities (97) admits a solution. We assume for simplicity that $\sigma_k = 1$ for every k. This can be achieved by replacing some pairs $(i,j) \in \mathcal{C}$ by (j,i); in fact the order is arbitrary, and all our assumptions are preserved by this relabeling. Then in view of the characterization of p_k, q_k,

$$\begin{aligned} (p_k \cdot v)^2 - (q_k \cdot v)^2 &= v^T p_k p_k^T v - v^T q_k q_k^T v = v^T A_{i_k} A_{i_k}^T v - v^T A_{j_k} A_{j_k}^T v \\ &= |A_{i_k} v|^2 - |A_{j_k} v|^2 \end{aligned}$$

so our task is to find some $v \in \mathbb{R}^n$ such that

$$(p_k \cdot v)^2 - (q_k \cdot v)^2 > 0 \tag{98}$$

for every $k \in \{1, \dots, n-1\}$. To do this, fix a nonzero vector v the subspace $\cup_{k=1}^{n-1} q_k^\perp$. We then only need to check that $v \cdot p_k \neq 0$ for every k. In fact, if $v \cdot p_k = 0$, then $v \in (\text{span}(p_k, q_1, q_2, \dots, q_{n-1}))^\perp$. Since $v \neq 0$, this would mean that $\{p_k, q_1, q_2, \dots, q_{n-1}\}$ are linearly dependent, and in view of our assumptions this is impossible. $\square$

We end this section by presenting a

Proof of Lemma 10 . First we claim that given arbitrary vectors $\tilde{p}, \tilde{q}$ we can find orthogonal p, q such that $\tilde{p}\tilde{p}^T - \tilde{q}\tilde{q}^T = pp^T - qq^T$. We write $M := \tilde{p}\tilde{p}^T - \tilde{q}\tilde{q}^T$. Note that M is a symmetric matrix of rank at most 2. This is obvious, since $\tilde{p}\tilde{p}^T$ and $\tilde{q}\tilde{q}^T$ are both symmetric with rank ≤ 1. The claim is clear if $\text{rank}(M) \leq 1$, so we assume that $\text{rank}(M) = 2$. Then $\tilde{p}$ and $\tilde{q}$ are linearly independent and in particular nonzero, so by considering vectors orthogonal to $\tilde{p}$ and $\tilde{q}$ respectively one sees that $\min_{|v|=1} v^T M v < 0 < \max_{|v|=1} v^T M v$. Thus M has one positive and one negative eigenvalue, and all other eigenvalues vanish (since $\text{rank}(M) = 2$). The claim

[2] Note that in general $\mathcal{H}_K$ is only a subgraph of $\mathcal{G}_K$

then follows by diagonalizing M. So in fact it suffices to prove $SO(n)A$ and $SO(n)B$ are rank-1 connected if and only if (93) holds true for any $p, q \in \mathbb{R}^n$.

A second simplification comes from noting since $SO(n)B$ is positive, B is invertible, and after multiplying A, B (on the right) and p, q (on the left) by B^{-1} we see that it suffices to prove the lemma for $B = Id$.

Now assume that (93) holds with $B = I$ and $p \cdot q = 0$. The eigenvalues of $A^T A = I + pp^T - qq^T$ are $\lambda_1 := 1 + |p|^2, \lambda_2 := 1 - |q|^2$, and $\lambda_3 = \ldots = \lambda_n = 1$. Let $v_1, \ldots, v_n$ be an associated orthonormal basis of eigenvectors. Since $A^T A$ is positive definite, it must be the case that $1 - |q|^2 > 0$. Let $A_0 = \sqrt{A^T A}$, so note that A_0 has an orthonormal basis of eigenvectors $v_1, \ldots, v_n$ and eigenvalues $\mu_i = \sqrt{\lambda_i}$ for $i = 1, \ldots, n$. Note that $0 < \mu_2 \le 1 \le \mu_1$, and that $\mu_1 < \mu_2$, since we have assumed that $A_0 \neq Id$. To show that A_0 is rank-1 connected to some $Q \in SO(n)$, it suffices to find $(n-1)$ orthonormal vectors $w_1, \ldots, w_{n-1}$ such that $\{A_0 w_i\}_{i=1}^{n-1}$ is also orthonormal, since then we can take Q to be the unique element of $SO(n)$ such that $Q w_i = A_0 w_i$ for $i = 1, \ldots, n-1$. Such a collection is provided by

$$w_1 = \left(\frac{\mu_2^2 - 1}{\mu_2^2 - \mu_1^2}\right)^{1/2} v_1 + \left(\frac{1 - \mu_1^2}{\mu_2^2 - \mu_1^2}\right)^{1/2} v_2, \qquad w_i = v_{i+1} \quad \text{for } i = 2, \ldots, n-1.$$

Now suppose that A is rank-1 connected to $SO(n)$, so that $A = Q + ab^T$ for some $Q \in SO(n)$ and nonzero column vectors a, b. We can also assume that $1 = |a|^2 = a^T a$; if not, replace a by $\frac{a}{|a|}$ and b by $|a|b$. Then

$$A^T A - Q^T Q = Q^T ab^T + ba^T Q + ba^T ab^T = \tilde{a} b^T + b \tilde{a}^T + bb^T$$

for $\tilde{a} = Q^T a$. If we define $\tilde{p} = \tilde{a} + b$, it follows that $A^T A - Q^T Q = \tilde{p}\tilde{p}^T - \tilde{a}\tilde{a}^T$.

And as we know we can find orthogonal vectors p, q such that $\tilde{p}\tilde{p}^T - \tilde{a}\tilde{a}^T = pp^T - qq^T$, so that (93) holds. This finishes the proof of the lemma. $\square$

8. Appendix: Auxiliary Lemmas

8.1. **Truncation Lemma.**

Lemma 11. *Let $n, m \ge 1$, and suppose that $\Omega \subset \mathbb{R}^n$ is a bounded Lipschitz domain. Suppose also that $f : \mathbb{R}^{m \times n} \to [0, \infty)$ is a function such that $f(v) \ge c_1 |v|^p - c_2$ for some $c_1 > 0, c_2 \ge 0$, and $p \ge 1$. Then for any $q \in [1, \infty)$ there exists a constant C such that, whenever $u \in W^{1,q}(\Omega; \mathbb{R}^m)$ satisfies $f(Du) \in W^{1,1}(\Omega; \mathbb{R})$, then for every $\lambda > 0$, there exists $w \in W^{1,\infty}(\Omega; \mathbb{R}^n)$ such that*

$$(i) \qquad \|Dw\|_{L^\infty(\Omega)} \le C\lambda,$$

$$(ii) \qquad \|Du - Dw\|_{L^q(\Omega)}^q \le \frac{C}{\lambda^q} \int_{\{x \in \Omega : Du(x) > \lambda\}} |Du|^q \, dx,$$

$$(iii) \qquad E := \{x \in \Omega : u(x) \neq w(x)\} \subset \{x \in \Omega \ : \ \sup_r \fint_{\Omega \cap B_r(x)} |Du| \, dy > \lambda\},$$

$$(iv) \qquad |E| \le \frac{C}{\lambda^q} \int_{\{x \in \Omega : Du(x) > \lambda\}} |Du|^q \, dx,$$

$$(v) \qquad \text{if} \quad c_1 \lambda^p - c_2 > 0, \ \text{then} \quad Cap_1(E) \le \frac{C}{c_1 \lambda^p - c_2} \|f(Du)\|_{W^{1,1}(\Omega)}.$$

Consequently, if $c_1 \lambda^p - c_2 > 0$, then there exists an open set E' with smooth perimeter such that $E \subset E'$ and $Per_\Omega(E') + (L^n(E'))^{\frac{n-1}{n}} \le \frac{C}{c_1 \lambda^p - c_2} \|f(Du)\|_{W^{1,1}(\Omega)}$.

We will apply the lemma with $f(Du) = d^p(Du, K)$ for some $p \ge 1$, where K is a compact subset of $\mathbb{R}^{m \times n}$.

Most of these conclusions are classical for $u \in W_0^{1,q}(\Omega;\mathbb{R}^m)$ if Ω is smooth enough, and *(i)* , *(ii)* , *(iv)* are proved in exactly the form stated above in Proposition A.1, [Fr-Ja-Mu 02]; hence we only sketch the proofs of these points below. (These conclusions do not require the hypothesis $f(Du) \in W^{1,1}$.) The main point is *(v)*: control over second derivatives of u yields an estimate on the capacity of the set $E = \{x \in \Omega : u(x) \neq w(x)\}$.

If we assume $f(Du) \in W^{1,s}$ for some $s > 1$, then by appealing to slightly different results from the literature but otherwise leaving the proof unchanged, we would obtain an estimate of $Cap_s(E)$. For example, if $u \in W^{2,s}(\Omega)$ then (taking $f(Du) = |Du|$) we would find that $Cap_s(E) \le \frac{C}{\lambda^s}\|Du\|^s_{W^{1,s}}$.

Proof. For any integrable function v on any open subset $U \subset \mathbb{R}^n$, we use the notation

$$M_U(v)(x) := \sup_{r>0} \fint_{B_r(x)\cap U} |v(y)|\ dy.$$

For $u \in W^{1,1}(\Omega;\mathbb{R}^m)$ and $U \subset \Omega$ we will write

$$R^\lambda(u;U) := \{x \in \Omega : M_U(Du)(x) \le \lambda\}.$$

We first assert that for any bounded Lipschitz domain Ω, there exists a constant C such that for every $\lambda > 0$ and $u \in W^{1,1}(\Omega;\mathbb{R}^m)$,

$$|u(x) - u(y)| \le C\lambda|x-y| \qquad \text{for all } x, y \in R^\lambda(u;\Omega). \tag{99}$$

This is well-known if $\Omega = \mathbb{R}^n$ and is essentially proved in [Fr-Ja-Mu 02] for bounded Lipschitz domains; we recall the argument at the end of the proof for the convenience of the reader. Once (99) is known, standard extension theorems assert the existence of a function $w : \Omega \to \mathbb{R}^m$ that satisfies the Lipschitz bound *(i)* and agrees with u on $R^\lambda(u;\Omega)$, so that *(iii)* holds. Then *(iv)* follows from *(iii)* by a covering argument, and *(ii)* is a consequence of *(i), (iv)*.

To prove *(v)*, we must estimate the 1-capacity of $\{x \in \Omega : M_\Omega(Du)(x) > \lambda\}$. To do this, note from Jensen's inequality and the assumptions on f that

$$\fint_{B_r(x)\cap\Omega} |Du|\ dy \le \left(\fint_{B_r(x)\cap\Omega} |Du|^p\ dy\right)^{1/p} \le \left(\fint_{B_r(x)\cap\Omega} \frac{1}{c_1}(f(Du) + c_2)\ dy\right)^{1/p}.$$

Thus $\{x \in \Omega : M_\Omega(Du)(x) > \lambda\} \subset \{x \in \Omega : M_\Omega(f(Du))(x) > (c_1\lambda^p - c_2)^+\}$. Hence *(v)* will follow once we check that

$$Cap_1\left(\{x \in \Omega : M_\Omega(F)(x) > \mu\}\right) \le \frac{C}{\mu}\|F\|_{W^{1,1}(\Omega)} \tag{100}$$

for all $F \in W^{1,1}(\Omega)$ and $\mu > 0$. This is well-known, see [Ev-Ga 92] Section 4.8 for example, (and requires only $\int |DF|$ on the right-hand side) if $\Omega = \mathbb{R}^n$ and for example F has compact support. To show that it remains valid in the present circumstances, recall that any bounded, Lipschitz domain is an extension domain (see for example [St 70], Theorem 5 in Section VI.3), so that there exists a function $\tilde{F} : \mathbb{R}^n \to \mathbb{R}$ with support in a fixed compact set (independent of F), such that

$$\tilde{F} = F \text{ on } \Omega, \text{ and } \quad \|\tilde{F}\|_{W^{1,1}(\mathbb{R}^n)} \le C\|F\|_{W^{1,1}(\Omega)}. \tag{101}$$

We may also take $\tilde{F}$ to be nonnegative (since if this does not hold, we may replace $\tilde{F}$ by $|\tilde{F}|$). Classical results mentioned above imply that

$$Cap_1\left(\left\{x \in \mathbb{R}^n : M_{\mathbb{R}^n}(\tilde{F})(x) > \mu\right\}\right) \le \frac{C}{\mu}\int_{\mathbb{R}^n} |D\tilde{F}|,$$

so in view of (101), to prove (100) it suffices to verify that

$$M_\Omega(F)(x) \le C M_{\mathbb{R}^n}(\tilde{F})(x) \text{ for all } x \in \Omega. \tag{102}$$

Fix a number $R > \mathrm{diam}(\Omega)$, so that $\Omega \cap B_r(x) = \Omega$ if $r \geq R$, for every $x \in \Omega$. Then for $x \in \Omega$,

$$\begin{aligned} M_\Omega(F)(x) &= \sup_{0<r<R} \fint_{\Omega \cap B_r(x)} F\, dy \\ &\leq \sup_{0<r<R} \frac{1}{|\Omega \cap B_r(x)|} \int_{B_r(x)} \tilde{F}\, dy \\ &= \sup_{0<r<R} \frac{|B_r(x)|}{|\Omega \cap B_r(x)|} \fint_{B_r(x)} \tilde{F}\, dy \leq (\sup_{0<r<R} \frac{|B_r(x)|}{|\Omega \cap B_r(x)|}) M_{\mathbb{R}^n}(\tilde{F})(x). \end{aligned}$$

And the fact that Ω is Lipschitz implies that $\sup_{0<r<R} \frac{|B_r(x)|}{|\Omega \cap B_r(x)|} < \infty$; if this were false, we could find a sequence of balls $B_k = B_{r_k}(x_k)$, with r_k necessarily tending to zero, with $x_k \in \Omega$, and such that the density ratios $\frac{|\Omega \cap B_k|}{|B_k|}$ tend to zero, and in a bounded Lipschitz domain this is easily seen to be impossible. Thus we have proved (100), and hence conclusion *(v)* as well.

To prove the final assertion about the existence of the set E', note that by the definition of capacity, there exists a function $h \in C_c^\infty(\mathbb{R}^n)$ such that

$$E \subset \mathrm{int}\,\{x : h(x) \geq 1\} \;\text{ and }\; \int_{\mathbb{R}^n} |Dh| \leq 2\, Cap_1(E).$$

By the coarea formula

$$\int_0^1 H^{n-1}\left(h^{-1}(t)\right) dL^1 t \leq \int_{\mathbb{R}^n} |Dh|.$$

Thus we must be able to find $t_0 \in (1/2, 1)$ with the property that $H^{n-1}\left(h^{-1}(t_0)\right) \leq 4\, Cap_1(E)$. We take $E' = \{x \in \Omega : h(x) > t_0\}$, so that the perimeter estimate is satisfied. As in the proof of Lemma 2, we can assume t_0 is one of the a.e. numbers in $(\frac{1}{2}, 1)$ such that by Sard's theorem, E' has smooth boundary. And by Chebyshev and Sobolev inequalities, we know that

$$|E'|^{\frac{n-1}{n}} \leq C\|h\|_{L^{n/(n-1)}} \leq C\|Dh\|_{L^1}.$$

Finally, we sketch the proof of (99). If Ω is the unit cube Q, then as noted by [Fr-Ja-Mu 02], one can deduce (99) by minor modifications of classical arguments, as expounded for example in [Ev-Ga 92] chapter 6. Next, suppose that Ω is a standard Lipschitz domain, or in other words, the image of the unit cube under a map of the form $x = (x', x_n) \mapsto \phi(x) = (x', q(x'))$, for $q : \mathbb{R}^{n-1} \to \mathbb{R}$ Lipschitz, note that ϕ is a biLipschitz mapping. Then given any $u \in W^{1,1}(\Omega; \mathbb{R}^m)$, we define $\tilde{u} : Q \to \mathbb{R}^m$ by $\tilde{u} = u \circ \phi$. It is straightforward to check that

$$M_Q(D\tilde{u})(x) \leq C M_\Omega(Du)(\phi(x)),$$

and hence that (99) in this case follows from applying the previous case to $\tilde{u}$. Finally, we note as in [Fr-Ja-Mu 02] that a bounded Lipschitz domain Ω can always be written as a finite union $\Omega = \cup_{i=1}^k \Omega_i$, where each Ω_i is (up to a change of variables) a standard Lipschitz domain, so that (99) holds for each Ω_i. This can be done in such a way that there exists some $r_1 > 0$ with the property that there exists some $r_1 > 0$ such that, for any $x, y \in \Omega$ such that $|x - y| < r_1$, there exists some i such that Ω_i contains both x and y. Since Ω is bounded, it clearly suffices to prove (99) for pairs x, y such $|x - y| < r_1$, so we need only show that for every $i = 1, \ldots, k$, there exists some C such that if $x, y \in \Omega_i \cap R^\lambda(u; \Omega)$, then $|u(x) - u(y)| \leq C\lambda|x - y|$.

To do this, we fix some i and argue as in the proof of (102) above to find that

$$M_{\Omega_i}(Du)(x) \leq C M_\Omega(Du)(x). \qquad \text{for all } x \in \Omega_i \text{ and } u \in W^{1,1}(\Omega; \mathbb{R}^m).$$

Thus if $x, y \in \Omega_i \cap R^\lambda(u; \Omega)$ then $x, y \in R^{C\lambda}(u; \Omega_i)$, and so the estimate $|u(x) - u(y)| \leq C\lambda|x - y|$ follows from the case of standard Lipschitz domains.

□

8.2. **Paths in the inverse of segments.**

Lemma 12. *Let $w : B_1 \subset \mathbb{R}^n \to \mathbb{R}^n$ be a Lipschitz function. Given a convex open set $\Lambda \subset \mathbb{R}^n$ such that $\Lambda \subset w(B_1) \setminus w(\partial B_1)$, let*

$$\deg(w, B_1, \xi) = d_0 \neq 0 \quad \text{for a.e. } \xi \in \Lambda \tag{103}$$

Then for L^{2n} a.e. $(\eta, \zeta) \in \Lambda \times \Lambda$,

$$\exists \ \ b > 0 \text{ and an injective Lipschitz function } g : [0, b] \to B_1 \text{ such that}$$
$$w(g(0)) = \eta, w(g(b)) = \zeta, \text{ and } w(g(t)) \in [\eta, \zeta] \quad \forall t \in [0, b]. \tag{104}$$

We will employ the framework of geometric measure theory, so that we work with integral k-currents. One can think of such a current as a k-submanifold of a Euclidean space that is described by specifying how it acts (via integration) on k-forms. We will write $\int_T \phi$ to indicate the action of a current T on a form ϕ. We will appeal to a number of classical facts about slicing of currents. The basic reference for this material is [Fed 69] Chapter 4.3, and a more accessible discussion, albeit without complete proofs, can be found in [Gi-Mo-So 98] section 2.5 of Chapter 2.

Proof. Step 1. We will write $W(x) = (x, w(x)) \in \mathbb{R}^n \times \mathbb{R}^n$ for $x \in B_1$, and $p_2((x, \xi)) = \xi \in \mathbb{R}^n$ for $(x, \xi) \in \mathbb{R}^n \times \mathbb{R}^n$. Note that $w = p_2 \circ W$.

We write G_w to denote the (current associated with the) graph of w, defined by

$$\int_{G_w} \phi \ := \ \int_{B_1} W^{\#} \phi$$

for an n-form ϕ in $\mathbb{R}^n \times \mathbb{R}^n$, where $W^{\#}$ denotes the pullback via W. (One can see G_w as an example of a Cartesian current, and an explicit expression for G_w can be found on page 230, [Gi-Mo-So 98].) The boundary ∂G_w of G_w is defined by $\int_{\partial G_w} \phi := \int_{G_w} d\phi$, and then the definition of G_w implies that $\int_{\partial G_w} \phi = \int_{\partial B_1} W^{\#} \phi$. These formulas imply that

$$\operatorname{Spt} G_w \ = \ \{(x, w(x)) : x \in \bar{B}_1\}, \qquad \operatorname{Spt} \partial G_w \ = \ \{(x, w(x)) : x \in \partial B_1\}.$$

We are using the fact that w is Lipschitz, so that $\{(x, w(x)) : x \in$ compact set $S\}$ is closed.

Step 2. For $\nu \in S^{n-1}$ we define the functions

$$q_\nu(\xi) := \xi - (\xi \cdot \nu)\nu = \text{ orthogonal projection onto } \nu^\perp \subset \mathbb{R}^n, \qquad Q_\nu := q_\nu \circ p_2.$$

We will write ξ' to denote a generic point in Image$(q_\nu) = \nu^\perp$. We will need some classical results about slices of integral currents. Recall that $\langle G_w, Q_\nu, \xi' \rangle$ denotes the slice of G_w by $Q_\nu^{-1}(\xi')$, which for H^{n-1} a.e. $\xi' \in \nu^\perp$ is a integral 1-current satisfying

$$\operatorname{Spt} \langle G_w, Q_\nu, \xi' \rangle \subset \operatorname{Spt} G_w \cap Q_\nu^{-1}(\xi'), \qquad \operatorname{Spt} \partial \langle G_w, Q_\nu, \xi' \rangle \subset \operatorname{Spt} \partial G_w \cap Q_\nu^{-1}(\xi')$$

(see [Fed 69] 4.3.8 (2) for the first inclusion, 4.3.1 p437 together with 4.3.8 (2) for the second inclusion, alternatively Section 2.5 [Gi-Mo-So 98] for a more readable presentation). The fact that a.e. slice $\langle G_w, Q_\nu, \xi' \rangle$ is an integral 1-current implies (see [Fed 69] 4.2.25) that we can write

$$\langle G_w, Q_\nu, \xi' \rangle = \sum_i R_i \qquad \text{for every } \nu \text{ and } H^{n-1} \text{ a.e. } \xi' \tag{105}$$

where each $R_i = R_i(\nu, \xi')$ is the image of an injective Lipschitz map $\gamma_i : I_i \subset \mathbb{R} \to \mathbb{R}^n \times \mathbb{R}^n$, and $I_i = (a_i, b_i) \subset \mathbb{R}$ is a bounded interval. That is, $\int_{R_i} \phi = \int_{\gamma_i(I_i)} \phi = \int_{I_i} \gamma_i^{\#} \phi$ for every 1-form ϕ in $\mathbb{R}^n \times \mathbb{R}^n$. The decomposition (105) is such that

$$\operatorname{Spt} R_i \subset \operatorname{Spt} \langle G_w, Q_\nu, \xi' \rangle \subset \operatorname{Spt} G_w \cap Q_\nu^{-1}(\xi') \tag{106}$$

$$\operatorname{Spt} \partial R_i \subset \operatorname{Spt} \partial \langle G_w, Q_\nu, \xi' \rangle \subset \operatorname{Spt} \partial G_w \cap Q_\nu^{-1}(\xi') \tag{107}$$

for every i. It follows from (106) that each γ_i has the form $\gamma_i(t) = (X_i(t), w(X_i(t)))$ for some Lipschitz path $X_i : I_i \to B_1$ such that $w(X_i(t)) \in q_\nu^{-1}(\xi')$ for every t.

The 0-current ∂R_i appearing in (107) is defined by $\partial R_i(\phi) = \phi(\gamma_i(b_i^-)) - \phi(\gamma_i(a_i^+))$ for every function smooth function ϕ on $\mathbb{R}^n \times \mathbb{R}^n$, so (107) asserts that if $\gamma_i(a_i^+) \neq \gamma_i(b_i^-)$ — that is, if $\partial R_i \neq 0$ — then $\gamma_i(a_i^+), \gamma_i(b_i^-) \in \{(x, w(x)) \ : \ x \in \partial B_1\}$. In particular

$$\text{if } \partial R_i \neq 0 \text{ then } w(X_i(a_i^+)), w(X_i(b_i^-)) \notin \Lambda. \tag{108}$$

Step 3. It T is any k-current in $\mathbb{R}^n \times \mathbb{R}^n$, we define $p_{2\#}G_w$ to be the k-current in $\mathbb{R}^n = \text{Image}(p_2)$ characterized by $\int_{p_{2\#}T} \phi = \int_T p_2^\# \phi$, and we write $T \llcorner \Lambda$ to denote the restriction of T to Λ. We claim that

$$(p_{2\#}\langle G_w, Q_\nu, \xi' \rangle) \llcorner \Lambda = d_0 \langle \Lambda, q_\nu, \xi' \rangle \tag{109}$$

for every ν and H^{n-1} a.e. ξ', for d_0 as in (103). It follows from basic properties of slicing that the current on the right-hand side is just the line segment $\Lambda \cap q_\nu^{-1}(\xi')$, with orientation and (nonzero) multiplicity.

Since $Q_\nu = q_\nu \circ p_2$,

$$p_{2\#}\langle G_w, Q_\nu, \xi' \rangle = p_{2\#}\langle G_w, q_\nu \circ p_2, \xi' \rangle = \langle p_{2\#}G_w, q_\nu, \xi' \rangle$$

for a.e. ξ', see [Fed 69] 4.3.2(7) for the last identity. It follows that

$$(p_{2\#}\langle G_w, Q_\nu, \xi' \rangle) \llcorner \Lambda = \langle (p_{2\#}G_w) \llcorner \Lambda, q_\nu, \xi' \rangle.$$

Thus to prove (109), it suffices to verify that $(p_{2\#}G_w) \llcorner \Lambda = d_0 \Lambda$. To prove this, we first note from the definitions that

$$\int_{p_{2\#}G_w} \phi = \int_{B_1} W^\# p_2^\# \phi = \int_{B_1} (p_2 \circ W)^\# \phi = \int_{B_1} w^\# \phi.$$

In particular, if we write $\phi = \phi(\xi) d\xi$, where $d\xi$ denotes the standard volume form on $\mathbb{R}^n$, then $w^\# \phi = \phi(w(x)) \det Dw(x)\, dx$, and so the change of variables degree formula implies that

$$\int_{p_{2\#}G_w} \phi \ = \int_{B_1} \phi(w(x)) \det Dw(x) \ dx = \ \int_{\mathbb{R}^n} \phi(\xi) \deg(w, B_1, \xi)\, d\xi.$$

We conclude from (103) that $\int_{p_{2\#}G_w} \phi \ = \ d_0 \int_\Lambda \phi$ if $\operatorname{Spt} \phi \subset \Lambda$. This says exactly that $(p_{2\#}G_w) \llcorner \Lambda = d_0 \Lambda$, which is what we needed to prove.

Step 4. We next claim that for every ν, for H^{n-1} a.e. $\xi' \in \nu^\perp$ and every i in the decomposition (105),

$$(p_{2\#}R_i) \llcorner \Lambda = d_i \langle \Lambda, q_\nu, \xi' \rangle \quad \text{ for some } d_i \in \mathbb{Z}. \tag{110}$$

To see this, let us write $\Xi_i(t) = p_2 \circ \gamma_i(t) = w(X_i(t))$. Then it follows from the definitions that $\int_{p_{2\#}R_i} \phi = \int_{I_i} \Xi_i^\# \phi$. In view of properties of $\Xi_i = w \circ X_i$ recorded in Step 2, this implies that $p_{2\#}R_i$ is supported in the line segment $q_\nu^{-1}(\xi')$, and moreover (108) implies that $\partial(p_{2\#}R_i) = 0$ in $\Lambda \cap q_\nu^{-1}(\xi')$. Then (110) follows from the Constancy Theorem, see for example [Fed 69] 4.1.7. (One can also deduce (110) by elementary arguments from the fact that $\int_{p_{2\#}R_i} \phi = \int_{I_i} \Xi_i^\# \phi$, together with the properties of Ξ_i used above.)

Step 5. It follows from (105) that for every ν and a.e. $\xi' \in \nu^\perp$, $p_{2\#}\langle G_w, Q_\nu, \xi' \rangle = \sum_i p_{2\#}R_i$, In view of Steps 3 and 4, this implies that the integer d_i in (110) is nonzero for at least one i. Then the fact that $p_{2\#}R_i = d_i \langle \Lambda, q_\nu, \xi' \rangle$ implies that the for the corresponding curve X_i, the image of $w \circ X_i$ covers $\Lambda \cap q^{-1}(\xi')$ and is contained in $q_\nu^{-1}(\xi')$. So for any two points η, ζ in $\Lambda \cap q^{-1}(\xi')$, we can find a path $g : [0, b] \to B_1$ satisfying (104) by defining g to be a reparametrization of the restriction of X_i to a suitable subinterval of I_i.

Step 6. Let $\mathcal{B} := \{(\eta, \zeta) \in \Lambda \times \Lambda \ : \ (104) \text{ does not hold.}\}$ Our goal is to show that $L^{2n}(\mathcal{B}) = 0$. Note that from Step 5

$$H^2 \left(\left\{ (\eta, \zeta) \in \mathcal{B} : \eta, \zeta \text{ both belong to } \Lambda \cap q_\nu^{-1}(\xi') \right\} \right) = 0 \tag{111}$$

for every $\nu \in S^{n-1}$ and H^{n-1} a.e. $\xi' \in \nu^\perp$.

Let $f = \mathbb{1}_{\mathcal{B}}$. By Fubini's theorem

$$\begin{aligned}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x,y)\,dydx &= \int_{\mathbb{R}^n}\int_{\eta\in S^{n-1}}\int_{t>0} f(x,x+t\eta)\,t^{n-1}dtdH^{n-1}\eta dx\\ &= \int_{\eta\in S^{n-1}}\int_{t>0}\int_{\mathbb{R}^n} f(x,x+t\eta)\,t^{n-1}dxdH^{n-1}\eta dt\\ &= \int_{\eta\in S^{n-1}}\int_{y\in\eta^{\perp}}\int_{s>0}\int_{t>0} f(y+s\eta,y+(t+s)\eta)\,dsdtdH^{n-1}ydH^{n-1}\eta.\end{aligned}$$

¿From (111) for any $\eta \in S^{n-1}$, and H^{n-1} a.e. $y \in \eta^{\perp}$ $\int_{s>0}\int_{r>0} f(y+s\eta, y+(t+s)\eta)\,t^{n-1}dsdt = 0$ and thus we have shown $L^n(\mathcal{B}) = 0$. $\square$

8.3. **A linear algebra lemma.**

Lemma 13. *Suppose that A is an invertible $n\times n$ matrix, and that $z_0, z_1, \dots z_n \in B_1(0) \subset \mathbb{R}^n$ and $\zeta_0, \zeta_1, \dots \zeta_n \in \mathbb{R}^n$ are points such that*

$$B_b(y) \subset \mathrm{conv}(z_0, z_1, \dots z_n) \text{ for some } b>0, y\in B_1 \tag{112}$$

and

$$||\zeta_i - \zeta_j| - |A(z_i - z_j)|| \le \epsilon \text{ for all } i\neq j \in \{0,1,\dots n\}. \tag{113}$$

Then there exists an affine function l_O with $Dl_O = O \in O(n)A$ and constant $C = C(b,n,A)$ such that

$$|\zeta_i - l_O(z_i)| \le C\epsilon \text{ for all } i \in \{0,1,\dots n\}. \tag{114}$$

Furthermore, if $z \in B_1$ and $\zeta\in\mathbb{R}^n$ are any other points such that

$$||\zeta_i - \zeta| - |A(z_i - z)|| \le \epsilon \text{ for all } i \in \{0,1,\dots n\} \tag{115}$$

then $|\zeta - l_O(z)| \le C\epsilon$ for the same l_O as in (114)*, and with $C = C(n,b,A)$.*

Proof of Lemma 13. By a translation we can assume that $z_0 = \zeta_0 = 0$. We can also assume that A is the identity; if not, simply replace each z_i by $\tilde{z}_i = Az_i$ and drop the tildes, so that (113) becomes

$$||\zeta_i - \zeta_j| - |z_i - z_j|| \le C\epsilon \text{ for all } i \neq j \in \{0,1,\dots n\}. \tag{116}$$

After these changes, $|z_i|, |\zeta_i| \le C$ for all i.

We define $l_{\widetilde{O}} : \mathbb{R}^n \to \mathbb{R}^n$ to be the unique linear map satisfying $l_{\widetilde{O}}(z_i) = \zeta_i$ for $i = 1,2,\dots n$. We will identify $l_{\widetilde{O}}$ with the matrix $\widetilde{O} = Dl_{\widetilde{O}}$. It follows from (112) that $\{z_1, z_2, \dots z_n\}$ are linear independent, and hence that $\widetilde{O}$ is well defined.

Step 1. We first show that

$$\left|\widetilde{O}(z_i)\cdot\widetilde{O}(z_j) - z_i\cdot z_j\right| \le C\epsilon. \tag{117}$$

Toward this goal, note that since $|\,|\zeta_i - \zeta_j| + |z_i - z_j|\,| \le C$ for all i,j,

$$\begin{aligned}\left||\zeta_i-\zeta_j|^2 - |z_i - z_j|^2\right| &\le c\,||\zeta_i - \zeta_j| - |z_i - z_j||\\ &\overset{(113)}{\le} C\epsilon.\end{aligned} \tag{118}$$

As a result,

$$2\,|\zeta_i\cdot\zeta_j - z_i\cdot z_j| \overset{(118)}{\le} \left|\,|\zeta_i|^2 + |\zeta_j|^2 - |z_i|^2 - |z_j|^2\,\right| + C\epsilon.$$

However, since $z_0 = \zeta_0 = 0$, the $j=0$ case of (116) implies that $\left|\,|\zeta_i|^2 - |z_i|^2\,\right| \le C\epsilon$, and similarly $\left|\,|\zeta_j|^2 - |z_j|^2\,\right| \le C\epsilon$, so (117) follows from the above.

Step 2. We next claim that for any $v \in S^{n-1}$ there exist $\gamma_1, \gamma_2, \dots \gamma_n$ with $|\gamma_i| \le \frac{2}{b}$ for each $i = 1,2\dots n$ such that $\sum_{i=1}^n \gamma_i z_i = v$.

Proof of Claim. Note that

$$\begin{aligned} B_b(-y) \cup B_b(y) &\subset \operatorname{conv}(z_0, z_1 \dots z_n) \cup \operatorname{conv}(z_0, -z_1 \dots - z_n) \\ &\subset \operatorname{conv}(z_1, \dots z_n, -z_1, \dots - z_n) \end{aligned}$$

which implies $B_b \subset \operatorname{conv}(z_1, \dots z_n, -z_1, \dots - z_n)$. So there exist positive $\beta_0, \beta_1, \dots \beta_{2n}$ such that $\sum_{i=0}^{2n} \beta_i = 1$ and $\sum_{i=1}^{n} (\beta_i - \beta_{i+n}) z_i = vb$. Since $\frac{|\beta_i - \beta_{i+n}|}{b} \leq \frac{2}{b}$ this completes Step 2.

Step 3. Let $\{e_1, e_2, \dots e_n\}$ be an orthonormal basis of $\mathbb{R}^n$. We claim that

$$\left| \widetilde{O}(e_i) \cdot \widetilde{O}(e_j) - \delta_{ij} \right| \leq C\epsilon \text{ for any } i, j \in \{1, 2, \dots n\}. \tag{119}$$

Proof of Claim. By Step 2 we can find coefficients $\alpha_j^i \in \mathbb{R}$ such that $\sum_{j=1}^n \alpha_j^i z_j = e_i$ and $\left|\alpha_j^i\right| \leq \frac{2}{b}$ for $i, j \in \{1, 2, \dots n\}$. Note

$$\sum_{k,l=1}^{n} \alpha_k^i \alpha_l^j z_k \cdot z_l = \delta_{ij} \text{ for any } i, j \in \{1, 2, \dots n\}. \tag{120}$$

Now

$$\begin{aligned} \left| \widetilde{O}(e_i) \cdot \widetilde{O}(e_j) - \delta_{ij} \right| &\overset{(120)}{=} \left| \widetilde{O}(e_i) \cdot \widetilde{O}(e_j) - \sum_{k,l=1}^{n} \alpha_k^i \alpha_l^j z_k \cdot z_l \right| \\ &\leq \left| \sum_{k,l=1}^{n} \alpha_k^i \alpha_l^j \left(\widetilde{O}(z_k) \cdot \widetilde{O}(z_l) - z_k \cdot z_l \right) \right| \\ &\leq \sum_{k,l=1}^{n} \left| \alpha_k^i \alpha_l^j \right| \left| \widetilde{O}(z_k) \cdot \widetilde{O}(z_l) - z_k \cdot z_l \right| \\ &\overset{(117)}{\leq} C\epsilon. \end{aligned}$$

Thus (119) is established.

Step 4. We now define $\{\xi_1, \xi_2, \dots \xi_n\}$ to be the orthonormal basis of $\mathbb{R}^n$ obtained via a Gram-Schmidt orthognalisation of the set of vectors $\left\{ \widetilde{O}(e_1), \widetilde{O}(e_2), \dots \widetilde{O}(e_n) \right\}$. Then an easy induction argument using (119) shows that

$$\left| \widetilde{O}(e_i) - \xi_i \right| \leq C\epsilon. \tag{121}$$

We define $l_O : \mathbb{R}^n \to \mathbb{R}^n$ to be the linear map such that $l_O(e_i) := \xi_i$ for $i = 1, 2, \dots n$. Note $O := Dl_O \in O(n)$. Also, by (121) we have $|\widetilde{O} - O| \leq C\epsilon$. In particular $|\zeta_i - O(z_i)| = |\widetilde{O}(z_i) - O(z_i)| \leq C\epsilon$, so that we have proved (114).

Step 5. Finally, suppose that $\zeta \in \mathbb{R}^n$ and $z \in B_1$ satisfy $||\zeta_i - \zeta| - |z_i - z|| \leq \epsilon$ for all $i \in \{0, 1, \dots n\}$. Then using (114) and the fact $O \in O(n)$, we find that

$$\begin{aligned} \left| |z_i - l_O^{-1}(\zeta)| - |z_i - z| \right| &\leq \left| |l_O^{-1}(\zeta_i - \zeta)| - |z_i - z| \right| + |l_O^{-1}(\zeta_i) - z_i| \\ &\leq C\epsilon \end{aligned}$$

for all $i \in \{0, 1, \dots n\}$. Arguing exactly as in the proof of (117) in Step 1, we deduce from the above that

$$|z_i \cdot (l_O^{-1}(\zeta) - z)| \leq C\epsilon$$

for every i. And this implies that $|\zeta - l_O(z)| = |l_O^{-1}(\zeta) - z| \leq C\epsilon$; this is proved in Lemma 15 in the next subsection. $\square$

8.4. **Coarea formula into S^{n-1} and bounding the diameter of a simplex.**

Lemma 14. *Let $\Theta_x : \mathbb{R}^n \to S^{n-1}$ be defined by $\Theta_x(z) = \frac{z-x}{|z-x|}$. Then for any function $h : \mathbb{R}^n \to \mathbb{R}$ such $H(z) := h(z)\ |x-z|^{1-n}$ is integrable,*

$$\int_{\psi \in S^{n-1}} \int_{\Theta_x^{-1}(\psi)} h(z)\ dH^1 z\ dH^{n-1}\psi = \int_{\mathbb{R}^n} \frac{h(z)}{|x-z|^{n-1}} dL^n z. \tag{122}$$

Proof. By a change of variables and Fubini's Theorem,

$$\begin{aligned}\int_{\mathbb{R}^n} \frac{h(z)}{|x-z|^{n-1}} dz &= \int_0^\infty \int_{z\in\partial B_s(x)} \frac{h(z)}{|x-z|^{n-1}} dH^{n-1} z\, ds \\ &= \int_0^\infty \int_{\psi\in S^{n-1}} h(s\psi + x)\ dH^{n-1}\psi\, ds \\ &= \int_{\psi\in S^{n-1}} \int_{\Theta_x^{-1}(\psi)} h(z)\ dH^1 z\ dH^{n-1}\psi.\end{aligned}$$

□

Lemma 15. *Let $z_0, z_1, \ldots z_n$ be vectors with the property that $B_b \subset \operatorname{conv}(z_0, z_1, \ldots z_n)$, and let $S := \{x : x \cdot z_i \leq 1$ for $i = 0, 1, \ldots n\}$. Then $S \subset B_{n/b}$.*

Proof of Lemma 15. Fix any $x_0 \in S$. Since $b\frac{x_0}{|x_0|} \in \bar{B}_b \subset \operatorname{conv}(z_0, z_1, \ldots z_n)$ there exists $\lambda_0, \lambda_1, \ldots \lambda_n \in [0,1]$ with $\sum_{i=0}^n \lambda_i z_i = b\frac{x_0}{|x_0|}$. So there must exist $i_0 \in \{0, 1, \ldots n\}$ such that $z_{i_0} \cdot \frac{x_0}{|x_0|} \geq \frac{b}{n}$. However as $x_0 \in S$ we have $x_0 \cdot z_{i_0} \leq 1$ this gives $|x_0| \leq \frac{n}{b}$. □

References

[Ad-Fo 00] R.A. Adams; J.F. Fournier. Sobolev Spaces, Second ed., Pure and Applied Mathematics, vol. 140, Academic Press, Boston, 2003.

[Co-Sc 06a] S. Conti, B. Schweizer. Rigidity and Gamma convergence for solid-solid phase transitions with $SO(2)$-invariance. Comm. Pure Appl. Math. 59 (2006), no. 6, 830–868.

[Co-Sc 06b] S. Conti, B. Schweizer. A sharp-interface limit for a two-well problem in geometrically linear elasticity. Arch. Rat. Mech. Anal. 179 (2006), 413-452.

[Cha-Mu 03] N. Chaudhuri, S. Müller. Rigidity Estimate for Two Incompatible Wells. Calc. Var. Partial Differential Equations 19 (2004), no. 4, 379–390.

[De-Se 06] C. De Lellis, L. J. Szekelyhidi. Simple proof of two well rigidity. C. R. Math. Acad. Sci. Paris 343 (2006), no. 5, 367–370.

[Ev-Ga 92] L.C. Evans. R.F. Gariepy. Measure theory and fine properties of functions. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[Fa-Zh 05] D. Faraco; X. Zhong, Geometric rigidity of conformal matrices. Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 4 (2005), no. 4, 557–585.

[Fed 69] H. Fedefer. Geometric measure theory. Die Grundlehren der mathematischen Wissenschaften, Band 153 Springer-Verlag New York Inc., New York 1969.

[Fo-Ga 95] I. Fonseca, W. Gangbo. Degree theory in analysis and applications. Oxford Lecture Series in Mathematics and its Applications, 2. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1995.

[Fr-Ja-Mu 02] G. Friesecke, R. D. James and S. Müller. A theorem on geometric rigidity and the derivation of nonlinear plate theory from three dimensional elasticity. Comm. Pure Appl. Math. 55 (2002) no. 11, 1461–1506.

[Fr-Ja-Mu 06] G. Friesecke; R. D. James; S. Müller. A hierarchy of plate models derived from nonlinear elasticity by gamma-convergence. Arch. Ration. Mech. Anal. 180 (2006), no. 2, 183–236.

[Gi-Mo-So 98] M. Giaquinta; G. Modica; J. Souvcek. Cartesian currents in the calculus of variations. I. Cartesian currents. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge, 37. Springer-Verlag, Berlin, 1998.

[Jo 61] F. John, Rotation and strain. Comm. Pure Appl. Math. 14 1961 391–413.

[Ko 82] R. V. Kohn. New integral estimates for deformations in terms of their nonlinear strains. Arch. Rational Mech. Anal. 78 (1982), no. 2, 131–172.

[Lio 50] J. Liouville Théoréme sur l'équation $dx^2 + dy^2 + dz^2 = \lambda \left(d\alpha^2 + d\beta^2 + d\gamma^2 \right)$ J. Math. Pures Appl, 1, (15) (1850), 103.

[Lor 05] A. Lorent. A two well Liouville Theorem. ESAIM Control Optim. Calc. Var. 11 (2005), no. 3, 310-356.

[Lor pr 06] A. Lorent. An L^p two well Liouville theorem. MIS-MPG preprint 72/2006. To appear Ann. Acad. Sci. Fenn. Math.

[Mat 92] J. Matos. Young measures and the absence of fine microstructures in a class of phase transitions. European J. Appl. Math. 3 (1992), no. 1.

[Re 67] Yu. G. Reshetnyak. Liouville's conformal mapping theorem under minimal regularity hypotheses. (Russian) Sibirsk. Mat. Ž. 8 1967 835–840.

[St 70] E. M. Stein. Singular integrals and Differentiability Properties of Functions. Princeton University Press, Princeton, New Jersey, 1970.

Department of Mathematics, University of Toronto, Toronto, Ontario, Canada
E-mail address: `rjerrard@math.toronto.edu`

Centro De Giorgi, Scuola Normale Superiore, Piazza dei Cavalieri, Pisa, Italy
E-mail address: `andrew.lorent@sns.it`