Concurrent normals to convex bodies and spaces of Morse functions

John Pardon

17 December 2008

Abstract

It is conjectured that if $K \subset \mathbb{R}^n$ is a convex body, then there exists a point in the interior of K which is the point of concurrency of normals from $2n$ points on the boundary of K. We present a topological proof of this conjecture in dimension four assuming ∂K is $C^{1,1}$. From the assumption that the conjecture fails for $K \subset \mathbb{R}^4$, we construct a retraction from $\overline{K}$ to ∂K. We apply the same strategy to the problem for lower n, assuming no regularity on ∂K, and show that it provides very simple proofs for the cases of two and three dimensions (the dimension three case was first proved by Erhard Heil). A connection between our approach to this problem and the homotopy type of some function spaces is also explored, and some conjectures along those lines are proposed.

1 Introduction

Conjecture 1.1 (Problem A3 in [3])**.** If $K \subset \mathbb{R}^n$ is a bounded open convex set, then there exists a point $p \in K$ which is the point of concurrence of $2n$ normals from points on ∂K.

In this paper, we present a proof of Conjecture 1.1 when $n = 4$ and ∂K is $C^{1,1}$ (Theorem 4.1). We also present a proof along the same lines for $n = 3$ (Theorem 3.2), a case which was proved using more geometric methods by Heil [6] [7] [5]. The method of our proofs is notable in that it is entirely topological: we use essentially no geometry. Instead, we rely on the decomposition of ∂K given by thinking of the distance squared function $d : \overline{K} \times \partial K \to \mathbb{R}$ as a Morse function on ∂K. Heil's proof for $n = 3$ on the other hand relies on the geometric fact of the existence of a *minimal spherical shell* for a convex body.

Our strategy is proof by contradiction: we assume that $d(p, \cdot)$ has fewer than eight critical points for all $p \in K \subset \mathbb{R}^4$ (critical points of $d(p, \cdot)$ correspond to normals passing through p), and construct a retraction from $\overline{K}$ to ∂K. The basic idea is that if a function on $\mathbb{S}^{n-1}$ has only a small number of critical points, it can't be far from a height function $p \mapsto x \cdot p$.

The fact that we use only these functions $d(p, \cdot)$ suggests that something deeper is going on, and we go on to interpret our approach in terms of homotopy groups of a space of functions on $\mathbb{S}^{n-1}$ with $\leq k$ critical points. Define $G_k^\circ(\mathbb{S}^{n-1}) \subset C^1(\mathbb{S}^{n-1})$ to be the set of those functions f for which $\nabla f = 0$ at no more than k points. Let $G_k(\mathbb{S}^{n-1}) \subset G_k^\circ(\mathbb{S}^{n-1})$

be the closure in G_k° of the set of Morse functions in G_k°. Now note that there is a natural embedding $e : \mathbb{S}^{n-1} \hookrightarrow G_2(\mathbb{S}^{n-1})$ which sends a point x to the function $p \mapsto x \cdot p$.

Conjecture 1.2 (see Section 5)**.** There exists $r : G_{2n-1}(\mathbb{S}^{n-1}) \to \mathbb{S}^{n-1}$ so that $r \circ e$ is the identity on $\mathbb{S}^{n-1}$. In other words, $G_{2n-1}(\mathbb{S}^{n-1})$ retracts to $\operatorname{Im} e$.

We will show (Lemma 5.7) that Conjecture 1.2 implies a weak form of Conjecture 1.1. The converse implication, i.e. that Conjecture 1.1 implies Conjecture 1.2, seems impossible unless one proves *a priori* that Conjecture 1.2 is true. Hence it is especially interesting that we prove Conjecture 1.1 (for $2 \le n \le 4$) using methods which in fact suffice to prove Conjecture 1.2 (for $2 \le n \le 4$). Such a generalization of Conjecture 1.1 does not seem possible using Heil's geometric methods. Conjecture 5.5 asserts that for $k \ge 2n$, $G_k(\mathbb{S}^{n-1})$ is contractible. In fact, any result concerning the homotopy or homology groups of following filtration would be new and interesting:

$$\mathbb{S}^{n-1} \overset{e}{\hookrightarrow} G_2(\mathbb{S}^{n-1}) \hookrightarrow G_4(\mathbb{S}^{n-1}) \hookrightarrow G_6(\mathbb{S}^{n-1}) \hookrightarrow G_8(\mathbb{S}^{n-1}) \hookrightarrow \cdots \tag{1.1}$$

We include here a brief review of literature on problems concerning normals to convex bodies. For one of the many proofs of Conjecture 1.1 in two dimensions, see [4]. In [12], it is shown that a smooth hypersurface in $\mathbb{R}^n$ has four concurrent normals. Conjecture 1.1 was proved for centrally symmetric bodies by Lusternik and Schnirelmann [10]. Kuiper [9] has generalized their result and shown that any convex body in $\mathbb{R}^n$ has n double normals. Zamfirescu [13] has shown that almost all (in the Baire category sense) convex bodies in $\mathbb{R}^n$ have the property that almost all (in the Baire category sense) points in $\mathbb{R}^n$ are the concurrence point of infinitely many normals.

For bodies of constant width, one conjectures that there are $4n-2$ normals which concur in the interior of K [3]. This is straightforward for $n = 2$, was proved by Heil [5] for $n = 3$, and is open for $n \ge 4$. Conjecture 1.1 is open for $n \ge 5$. We remark that for Conjecture 1.1, $2n$ normals is the maximum one can expect to find in the worst case: an ellipsoid in $\mathbb{R}^n$ with axes of length 1, $1+\epsilon$, $\dots$, $1+(n-1)\epsilon$ shows this.

1.1 Definition of a normal to a nonsmooth convex set

It should be noted that a small generalization (which many authors simply omit) of the word normal must be made for Conjecture 1.1 to not be trivially false in every dimension. For example, take a regular n-simplex $\Delta^n \subset \mathbb{R}^n$. Then every point in the interior of Δ^n has a unique normal to each of the faces of Δ^n, making for $n+1$ normals. As is familiar to those who work in convexity, we instead define a hyperplane P to be tangent to an open convex set K if and only if it satisfies $P \cap K = \varnothing$ and $P \cap \overline{K} \neq \varnothing$. A normal to ∂K is then a line normal to P passing through some point in $P \cap \overline{K}$. Thus, for example, every point in the interior of an equilateral triangle in $\mathbb{R}^2$ is the concurrence point of six normals: one from each edge and one from each vertex. In essence we allow normals from points where ∂K is not differentiable. The reader may also think of this as instead considering normals to the parallel surface $K + B_\epsilon$, whose normals coincide with the normals of K, and whose boundary is of class $C^{1,1}$.

In fact, ∂K is of class $C^{1,1}$ if and only if we can write $K = K_{-\epsilon} + B_\epsilon$ for some $\epsilon > 0$ and some convex set $K_{-\epsilon}$ (see [2, p475], or [9, p74] where the condition $C^{1,1}$ is called C^{2-}).

2 The case $n = 2$

Lemma 2.1. *If $K \subset \mathbb{R}^2$ is a bounded open convex set, then there exists a point $p \in K$ which is the point of concurrence of four normals from points on ∂K.*

Proof. Consider the function $d_p : \partial K \to \mathbb{R}$ ($p \in \overline{K}$) which computes the distance squared to p. Any local extremum of d_p corresponds to a normal passing through p. Thus it suffices to find $p \in K$ so that d_p has at least four local extrema. Suppose for sake of contradiction that there is no such p. Now a continuous function $f : \mathbb{S}^1 \to \mathbb{R}$ which has fewer than four local extrema must have a unique local minimum. Thus the function $r : \overline{K} \to \partial K$ defined by $r(p) = \operatorname{argmin}_q d_p(q)$ is continuous (by $\operatorname{argmin}_x f(x)$ we mean that value of x for which $f(x)$ attains its minimum; if multiple such x exist (or none does), then $\operatorname{argmin}_x f(x)$ is undefined). Furthermore, $r = \mathrm{id}_{\partial K}$ on ∂K. Thus r is a retraction from $\overline{K}$ to ∂K, a contradiction. □

3 The case $n = 3$

For the higher dimensions, we need $d_p : \partial K \to \mathbb{R}$ to be Morse for a dense set of p. The following lemma proved by Heil guarantees that this will be the case.

Lemma 3.1 (Heil [5, p176 Lemma 2′])**.** *If $K \subset \mathbb{R}^n$ is a bounded open convex set and $\epsilon > 0$, then the subset of points $p \in \mathbb{R}^n$ for which $d_p^\epsilon : \partial K_\epsilon \to \mathbb{R}$ is not Morse has measure zero. Here K_ϵ is the parallel body $K + B_\epsilon$ (which shares the same normals as K) and d_p^ϵ is the distance squared function.*

We say that $f : \partial K_\epsilon \to \mathbb{R}$ is Morse if f is C^1 and at every point where $\nabla f = 0$, we have:

- ∂K_ϵ is twice differentiable
- f is twice differentiable
- the Hessian of f is nondegenerate

Proof. Since ∂K_ϵ is $C^{1,1}$, the set of points A where ∂K_ϵ is not second differentiable has measure zero. Let $j : N(\partial K_\epsilon) \to \mathbb{R}^n$ be the natural map from the normal bundle. The map j is locally Lipschitz since ∂K_ϵ is $C^{1,1}$. Now d_p^ϵ is Morse if p is not in $j(A)$ and p is not a critical value of j. We see that $j(A)$ is of measure zero since A is, and that the set of critical values of j has measure zero by a generalization Sard's Theorem (see [5, p175]). □

Theorem 3.2. *If $K \subset \mathbb{R}^3$ is a bounded open convex set, then there exists a point $p \in K$ which is the point of concurrence of six normals from points on ∂K.*

Proof. Let $V \subseteq \overline{K}$ be the set of points $p \in \overline{K}$ such that d_p on ∂K has a unique local minimum. Similarly define $W \subseteq \overline{K}$ to be the set of points where d_p has a unique local maximum. It is easy to see that both V and W are closed.

Now if $\overline{K} \neq \partial K \cup V \cup W$ then we are done. Since $\overline{K} - \partial K \cup V \cup W$ is open, by Lemma 3.1, it contains a point p such that $d_p^\epsilon : \partial K_\epsilon \to \mathbb{R}$ is Morse. Now since d_p has ≥ 2 local minima and ≥ 2 local maxima, we know that d_p^ϵ does as well. Any Morse function on $\mathbb{S}^2$

with ≥ 2 local minima and ≥ 2 local maxima must have ≥ 6 critical points. Thus there are six normals from K_ϵ passing through p. Normals to K_ϵ are normal to K, so K has six concurrent normals and we are done. Thus in the remainder of the proof we may assume $\overline{K} = \partial K \cup V \cup W$.

Define $r : V \cup \partial K \to \partial K$ and $R : W \to \partial K$ by $r(p) = \operatorname{argmin} d_p$ and $R(p) = \operatorname{argmax} d_p$. Clearly r and R are continuous, r is the identity on ∂K, and $r \neq R$ on $(V \cup \partial K) \cap W$. These functions contradict Lemma 3.3 (below), so we are done. ☐

Lemma 3.3. *Let $K \subset \mathbb{R}^n$ be a bounded open convex set. Let $A, B \subseteq \overline{K}$ be two closed sets so that $\partial K \subseteq A$. Then there does not exist any pair of functions with the following properties:*

- *$r : A \to \partial K$ and $R : B \to \partial K$ are continuous*
- *$r \neq R$ on $A \cap B$*
- *$r = id_{\partial K}$ on ∂K.*

Proof. Let $E = \{(p, q) \in B \times \partial K \mid q \neq R(p)\}$ in the subspace topology. Then E is a bundle over B with fiber D^2.

Since $r \neq R$ on $A \cap B$, the function r on $A \cap B \subseteq B$ forms an incomplete (that is, partially defined) section of the bundle E. Since the fiber D^2 is contractible, r extends continuously to a complete (that is, everywhere defined) section of E [11, p55]. After this extension, we see that r is a continuous map $\overline{K} \to \partial K$ and by assumption, r is the identity on ∂K. This is a contradiction, so we are done. ☐

4 The case $n = 4$

Theorem 4.1. *If $K \subset \mathbb{R}^4$ is a bounded open convex set whose boundary ∂K is of class $C^{1,1}$, then there exists a point $p \in K$ which is the point of concurrence of eight normals from points on ∂K.*

If ∂K is of class C^1, then either there exists a point $p \in K$ where eight normals concur or there exists a point $p \in \partial K$ where infinitely many normals concur.

Proof. The first statement trivially follows for $K \subset \mathbb{R}^4$ of class $C^{1,1}$ by applying the second statement to a parallel body $K_{-\epsilon}$ (with $K_{-\epsilon} + B_\epsilon = K$, compare Section 1.1). Below we prove the second statement.

Our strategy is the same as in the proof of Theorem 3.2: if d_p has ≤ 7 critical points for all $p \in K$, and finitely many critical points for $p \in \partial K$, then we will eventually produce functions r and R which contradict Lemma 3.3.

Let $V_0 \subseteq \overline{K}$ be the set of those p such that d_p has a unique local minimum. Define $r_0 : V_0 \to \partial K$ to be $r_0(p) = \operatorname{argmin} d_p$. Similarly define $W_0 \subseteq \overline{K}$ and $R_0 : W_0 \to \partial K$, this time concerning the maximum. As before, V_0 and W_0 are closed and r_0 and R_0 are continuous. Let $U_0 = \overline{K} - V_0 - W_0$.

Let $m_2 \subseteq \overline{K}$ be the set of points p such that d_p has exactly two local minima. Define $\lambda : m_2 \to \mathbb{R}$ as follows: $\lambda(p)$ is the minimum value a_0 such that for all $a \geq a_0$, $d_p^{-1}((-\infty, a])$ is connected (if d_p is Morse and has only one index 1 critical value, then $\lambda(p)$ is that critical value). Observe that λ is continuous (although it need not extend continuously to $\overline{m_2}$). Similarly define M_2 and Λ, this time concerning maxima.

Now by definition, $m_2 \cap M_2 \subseteq U_0$. If this inclusion is proper, then we are done. We can find a p so that d_p has ≥ 3 local minima and ≥ 2 local maxima (or ≥ 2 local minima and ≥ 3 local maxima). Using Lemma 3.1, we can find a nearby p so that d_p^ϵ is Morse, and also has ≥ 3 local minima and ≥ 2 local maxima (or ≥ 2 local minima and ≥ 3 local maxima). But any Morse function on $\mathbb{S}^3$ with ≥ 3 local minima and ≥ 2 local maxima must have ≥ 8 critical points, so we have found eight concurrent normals. Thus in the remainder of the proof we may assume $U_0 = m_2 \cap M_2$.

Define $\ell(p) = \max d_p - \min d_p$. We can assume $\min_{p \in \overline{K}} \ell(p) > 0$, since if $\ell(p) = 0$ then K is a sphere with center p and the conclusion is trivial.

Let $H \subseteq U_0$ be those points $p \in U_0$ for which $\lambda(p) \geq \Lambda(p)$ (if d_p is Morse and has unique index 1 and index 2 critical points, then $\lambda(p) \geq \Lambda(p)$ means that the index 1 critical value of d_p is greater than or equal to the index 2 critical value). Note $\overline{H} \cap U_0 = H$ since λ and Λ are continuous.

If H turns out to be empty, then the proof becomes much easier. We can skip to Remark 4.5 below, from which point the proof is in the same vein as the proof of Theorem 3.2. If H is empty, then the list of properties in Remark 4.5 all hold if one replaces the subscript 1's with subscript 0's (and define $r = \mathrm{id}_{\partial K}$ on ∂K). The reader is at this point encouraged to read the short argument remaining after Remark 4.5, keeping in mind that this argument is all that is necessary when H is empty.

Thus, what we need to show between here and Remark 4.5, is how to extend r_0 and R_0 continuously to larger domains whose union contains H (these new functions will then be r_1 and R_1). The first step is to define the functions π and Π on H which will serve to extend r_0 and R_0 respectively.

The following two lemmas are the crux of the proof.

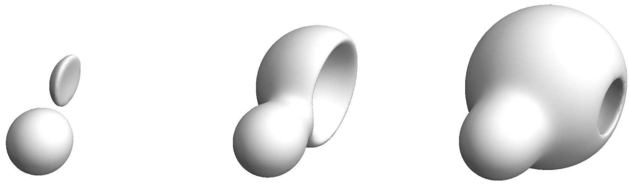
Lemma 4.2 (Structure of d_p for $p \in H$)**.** *Suppose d_p has ≤ 7 critical points for $p \in H$. Then for $p \in H$, there are distinguished local minima and maxima, $\pi(p)$ and $\Pi(p)$, of d_p. Furthermore, $\pi, \Pi : H \to \partial K$ are continuous, and $d_p(\pi(p)) < d_p(\Pi(p))$.*

Proof. The reader may find it helpful to refer to Figure 1 for a schematic of the sublevel set evolution of a function $d_p : \partial K \to \mathbb{R}$ in the case that $\lambda(p) < \Lambda(p)$. Figure 2 shows the case $\lambda(p) \geq \lambda(p)$.

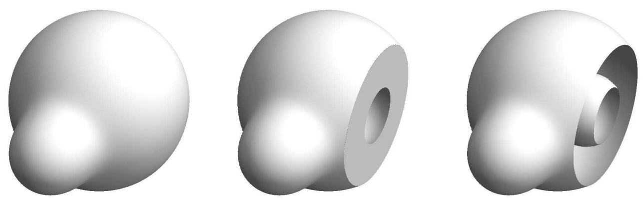
Observe that in the case $\lambda(p) \geq \Lambda(p)$, the sublevel sets must form a second homology group *before* the two components of the sublevel sets merge. This should be especially clear from Figure 2. Thus the two minima of d_p are *distinguishable* when $\lambda(p) \geq \Lambda(p)$, whereas they are not if $\lambda(p) < \Lambda(p)$. We define $\pi(p)$ to be the local minimum of d_p which is in the component of the sublevel set $d_p^{-1}((-\infty, \Lambda(p)])$ which has nonzero H_2.

Let us put this on more rigorous footing as follows. For ease of notation below, we write λ for $\lambda(p)$ and similarly for Λ.

The argument is easiest when $\lambda > \Lambda$. In this case, we see that $d_p^{-1}([\frac{1}{2}(\lambda+\Lambda), \infty))$ has two connected components. Thus $d_p^{-1}((-\infty, \frac{1}{2}(\lambda+\Lambda)))$ has nonzero H_2. Now $d_p^{-1}((-\infty, \frac{1}{2}(\lambda+\Lambda)))$ also has two connected components (one for each local minimum of d_p). If both components had nonzero H_2, then d_p would have ≥ 3 local maxima, which we know is not the case. Thus

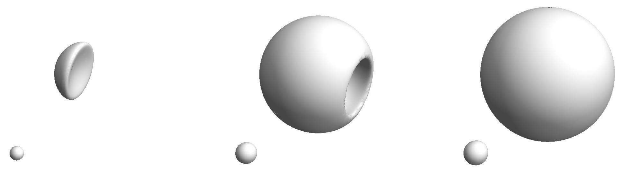


(a) $v < \lambda$ (b) $\lambda < v < \Lambda$ (c) $\lambda < v < \Lambda$

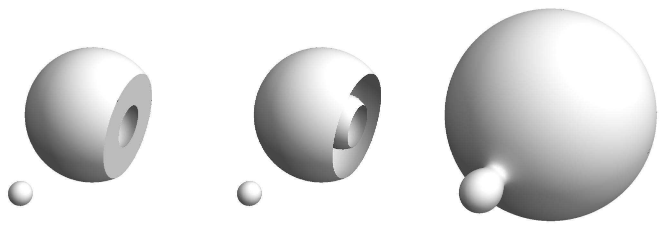


(d) $\Lambda < v$ (e) $\Lambda < v$ sublevel set cross section (f) $\Lambda < v$ level set cross section

Figure 1: This the sublevel set evolution schematic of the case $\lambda < \Lambda$. The subcaptions show the value generating the sublevel set. The sublevel sets are subsets of $\partial K \cong \mathbb{S}^3$, which we have represented as $\mathbb{R}^3$ (with a point at infinity).



(a) $v < \Lambda$ (b) $v < \Lambda$ (c) $\Lambda < v < \lambda$



(d) $\Lambda < v < \lambda$ sublevel set cross section

(e) $\Lambda < v < \lambda$ level set cross section

(f) $\lambda < v$

Figure 2: This the sublevel set evolution schematic of the case $\lambda \geq \Lambda$. The subcaptions show the value generating the sublevel set. The sublevel sets are subsets of $\partial K \cong \mathbb{S}^3$, which we have represented as $\mathbb{R}^3$ (with a point at infinity).

exactly one component of $d_p^{-1}((-\infty, \frac{1}{2}(\lambda+\Lambda)])$ has nonzero H_2, and we define $\pi(p)$ to be the minimum corresponding to that component.

Some care must be taken in the case that $\lambda = \Lambda$. In this case, we see $d_p^{-1}((-\infty, \Lambda])$ is actually connected. Now let us consider the points in $d_p^{-1}((-\infty, \Lambda])$ where the two components of $d_p^{-1}((-\infty, \Lambda))$ merge (call these the wedge points). By this we mean points in $d_p^{-1}((-\infty, \Lambda])$ which are in the closure of both of the components of $d_p^{-1}((-\infty, \Lambda))$. Now it is easy to see that these wedge points are critical points of d_p. This is because in a neighborhood of any noncritical point of d_p, we can find local coordinates in which d_p is just $(u, v, w) \mapsto u$. Now if the set of wedge points is ever infinite, we are done, since we have found infinitely many critical points of d_p, and thus infintely many normals to K concurrent at p.

In fact, we are done if there is ever more than one wedge point. To see this, suppose there are finitely many wedge points. Let A and B denote the closures of the two components of $d_p^{-1}((-\infty, \Lambda))$. Let $a_1, \dots, a_k \in A$ and $b_1, \dots, b_k \in B$ denote the k wedge points, so that $d_p^{-1}((-\infty, \Lambda]) = A \vee_{(a_1,b_1),\dots,(a_k,b_k)} B$ where we have identified a_i with b_i. We see from this that $\operatorname{rk} H_1(d_p^{-1}((-\infty, \Lambda]))) \geq k-1$. Now if we take $p_1 \in K$ sufficiently close to p, and pick a small $\epsilon > 0$, so that $d_{p_1}^{\epsilon}$ is Morse, then we see that $d_{p_1}^{\epsilon}$ has ≥ 2 local minima and ≥ 2 local maxima. Also, if $k \geq 2$, then there is some sublevel set of $d_{p_1}^{\epsilon}$ with nonzero H_1. Thus $d_{p_1}^{\epsilon}$ also has ≥ 2 index 1 critical points. Hence $d_{p_1}^{\epsilon}$ has ≥ 8 total critical points so we are done.

Thus we can assume that the wedge point is unique. Hence $d_p^{-1}((-\infty, \Lambda])$ is $A \vee_{(a,b)} B$. Thus we see that $H_2(d_p^{-1}((-\infty, \Lambda]) = H_2(A) \oplus H_2(B)$. Now from this we can see how to finish as above. Since $d_p^{-1}((-\infty, \Lambda])$ separates ∂K into two and only two components, we see that exactly one of $H_2(A)$ and $H_2(B)$ is nonzero. We then take $r(p)$ to be the minimum corresponding to the set with nonzero H_2.

Now it is also easy to see that π cannot switch between the two local minima, and each local minimum depends continuously on p, so π is continuous.

We define Π in an exactly symmetric way to π. It is evident that $d_p(\pi(p)) < d_p(\Pi(p))$. ☐

As is remarked in the preceding proof, we may assume the truth of Lemma 4.2 in the remainder of the proof since otherwise we are already done.

Lemma 4.3. *If any of the following is false, then K has infinitely many normals that concur in $\overline{K}$. For $p \in V_0 - W_0$, we have $\lim_{H \ni p_1 \to p} \pi(p_1) = r_0(p)$, and for $p \in W_0 - V_0$, we have $\lim_{H \ni p_1 \to p} \Pi(p_1) = R_0(p)$. Also, for every $p \in V_0 \cap W_0$ and every $\epsilon > 0$, there exists $\delta > 0$ so that $|p_1 - p| < \delta$ implies that either $|\pi(p_1) - r_0(p)| < \epsilon$ or $|\Pi(p_1) - R_0(p)| < \epsilon$.*

Proof. Let us show that for $p \in V_0 - W_0$, $\lim_{p_1 \to p} \pi(p_1) = r_0(p)$. Suppose that there were a sequence of $p_n \in H$ so that $p_n \to p \in V_0 - W_0$ and $\pi(p_n) \not\to r_0(p)$; we will show that then K has infinitely many concurrent normals which concur in $\overline{K}$. We can assume WLOG that $\pi(p_n)$ is convergent. Let us now show that since $\pi(p_n) \not\to r_0(p)$, we have $\lambda(p_n) - d_{p_n}(\pi(p_n)) \to 0$. First of all, $\lambda(p_n) > d_{p_n}(\pi(p_n))$ for all n. Now suppose for sake of contradiction that $\lambda(p_n) - d_{p_n}(\pi(p_n)) > \epsilon$ for infinitely many n. We can assume WLOG that $\lambda(p_n) - \epsilon$ is larger than $d_p(r_0(p))$ for sufficiently large n. Now consider $d_p^{-1}((-\infty, \liminf_{n\to\infty} \lambda(p_n) - \epsilon/2))$. We see that $r_0(p)$ and $\lim_{n\to\infty} \pi(p_n)$ are in two different components of this set, which contradicts the fact that d_p has exactly one local minimum. Now since $\Lambda \leq \lambda$ for $p_n \in H$, certainly we also have $\Lambda(p_n) - d_{p_n}(\pi(p_n)) \to 0$.

Consider the sets $S_n = d_{p_n}^{-1}((-\infty, \Lambda(p_n)])$. The reader is encouraged to refer to Figure 2(d) and 2(e). If $\lambda(p_n) > \Lambda(p_n)$, let S'_n be the component of S_n containing $\pi(p_n)$. If $\lambda(p_n) = \Lambda(p_n)$, then there is a unique point whose removal disconnects S_n (see the proof of Lemma 4.2), and we take S'_n to be the union of that unique point and the $\pi(p_n)$ component of S_n minus that point.

The sets S'_n separate the two maxima of d_{p_n} (that is, they are in different connected components of $\partial K - S'_n$). Since $p \notin W_0$, the two maxima of d_{p_n} approach the two maxima of d_p as $n \to \infty$. We know that S'_n stays away from both maxima of d_{p_n}, so for sufficiently large n, S'_n separates the two maxima of d_p. Thus we see that *the set of limit points of S'_n as $n \to \infty$ is infinite.* But let us now observe that any limit point of S'_n as $n \to \infty$ must be a critical point of d_p. To see this, it is easiest to prove that a regular point of d_p cannot be a limit point of S'_n. First observe that the values of d_{p_n} at points in S'_n are contained in $[\pi(p_n), \Lambda(p_n)]$, and interval whose length is going to zero. If $x \in \partial K$ is a regular point of d_p, then we can find local coordinates in which d_p is just $(u, v, w) \mapsto u$. Then if we pick n large enough so that the length of the interval $[\pi(p_n), \Lambda(p_n)]$ is much smaller than the range of d_p in our neighborhhod of x, then we see that S_n just corresponds to the set where the coordinate u is $\leq a$, contradicting our fact that the values of d_p in S'_n are sandwiched in an interval of much smaller length. Hence x cannot be a limit point of S'_n Since d_p has infinitely many critical points, there are infinitely many normals that concur at $p \in \overline{K}$.

The statement that for $p \in W_0 - V_0$, $\lim_{p_1 \to p} \Pi(p_1) = R_0(p)$ follows from an exactly symmetric argument.

To demonstrate the last part of the lemma, we argue as before and again consider the sets S'_n. However this time, we use the fact that they separate the maximum of d_{p_n} other than $\Pi(p_n)$ from the minimum of d_{p_n} other than $\pi(p_n)$. Since by assumption $\pi(p_n)$ stays away from the minimum of d_p, we see that the other minimum of d_{p_n} must approach the minimum of d_p. The same is true for the maximum. Thus we see that we get a similar situation as above, namely that for sufficiently large n, S'_n separates the minimum and the maximum of d_p (this is true since S'_n stays away from both the minimum other than $\pi(p_n)$ and the maximum other than $\Pi(p_n)$ of d_{p_n}). Now by the same reasoning as above, we see that S'_n has infinitely many limit points, all of which must be critical points of d_p, showing again that there are infinitely many normals concurrent at p. ☐

In the remainder of the proof we may assume the truth of the statements in Lemma 4.3 since otherwise we are already done.

As stated above, the two preceding lemmas are the crux of the proof. I do not have a good conceptual reason why they should be true, though it seems key that $d_p(\pi(p)) < d_p(\Pi(p))$.

The next conceptual step is to stitch r_0 and π together so they form a continuous function (and do the same with R_0 and Π). By Lemma 4.3, we have $\lim_{p_1 \to p} \pi(p_1) = r_0(p)$ when $p \in V_0 - W_0$. However this can fail for $p \in V_0 \cap W_0$. Thus our first step is to take care of continuity at points $p \in V_0 \cap W_0$. We will define sets $v_1, w_1 \subseteq H$ whose union is H so that $\lim_{v_1 \ni p_1 \to p} \pi(p_1) = r_0(p)$ for $p \in V_0 \cap W_0$ and similarly for R_0 and w_1.

Let $\eta > 0$. Let v_1° be the set of $p \in H$ where $d_p(\pi(p)) - \min d_p \leq \eta\ell(p)$. Let w_1° be the set of $p \in H$ where $d_p(\Pi(p)) - \min d_p \geq \eta\ell(p)$. Any sufficiently small $\eta > 0$ will do; we will discuss how to choose it later. Observe that $v_1^\circ \cup w_1^\circ = H$ (since $d_p(\pi(p)) < d_p(\Pi(p))$) and that v_1° and w_1° are closed as subsets of H.

We will now construct sets $v_1 \subseteq v_1^\circ$ and $w_1 \subseteq w_1^\circ$. If $V_0 \cap W_0 = \varnothing$, then $v_1 = v_1^\circ$ and $w_1 = w_1^\circ$ and we pick up again after the definition of v_1 and w_1 in the case $V_0 \cap W_0 \neq \varnothing$ is described below.

We define:

$$\epsilon^\circ(\delta) := \max\Bigg\{ \max_{\substack{p \in V_0 \cap W_0 \\ p_2 \in H \\ |p-p_2| \leq \delta}} \left[\min\left(\min_{p_2 \in v_1^\circ} |\pi(p_2) - r_0(p)|, \min_{p_2 \in w_1^\circ} |\Pi(p_2) - R_0(p)|\right)\right], \\ \max_{\substack{p,q \in V_0 \\ |p-q| \leq 2\delta}} 2|r_0(p) - r_0(q)|, \max_{\substack{p,q \in W_0 \\ |p-q| \leq 2\delta}} 2|R_0(p) - R_0(q)|, \delta\Bigg\} \tag{4.1}$$

The expression $\min(\min_{p_2 \in v_1^\circ} |\pi(p_2) - r_0(p)|, \min_{p_2 \in w_1^\circ} |\Pi(p_2) - R_0(p)|)$ above requires some explanation. What we mean is for this to equal:

$$\begin{cases} \min\left(|\pi(p_2) - r_0(p)|, |\Pi(p_2) - R_0(p)|\right) & p_2 \in v_1^\circ \cap w_1^\circ \\ |\pi(p_2) - r_0(p)| & p_2 \in v_1^\circ - w_1^\circ \\ |\Pi(p_2) - R_0(p)| & p_2 \in w_1^\circ - v_1^\circ \end{cases} \tag{4.2}$$

Clearly $\epsilon^\circ(\delta)$ is increasing and satisfies $\epsilon^\circ(\delta) \to 0$ as $\delta \to 0$ (this is a consequence of Lemma 4.3 and the compactness of $V_0 \cap W_0$). For technical reasons, we need to replace ϵ° with a function that satisfies an additional property: we need it to be *upper semicontinuous* (so that v_1 and w_1 defined below are closed). To achieve this, all that is necessary is to change the value of ϵ° at each jump discontinuity so that it assumes the larger value (i.e. the limit from the right). We call this modified function $\epsilon(\delta) := \limsup_{\delta_1 \to \delta} \epsilon^\circ(\delta_1)$, which the reader can check gives us what we want. Clearly we have $\epsilon(\delta)$ is increasing and satisfies $\epsilon(\delta) \to 0$ as $\delta \to 0$. Also observe that $\epsilon(\delta) \geq \epsilon^\circ(\delta)$ and that ϵ is upper semicontinuous.

Define:

$$v_1 = \Big\{p_1 \in v_1^\circ \;\Big|\; |\pi(p_1) - r_0(p)| \leq \epsilon(|p_1 - p|) \text{ for} \\ \text{at least one of the } p \in V_0 \cap W_0 \text{ closest to } p_1\Big\} \tag{4.3}$$

$$w_1 = \Big\{p_1 \in w_1^\circ \;\Big|\; |\Pi(p_1) - R_0(p)| \leq \epsilon(|p_1 - p|) \text{ for} \\ \text{at least one of the } p \in V_0 \cap W_0 \text{ closest to } p_1\Big\} \tag{4.4}$$

Observe that $v_1 \cup w_1 = H$ (by the definition of ϵ and since $v_1^\circ \cup w_1^\circ = H$) and that v_1 and w_1 are closed as subsets of H (since ϵ is upper semicontinuous).

Let us now observe that the following two functions are continuous:

$$p \mapsto \begin{cases} r_0(p) & p \in V_0 \\ \pi(p) & p \in v_1 \end{cases} \qquad p \mapsto \begin{cases} R_0(p) & p \in W_0 \\ \Pi(p) & p \in w_1 \end{cases} \tag{4.5}$$

Consider for instance the first function. Continuity at points in $V_0 - W_0$ is confirmed by Lemma 4.3. Continuity at points in $V_0 \cap W_0$ follows from the definition of v_1, using the fact that $\lim_{\delta\to 0^+}\epsilon(\delta)=0$. Continuity on v_1 is clear since π is continuous and V_0 is closed. Continuity of the second function is similarly demonstrated.

It is no good however if either of these functions has no continuous extensions, which can be the case if $V_0 \cup v_1$ or $W_0 \cup w_1$ is not closed! Thus we proceed to restrict these functions to closed domains.

The only way that $V_0 \cup v_1$ can fail to be closed is if there is a sequence of points in v_1 which approachess some point in $W_0 - V_0$. Thus we will remove an open disc of radius $\delta(p)$ centered at each point $p \in W_0 - V_0$ from v_1. We define $\delta(p)$ for $p \in W_0 - V_0$ to be the largest $\delta > 0$ with the following properties:

- $2\delta \le \operatorname{dist}(p, V_0)$
- The open δ neighborhood of p intersect H is contained in w_1

Lemma 4.4. *Given $p \in W_0 - V_0$, there exists a $\delta > 0$ satisfying the above.*

Proof. If $V_0 \cap W_0 = \varnothing$, then is nothing to prove. Otherwise, let L be the distance from p to the closest point in $V_0 \cap W_0$ to p. Now it suffices to find $\delta > 0$ so that:

$$\epsilon(L-\delta) \ge |\Pi(p_1) - R_0(p_3)| \tag{4.6}$$

for $|p_1 - p| \le \delta$ and $p_1 \in H$, where $p_3 = p_3(p_1)$ denotes the closest point in $V_0 \cap W_0$ to p_1. Now the right hand side of equation (4.6) is bounded above by $|\Pi(p_1) - R_0(p)| + |R_0(p) - R_0(p_3)|$. First restrict $\delta < L/3$ so that it suffices to make $|\Pi(p_1) - R_0(p)| + |R_0(p) - R_0(p_3)| \le \epsilon(2L/3)$. Now the first term can be made arbitrarily small by making δ sufficiently small. We see that the distance from p to p_3 is at most $2\delta + L$ and $p, p_3 \in W_0$. Thus we find that $\epsilon(2L/3) \ge 2|R_0(p) - R_0(p_3)|$ as long as $2\delta + L \le 4L/3$ (by the definition of ϵ), and we can make this true by picking δ sufficiently small. ☐

Now we remove any point $q \in v_1$ with $|q - p| < \delta(p)$ for some $p \in W_0 - V_0$ (that is, we remove the open set $\bigcup_{p\in W_0 - V_0} B_{\delta(p)}(p)$ from v_1). Call the resulting set v_1'. Do the analogous definition and procedure to obtain w_1'. From the definition of $\delta(p)$, it is clear that since $v_1 \cup w_1 = H$, we in fact have $v_1' \cup w_1' = H$.

Now we observe that the following functions are continuous and defined on closed domains:

$$r_1^\circ(p) := \begin{cases} r_0(p) & p \in V_0 \\ \pi(p) & p \in v_1' \end{cases} \qquad R_1(p) := \begin{cases} R_0(p) & p \in W_0 \\ \Pi(p) & p \in w_1' \end{cases} \tag{4.7}$$

Let $V_1^\circ = V_0 \cup v_1'$ and $W_1 = W_0 \cup w_1'$ be the closed domains of r_1° and R_1 respectively. Observe that $\overline{K} - V_1^\circ - W_1 = U_0 - H =: U_1^\circ$ is open (as a subset of $\overline{K}$), and is the set of those points in $U_0 = m_2 \cap M_2$ where $\lambda < \Lambda$. Observe also that $r_1^\circ(p)$ is always a local minimum of d_p, and $R_1(p)$ is always a local maximum of d_p. Since ∂K is of class C^1, we can choose $\eta > 0$ small enough so that for $p \in \partial K$, the only local minimum of d_p within distance η of p is p itself. Hence with this choice of η, we can extend the definition of r_1° to include ∂K, define it to be the identity there, and it is still continuous, defined on a closed set, and always equal to a local minimum:

$$r_1(p) := \begin{cases} r_1^\circ(p) & p \in V_1^\circ \\ p & p \in \partial K \end{cases} \tag{4.8}$$

Let $V_1 = V_1^\circ \cup \partial K$ and $U_1 = U_1^\circ - \partial K$.

Remark 4.5. Let us now catalogue our functions and their properties:

- $r_1 : V_1 \to \partial K$ and $R_1 : W_1 \to \partial K$ are continuous
- V_1 and W_1 are closed
- For all $p \in V_1$, $r_1(p)$ is a local minimum of d_p
- For all $p \in W_1$, $R_1(p)$ is a local maximum of d_p
- $\partial K \subseteq V_1$ and r_1 is the identity on ∂K
- $\overline{K} - (V_1 \cup W_1) =: U_1$ is open
- $U_1 \subseteq m_2 \cap M_2$ and $\lambda(p) < \Lambda(p)$ for all $p \in U_1$

From this, the rest of the proof is simply an elaboration on our strategy for Theorem 3.2. Specifically, we may use level sets of d_p to divide ∂K into two balls, one containing the two local minima and one containing the two local maxima. Such a division is possible exactly because $\lambda(p) < \Lambda(p)$.

We define the cutoff functions $c_1, c_2 : U_1 \to \mathbb{R}$ as $c_1(p) = \frac{2}{3}\lambda(p) + \frac{1}{3}\Lambda(p)$ and $c_2(p) = \frac{1}{3}\lambda(p) + \frac{2}{3}\Lambda(p)$.

Lemma 4.6. *For $p \in V_1$, $\liminf_{p_1 \to p} c_1(p) \geq d_p(r_1(p))$; if equality holds, then d_p has only one local minimum (and therefore $r_1(p)$ is that minimum). We have an analogous result for c_2 and R_1 as well.*

Proof. It is easy to see that if d_p had multiple local minima, then c_1 would be bounded below by $\lambda(p) > d_p(r_1(p))$ as $p_1 \to p$. □

Let us consider the structure of the sublevel sets of d_p for some $p \in U_1$ where d_p^ϵ is Morse. The Morse function d_p^ϵ has two local minima and two local maxima. If for some p it had > 1 index 1 or > 1 index 2 critical points, it would have ≥ 8 total critical points and we would be done. Thus we may assume each such Morse function d_p^ϵ has one index 1 critical point and one index 2 critical point. Now since $p \in U_1$, we have $\lambda(p) < \Lambda(p)$, so we refer to Figure 1 for a schematic of the sublevel set evolution. In particular note 1(b). One observes that ${d_p^\epsilon}^{-1}((c_1(p), c_2(p)))$ is homeomorphic to $\mathbb{S}^2 \times (0, 1)$. In particular, we can pick a C^1 embedded sphere that represents the fundamental second homology class of $d_p^{-1}((c_1(p), c_2(p)))$. Such a sphere then separates ∂K into two balls, one of which contains the two local minima of d_p and one of which contains the two local maxima of d_p. The following lemma, whose proof may be omitted on a first reading, shows that in fact we can pick such a C^1 embedded sphere for all $p \in U_1$.

Lemma 4.7. *There exists a family of C^1 embedded spheres:*

$$\mathcal{S}(p) \subset d_p^{-1}((c_1(p), c_2(p))) \tag{4.9}$$

varying continuously with $p \in U_1$ which separate the two local minima of d_p from the two local maxima of d_p.

Proof. For $p \in U_1$, d_p has the following six critical points: two local minima, two local maxima, one critical point with critical value $\lambda(p)$, and one critical point with critical value $\Lambda(p)$. Now we only need to take care of the case that there is never more than one additional critical point. Let $G \subseteq U_1$ be the set on which d_p has a seventh critical point whose critical value is in $[c_1(p), c_2(p)]$. Observe that $\overline{G} \cap U_1 = G$. Now the function $\gamma : G \to [c_1(p), c_2(p)]$ giving the critical value is continuous. Extend γ continuously to all of U_1 so that it still satisfies $\gamma : U_1 \to [c_1(p), c_2(p)]$.

Thus we see that for $p \in U_1$ and $g \in (c_1(p), c_2(p)) - \{\gamma(p)\}$, $d_p^{-1}(g)$ forms a C^1 embedded sphere in ∂K with the desired property. One may ask why $d_p^{-1}(g)$ has this property. One may choose a p_1 close to p and an $\epsilon > 0$ so that $d_{p_1}^\epsilon$ is Morse, and thus ${d_{p_1}^\epsilon}^{-1}(g)$ is a C^1 embedded sphere with the desired property. Letting $p_1 \to p$ and $\epsilon \to 0$, we see that $d_p^{-1}(g)$ is also a C^1 embedded sphere, and that it separates the two local minima of d_p from the two local maxima of d_p.

Define $\alpha(p) = \frac{3}{5}\lambda(p) + \frac{2}{5}\Lambda(p)$ and $\beta(p) = \frac{2}{5}\lambda(p) + \frac{3}{5}\Lambda(p)$, so that:

$$c_1(p) < \alpha(p) < \beta(p) < c_2(p) \tag{4.10}$$

Let $Y \subseteq U_1$ be the set of $p \in U_1$ where $\gamma(p) \in [\alpha(p), \beta(p)]$. Over Y, let us consider two families of spheres given by $d_p^{-1}(\frac{1}{2}c_1(p) + \frac{1}{2}\alpha(p))$ and $d_p^{-1}(\frac{1}{2}\beta(p) + \frac{1}{2}c_2(p))$. This is a continuous family of C^1 spheres. They are disjoint and divide ∂K into three components: a B^3 containing the two local minima of d_p, a B^3 containing the two local maxmima of d_p, and an $\mathbb{S}^2 \times [0, 1]$ which is the set $d_p^{-1}([\frac{1}{2}c_1(p) + \frac{1}{2}\alpha(p), \frac{1}{2}\beta(p) + \frac{1}{2}c_2(p)])$. Now let us fix diffeomorphisms from $\mathbb{S}^2 \times [0, 1]$ to this last component, the diffeomorphisms varying continuously with $p \in Y$. Thus for $(p, t) \in Y \times [0, 1]$, we have a continuously varying family of C^1 embedded spheres $\mathcal{S}(p, t)$, where in particular $\mathcal{S}(p, 0) = d_p^{-1}(\frac{1}{2}c_1(p) + \frac{1}{2}\alpha(p))$ and $\mathcal{S}(p, 1) = d_p^{-1}(\frac{1}{2}\beta(p) + \frac{1}{2}c_2(p))$. Now it is easy to see that the following definition suffices:

$$\mathcal{S}(p) = \begin{cases} d_p^{-1}(\frac{1}{2}\beta(p) + \frac{1}{2}c_2(p)) & \gamma(p) \leq \alpha(p) \\ \mathcal{S}(p, \mathrm{LIN}_p(\gamma(p))) & \alpha(p) \leq \gamma(p) \leq \beta(p) \\ d_p^{-1}(\frac{1}{2}c_1(p) + \frac{1}{2}\alpha(p)) & \beta(p) \leq \gamma(p) \end{cases} \tag{4.11}$$

where $\mathrm{LIN}_p : [\alpha(p), \beta(p)] \to [0, 1]$ is the order reversing linear homeomorphism $x \mapsto 1 - \frac{x-\alpha}{\beta-\alpha}$. Clearly when p satisfies more than one of the criteria on the right, the two definitions of $\mathcal{S}(p)$ coincide. □

Consider two bundles over U_1, namely those whose fibers over $p \in U_1$ respectively are the two open balls that ∂K is separated into by $\mathcal{S}(p)$. Restrict the bundle containing the two local minima of d_p to the base space of those $p \in U_1$ satisfying $\mathrm{dist}(p, V_1) \leq \mathrm{dist}(p, W_1)$ and call the resulting bundle ω and the resulting base space $|\omega|$. Similarly define the bundle Ω over the base space $|\Omega|$. Observe $|\omega| \cup |\Omega| = U_1$.

We now wish to extend r_1 continuously from V_1 to $V_1 \cup |\omega|$ (observe $V_1 \cup |\omega|$ is closed). We do this as follows. Extend r_1 arbitrarly to a closed neighborhood of V_1. Remove any extension which is not in $V_1 \cup |\omega|$. Additionally, remove from this extension any point in $|\omega|$ where $d_p(r_1(p)) > c_1(p)$ (this is necessary since at such points, the extended function may fail to be a section of ω). Call the resulting function $\tilde{r}_1$ (which is defined on a closed set containing V_1 and contained in $V_1 \cup |\omega|$). Now the fiber of ω is contractible. Viewing

$\tilde{r}_1$ restricted to $|\omega|$ as a partial section of ω, we see that we may extend $\tilde{r}_1$ continuously to a section of ω. Call this extended function r_2, which is continuous on $|\omega|$. Let us now see that in fact r_2 is continuous at points in V_1 as well. Clearly, one only need check continuity at those points of V_1 whose every neighborhood contains points which were cutoff (that is, cutoff when we removed points where $d_p(r_1(p)) > c_1(p)$). But such points satisfy $\liminf_{p_1 \to p} c_1(p) = d_p(r_1(p))$. Thus at such points, $r_1(p)$ must be the unique global minimum of d_p. Then we see that $d_{p_1}(r_2(p_1)) \to d_p(r_1(p)) = \min d_p$ as $p_1 \to p$ implies $r_2(p_1) \to r_1(p)$, so we are done. Hence $r_2 : V_1 \cup |\omega| \to \partial K$ is continuous and is a section of the bundle ω.

Perform the analogous extension for R_1 to a section of Ω to obtain R_2.

Now we have two continuous functions (r_2 and R_2) to ∂K defined on closed subsets of $\overline{K}$ ($V_1 \cup |\omega|$ and $W_1 \cup |\Omega|$) whose union is $\overline{K}$, which are never equal to each other, and one of which is the identity on ∂K. This contradicts Lemma 3.3, so we are done. ☐

5 Interpretation in terms of homotopy groups of function spaces

Recall that very little geometry was used in the proof of Theorem 4.1 (or in Lemma 2.1 or Theorem 3.2). What we have really shown is that $\partial K \ni p \mapsto d_p \in F_{2n-1}$ cannot be extended to all $p \in \overline{K}$, where F_{2n-1} is the (suitably defined) set of functions on ∂K with $\leq 2n-1$ critical points. We now give a rigorous definition of such a space of functions and state a conjecture (Conjecture 1.2) which is closely related to Conjecture 1.1, though not *a priori* equivalent to it (see, however, Lemma 5.7 below). We then note that the proofs given above apply directly to prove this conjecture for $2 \leq n \leq 4$.

Definition 5.1. Let M be a closed manifold. Let $G_k^\circ(M)$ denote the set of functions $f \in C^1(M)$ with $\leq k$ critical points (points where $\nabla f = 0$). We topologize $G_k^\circ(M)$ using the topology of $C^1(M)$. Define $G_k(M) \subset G_k^\circ(M)$ to be the closure in $G_k^\circ(M)$ of the Morse functions in $G_k^\circ(M)$.

One can also imagine taking $G_k^\ell(M)$ to be as above but instead using $C^\ell(M)$. Here we just stick to the simplest case.

Definition 5.2. We define the *canonical embedding* $e : \mathbb{S}^{n-1} \hookrightarrow C^1(\mathbb{S}^{n-1})$ by sending $x \in \mathbb{S}^{n-1}$ to the function $p \mapsto x \cdot p$.

Problem 5.3. Study the stable and unstable homology and homotopy groups of the filtration:

$$\mathbb{S}^{n-1} \overset{e}{\hookrightarrow} G_2(\mathbb{S}^{n-1}) \hookrightarrow G_3(\mathbb{S}^{n-1}) \hookrightarrow G_4(\mathbb{S}^{n-1}) \hookrightarrow G_5(\mathbb{S}^{n-1}) \hookrightarrow \cdots \tag{5.1}$$

One expects that $G_{2k}(\mathbb{S}^{n-1}) \hookrightarrow G_{2k+1}(\mathbb{S}^{n-1})$ is a homotopy equivalence since every Morse function on $\mathbb{S}^{n-1}$ has an even number of critical points. In this case, one then studies:

$$\mathbb{S}^{n-1} \overset{e}{\hookrightarrow} G_2(\mathbb{S}^{n-1}) \hookrightarrow G_4(\mathbb{S}^{n-1}) \hookrightarrow G_6(\mathbb{S}^{n-1}) \hookrightarrow G_8(\mathbb{S}^{n-1}) \hookrightarrow \cdots \tag{5.2}$$

Conjecture 5.4. There exists $r : G_{2n-1}(\mathbb{S}^{n-1}) \to \mathbb{S}^{n-1}$ so that $r \circ e$ is the identity on $\mathbb{S}^{n-1}$. In other words, $G_{2n-1}(\mathbb{S}^{n-1})$ retracts to $\operatorname{Im} e$.

Conjecture 5.5. For all $k \geq 2n$, $G_k(\mathbb{S}^{n-1})$ is contractible.

A weaker version of Conjecture 5.5 is that for all sufficiently large k, $G_k(\mathbb{S}^{n-1})$ is contractible. In fact, finding any k such that $G_k(\mathbb{S}^{n-1})$ is contractible would be interesting.

5.1 The connection with Conjecture 1.1

Lemma 5.6. *The canonical embedding* $e : \mathbb{S}^{n-1} \hookrightarrow G_{2n}(\mathbb{S}^{n-1})$ *is null homotopic.*

Proof. We define an ellipsoid $E \subset \mathbb{R}^n$ with axes of length 1, $1+\epsilon, \dots, 1+(n-1)\epsilon$. Now this convex body has no point in its interior with more than $2n$ normals to the boundary. Thus the functions d_p for $p \in \overline{E}$ are all in $G_{2n}(\partial E)$. Hence $q \mapsto d_q$ for $q \in \partial E$ is null homotopic. But it is also easy to see that this map is homotopic to the canonical embedding e. ☐

Lemma 5.7. *Suppose Conjecture 1.2 is true for some n. Then if $K \subset \mathbb{R}^n$ is a bounded open convex set, it has $2n$ normals which are concurrent somewhere in $\mathbb{R}^n$ (i.e. not necessarily inside K). Additionally, Conjecture 1.1 is true for $K \subset \mathbb{R}^n$ which are C^2 close to $\mathbb{S}^{n-1}$.*

Proof. Given $K \subset \mathbb{R}^n$, suppose it is not the case that some $2n$ normals concur somewhere in $\mathbb{R}^n$. Then $B_1 + tK$ certainly does not have $2n$ concurrent normals for $t \in (0, 1]$. Now we see that varying t from 0 to 1 gives a homotopy between $\partial(B_1+K) \ni p \mapsto d_p \in G_{2n-1}(\partial(B_1+K))$ and $\partial B_1 \ni p \mapsto d_p \in G_{2n-1}(\mathbb{S}^{n-1})$. Since the latter is not null homotopic (by Conjecture 1.2 since it is homotopic to e), the former also cannot be null homotopic. But since $B_1 + K$ does not have $2n$ concurrent normals, we see that the former map is in fact null homotopic, a null homotopy given simply by $\overline{B_1 + K} \ni p \mapsto d_p \in G_{2n-1}(\partial(B_1 + K))$ (note the use of Lemma 3.1 to show that $d_p \in G_{2n-1}(\partial(B_1 + K))$!). This is a contradiction, so we are done.

If $K \subset \mathbb{R}^n$ is C^2 close to $\mathbb{S}^{n-1}$, then $\partial K \ni p \mapsto d_p \in G_2(\partial K)$ is homotopic to e, so in particular is not null homotopic in G_{2n-1}. But if Conjecture 1.1 failed for K, then $\partial K \ni p \mapsto d_p \in G_2(\partial K)$ would be null homotopic in G_{2n-1}, since every $d_p \in \overline{K}$ would have $\leq 2n-1$ critical points, and d_p is Morse for a dense set of $p \in \overline{K}$. ☐

Theorem 5.8. *Conjecture 1.2 is true for* $2 \leq n \leq 4$.

Proof. In the notation of our proofs of Conjecture 1.1 for $2 \leq n \leq 4$, e is the analogue of $\partial K \ni p \mapsto d_p$, and $r : G_{2n-1}(\mathbb{S}^{n-1}) \to \mathbb{S}^{n-1}$ is analogous to $r : \overline{K} \to \partial K$. With this correspondence, the proofs of Lemma 2.1, Theorem 3.2, and Theorem 4.1 work essentially as written.

For example, let us sketch the $n = 3$ case, that is, we show that $G_5(\mathbb{S}^2)$ retracts to the image of $e : \mathbb{S}^2 \hookrightarrow G_5(\mathbb{S}^2)$. Let $V \subseteq G_5(\mathbb{S}^2)$ consist of those functions with a unique local minimum. Let $W \subseteq G_5(\mathbb{S}^2)$ be those with a unique local maximum. Let $r : V \to \mathbb{S}^2$ be defined by $r(f) = \text{argmin}_p f(p)$, and similarly define $R : W \to \mathbb{S}^2$.

Now let us see that $V \cup W = G_5(\mathbb{S}^2)$. If this were not the case, then there would be $f \in G_5(\mathbb{S}^2)$ with ≥ 2 local minima and ≥ 2 local maxima. By definition, Morse functions are dense in $G_5(\mathbb{S}^2)$, so there exists a Morse function $f_1 \in G_5(\mathbb{S}^2)$ that is close to f. By requiring f_1 to be sufficiently close to f, we can force f_1 also to have ≥ 2 local minima and ≥ 2 local maxima. Since f_1 is Morse, it must have ≥ 6 critical points, contradicting the fact that it is in $G_5(\mathbb{S}^2)$. Thus $V \cup W = G_5(\mathbb{S}^2)$.

Now consider $E = \{(f,p) \in W \times \mathbb{S}^2 \mid p \neq R(f)\}$. We see that E is a bundle over W with fiber D^2. Since $r \neq R$ on $V \cap W$, r forms a partial section of E. Thus since the fiber is contractible, r extends to an entire section of E. Hence we have a continuous function $r : G_5(\mathbb{S}^2) \to \mathbb{S}^2$. It is easy to see that $(-r) \circ e = \mathrm{id}_{\mathbb{S}^2}$ ($r \circ e$ is the antipodal map), since $\operatorname{Im} e \subseteq V$, where we have defined r explicitly. □

5.2 Morse-Smale functions

If we restrict ourselves to Morse-Smale functions, one observes the following.

Lemma 5.9. *If $f : \mathbb{S}^{n-1} \to \mathbb{R}$ is a Morse function whose gradient flow is Morse-Smale, and f has $\leq 2n-1$ critical points, then there exists $k < n-1$ such that the union of the stable manifolds of all critical points of index $\leq k$ is contractible.*

Proof. Since there are n possible indices and $< 2n$ critical points, there is some k_0 for which there is a unique critical point of index k_0. Call this critical point c_0. Set:

$$k = \begin{cases} 0 & k_0 = 0 \\ k_0 & k_0 > 0 \text{ and } \partial c_0 \neq 0 \\ k_0 - 1 & k_0 > 0 \text{ and } \partial c_0 = 0 \end{cases} \tag{5.3}$$

Let X_k be the union of the stable manifolds of all the critical points of index $\leq k$. Now by examining the cell decomposition of $\mathbb{S}^{n-1}$ induced by f, one observes that $H_*(X_k) = 0$ and $\pi_1(X_k) = 0$, so X_k is contractible as claimed. □

Given the canonical nature of their construction, it is tempting to expect that these contractible subsets may in some weak sense form a bundle over $G_{2n-1}(\mathbb{S}^{n-1})$. Specifically, we might expect it to be a *stratified bundle* (see [1]), the strata of the base space being:

$$S_k = \Big\{ f \in G_{2n-1}(\mathbb{S}^{n-1}) \;\Big|\; k \text{ is the smallest integer satisfying the conclusion of Lemma 5.9} \Big\} \tag{5.4}$$

Now one can try to extend the partial section e^{-1} defined on $\operatorname{Im} e$ to a section of the entire bundle, using the contractibility of the fibers. This would then give a suitable function r satisfying the conclusion of Conjecture 1.2.

It is the author's opinion, however, that this approach as stated above fails even when $n = 4$. Thus in our proof of Theorem 4.1 we have used more intricate techniques to deal with the existence of functions that are not Morse-Smale. Recall that the bulk of the proof was spent dealing with points where $\lambda(p) \geq \Lambda(p)$. This is because it is in this case that d_p can fail to be Morse-Smale: a flow line descending from the index one critical point can flow to the index two critical point. In classical Morse theory, one often assumes the function is self indexing, so this situation cannot occur, but it is not possible to deform our entire parameter space of functions to be self indexing since we have critical points which cancel. We expect that the relative heights of the critical points play a crucial role in determining the homotopy properties of $G_k(M)$.

A result of Klein [8] may be useful in approaching the problem in higher dimensions. It gives a construction of a canonical contractible parameter space of quasi cell decompositions of a manifold given some fixed Morse function.

References

[1] Hans-Joachim Baues and Davide L. Ferrario. Stratified fibre bundles. *Forum Math.*, 16(6):865–902, 2004.

[2] Luis A. Caffarelli. Boundary regularity of maps with convex potentials. II. *Ann. of Math. (2)*, 144(3):453–496, 1996.

[3] Hallard T. Croft, Kenneth J. Falconer, and Richard K. Guy. *Unsolved problems in geometry.* Problem Books in Mathematics. Springer-Verlag, New York, 1991. Unsolved Problems in Intuitive Mathematics, II.

[4] Narsingh Deo and M. S. Klamkin. Existence of four concurrent normals to a smooth closed curve. *Amer. Math. Monthly*, 77:1083–1084, 1970.

[5] E. Heil. Concurrent normals and critical points under weak smoothness assumptions. In *Discrete geometry and convexity (New York, 1982)*, volume 440 of *Ann. New York Acad. Sci.*, pages 170–178. New York Acad. Sci., New York, 1985.

[6] Erhard Heil. Existenz eines 6-Normalenpunktes in einem konvexen Körper. *Arch. Math. (Basel)*, 32(4):412–416, 1979.

[7] Erhard Heil. Korrektur zu: "Existenz eines 6-Normalenpunktes in einem konvexen Körper" [Arch. Math. (Basel) **32** (1979), no. 4, 412–416; MR 80j:52002a]. *Arch. Math. (Basel)*, 33(5):496, 1979/80.

[8] John R. Klein. Coordinate free Morse theory. *Manuscripta Math.*, 71(3):283–294, 1991.

[9] Nicolaas H. Kuiper. Double normals of convex bodies. *Israel J. Math.*, 2:71–80, 1964.

[10] L. Lusternik and L. Schnirelmann. *Méthodes topologiques dans les problèmes variationnels. I. Pt. Espaces à un nombre fini de dimensions. Traduit du russe par J. Kravtchenko.* Paris: Hermann & Cie. 51 S., 5 Fig., 1934.

[11] Norman Steenrod. *The Topology of Fibre Bundles.* Princeton Mathematical Series, vol. 14. Princeton University Press, Princeton, N. J., 1951.

[12] Bernd Wegner. Existence of four concurrent normals to a smooth closed hypersurface of E^n. *Amer. Math. Monthly*, 80:782–785, 1973.

[13] Tudor Zamfirescu. Points on infinitely many normals to convex surfaces. *J. Reine Angew. Math.*, 350:183–187, 1984.