Communications in
Analysis and Geometry
Volume 26, Number 6, 1251–1316, 2018

Compactness and isotopy finiteness for submanifolds with uniformly bounded geometric curvature energies

Sławomir Kolasiński, Paweł Strzelecki,
and Heiko von der Mosel

In this paper, we establish compactness for various geometric curvature energies including integral Menger curvature, and tangent-point repulsive potentials, defined a priori on the class of compact, embedded m-dimensional Lipschitz submanifolds in $\mathbb{R}^n$. It turns out that due to a smoothing effect any sequence of submanifolds with uniformly bounded energy contains a subsequence converging in C^1 to a limit submanifold.

This result has two applications. The first one is an isotopy finiteness theorem: there are only finitely many isotopy types of such submanifolds below a given energy value, and we provide explicit bounds on the number of isotopy types in terms of the respective energy. The second one is the lower semicontinuity — with respect to Hausdorff-convergence of submanifolds — of all geometric curvature energies under consideration, which can be used to minimise each of these energies within prescribed isotopy classes.

1. Introduction

1.1. Introduction and main results

In this paper, we prove compactness and isotopy finiteness for several functionals $\mathcal{E} : \mathscr{C}^{0,1}_{m,n} \to [0,\infty]$ — we refer to them as *geometric curvature energies* — defined on the class $\mathscr{C}^{0,1}_{m,n}$ of all compact, m-dimensional embedded Lipschitz submanifolds of $\mathbb{R}^n$. To reach this goal, we use previously established uniform $C^{1,\alpha}$-a priori estimates on local graph representations to not only prove compactness, but also to gain sufficient geometric rigidity, such that two submanifolds of finite energy that have small Hausdorff-distance are ambiently isotopic. As a consequence of this two-fold regularisation of these energies, we obtain isotopy finiteness: each sub-level set

$$\mathcal{A}^{\mathcal{E}}_{m,n}(E,d) := \big\{ \Sigma \in \mathscr{C}^{0,1}_{m,n} : \ \mathcal{E}(\Sigma) \leq E, \ \operatorname{diam} \Sigma \leq d \big\} \tag{1}$$

contains only finitely many manifolds up to diffeomorphism but also up to *isotopy*. We also give a crude yet explicit estimate of the number of these isotopy classes. In addition, we prove lower semicontinuity of all geometric curvature energies with respect to Hausdorff-convergence, which can be combined with compactness to minimise each energy in a fixed isotopy class.

The compactness and finiteness theorems for abstract (smooth) Riemannian manifolds, in different guises and under several sets of assumptions, date back at least to J. Cheeger's paper [10]. In particular, [10, Thm. 3.1] states that for $n \neq 4$ and any given constant $C < \infty$ there are only finitely many diffeomorphism types of Riemannian manifolds M such that

$$\big\| \, |K_M| \, \big\|^{1/2}_{L^\infty} \cdot \operatorname{Vol}(M)^{1/n} + \frac{\operatorname{diam} M}{\operatorname{Vol}(M)^{1/2}} < C \, .$$

The left-hand side is bounded if, for example, the sectional curvature satisfies $|K_M| \leq 1$, the diameter of M is at most d and the volume — at least v; the lower bound on $\operatorname{Vol} M$ can be replaced by a lower bound on the injectivity radius. Later on, M. Gromov, see [19] and [20, Thm. 8.28], generalised Cheeger's work and introduced the powerful concept of Gromov–Hausdorff convergence, enabling the study of collapse of sequences of manifolds with

bounded curvature, where in the absence of bounds on the injectivity radius singularities can appear in the limit. For a proof of Gromov's compactness theorem with an improvement on the regularity of the limiting metric, we refer to S. Peters [30]. Anderson and Cheeger [4] prove that the space of all Riemannian manifolds with uniform lower bounds on the Ricci curvature Ric_M and the injectivity radius, and uniform upper bounds on the volume, is compact in the C^α topology for any $\alpha < 1$ (meaning C^α convergence of the Riemannian metrics). The same authors in [3] obtain a finiteness theorem for m-dimensional Riemannian manifolds with uniform upper bounds for the diameter and $|\mathrm{Ric}_M|$, uniform lower bounds for the volume, and uniform bounds for the scalar curvature in $L^{m/2}$. Newer developments include the papers by A. Petrunin and W. Tuschmann [31], W. Tuschmann [43], and V. Kapovich, A. Petrunin and W. Tuschmann [23]. In particular, Tuschmann [43] proves that the class $\mathcal{M}(n, C, D)$ of simply connected closed n-dimensional Riemannian manifolds with sectional curvature $|K| \le C$ and diameter $\le D$ contains finitely many diffeomorphism types provided $n \le 6$ (surprisingly, this result fails in each dimension $n \ge 7$).

As Cheeger writes in his survey [11, p. 235], in a passage commenting on one of the versions of his own finiteness theorem, *Intuitively, the idea is that these manifolds can be constructed from a definite numbers of standard pieces.* The same comment applies in the present paper, with one notable difference: all the submanifolds we deal with are embedded in the same $\mathbb{R}^n$, but their Riemannian metrics g induced by this embedding are, typically, *only* of class C^α and not better, so that there is no way to define the classic curvature tensor of g. Instead of that, we pick up a family of geometric 'energies' that can be defined without relying on the C^2 (or $C^{1,1}$) regularity of the underlying manifold; a bound on each of these energies, combined with a bound on the diameter, yields a bound on the number of ambient isotopy types (which is stronger than bounding the number of diffeomorphism types). Each of these energies can, in fact, be defined also for non-smooth sets, more general than Lipschitz submanifolds. It also can be minimised in a given isotopy class.

To state our results precisely, let us introduce the appropriate definitions first. For an $(m+2)$-tuple $(x_0, x_1, \ldots, x_{m+1})$ of points of $\mathbb{R}^n$, we denote the $(m+1)$-dimensional simplex with vertices at the x_i's by $\triangle(x_0, x_1, \ldots, x_{m+1})$. The *discrete Menger curvature* of $(x_0, x_1, \ldots, x_{m+1})$ is defined by

$$\mathcal{K}(x_0, \ldots, x_{m+1}) = \frac{\mathcal{H}^{m+1}(\triangle(x_0, \ldots, x_{m+1}))}{\mathrm{diam}(\{x_0, \ldots, x_{m+1}\})^{m+2}}\,. \tag{2}$$

For $m = 1$, $n = 3$ we have

$$\mathcal{K}(x_0, x_1, x_2) = \frac{\operatorname{Area}(\triangle(x_0, x_1, x_2))}{\max(|x_0 - x_1|, |x_1 - x_2|, |x_2 - x_0|)^3} \leq \frac{1}{4R(x_0, x_1, x_2)},$$

where $R(x_0, x_1, x_2)$ stands for the circumradius[1] of the triangle $\triangle(x_0, x_1, x_2)$.

For a Lipschitz manifold $\Sigma \in \mathscr{C}^{0,1}_{m,n}$, a number $l \in \{1, \ldots, m+2\}$, and $p > 0$ we set

$$\mathcal{E}^l_p(\Sigma) = \int_{\Sigma^l} \sup_{x_l, \ldots, x_{m+1} \in \Sigma} \mathcal{K}(x_0, \ldots, x_{m+1})^p \, d\mathcal{H}^{ml}_{x_0, \ldots, x_{l-1}} \,. \tag{3}$$

The integration in (3) is performed over the product $\Sigma^l = \Sigma \times \ldots \times \Sigma$ of l copies of Σ, with respect to the m-dimensional Hausdorff measure $\mathcal{H}^m$ on each copy, i.e., with respect to $\mathcal{H}^{ml}$ on Σ^l. Before that, one takes the supremum of $\mathcal{K}(x_0, \ldots, x_{m+1})$ with respect to all variables x_i with indices $i \geq l$. (For $l = m + 2$, no supremum is being taken). Please note that formally the integrand is undefined on the diagonal of the product. However, we tacitly omit this issue: the choice of values of the integrand on the diagonal does not affect the value of the integral in (3), as the diagonal is of measure zero in the product.

In particular, the functional $\mathcal{E}^{m+2}_p$ is called the *integral Menger curvature* of Σ.

Besides all the $\mathcal{E}^l_p$ energies, we consider also two other functionals that are defined via averaging the inverse powers of the radii of spheres that are tangent to Σ at one point and pass through another point of Σ. Namely, we write

$$R_{\mathrm{tp}}(x, y) = \frac{|x - y|^2}{2 \operatorname{dist}(y, x + T_x\Sigma)} \tag{4}$$

to denote the radius of the smallest sphere which passes through $y \in \Sigma$ and is tangent to the m-dimensional affine plane $x + T_x\Sigma$. (Note that for a Lipschitz manifold $\Sigma \in \mathscr{C}^{0,1}_{m,n}$ the tangent plane $T_x\Sigma \in G(n, m)$ is indeed well-defined for $\mathcal{H}^m$-almost all $x \in \Sigma$ due to the classic Rademacher theorem.)

[1]The triple integral over the inverse squared circumradius also known as *total Menger curvature* was an essential tool in G. David's proof [12] of the famous Vitushkin conjecture on characterising the one dimensional compact subsets of the complex plane that are removable for bounded analytic functions; see, e.g., [42]. The most obvious generalisation to inverse powers of circumsphere radii of simplices turns out to be too singular for our purposes here; see the discussion in [40, Appendix B].

Set

$$\mathcal{T}_p(\Sigma) = \int_\Sigma \int_\Sigma R_{\mathrm{tp}}(x,y)^{-p}\, d\mathcal{H}^m_x\, d\mathcal{H}^m_y, \tag{5}$$

$$\mathcal{T}^G_p(\Sigma) = \int_\Sigma \sup_{y\in\Sigma} \left(R_{\mathrm{tp}}(x,y)^{-p}\right) d\mathcal{H}^m_x\,. \tag{6}$$

(Again, in (5) it does not matter how $1/R_{\mathrm{tp}}$ is defined on the diagonal in $\Sigma \times \Sigma$.)

For each of the energies $\mathcal{E} \in \{\mathcal{E}^l_p, \mathcal{T}_p, \mathcal{T}^G_p\}$, we write $p_0(\mathcal{E})$ to denote the scaling invariant exponent for $\mathcal{E}$. Since $\mathcal{K}$ and $1/R_{\mathrm{tp}}$ scale like the inverse of length, it is easy to see that $p_0(\mathcal{E})$ equals the product of m and the number of integrals of Σ in the definition of $\mathcal{E}$. Thus,

$$p_0(\mathcal{E}^l_p) = ml, \qquad p_0(\mathcal{T}_p) = 2m, \qquad p_0(\mathcal{T}^G_p) = m. \tag{7}$$

Theorem 1 (finiteness of isotopy types). *Let $E, d > 0$ be some numbers. Assume that $\mathcal{E} \in \{\mathcal{E}^l_p, \mathcal{T}_p, \mathcal{T}^G_p\}$ and $p > p_0(\mathcal{E})$. There are at most $K = K(E,d,m,n,l,p)$ different (ambient) isotopy types in $\mathcal{A}^{\mathcal{E}}_{m,n}(E,d)$.*

Recall that two topological embedded submanifolds Σ_1, Σ_2 of $\mathbb{R}^n$ are *ambient isotopic* if and only if there exists a continuous map $H : \mathbb{R}^n \times [0,1] \to \mathbb{R}^n$ such that

$$\begin{gathered} H_t = H(\cdot,t) \quad \text{is an embedding for each } t \in [0,1]\,, \\ H(x,1) = x \quad \text{for all } x \in \mathbb{R}^n, \quad \text{and} \quad H(\Sigma_2 \times \{0\}) = \Sigma_1\,. \end{gathered} \tag{8}$$

(Note that the inclusion mapping $f_2 : \Sigma_2 \to \mathbb{R}^n$ yields one embedding of Σ_2 in $\mathbb{R}^n$, and $f_1 := H_0 \circ f_2 : \Sigma_2 \to \mathbb{R}^n$ yields another one, so that (8) agrees with the definition of *ambient isotopy* $\tilde{H} : \mathbb{R}^n \times [0,1] \to \mathbb{R}^n \times [0,1]$, $\tilde{H}(x,t) := (H(x,t),t)$ *between the embeddings f_2 and f_1* as in Burde–Zieschang [9, p.2].) However, because of the smoothing properties of all geometric curvature energies described in more detail in Section 1.2 all Lipschitz submanifolds with finite energy are actually of class C^1, so that Theorem 1 is stronger than diffeomorphism finiteness since it even bounds the C^1-isotopy types.

Remark 1.1. We do not have an optimal estimate for the number $K = K(E,d,m,n,l,p)$ in the above theorem. However, a crude check of constants

involved in the argument yields

$$\log\log K \leq c(m,n,l,p)\Big(|\log d| + \log\left(E^{1/p}+1\right)+1\Big) \tag{9}$$

with a constant $c(m,n,l,p)$ that blows up for $p \searrow p_0(\mathcal{E})$. (See Section 6, Remark 6.2). Thus, as expected, for fixed dimensions m and n, K blows up for $E \to \infty$ (with $\mathcal{E}$ and d fixed), and for $p \searrow p_0(\mathcal{E})$ (with E and d fixed).

It is also worth noting that no lower bounds for the volume (or lower bounds for the injectivity radius) are needed in our work. Intuitively, the reason is that the onset of thin tubes or narrow tentacles is being penalised by each of the energies we consider. The same penalisation effect takes care of a quantitative embeddedness: if two roughly parallel sheets of an embedded manifold Σ are too close to each other preventing Σ to be described locally as *one* graph, then there are lots of roughly regular very small simplices with vertices on Σ (and of small tangent spheres passing through a second point of Σ), causing the integrands $\mathcal{K}$ and $1/R_{\mathrm{tp}}$ in (2) and (4) to be very large on a set of positive measure. We will describe the energies' quantitative control on local graph patches more precisely in Section 1.2.

Our next result states that for any geometric curvature energy all sublevel sets are sequentially closed and compact with respect to Hausdorff convergence, and that all these energies are sequentially lower semicontinuous.

Theorem 2 (lower semicontinuity and compactness). *For $E, d \in (0,\infty)$ and for a geometric curvature energy $\mathcal{E} \in \{\mathcal{E}^l_p, \mathcal{T}_p, \mathcal{T}^G_p\}$, where $p > p_0(\mathcal{E})$ and $l \in \{1,2,\ldots,m+2\}$, the following holds.*

(i) *If $\Sigma_j \in \mathcal{A}^{\mathcal{E}}_{m,n}(E,d)$ for all $j \in \mathbb{N}$ and if the Σ_j converge to a compact set $\Sigma \subset \mathbb{R}^n$ with respect to the Hausdorff-metric as $j \to \infty$, then $\Sigma \in \mathcal{A}^{\mathcal{E}}_{m,n}(E,d)$, and moreover,*

$$\mathcal{E}(\Sigma) \leq \liminf_{j\to\infty} \mathcal{E}(\Sigma_j).$$

(ii) *For any sequence $(\Sigma_j)_j \subset \mathcal{A}^{\mathcal{E}}_{m,n}(E,d)$ with $0 \in \Sigma_j$ for all j there is a submanifold $\Sigma \in \mathcal{A}^{\mathcal{E}}_{m,n}(E,d)$ and a subsequence $(\Sigma_{j_k})_k \subset (\Sigma_j)_j$ such that $d_{\mathcal{H}}(\Sigma_{j_k}, \Sigma) \to 0$ as $k \to \infty$.*

With the energies' quantitative control over local graph patches described in more detail in Section 1.2 one actually obtains C^1-compactness of these graph patches, which considerably improves the Hausdorff-convergence

to C^1-convergence in both parts of Theorem 2. Since, roughly speaking, isotopy types stabilise under C^1-convergence we can use Theorem 2 to deduce the following existence result by means of the direct method in the calculus of variations.

Corollary 1 (existence of minimisers in isotopy classes). *Let $\mathcal{E}$, p, E and d be as in Theorem 2. For each reference manifold $M_0 \in \mathcal{A}^{\mathcal{E}}_{m,n}(E,d)$ there exists $\Sigma_0 \in \mathcal{A}^{\mathcal{E}}_{m,n}(E,d)$ such that*

$$\mathcal{E}(\Sigma_0) = \inf\left\{\mathcal{E}(\Sigma) : \Sigma \in \mathcal{A}^{\mathcal{E}}_{m,n}(E,d) \text{ and } \Sigma \text{ is ambient isotopic to } M_0\right\}.$$

Remark. It is easy to see, via simple covering arguments, that Theorems 1 and 2 hold under another set of assumptions, with the diameter bounds replaced by volume bounds, i.e. with classes $\mathcal{A}^{\mathcal{E}}_{m,n}(E,d)$ replaced by

$$\tilde{\mathcal{A}}^{\mathcal{E}}_{m,n}(E,H) := \{\Sigma \in \mathscr{C}^{0,1}_{m,n} :\ \mathcal{E}(\Sigma) \le E,\ \mathcal{H}^m(\Sigma) \le H\}.$$

(For Σ's with $\mathcal{E}(\Sigma) \le E$, the diameter bounds are equivalent to volume bounds).

There are numerous papers in the literature dealing with compactness and finiteness results for *immersed* manifolds, starting with J. Langer [26] who considers immersed surfaces in $\mathbb{R}^3$ of class $W^{2,p}$ for $p > 2$, with L^p bounds on the second fundamental form; for a generalization to immersed hypersurfaces in $\mathbb{R}^n$ see [13]. G. Smith [34, 35] considers compactness of immersed complete submanifolds of class $C^{k,\alpha}$ with $k \ge 2$, $\alpha \in (0,1)$, assuming uniform bounds on the second fundamental forms (and their derivatives). Recently, P. Breuning [8] has studied compactness for a wide class of (r,λ)–immersions, i.e., C^1 immersions that can be represented as λ-Lipschitz graphs at a uniform length scale r.

Our work differs from all these papers in that we deal only with *embedded* objects. The upper bounds on any of the geometric curvature energies we consider do guarantee that the limit of a convergent sequence of submanifolds — even if the convergence, a priori, takes place only in the Hausdorff distance — is again an embedded submanifold; this is due to the penalisation effects mentioned before. It is easy to see that under the assumptions of [26], [34] or [8] a sequence of embedded submanifolds might converge to a limit which is only immersed, not embedded. In our case, one might later try and use the additional information on the structure of converging (finite energy) submanifolds to prove that the limit is not just immersed but also embedded. However, we choose a different and clearer way: first, we use Blaschke's

selection theorem to obtain a candidate compact set for the limit, and then, applying a graph tilting lemma and the Arzelà—Ascoli compactness theorem, we (a) prove that this limit is in fact a $C^{1,\alpha}$ embedded submanifold and (b) obtain the desired type of convergence.

The only comparable result that we are aware of is due to O. Durumeric: [14, Thm. 2] ascertains that there are only finitely many diffeomorphism and isotopy classes of connected $C^{1,1}$ manifolds with lower bounds on a very specific functional, namely on the normal injectivity radius r_{ni}, combined roughly speaking with bounds on volume and diameter. Our estimate (9) on the number of isotopy types is similar in spirit to Durumeric's [14, Sec. 5] where $E^{1/p}$ (which controls the *bending* of Σ in a single coordinate chart, cf. Section 1.2 below) is replaced by the inverse of the injectivity radius. (In [28], A. Nabutovsky studies the intriguing 'energy landscape' of $\mathcal{E}(\Sigma) = \operatorname{vol}^{1/n}(\Sigma)/r_{ni}(\Sigma)$ on the set of $C^{1,1}$ topological hyperspheres $\Sigma = \Sigma^n \subset \mathbb{R}^{n+1}$; in particular his energy $\mathcal{E}$ has infinitely many distinct local minima.)

Let us note that for *curves* in $\mathbb{R}^3$, J. O'Hara, see [1, pp. 1340–1343], mentions a few results that use the same analytic mechanism of proof that we deal with: sequences of knots $f\colon \mathbb{S}^1 \to \mathbb{R}^3$ that are uniformly bilipschitz and remain bounded in a fixed $C^{1,\alpha}$ space, are precompact in C^1; moreover, the knot class has to be preserved in the limit. A bound for the bilipschitz constant of f and for its $C^{1,\alpha}$-norm translates into a bound on the number of possible knot types parameterised by f. For $m = 1$, all our energies are valid knot energies of that type: upon fixing the length, upper bounds on the energy yield bilipschitz and $C^{1,\alpha}$ bounds on the arc-length parameterisations of knots, resulting in bounds on the number of knot classes, on the average crossing number, stick number etc.; see [38] for more details. The results of the present paper open several questions in higher dimensional geometric knot theory, concerning, e.g., the possible relations between bounds on the energies of 'knot conformations' and bounds on higher dimensional knot invariants.

1.2. The strategy of proofs and more general results for $C^{1,\alpha}$-submanifolds

Let us start by explaining why we assume integrability above scale-invariance for each of the geometric curvature energies. A simple scaling argument shows that if $\Sigma \subset \mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R}$ is a cone over an $(m-1)$-dimensional smooth manifold $\Sigma_0 \subset \mathbb{R}^{n-1}$, with vertex $v = (0, \ldots, 0, 1)$ and $\mathcal{E} \in \{\mathcal{E}_p^l, \mathcal{T}_p, \mathcal{T}_p^G\}$,

then $\mathcal{E}(\Sigma) = \infty$ whenever $p \geq p_0(\mathcal{E})$. In fact, for such p the sequence of energies of disjoint pieces of Σ,

$$\epsilon_j := \mathcal{E}\big(\Sigma \cap \{x \in \mathbb{R}^n : \quad 2^{-j-1} < |x - v| \leq 2^{-j}\}\big), \qquad j = 0, 1, 2, \ldots,$$

is nondecreasing (by scaling!), and we have $\mathcal{E}(\Sigma) \geq \epsilon_0 + \epsilon_1 + \cdots$.

Intuitively, each fold or cusp of Σ should introduce even 'more' small simplices (or small tangent spheres that contain another point of Σ), and thereby should lead to an increase of energy. This strongly suggests that for any $p > p_0(\mathcal{E})$ the functional $\mathcal{E}$ should have nice smoothing properties. This is indeed true; if $\mathcal{E}(\Sigma)$ is finite for some p above the critical exponent $p_0(\mathcal{E})$, then Σ is an embedded manifold of class $C^{1,\alpha}$. Moreover, Σ can be assembled from a finite number of *standard graph patches* — corresponding to the *standard pieces* in Cheeger's words quoted in Section 1.1 above — with the size and the graph norms (controlling how much Σ can bend at length scales determined by the energy) explicitly controlled in terms of $\mathcal{E}(\Sigma)$. This is the reason why we can obtain compactness and finiteness results, along with semicontinuity of all these energies.

Here is a precise description of what we mean by *standard graph patches*.

For $\alpha \in (0, 1]$, let $\mathscr{C}^{1,\alpha}_{m,n}$ denote the set of all compact, $C^{1,\alpha}$-smooth, m-dimensional embedded submanifolds of $\mathbb{R}^n$.

Definition 1 ($C^{1,\alpha}$-graph patches). For $R > 0, L > 0$, $d > 0$, and $\alpha \in (0, 1]$ we define $\mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$ to be the class of those submanifolds $\Sigma \in \mathscr{C}^{1,\alpha}_{m,n}$ that satisfy the following three conditions:

(i) *diameter bounds:* $\Sigma \subset \overline{\mathbb{B}}^n(0, d)$;

(ii) *size of graph patches:* for each point $x \in \Sigma$ there exists a function $f_x : T_x\Sigma \to T_x\Sigma^\perp$ of class $C^{1,\alpha}$ such that $\Sigma \cap \mathbb{B}(x, R) = (x + \mathrm{graph}(f_x)) \cap \mathbb{B}(x, R)$, $f_x(0) = 0$, and $Df_x(0) = 0$;

(iii) *controlled bending:* for each $x \in \Sigma$, we have $\|Df_x(\xi) - Df_x(\eta)\| \leq L|\xi - \eta|^\alpha$ for all $\xi, \eta \in T_x\Sigma$, and $\mathrm{Lip}(f_x) \leq 1$.

We state below — in a version that is adapted for our needs in this paper — a general regularity result which has been proved in our earlier works, see [41] for the case of $\mathcal{T}_p$, [25] for $\mathcal{T}^G_p$, and [24] for all the $\mathcal{E}^l_p$-energies, $l = 1, \ldots, m + 2$. (The case of $\mathcal{E}^{m+2}_p$ for $m = 2$, $n = 3$ dates back to [40]; for curves in $\mathbb{R}^n$, see also [37], [36] and [39]).

Regularity Theorem. *Fix $\mathcal{E} \in \{\mathcal{E}^l_p, \mathcal{T}_p, \mathcal{T}^G_p\}$ and $p > p_0(\mathcal{E})$. Assume that a Lipschitz manifold $\Sigma \in \mathscr{C}^{0,1}_{m,n}$ satisfies $\mathcal{E}(\Sigma) \leq E < \infty$. If $\Sigma \subset \mathbb{B}^n(0,d)$, then $\Sigma \in \mathscr{C}^{1,\alpha}_{m,n}(R,L,d)$ for the exponent $\alpha = 1 - p_0(\mathcal{E})/p \in (0,1)$, with R and L depending only on m, n, l, p, p_0 and E. In fact, one can take*

$$R = c_1(m,n,l,p)E^{-1/(p-p_0(\mathcal{E}))}\,, \qquad L = c_2(m,n,l,p)E^{1/p} \tag{10}$$

for some constants c_1 and c_2 depending only on m, n, l, and p.

Let us note here that according to Blatt and Kolasiński [6], for an m-dimensional embedded C^1-submanifold $\Sigma \subset \mathbb{R}^n$ and $p_1 = m(l-1) < p$ (note that $p_0(\mathcal{E}^l_p) = ml > p_1$) the condition $\mathcal{E}^l_p(\Sigma) < \infty$ is *equivalent* to Σ being locally a graph of a function in the fractional Sobolev space $W^{1+s,p}$ with $s = 1 - m(l-1)/p \in (0,1)$. In combination with the Sobolev imbedding, this implies that the exponent α in the Regularity Theorem (which is one of the key technical tools for the present paper) is *best possible*. Moreover, there are m-dimensional graphs in $\mathbb{R}^n$ with finite curvature energy $\mathcal{E}^l_p$ for which the graph function is *nowhere* twice differentiable, so we definitely cannot work with classic curvatures in our setting!

Nevertheless, the Regularity Theorem paves the way to our results on compactness, finiteness and semicontinuity of geometric curvature energies with respect to sequences $\Sigma_j \subset \mathcal{A}^{\mathcal{E}}_{m,n}(E,d)$, but also in the more general subclass of $\mathscr{C}^{1,\alpha}_{m,n}$ introduced in Definition 1. The key idea is the following: due to the regularity estimates, energy and diameter bounds present in the definition of $\mathcal{A}^{\mathcal{E}}_{m,n}(E,d)$ force all the Σ_j to be in the same, *fixed*, class $\mathscr{C}^{1,\alpha}_{m,n}(R,L,d)$ up to translations. Then, the *controlled bending* condition satisfied by all the Σ_j enters the scene: after a technical preparation involving some *graph tilting*, it enables applications of the Arzelà–Ascoli compactness theorem in all graph patches. Thus, in fact much more can be said about the convergence of Σ_j, at least along a subsequence. Let us make this more precise.

Definition 2 ($C^{1,\alpha}$-convergence of graph patches). A sequence $(\Sigma_j)_{j\in\mathbb{N}} \subset \mathscr{C}^{1,\alpha}_{m,n}$ is said to *converge in* $\mathscr{C}^{1,\alpha}_{m,n}$ to the set $\Sigma_0 \subset \mathbb{R}^n$ if

(i) $d_{\mathcal{H}}(\Sigma_j, \Sigma_0) \xrightarrow{j\to\infty} 0$;

(ii) Σ_0 is a $C^{1,\alpha}$-smooth embedded submanifold of $\mathbb{R}^n$;

(iii) there is an index $j_0 \in \mathbb{N}$ and a radius $\rho > 0$ such that for each $x \in \Sigma_0$ and for each $j \in \mathbb{N}$ with $j \geq j_0$ or $j = 0$ there exists a function $f_{x,j} \in$

$C^{1,\alpha}(T_x\Sigma_0, T_x\Sigma_0^\perp)$ such that

$$\Sigma_j \cap \mathbb{B}(x,\rho) = (x + \operatorname{graph}(f_{x,j})) \cap \mathbb{B}(x,\rho)$$

and

$$\|f_{x,j} - f_{x,0}\|_{C^{1,\alpha'}(T_x\Sigma_0, T_x\Sigma_0^\perp)} \xrightarrow{j\to\infty} 0 \qquad \text{for each } \alpha' \in (0,\alpha).$$

As suggested above and mentioned in Section 1.1, the following stronger compactness result on $\mathscr{C}^{1,\alpha}_{m,n}(R,L,d)$ holds, from which part (ii) of Theorem 2 follows directly, but which is also essential to prove the other results stated in Section 1.1.

Theorem 3 (compactness). *Let $R, L, d \in (0,\infty)$ and $\alpha \in (0,1]$. Any sequence of submanifolds $(\Sigma_j)_{j\in\mathbb{N}} \subset \mathscr{C}^{1,\alpha}_{m,n}(R,L,d)$ contains a subsequence which converges in $\mathscr{C}^{1,\alpha}_{m,n}$ to some submanifold $\Sigma_0 \in \mathscr{C}^{1,\alpha}_{m,n}(R,L,d)$.*

The convergence in $\mathscr{C}^{1,\alpha}_{m,n}$ is strong enough to make all the Σ_j with j large enough ambient isotopic to the limiting manifold Σ_0. It also turns out that there is a diffeomorphism of the ambient space $J_j : \mathbb{R}^n \to \mathbb{R}^n$ which is close to the identity in the bilipschitz sense and maps Σ_j to Σ_0.

Theorem 4 (isotopy and diffeomorphism of ambient space). *Let $R, L, d \in (0,\infty)$, $\alpha \in (0,1]$ and let $(\Sigma_j)_{j\in\mathbb{N}} \subset \mathscr{C}^{1,\alpha}_{m,n}(R,L,d)$ be a sequence of submanifolds which converges in $\mathscr{C}^{1,\alpha}_{m,n}$ to $\Sigma_0 \in \mathscr{C}^{1,\alpha}_{m,n}(R,L,d)$. Then there exists $j_0 \in \mathbb{N}$ such that for each $j \geq j_0$ the manifolds Σ_j and Σ_0 are ambient isotopic. Moreover, for each $j \geq j_0$ there exists a diffeomorphism of the ambient space $J_j : \mathbb{R}^n \to \mathbb{R}^n$ such that*

$$J_j(\Sigma_j) = \Sigma_0 \quad \textit{and} \quad \operatorname{biLip}(J_j) \leq 1 + C_J d_{\mathcal{H}}(\Sigma_0, \Sigma_j)^{\alpha/2},$$

where $C_J = C_J(R, L, \alpha, m, n)$.

Here, $\operatorname{biLip}(f)$ denotes the bilipschitz constant of an injective map $f : X \to f(X) \subset Y$ between two metric spaces (X, d_X) and (Y, d_Y), i.e. whenever $A \subset X$,

$$\operatorname{biLip}(f, A) := \max\{\operatorname{Lip}(f, A), \operatorname{Lip}(f^{-1}, f(A))\}, \quad \operatorname{biLip}(f) := \operatorname{biLip}(f, X).$$

As usual,

$$\operatorname{Lip}(f, A) = \sup_{x,y\in A,\ x\neq y} \frac{d_Y(f(x), f(y))}{d_X(x,y)} \quad \text{and} \quad \operatorname{Lip}(f) = \operatorname{Lip}(f, X)$$

denotes the Lipschitz constant of $f : X \to Y$.

We actually establish an upper bound ρ on the Hausdorff-distance $d_{\mathcal{H}}(\Sigma_1, \Sigma_2)$ depending only on the parameters R, L, α, m, and n, such that if $d_{\mathcal{H}}(\Sigma_1, \Sigma_2) \leq \rho$, then Σ_1 and Σ_2 are ambiently isotopic, and such that a global bilipschitz diffeomorphism J on $\mathbb{R}^n$ with $J(\Sigma_2) = \Sigma_1$ exists (see Corollary 4.9 and Lemma 4.10). This uniform bound leads not only to the proof of Theorem 4 but allows us also in the end to give a quantitative estimate on the number of isotopy types in Theorem 1.

The rest of the paper is organised as follows. In Section 2, we gather simple preliminary material. In Section 3, after a technical preparation devoted to graph tilting for functions of class $C^{1,\alpha}$, we prove Theorem 3. In Section 4, we construct the isotopies between the submanifolds Σ_j which converge in $\mathscr{C}^{1,\alpha}_{m,n}$, employing a C^1-version of the tubular neighbourhood theorem; parts of this material seem to be 'folklore' but we give the details for the sake of completeness. This leads to the proof of Theorem 4. Section 5 contains the proof of semicontinuity and compactness, Theorem 2, and of Corollary 1, and the final Section 6 lays out the explicit estimate for the number of isotopy types (stated in Theorem 1). The whole exposition is more or less self-contained.

Remark. The letter C will denote a constant whose value may change even in a single string of estimates. Subscripted constants (e.g. C_l, C_{ang} etc.) have global meaning and their value is fixed. We write $C = C(\alpha, \beta, \gamma)$ when C depends *only* on α, β and γ.

2. Preliminaries

Most of the notation in the paper is standard. In particular, we use the usual $\|\cdot\|_{C^{1,\alpha}}$ norms, and

$$d_{\mathcal{H}}(E, F) := \sup\{\operatorname{dist}(y, F) : y \in E\} + \sup\{\operatorname{dist}(z, E) : z \in F\}$$

denotes the Hausdorff distance of sets in $\mathbb{R}^n$.

For a measure μ, we write $f_*\mu$ to denote its push-forward, and $\operatorname{spt}(\mu)$ to denote its support, cf. Federer [17, Chapter 2] or Matilla [27, Chapter 1] for definitions.

2.1. The Grassmannian.

Throughout the paper, $G(n,m)$ stands for the Grassmannian of all m-dimensional linear subspaces[2] of $\mathbb{R}^n$.

For an m-plane $U \in G(n,m)$ we let $U^\perp$ be the orthogonal complement of U. The symbol $U_\natural$ denotes the orthogonal projection of $\mathbb{R}^n$ onto U. For $U, V \in G(n,m)$ we set

$$\measuredangle(U,V) := \|U_\natural - V_\natural\|. \tag{11}$$

This is a metric, and $G(n,m)$ endowed with this metric is compact.

Remark 2.1. Using [2, 8.9(3)] we have for $U, V \in G(n,m)$

$$\begin{aligned}\measuredangle(U,V) &= \|U_\natural - V_\natural\| = \|U_\natural^\perp - V_\natural^\perp\| = \|U_\natural^\perp \circ V_\natural\| \\ &= \|U_\natural \circ V_\natural^\perp\| = \|V_\natural^\perp \circ U_\natural\| = \|V_\natural \circ U_\natural^\perp\|\,.\end{aligned}$$

In particular

$$\measuredangle(U,V) = \sup_{e \in U \cap \mathbb{S}} |V_\natural^\perp e| = \sup_{e \in V \cap \mathbb{S}} |U_\natural^\perp e|\,.$$

Here and later $\mathbb{S} = \{x \in \mathbb{R}^n : |x| = 1\}$ denotes the unit sphere in $\mathbb{R}^n$.

Lemma 2.2. *Assume $U, V \in G(n,m)$. If $\measuredangle(U,V) < 1$, then*

- *$U_\natural|_V : V \to U$ is a linear isomorphism,*
- *$U^\perp \cap V = \{0\}$,*
- *setting $L = (U_\natural|_V)^{-1} : U \to V$ we have*

$$\|L\| = (1 - \measuredangle(U,V)^2)^{-1/2}\,.$$

---

[2]Formally, $G(n,m)$ is defined as the homogeneous space

$$G(n,m) := O(n)/(O(m) \times O(n-m))\,,$$

where $O(n)$ is the orthogonal group; see e.g. A. Hatcher's book [21, Section 4.2, Examples 4.53–4.55]. Thus $G(n,m)$ becomes a topological space with the quotient topology. We work with the angular metric, cf. (11).

Proof. If $\measuredangle(U,V) < 1$, then, by Remark 2.1, for each $v \in V$, $v \neq 0$

$$|U_\natural v|^2 = |v|^2(1 - |U_\natural^\perp(v/|v|)|^2 > 0\,.$$

Hence, $\ker U_\natural|_V = \{0\}$ and, since $\dim U = \dim V$, $U_\natural|_V$ is a linear isomorphism. In particular,

$$\ker U_\natural \cap V = U^\perp \cap V = \{0\}.$$

Observe that, by Remark 2.1,

$$\inf_{e\in V\cap\mathbb{S}} |U_\natural e|^2 = 1 - \sup_{e\in V\cap\mathbb{S}} |U_\natural^\perp e|^2 = 1 - \measuredangle(U,V)^2\,.$$

Set $L = (U_\natural|_V)^{-1}$ and compute

$$\begin{aligned}\|L\| &= \sup_{u\in U,u\neq 0} |Lu||u|^{-1} = \sup_{v\in V,v\neq 0} |LU_\natural v||U_\natural v|^{-1}\\ &= \sup_{v\in V,v\neq 0} |v||U_\natural v|^{-1} = \left(\inf_{v\in V,v\neq 0} |U_\natural(v/|v|)|\right)^{-1}\\ &= (1-\measuredangle(U,V)^2)^{-1/2}\,.\end{aligned}$$

□

Remark 2.3. Let $X \in G(n,m)$ and $Y \in G(n,n-m)$ be such that

$$\measuredangle(X^\perp, Y) < 1.$$

Then, by Lemma 2.2, $X \cap Y = \{0\}$ and we can define the *oblique projection* $P : \mathbb{R}^n \to X$ along Y, i.e., a linear map such that

$$P \circ P = P\,, \quad \ker P = Y \quad \text{and} \quad \operatorname{im} P = X\,. \tag{12}$$

Note that P can also be characterised by the requirement

$$\{Pv\} = (v+Y)\cap X\,. \tag{13}$$

Proposition 2.4. *Let $\theta \in [0,1]$, $\lambda \in [0,1)$ and $k \in \{1,\dots,n-1\}$. Let X, $Y \in G(n,k)$ and $Z \in G(n,n-k)$ be such that $\measuredangle(X,Y) \le \theta$ and $\measuredangle(Y,Z^\perp) \le \lambda$. For any $\mathbf{x} \in X$ and $\mathbf{z} \in Z$ with $\mathbf{z}-\mathbf{x} \in Y$ one has the estimate*

$$|\mathbf{z}| \le \frac{\theta}{1-\lambda}|\mathbf{x}|\,.$$

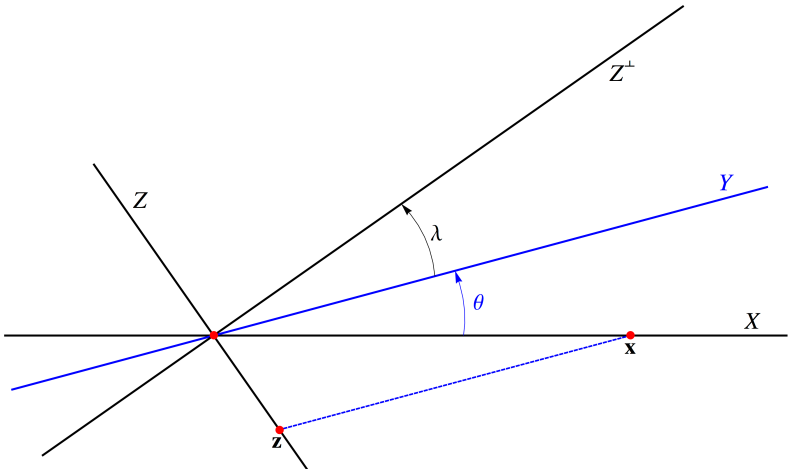

Figure 1: The situation in Proposition 2.4: if $\mathbf{x}$, X, Z are fixed, $\theta \to 0$, and $\mathbf{z} - \mathbf{x} \in Y$, then $\mathbf{z} \to 0$ and $|\mathbf{z}| \lesssim \theta$.

Proof. Since $Z_\natural^\perp \mathbf{u} = X_\natural^\perp \mathbf{x} = 0$, two applications of the triangle inequality lead to

$$\begin{aligned} |\mathbf{z}| \le |Y_\natural \mathbf{z}| + |Y_\natural^\perp \mathbf{z}| &\le |(Y_\natural - U_\natural^\perp)\mathbf{z}| + |Y_\natural^\perp(\mathbf{z} - \mathbf{x})| + |(Y_\natural^\perp - X_\natural^\perp)\mathbf{x}| \\ &\le \lambda |\mathbf{z}| + \theta |\mathbf{x}| \,. \end{aligned}$$

□

2.2. An elementary topological result

In a few proofs, we need to rely on the following standard topological result. For the sake of completeness, we present a proof using degree mod 2. (There are of course other proofs, relying on the non-existence of the retraction of a ball onto its boundary or, equivalently, on Brouwer's fixed point theorem.)

Proposition 2.5. *Let $\rho > 0$, $\sigma \in (0,1)$ and $F \in C^0(\overline{\mathbb{B}}^m(0,\rho), \mathbb{R}^m)$ be such that*

$$|F(x) - x| \le \sigma\rho \qquad \text{for all } x \in \overline{\mathbb{B}}^m(0,\rho)\,.$$

Then for each $y \in \mathbb{B}^m(0,(1-\sigma)\rho)$ there exists $x \in \overline{\mathbb{B}}^m(0,\rho)$ such that $F(x) = y$.

Proof. Fix $y \in \mathbb{B}^m(0,(1-\sigma)\rho)$. Assume that $y \notin F(\overline{\mathbb{B}}^m(0,\rho))$. Then,

$$G : \partial\mathbb{B}^m(0,\rho) \to \partial\overline{\mathbb{B}}^m(0,\rho)\,, \quad G(z) = \frac{F(z)-y}{|F(z)-y|}\rho \quad \text{for } z \in \partial\overline{\mathbb{B}}^m(0,\rho),$$

is well defined and continuous. Since $|F(x)-x| \leq \sigma\rho$ for all $x \in \overline{\mathbb{B}}^m(0,\rho)$,

$$|-tz+t(F(z)-y)| \leq t|F(z)-z|+t|y| < t\sigma\rho+t(1-\sigma)\rho < \rho = |z|$$

for all $z \in \partial\overline{\mathbb{B}}^m(0,\rho)$ and $t \in [0,1)$; hence $(1-t)z+t(F(z)-y) \neq 0$ for all $z \in \partial\mathbb{B}^m(0,\rho)$ and for all $t \in [0,1]$. Thus, the map

$$H : [0,1] \times \partial\overline{\mathbb{B}}^m(0,\rho) \to \partial\overline{\mathbb{B}}^m(0,\rho)\,, \quad H(t,z) = \frac{(1-t)z+t(F(z)-y)}{|(1-t)z+t(F(z)-y)|}\rho$$

yields a well defined homotopy of G and the identity map on the sphere $\partial\overline{\mathbb{B}}^m(0,\rho)$. Hence G has mod 2 degree 1. On the other hand, one can extend G to the continuous mapping

$$\tilde{G} : \overline{\mathbb{B}}^m(0,\rho) \to \partial\overline{\mathbb{B}}^m(0,\rho)\,, \qquad \tilde{G}(z) = \frac{F(z)-y}{|F(z)-y|}\rho\,.$$

Thus G has mod 2 degree 0, a contradiction. For the relevant results on the mod 2 degree one may, e.g., consult [22, pp. 124,125]. □

3. Compactness

Each manifold $\Sigma \in \mathscr{C}^{1,\alpha}_{m,n}(R,L,d)$ is assembled from standard graph patches that have controlled bending at length scales $\lesssim R$. Thus, intuitively, if two such manifolds $\Sigma_1, \Sigma_2 \in \mathscr{C}^{1,\alpha}_{m,n}(R,L,d)$ are sufficiently close in Hausdorff distance, their tangent planes at points $x \in \Sigma_1$, $y \in \Sigma_2$ with $|x-y| \lesssim d_{\mathcal{H}}(\Sigma_1,\Sigma_2)$ must be close, too, for otherwise the Hausdorff distance of the manifolds would be too large. Before giving the precise quantitative statement, let us mention two simple consequences of Definition 1 valid for each $\Sigma \in \mathscr{C}^{1,\alpha}_{m,n}(R,L,d)$: For any $r \in (0,R]$ one finds $x+v+f_x(v) \in \mathbb{B}^n(x,\sqrt{2}r)$ for all $x \in \Sigma$, $v \in T_x\Sigma \cap \mathbb{B}^n(0,r)$, since $|f_x(v)| = |f_x(v)-f_x(0)| \leq |v| < r$ so that

$$|x-(x+v+f_x(v))|^2 = |v|^2+|f_x(v)|^2 < 2r^2.$$

Secondly, one can improve the estimate for $|f_x(v)|$ for such x, v, and $r \in (0, R]$ as follows.

$$|f_x(v)| = \left| \int_0^1 \frac{d}{dt} f_x(tv)\, dt \right| = \left| \int_0^1 (Df_x(tv) - Df_x(0))v\, dt \right| \leq L|v|^{1+\alpha} < Lr^{1+\alpha}. \tag{14}$$

Lemma 3.1 (proximity of tangent planes). *Let $R, L, d > 0$, $\alpha \in (0, 1]$, $A \geq 1$, and $\Sigma_1, \Sigma_2 \in \mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$ such that*

$$d_{\mathcal{H}}(\Sigma_1, \Sigma_2) < \min\left\{2^{-6}A^{-2}R^2,\, L^{-2/\alpha},\, 1\right\},$$

and let $x \in \Sigma_1$ and $y \in \Sigma_2$ be such that $|x - y| \leq A d_{\mathcal{H}}(\Sigma_1, \Sigma_2)$. Then there exists a constant $C_{\mathrm{ang}} = C_{\mathrm{ang}}(L, A)$ such that

$$\sphericalangle(T_x\Sigma_1, T_y\Sigma_2) \leq C_{\mathrm{ang}} d_{\mathcal{H}}(\Sigma_1, \Sigma_2)^{\alpha/2}\,. \tag{15}$$

In fact, one can take $C_{\mathrm{ang}}(L, A) = L\big(1 + (4A)^2\big) + 2A$.

Proof. For $d_{\mathcal{H}}(\Sigma_1, \Sigma_2) = 0$ we have $x = y$ and $\Sigma_1 = \Sigma_2$ as C^1-manifolds, so that $T_x\Sigma_1 = T_y\Sigma_2$; hence both sides of (15) are zero. So, let us assume that $d_{\mathcal{H}}(\Sigma_1, \Sigma_2) > 0$. The following arguments hold for all $u \in T_x\Sigma_1$ with

$$0 < |u| = d_{\mathcal{H}}(\Sigma_1, \Sigma_2)^{1/2} < \min\left\{2^{-3}A^{-1}R,\, L^{-1/\alpha},\, 1\right\}. \tag{16}$$

Since $\Sigma_1 \in \mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$ we find

$$p := x + u + f_x(u) \in \Sigma_1 \cap \mathbb{B}(x, R)$$

by our remark preceding this lemma, since $|u| < R/(2^3 A) < R/\sqrt{2}$. We thus infer

$$(T_x\Sigma_1)_\natural(p - x) = (T_x\Sigma_1)_\natural(u + f_x(u)) = u,$$

and

$$\begin{aligned} \operatorname{dist}(p - x, T_x\Sigma_1) &= \operatorname{dist}(u + f_x(u), T_x\Sigma_1) \\ &= |f_x(u)| \overset{(14)}{\leq} L|u|^{1+\alpha} \overset{(16)}{<} L|u|\frac{1}{L} = |u|. \end{aligned} \tag{17}$$

In particular,

$$|p - x|^2 = |u|^2 + |f_x(u)|^2 = |u|^2 + \operatorname{dist}^2(p - x, T_x\Sigma_1) < 2|u|^2.$$

Next choose a point $q \in \Sigma_2$ such that $|p-q| \leq d_{\mathcal{H}}(\Sigma_1, \Sigma_2)$ and set $v := (T_y\Sigma_2)_\natural(q-y)$. Then one finds

$$\begin{aligned}|v| &\leq \|(T_y\Sigma_2)_\natural\||q-y| \ = |q-y| \\ &\leq |q-p| + |p-x| + |x-y| \leq d_{\mathcal{H}}(\Sigma_1,\Sigma_2) + \sqrt{2}|u| + A d_{\mathcal{H}}(\Sigma_1,\Sigma_2) \\ &\leq |u|\Big[(1+A)|u| + \sqrt{2}\Big] \leq A|u|\Big[2|u| + \sqrt{2}\Big] \overset{(16)}{<} 4A|u| \overset{(16)}{<} \frac{1}{\sqrt{2}}R,\end{aligned}$$

so that $y + v + f_y(v) \in \mathbb{B}^n(y,R)$ and $q - y = v + f_y(v) \in \operatorname{graph}(f_y)$, by virtue of our remark preceding this lemma. Employing the identities

$$\begin{aligned}&\operatorname{dist}(p-x, T_x\Sigma_1) = |f_x(u)| = |p-(x+u)| \quad \text{and} \\ &\operatorname{dist}(q-y, T_y\Sigma_2) = |f_y(v)| = |q-(y+v)|\end{aligned}$$

we can write

$$\begin{aligned}\operatorname{dist}(u, T_y\Sigma_2) \leq |u-v| &\leq |u+(x-p)| + |p-q| + |q-(y+v)| + |y-x| \\ &\leq |f_x(u)| + d_{\mathcal{H}}(\Sigma_1,\Sigma_2) + |f_y(v)| + A|u|^2 \\ &\overset{(14)}{\leq} L\big(|u|^{1+\alpha} + |v|^{1+\alpha}\big) + (1+A) d_{\mathcal{H}}(\Sigma_1,\Sigma_2) \\ &\overset{(16)}{<} \big[L\big(1+(4A)^2\big) + 2A\big]|u|^{1+\alpha} =: C_{\mathrm{ang}}(L,A)|u|^{1+\alpha}.\end{aligned}$$

Since $\operatorname{dist}(u, T_y\Sigma_2) = |(T_y\Sigma_2)^\perp_\natural(u)| = |u|\big|(T_y\Sigma_2)^\perp_\natural\big(\frac{u}{|u|}\big)\big| = |u| \operatorname{dist}\big(\frac{u}{|u|}, T_y\Sigma_2\big)$ we arrive at $\operatorname{dist}\big(\frac{u}{|u|}, T_y\Sigma_2\big) \leq C_{\mathrm{ang}}(L,A)|u|^\alpha = C_{\mathrm{ang}}(L,A) d_{\mathcal{H}}(\Sigma_1,\Sigma_2)^{\alpha/2}$. Since the requirement (16) on $u \in T_x\Sigma_1$ does not depend on the direction $e := u/|u| \in \mathbb{S}^{n-1} = \mathbb{S}$ we obtain by Remark 2.1

$$\begin{aligned}\measuredangle(T_x\Sigma_1, T_y\Sigma_2) &= \sup_{e \in T_x\Sigma_1 \cap \mathbb{S}} |(T_y\Sigma_2)^\perp_\natural e| \\ &= \sup_{e \in T_x\Sigma_1 \cap \mathbb{S}} \operatorname{dist}(e, T_y\Sigma_2) \leq C_{\mathrm{ang}} d_{\mathcal{H}}(\Sigma_1,\Sigma_2)^{\alpha/2}.\end{aligned}$$

□

To prove Theorem 3, one applies the Arzelà–Ascoli theorem to graph patches of the sequence Σ_j. To make this possible, it is necessary to tilt all the graphs (of a subsequence of the Σ_j's intersected with a fixed ball of radius $\approx R$) so that they are all defined over *the same* plane. Here is a technical lemma that we shall use.

Lemma 3.2 (graph tilting). *Let* $V \in G(n,m)$, $\alpha \in (0,1]$, $\vartheta \in (0, \frac{1}{100})$, $L > 0$, $v \in V$, $r \in (0,\infty]$ *and* $f \in C^{1,\alpha}(V, V^\perp)$ *is such that* $\mathrm{Lip}(f) \leq 1$ *and*

$$\|Df(x) - Df(y)\| \leq L|x-y|^\alpha \quad \text{for } x, y \in V \cap \mathbb{B}(v,r)\,.$$

Then the following holds. For each $U \in G(n,m)$ *with* $\measuredangle(U,V) \leq \vartheta$ *there exists a function* $g \in C^{1,\alpha}(U, U^\perp)$ *such that for* $\omega := U_\natural(v + f(v))$,

$$\begin{aligned} &\mathrm{graph}(f) = \mathrm{graph}(g) \quad \textit{and} \\ &\|Dg(\xi) - Dg(\eta)\| \leq L_g(L,\vartheta,\alpha)|\xi-\eta|^\alpha \quad \text{for } \xi, \eta \in U \cap \mathbb{B}(\omega, \tfrac{1}{1+3\vartheta} r)\,, \end{aligned}$$

where $L_g(L,\vartheta,\alpha) := L(1+12\vartheta)(1+3\vartheta)^\alpha \,/\, (1-4\vartheta)$. *Moreover,* $\mathrm{Lip}(g) \leq (1+2\vartheta)/(1-2\vartheta)$, *and* $g(0) = 0$ *if* $f(0) = 0$. *If* $f(0) = 0$ *and* $Df(0) = 0$ *then*

$$\|Dg(0)\|^2 \leq \frac{\vartheta^2}{1-\vartheta^2}. \tag{18}$$

Remark. By taking $r = \infty$ we mean $\mathbb{B}(v,r) = \mathbb{R}^n$.

Proof. If $U = V$ simply set $g := f$, and we are done. So assume $\measuredangle(U,V) \in (0,\vartheta]$ in the following.

Step 1: defining g. Set $\Sigma := \mathrm{graph}(f)$ and for $p_1, p_2 \in \Sigma$ define $x_1 := V_\natural(p_1)$, $x_2 := V_\natural(p_2)$, $z_1 := U_\natural(p_1)$, and $z_2 := U_\natural(p_2)$. Then $p_i = x_i + f(x_i)$ for $i = 1,2$, and since $\mathrm{Lip}(f) \leq 1$ and $\measuredangle(U,V) \leq \vartheta < \frac{1}{100}$ we have

$$\begin{aligned} |(x_2 - x_1) - (z_2 - z_1)| &= |(V_\natural - U_\natural)(p_2 - p_1)| \\ &\leq \|V_\natural - U_\natural\| |p_2 - p_1| \leq \vartheta |p_2 - p_1| \\ &= \vartheta \big|(x_2 - x_1) + (f(x_2) - f(x_1))\big| \\ &\leq 2\vartheta |x_2 - x_1| < \tfrac{1}{50}|x_2 - x_1|, \end{aligned} \tag{19}$$

where only the very last inequality is restricted to the case $x_1 \neq x_2$. If $z_1 = U_\natural(p_1) = U_\natural(p_2) = z_2$ then (19) implies $0 \leq (1-2\vartheta)|x_2 - x_1| \leq 0$, hence $x_1 = V_\natural(p_1) = V_\natural(p_2) = x_2$, so that

$$p_1 = x_1 + f(x_1) = x_2 + f(x_2) = p_2.$$

In other words, if $p_1 \neq p_2$ then $U_\natural(p_1) \neq U_\natural(p_2)$, or $U_\natural|_\Sigma : \Sigma \to U$ is injective.

Setting $q := U_\natural(0 + f(0)) \in U$,

$$\phi_1 : V \ni x \longmapsto x + f(x) \in \Sigma, \qquad \phi_2 : \Sigma \ni p \longmapsto U_\natural(p) - q,$$

we find that $\phi := \phi_2 \circ \phi_1 : V \to U$ is injective, since both ϕ_1 and ϕ_2 are injective. Moreover, ϕ is continuous, and $\phi(0) = \phi_2(\phi_1(0)) = \phi_2(0 + f(0)) = 0$. Letting $x_1 := 0$ in (19), and setting $x := x_2$, $z := z_2 = \phi(x_2) = \phi(x)$ we infer from (19)

$$(20)\quad |x - \phi(x)| \le 2\vartheta|x| \qquad \text{for all} \quad x = V_\natural(x + f(x)) \in V_\natural(\operatorname{graph}(f)) = V.$$

Notice that the restricted projection $V_\natural|_U : U \to V$ is bijective, since $\measuredangle(U, V) \le \vartheta < 1$. Indeed, since $\dim U = \dim V$ it suffices to check that $V_\natural|_U$ is injective. But $V_\natural(u_1) = V_\natural(u_2)$ for $u_1, u_2 \in U$ with $u_1 \ne u_2$ would imply that $0 \ne u_1 - u_2 \in U$ would be contained in the kernel of $V_\natural$, so that we would arrive at the contradictory inequality

$$\begin{aligned} \|V_\natural - U_\natural\| &\ge \left| V_\natural\left(\frac{u_1 - u_2}{|u_1 - u_2|}\right) - U_\natural\left(\frac{u_1 - u_2}{|u_1 - u_2|}\right)\right| \\ &= \left|\frac{u_1 - u_2}{|u_1 - u_2|}\right| = 1 > \|V_\natural - U_\natural\| > 0. \end{aligned}$$

To show that ϕ is also surjective (hence bijective) consider a linear isometry $I_V : V \to \mathbb{R}^m$, define $F \in C^0(\mathbb{R}^m, \mathbb{R}^m)$ to be $F := I_V \circ V_\natural|_U \circ \phi \circ I_V^{-1}$, and estimate for $\xi \in \mathbb{R}^m$

$$\begin{aligned} |\xi - F(\xi)| &= |I_V \circ V_\natural\big(I_V^{-1}(\xi)\big) - I_V \circ V_\natural \circ \phi\big(I_V^{-1}(\xi)\big)| \\ &= |V_\natural\big(I_V^{-1}(\xi)\big) - V_\natural \circ \phi\big(I_V^{-1}(\xi)\big)| \\ &\le |I_V^{-1}(\xi) - \phi\big(I_V^{-1}(\xi)\big)| \overset{(20)}{\le} 2\vartheta|I_V^{-1}(\xi)| = 2\vartheta|\xi|. \end{aligned}$$

Thus, F satisfies the assumptions of Proposition 2.5 for each $\rho > 0$ and $\sigma := 2\vartheta < 1$, which implies that $F : \mathbb{R}^m \to \mathbb{R}^m$ is surjective. Therefore $V_\natural|_U \circ \phi : V \to V$ is surjective, and finally also $\phi : V \to U$ is surjective, hence bijective.

We are now in the position to define $g := U_\natural^\perp \circ \phi_1 \circ \phi^{-1} \circ \tau_{-q} : U \to U^\perp$, where $\tau_q(x) = x + q$ for $x \in \mathbb{R}^n$ is the usual translation. Since ϕ_1 is of class $C^{1,\alpha}$ one finds $\phi \in C^{1,\alpha}$, and so $\phi^{-1} \in C^{1,\alpha}$, and hence $g \in C^{1,\alpha}(U, U^\perp)$; see, e.g., [7, Section 2.2] for a brief self-contained argument showing that the class $C^{1,\alpha}$ is closed under composition and inversion. Moreover, if $p \in \Sigma$, $V_\natural p = x$ and $U_\natural p = z$, then $z = \phi(x) + q$ and $x + f(x) = p = z + g(z)$; hence, $\operatorname{graph}(f) = \Sigma = \operatorname{graph}(g)$. In particular, if $f(0) = 0$ then $0 = 0 + f(0) = z +$

$g(z)$ for $z = U_\natural(0) = 0$; hence $g(0) = 0$. Since $x \mapsto x + f(x)$ parameterises $\Sigma = \mathrm{graph}(f)$, one has in the point $p = x + f(x) \in \Sigma$ the m-dimensional tangent plane $T_p\Sigma = (\mathrm{Id} + Df(x))(T_xV) = (\mathrm{Id} + Df(x))(V)$, and likewise, $T_p\Sigma = (\mathrm{Id} + Dg(z))(U)$ if $p = z + g(z)$. Thus, if $f(0) = 0$ and $Df(0) = 0$, then we find

$$\begin{aligned} &T_0\Sigma = (\mathrm{Id} + Df(0))(V) = V = (\mathrm{Id} + Dg(0))(U) \quad \text{and} \\ &\sphericalangle(U, (\mathrm{Id} + Dg(0)(U)) = \sphericalangle(U, V) \le \vartheta, \end{aligned}$$

so that we can apply [2, 8.9(5)] with $S \equiv S_2 := U$ and $S_1 := (\mathrm{Id} + Dg(0))(U)$, $\eta_2 := 0$, $\eta_1 := Dg(0)$, to obtain (18).

Step 2: Lipschitz continuity of g and oscillation of Dg. By definition, $z + g(z) = x + f(x)$ for $z = \phi(x) + q \in U$, $x \in V$, so that by (19) and by the assumption $\mathrm{Lip}(f) \le 1$,

$$\begin{aligned} |g(z_2) - g(z_1)| &= |(z_2 + g(z_2) - (z_1 + g(z_1)) - (z_2 - z_1)| \\ &= |(x_2 + f(x_2) - (x_1 + f(x_1)) - (\phi(x_2) - \phi(x_1)| \\ &\le |f(x_2) - f(x_1)| + |(x_2 - \phi(x_2)) - (x_1 - \phi(x_1))| \\ &\le |x_2 - x_1| + |U_\natural(f(x_1) - f(x_2)) + U_\natural^\perp(x_2 - x_1)| \\ &\overset{\text{Rem. 2.1}}{\le} |x_2 - x_1| + \sphericalangle(U, V)(|f(x_1) - f(x_2)| + |x_2 - x_1|) \\ &\overset{(19)}{\le} \frac{1 + 2\vartheta}{1 - 2\vartheta} |z_2 - z_1| \,. \end{aligned}$$

With $T_p\Sigma = (\mathrm{Id} + Df(x))(V)$ for $p := x + f(x) \in \Sigma$ we obtain for any $v \in V$, $v \ne 0$,

$$\left| V_\natural^\perp \left(\frac{v + Df(x)v}{|v + Df(x)v|} \right) \right|^2 = \frac{|Df(x)v|^2}{|v|^2 + |Df(x)v|^2} \le \frac{\|Df\|_\infty^2 |v|^2}{|v|^2 + \|Df\|_\infty^2 |v|^2} \le \frac{1}{2},$$

since $\|Df\|_\infty = \mathrm{Lip}(f) \le 1$, and by the fact that for $c > 0$ the function $\xi \mapsto \xi/(c + \xi)$ is non-decreasing on $[0, \infty)$. Thus, according to Remark 2.1,

$$\sphericalangle(T_p\Sigma^\perp, V^\perp) = \sphericalangle(T_p\Sigma, V) \le \frac{1}{\sqrt{2}} < 1, \tag{21}$$

which implies

$$\sphericalangle(T_p\Sigma^\perp, U^\perp) = \sphericalangle(T_p\Sigma, U) \le \sphericalangle(T_p\Sigma, V) + \sphericalangle(V, U) \le \frac{1}{\sqrt{2}} + \vartheta < 1. \tag{22}$$

Consequently, the oblique projections $F_p : \mathbb{R}^n \to X := T_p\Sigma$ along $V^\perp$ with $\ker F_p = V^\perp$, and $G_p : \mathbb{R}^n \to X$ along $U^\perp$ with $\ker G_p = U^\perp$ are well-defined, and satisfy $\|F_p\| = \max_{e\in\mathbb{S}} |F_p(e)| \leq \sqrt{2} < 2$, and $\|G_p\| < 2$, which can be seen as follows. Assume without loss of generality $\|F_p\| > 0$. Since

$$\begin{aligned} |F_p(e)|^2 &= |F_p(V_\natural(e)) + F_p(V_\natural^\perp(e))|^2 \\ &= |F_p(V_\natural(e))|^2 < \|F_p\|^2\Big(|V_\natural(e)|^2 + |V_\natural^\perp(e)|^2\Big) = \|F_p\|, \end{aligned} \tag{23}$$

if $V_\natural^\perp(e) \neq 0$. But $\mathbb{S} = \mathbb{S}^{n-1}$ is compact, so that there exists $e^* \in \mathbb{S}$ (not necessarily unique) with $\|F_p\|^2 = |F_p(e^*)|^2$, which then necessarily means that $V_\natural^\perp(e^*) = 0$, i.e., $e^* \in \mathbb{S} \cap V$. For any such e^* we can write

$$|F_p(e^*)|^2 = |e^* + F_p(e^*) - e^*|^2 \overset{(13)}{=} 1 + |F_p(e^*) - e^*|^2 = 1 + |V_\natural^\perp(F_p(e^*))|^2,$$

since $F_p(e^*) - e^* \in V^\perp$ and $e^* \in V$; see (13). Now, with

$$|V_\natural^\perp(F_p(e^*))| \leq \measuredangle(V, X)|F_p(e^*)| \overset{(21)}{\leq} \frac{1}{\sqrt{2}}|F_p(e^*)|$$

one finds $|F_p(e^*)|^2 \leq 1 + \frac{1}{2}|F_p(e^*)|^2$, which immediately gives

$$\|F_p\| = |F_p(e^*)| \leq \sqrt{2}.$$

A similar argument for G_p using (22) instead of (21) leads to $\|G_p\|^2 \leq 1 + ((1/\sqrt{2}) + \vartheta)^2\|G_p\|^2$, and hence

$$\|G_p\| \leq \frac{1}{\sqrt{1 - ((1/\sqrt{2}) + \vartheta)^2}} < \frac{1}{\sqrt{1 - ((1/\sqrt{2}) + (1/100))^2}} < 2.$$

For $z_1, z_2 \in U$ and $p_i := z_i + g(z_i)$, $i = 1, 2$, let $x_1, x_2 \in V$ be those unique points with $p_i = x_i + f(x_i)$ for $i = 1, 2$. With $T_{p_i}\Sigma = (\mathrm{Id} + Df(x_i))(V) = (\mathrm{Id} + Dg(z_i))(U)$, and $Df(x_i)(V) \subset V^\perp$, $Dg(z_i)(U) \subset U^\perp$ for $i = 1, 2$, one obtains for $v \in V$ and $u \in U$

$$v + Df(x_i)v\,, \quad u + Dg(z_i)u \in T_{p_i}\Sigma \qquad \text{for } \ i = 1, 2,$$

which implies $v + Df(x_i)v = F_{p_i}(v)$ and $u + Dg(z_i)u = G_{p_i}(u)$ for $i = 1, 2$. Thus, it suffices to estimate

$$\|Dg(z_1) - Dg(z_2)\| = \sup_{e\in U\cap\mathbb{S}} |(G_{p_1} - G_{p_2})(e)|.$$

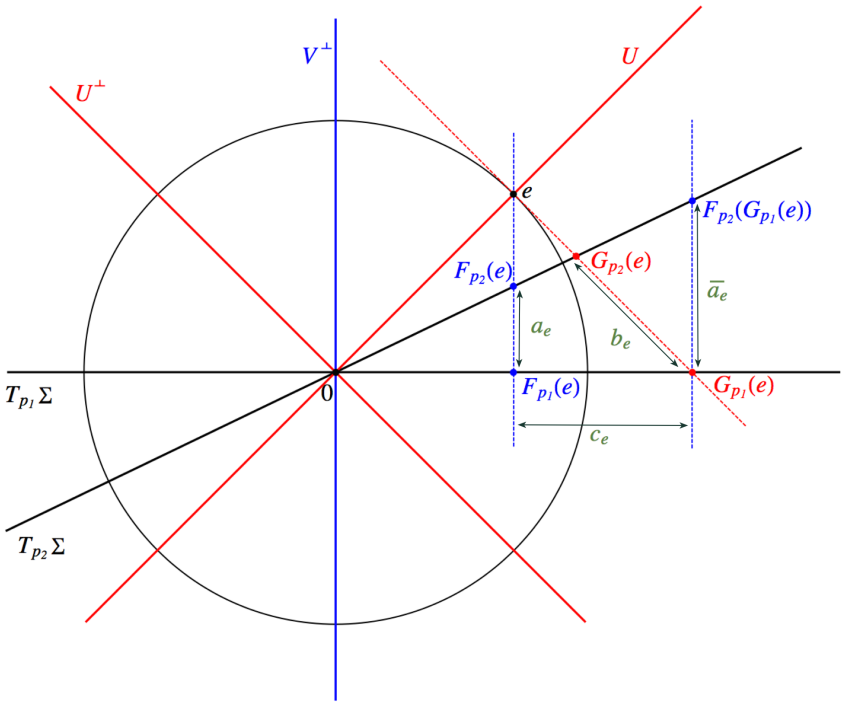

Figure 2: A unit vector $e \in U \cap \mathbb{S}$ and its corresponding a_e, b_e, c_e and $\bar{a}_e$.

For any given unit vector $e \in U \cap \mathbb{S}$, we set

$$\begin{aligned} &a_e = F_{p_1}e - F_{p_2}e \in V^\perp\,, \quad b_e = G_{p_1}e - G_{p_2}e \in U^\perp\,,\\ &c_e = F_{p_1}e - G_{p_1}e \in T_{p_1}\Sigma\,,\\ &\text{and} \quad \bar{a}_e = F_{p_1}(G_{p_1}e) - F_{p_2}(G_{p_1}e) = G_{p_1}e - F_{p_2}(G_{p_1}e) \in V^\perp\,. \end{aligned}$$

Since $F_{p_1}e \in e + V^\perp$, we have $e + V^\perp = F_{p_1}e + V^\perp$, which means that $F_{p_2}(F_{p_1}e) = F_{p_2}e$. In consequence we may write

$$\begin{aligned} (24) \qquad |\bar{a}_e| &\le |(F_{p_1} - F_{p_2})(G_{p_1}e - F_{p_1}e)| + |F_{p_1}(F_{p_1}e) - F_{p_2}(F_{p_1}e)|\\ &\le \|F_{p_1} - F_{p_2}\|\,|c_e| + |(F_{p_1} - F_{p_2})e|\,. \end{aligned}$$

Recall that $\sphericalangle(U^\perp, V^\perp) \le \vartheta$ by assumption, and by (22) for $p := p_1 \in \Sigma$ that

$$\sphericalangle(U^\perp, T_{p_1}\Sigma^\perp) = \sphericalangle(U, T_{p_1}\Sigma) \le \frac{1}{\sqrt{2}} + \vartheta < \frac{3}{4},$$

so that we can apply Proposition 2.4 to the subspaces $X := V^\perp$, $Y := U^\perp$, $Z := T_{p_1}\Sigma$, and to the points $x := F_{p_1}(e) - e \in V^\perp$, and $z := c_e \in T_{p_1}\Sigma$ (hence $z - x = -(G_{p_1}(e) - e) \in U^\perp$) to arrive at

$$|c_e| \leq 4\vartheta |F_{p_1}(e) - e| \leq 4\vartheta \big(\|F_{p_1}\| + |e|\big) < 12\vartheta. \tag{25}$$

Combining (25) and (24) we get

$$|\bar{a}_e| < \|F_{p_1} - F_{p_2}\|(1 + 12\vartheta)\,.$$

Observe that $\bar{a}_e - b_e = G_{p_2}(e) - F_{p_2}\big(G_{p_1}(e)\big) \in T_{p_2}\Sigma$. Applying Prop. 2.4 to $z := b_e - \bar{a}_e \in T_{p_2}\Sigma =: Z$, $x := b_e \in U^\perp =: X$ (with $z - x = -\bar{a}_e \in V^\perp =: Y$, so that $\sphericalangle(X, Y) = \sphericalangle(U^\perp, V^\perp) \leq \vartheta =: \theta$, and $\sphericalangle(Y, Z^\perp) = \sphericalangle(V, T_{p_2}\Sigma) \leq 1/\sqrt{2} =: \lambda < 3/4$ by (21)) yields $|b_e - \bar{a}_e| < 4\vartheta |b_e|$, and in consequence, $|G_{p_1}(e) - G_{p_2}(e)| = |b_e| \leq |b_e - \bar{a}_e| + |\bar{a}_e| \leq 4\vartheta |b_e| + |\bar{a}_e|$, i.e.,

$$|b_e| \leq \tfrac{1}{1-4\vartheta}|\bar{a}_e| \leq \tfrac{1+12\vartheta}{1-4\vartheta}\|F_{p_1} - F_{p_2}\|.$$

Since $e \in U \cap \mathbb{S}^{n-1}$ was arbitrary, we conclude

$$\|Dg(z_1) - Dg(z_2)\| \leq \tfrac{1+12\vartheta}{1-4\vartheta}\|F_{p_1} - F_{p_2}\| = \tfrac{1+12\vartheta}{1-4\vartheta}\|Df(x_1) - Df(x_2)\|\,. \tag{26}$$

At this point we know already that $\Sigma = \operatorname{graph}(f) = \operatorname{graph}(g)$ and that

$$\operatorname{Lip}(g) \leq \frac{1 + 2\vartheta}{1 - 2\vartheta}. \tag{27}$$

Exchanging U with V and f with g, and using (27) (instead of $\operatorname{Lip}(f) \leq 1$) in the derivation of (19) one shows

$$|(x_2 - x_1) - (z_2 - z_1)| \leq \vartheta \left(1 + \tfrac{1+2\theta}{1-2\theta}\right) |z_2 - z_1| \leq 3\vartheta |z_2 - z_1|\,, \tag{28}$$

whenever $x_i = V_\natural(p_i)$, $z_i = U_\natural(p_i)$ and $p_i = z_i + g(z_i) \in \Sigma$ for $i = 1, 2$. Set $s = \frac{1}{1+3\vartheta}r$. If $v \in U$, $x = V_\natural(v + g(v))$, $z_i \in U \cap \mathbb{B}(v, s)$ and $x_i = V_\natural(z_i + g(z_i))$ for $i = 1, 2$, then $|x_i - x| \leq (1 + 3\vartheta)|z_i - v| < (1 + 3\vartheta)s = r$ by (28). Hence, employing (26), (28), and our assumption on the oscillation of Df on $V \cap \mathbb{B}(v, r)$ one obtains

$$\begin{aligned} \|Dg(z_1) - Dg(z_2)\| &\overset{(26)}{\leq} \tfrac{1+12\vartheta}{1-4\vartheta}\|Df(x_1) - Df(x_2)\| \\ &\leq \tfrac{1+12\vartheta}{1-4\vartheta} L|x_1 - x_2|^\alpha \overset{(28)}{\leq} \tfrac{(1+12\vartheta)(1+3\vartheta)^\alpha}{1-4\vartheta} L|z_1 - z_2|^\alpha\,, \end{aligned}$$

for all $z_1, z_2 \in U \cap \mathbb{B}(v, \frac{1}{1+3\vartheta}r)$. □

Proof of Theorem 3. Applying Blaschke's selection theorem (cf. [33]) to the sequence $(\Sigma_j)_j$ contained in the class $\mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$ we obtain a subsequence (still denoted by $(\Sigma_j)_j$) that converges in the Hausdorff metric to a compact set $\Sigma_0 \subset \overline{\mathbb{B}}(0, d) \subset \mathbb{R}^n$. Up to possibly choosing another subsequence we can assume that $d_{\mathcal{H}}(\Sigma_0, \Sigma_{j+1}) \leq \frac{1}{2} d_{\mathcal{H}}(\Sigma_0, \Sigma_j)$ for each $j \in \mathbb{N} \sim \{0\}$. Then, by induction,

$$d_{\mathcal{H}}(\Sigma_0, \Sigma_{j+k}) \leq 2^{-k} d_{\mathcal{H}}(\Sigma_0, \Sigma_j)$$

and in consequence, by the triangle inequality,

$$d_{\mathcal{H}}(\Sigma_j, \Sigma_{j+k}) \geq (1 - 2^{-k}) d_{\mathcal{H}}(\Sigma_0, \Sigma_j) \tag{29}$$

for $j, k \in \mathbb{N} \sim \{0\}$.

We will prove in the first step that $\Sigma_0 \in \mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$, that is, Σ_0 satisfies Definition 1, and then in the second step that the sequence $(\Sigma_j)_{j\in\mathbb{N}}$ converges to Σ_0 in $\mathscr{C}^{1,\alpha}_{m,n}$, i.e., converges in the sense of Definition 2. Note that condition (i) of Definition 1 for Σ_0 and condition (i) of Definition 2 for $(\Sigma_j)_{j\in\mathbb{N}}$ are automatically satisfied.

Step 1: $\Sigma_0 \in \mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$. The convergence of the Σ_j in the Hausdorff metric implies that there is an index $j_0 \in \mathbb{N}$ such that

$$d_{\mathcal{H}}(\Sigma_j, \Sigma_l) < \min\left\{2^{-10} R^2\,,\, L^{-2/\alpha}\,,\, 1\right\} \quad \forall j, l \geq j_0\,. \tag{30}$$

Fix a point in the limit set $x_0 \in \Sigma_0$. Choose points $x_j \in \Sigma_j$ which realise the distance from x_0, i.e.

$$|x_0 - x_j| = \operatorname{dist}(x_0, \Sigma_j) \leq d_{\mathcal{H}}(\Sigma_0, \Sigma_j) \quad \text{for } j \in \mathbb{N}. \tag{31}$$

Then $x_j \to x_0$ as $j \to \infty$ and

$$\begin{aligned} |x_j - x_{j+k}| &\leq |x_j - x_0| + |x_0 - x_{j+k}| \\ &\overset{(31)}{\leq} 2 d_{\mathcal{H}}(\Sigma_0, \Sigma_j) \\ &\overset{(29)}{\leq} 4 d_{\mathcal{H}}(\Sigma_j, \Sigma_{j+k}) \qquad \text{for all } \ j, k \geq 1. \end{aligned} \tag{32}$$

Recalling (30), we may apply Lemma 3.1 with $A := 4$ to deduce, for the above $x_j \in \Sigma_j$ and $x_l \in \Sigma_l$ with $j, l \geq j_0$, the angle estimate

$$\sphericalangle(T_{x_j}\Sigma_j, T_{x_l}\Sigma_l) \leq C_{\mathrm{ang}}(L, 4) d_{\mathcal{H}}(\Sigma_j, \Sigma_l)^{\alpha/2}\,. \tag{33}$$

Consequently, $(T_{x_j}\Sigma_j)_j$ is a Cauchy sequence in $G(n,m)$ that converges to some $T \in G(n,m)$, i.e.,

$$\vartheta_j := \measuredangle(T_{x_j}\Sigma_j, T) \to 0 \text{ as } j \to \infty\,. \tag{34}$$

Recall that by Definition 1 we find for each $j \in \mathbb{N}$ a function $f_j \in C^{1,\alpha}(T_{x_j}\Sigma_j, T_{x_j}\Sigma_j^\perp)$ with $f_j(0) = 0$, $Df_j(0) = 0$, such that

$$\Sigma_j \cap \mathbb{B}(x_j, R) = (x_j + \operatorname{graph}(f_j)) \cap \mathbb{B}(x_j, R), \tag{35}$$

with the uniform estimates $\operatorname{Lip}(f_j) \le 1$ and $\|Df_j(x) - Df_j(y)\| \le L|x-y|^\alpha$ for all $x, y \in T_{x_j}\Sigma_j$. We can assume by (33) that $\vartheta_j \in (0, 1/100)$ for all $j \ge j_0$, so that Lemma 3.2 applied to the radius $r = \infty$ leads to functions $g_j \in C^{1,\alpha}(T, T^\perp)$ such that $\operatorname{graph}(f_j) = \operatorname{graph}(g_j)$, $g_j(0) = 0$ and

$$\|Dg_j(\xi) - Dg_j(\eta)\| \le L_j|\xi - \eta|^\alpha \quad \text{for all } \xi, \eta \in T \text{ and } j \ge j_0, \tag{36}$$

where $L_j := L_{g_j}(L, \vartheta_j, \alpha) \to L$ as $j \to \infty$. Moreover,

$$\operatorname{Lip}(g_j) \le \frac{1+2\vartheta_j}{1-2\vartheta_j} \to 1 \quad \text{on } T, \text{ and } \|Dg_j(0)\|^2 \le \frac{\vartheta_j^2}{1-\vartheta_j^2} \to 0 \quad \text{as } j \to \infty. \tag{37}$$

In addition, (35) translates into

$$\Sigma_j \cap \mathbb{B}(x_j, R) = (x_j + \operatorname{graph}(g_j)) \cap \mathbb{B}(x_j, R) \quad \text{for all } j \ge j_0. \tag{38}$$

Because of the uniform estimates (36) and (37) we can repeatedly apply Arzela-Ascoli's theorem to successively choose subsequences $(j_{i+1})_{i+1} \subset (j_i)_i$ for $i \in \mathbb{N}$, such that g_{j_i} converges in C^1 to a function $G_i \in C^{1,\alpha}(T \cap \overline{\mathbb{B}}(0,i), T^\perp)$ with $G_i(0) = 0$, such that

$$\|DG_i(\xi) - DG_i(\eta)\| \le L|\xi - \eta|^\alpha \quad \text{for all } \xi, \eta \in T \cap \overline{\mathbb{B}}(0,i), \tag{39}$$

and with

$$\operatorname{Lip}(G_i) \le 1 \quad \text{on } T \cap \overline{\mathbb{B}}(0,i) \text{ and } DG_i(0) = 0. \tag{40}$$

In addition, one has $G_{i+1}|_{\overline{\mathbb{B}}(0,i)} = G_i$ for all $i \in \mathbb{N}$. Then the diagonal sequence g_{j_j} converges in $C^1_{\mathrm{loc}}(T, T^\perp)$ to some limit function $G \in C^{1,\alpha}(T, T^\perp)$

satisfying $G(0) = 0$, $DG(0) = 0$, $\mathrm{Lip}(G) \le 1$ on T, and the estimate

$$\|DG(\xi) - DG(\eta)\| \le L|\xi - \eta|^\alpha \quad \text{for all } \xi, \eta \in T. \tag{41}$$

Applying (38) to the diagonal sequence $(g_{j_j})_{j_j} \subset (g_j)_j$ combined with (31) one finds

$$\Sigma_0 \cap \mathbb{B}(x_0, R) = (x_0 + \mathrm{graph}(G)) \cap \mathbb{B}(x_0, R), \tag{42}$$

which concludes Step 1 since Σ_0 is represented near the arbitrarily chosen point $x_0 \in \Sigma_0$ as a graph of the $C^{1,\alpha}$ function $G : T \to T^\perp$ satisfying all the requirements in Definition 1 observing, in addition, that since $x \mapsto x_0 + x + G(x)$ parameterises Σ_0 locally near x_0 one has $T_{x_0}\Sigma_0 = (\mathrm{Id} + DG(0))(T) = T$ since $DG(0) = 0$, which a posteriori shows that the m-plane T does not depend on the sequence $x_j \to x_0$.

Step 2: Σ_j converges in $\mathscr{C}^{1,\alpha}_{m,n}$ to Σ_0. It suffices to check condition (iii) of Definition 2. Let $j_2 \in \mathbb{N}$, $j_2 \ge 200$ be such that

$$d_{\mathcal{H}}(\Sigma_j, \Sigma_0) < \min\left\{2^{-10}R^2,\, L^{-2/\alpha},\, 2^{-8}R,\, \left(2^{-7}C_{\mathrm{ang}}(L,4)^{-1}\right)^{2/\alpha},\, 2^{-8}\right\} \quad \text{for } j \ge j_2\,. \tag{43}$$

Fix $x \in \Sigma_0$ and set $T = T_x\Sigma_0$. As before for each $j \in \mathbb{N}$ find $x_j \in \Sigma_j$ such that $|x - x_j| = \mathrm{dist}(x, \Sigma_j) \le d_{\mathcal{H}}(\Sigma_0, \Sigma_j)$ and let $f_j : T_{x_j}\Sigma_j \to T_{x_j}\Sigma_j^\perp$ and $f_0 := f_x : T \to T^\perp$ be the functions whose existence is guaranteed by condition (ii) of Definition 1. According to Lemma 3.1 (generously for $A = 4$), we get by (43)

$$\sphericalangle(T_x\Sigma_0, T_{x_j}\Sigma_j) \le C_{\mathrm{ang}}(L,4)d_{\mathcal{H}}(\Sigma_0, \Sigma_j)^{\alpha/2} \overset{(43)}{<} 2^{-7} < \tfrac{1}{100} \quad \text{for } j \ge j_2. \tag{44}$$

An application of Lemma 3.2 yields now functions $g_j \in C^{1,\alpha}(T, T^\perp)$ such that $\mathrm{graph}(g_j) = \mathrm{graph}(f_j)$ and $\mathrm{Lip}(g_j) \le \frac{51}{49} < 2$ for each $j \ge j_2$, or $j = 0$ where $g_0 = f_0$. Set $h_j(\eta) = g_j(\eta - T_\natural(x_j - x)) + T^\perp_\natural(x_j - x)$ for $\eta \in T$ and $j \ge j_2$, and for $j = 0$ we have $h_0 = g_0 = f_0$, so that $x_j + \mathrm{graph}(g_j) = x + \mathrm{graph}(h_j)$ and consequently, recalling (43),

$$\Sigma_j \cap \mathbb{B}(x, (1 - 2^{-8})R) = (x + \mathrm{graph}(h_j)) \cap \mathbb{B}(x, (1 - 2^{-8})R) \quad \text{for } j \ge j_2 \text{ or } j = 0. \tag{45}$$

Set $\rho := \min\{\frac{1}{12}R, \frac{1}{2}(2^{-7}/L)^{1/\alpha}\}$ and note that

$$\{x + \eta + h_j(\eta) : \eta \in T \cap \mathbb{B}(0, 3\rho)\} \subseteq \Sigma_j \quad \text{for } j \ge j_2 \text{ or } j = 0, \tag{46}$$

because $\mathrm{Lip}(g_j) = \mathrm{Lip}(h_j) < 2$. Let $\eta \in T \cap \mathbb{B}(0, 2\rho)$, $j \in \mathbb{N}$, $j \geq j_2$ and set $p := x + \eta + h_0(\eta) \in \Sigma_0$ and $q := x + \eta + h_j(\eta) \in \Sigma_j$. There exists $z \in \Sigma_j$ with $|z - p| \leq d_{\mathcal{H}}(\Sigma_0, \Sigma_j)$. By (43) we have $d_{\mathcal{H}}(\Sigma_0, \Sigma_j) < \rho$, and if we write $z = x + \xi + h_j(\xi)$, then $\eta - \xi = T_\natural(p - x) - T_\natural(z - x)$ so that $|\eta - \xi| \leq |z - p| \leq d_{\mathcal{H}}(\Sigma_0, \Sigma_j)$, and therefore $\xi \in T \cap \mathbb{B}(0, 3\rho)$. Since $\mathrm{Lip}(h_j) < 2$ we obtain

$$\begin{aligned} |h_0(\eta) - h_j(\eta)| = |p - q| &\leq |p - z| + |z - q| \\ &\leq d_{\mathcal{H}}(\Sigma_0, \Sigma_j) + |\eta - \xi| + |h_j(\eta) - h_j(\xi)| \\ &< 4 d_{\mathcal{H}}(\Sigma_0, \Sigma_j)\,. \end{aligned} \tag{47}$$

We already know that $\Sigma_0 \in \mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$ so employing Lemma 3.1 for $A = 4$, we get

$$\measuredangle(T_p\Sigma_0, T_q\Sigma_j) \leq C_{\mathrm{ang}}(L, 4) d_{\mathcal{H}}(\Sigma_0, \Sigma_j)^{\alpha/2} \quad \text{for } j \geq j_2 \text{ or } j = 0. \tag{48}$$

Apply [2, 8.9(5)] with $\eta_1 := Dh_j(\eta)$, $\eta_2 := Dh_0(\eta)$, $S_1 := T_q\Sigma_j$, $S_2 := T_p\Sigma_0$, and $S := T$ to obtain, recalling (44) and $\mathrm{Lip}(h_0) = \mathrm{Lip}(f_0) \leq 1$,

$$\begin{aligned} \|Dh_j(\eta) - Dh_0(\eta)\|^2 &\leq \frac{\measuredangle(T_q\Sigma_j, T_p\Sigma_0)^2}{1 - \measuredangle(T_q\Sigma_j, T)^2} \left(1 + \|Dh_0(\eta)\|^2\right) \\ &\leq \frac{2}{1 - \measuredangle(T_q\Sigma_j, T)^2} C_{\mathrm{ang}}(L, 4)^2 d_{\mathcal{H}}(\Sigma_j, \Sigma_0)^\alpha\,. \end{aligned} \tag{49}$$

To analyze the term in the denominator we estimate using (48) and (43)

$$\measuredangle(T_q\Sigma_j, T) \leq \measuredangle(T_q\Sigma_j, T_p\Sigma_0) + \measuredangle(T_p\Sigma_0, T) \overset{(48),(43)}{\leq} 2^{-7} + \measuredangle(T_p\Sigma_0, T).$$

For the last summand we again use [2, 8.9(5)], this time for $S := T$, $\eta_1 := Dh_0(\eta)$, $S_1 := T_p\Sigma_0 = (\mathrm{Id} + \eta_1)(T)$, $\eta_2 := 0$, and $S_2 := (\mathrm{Id} + \eta_2)(T) = T$, to deduce by virtue of $Dh_0(0) = Df_0(0) = 0$ the angle estimate

$$\measuredangle(T_p\Sigma_0, T) \leq \|Dh_0(\eta)\| \leq L|\eta|^\alpha \overset{(43)}{\leq} 2^{-7}$$

by our choice of ρ. Therefore we can insert the resulting estimate $\measuredangle(T_q\Sigma_j, T)^2 \leq 2^{-12}$ into (49) to obtain

$$\|Dh_j(\eta) - Dh_0(\eta)\|^2 \leq 3 C_{\mathrm{ang}}(L, 4)^2 d_{\mathcal{H}}(\Sigma_j, \Sigma_0)^\alpha. \tag{50}$$

Since $\eta \in T \cap \mathbb{B}(0, 2\rho)$ and $j \geq j_2$ were chosen arbitrarily, the estimates (47) and (50) hold for any $\eta \in T \cap \mathbb{B}(0, 2\rho)$ and $j \geq j_2$. Fix a smooth cut-off function $\varphi : T \to \mathbb{R}$ such that $\varphi(\eta) = 1$ for $\eta \in T \cap \mathbb{B}(0, \rho)$ and $\varphi(\eta) = 0$

for $\eta \in T \sim \mathbb{B}(0, 2\rho)$. For $j \geq j_2$ and for $j = 0$ define $f_{x,j} \in C^{1,\alpha}(T, T^\perp)$ by $f_{x,j}(\eta) := h_j(\eta)\varphi(\eta)$ for $\eta \in T \cap \mathbb{B}(0, 2\rho)$ and $f_{x,j}(\eta) = 0$ for $\eta \in T \sim \mathbb{B}(0, 2\rho)$. Estimates (47) and (50) show that the sequence $(f_{x,j})_{j\in\mathbb{N}}$ converges in $C^1(T, T^\perp)$ to the function $f_{x,0}$. Since the limit function $f_{x,0} = \varphi h_0$ is of class $C^{1,\alpha}$, it follows that $(f_{x,j})_{j\in\mathbb{N}}$ actually converges in $C^{1,\alpha'}$ for any $\alpha' \in (0, \alpha)$. Moreover, by (45) and (46) one sees that

$$\Sigma_j \cap \mathbb{B}(x, \rho) = (x + \operatorname{graph}(f_{x,j})) \cap \mathbb{B}(x, \rho) \quad \text{for } j \geq j_2 \text{ or } j = 0.$$

Therefore, $(\Sigma_j)_{j\in\mathbb{N}}$ satisfies condition (iii) of Definition 2 and the proof is complete. □

4. Isotopies, tubular neighbourhoods and diffeomorphisms

To prove Theorem 4 we proceed as in [22, Chapter 4, Section 5].

We assign to each $V \in G(n, n-m)$ an orthogonal projection $V_\natural \in \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^n)$ onto V. By [17, 3.1.19(2)] the set

$$\mathcal{G} = \{P \in \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^n) : P \circ P = P,\ P^* = P,\ \operatorname{trace} P = n - m\}$$

is a C^∞-submanifold of $\mathbb{R}^{n^2}$, and the mapping $V \mapsto V_\natural$ is a C^∞-diffeomorphism *and an isometry*.

Definition 4.1. Let $\Sigma \subset \mathbb{R}^n$ be an m-dimensional C^1-submanifold of $\mathbb{R}^n$ and $\varepsilon > 0$. A map $\Phi : \Sigma \to \mathcal{G}$ is called an *ε-normal map for* Σ if Φ is C^1-smooth, $\operatorname{Lip}(\Phi) < \infty$ and if

$$\|\Phi(x) - (T_x\Sigma)^\perp_\natural\| \leq \varepsilon \quad \forall x \in \Sigma\,.$$

Lemma 4.2 (nearly normal spaces of class C^1). *Let $L, R, d > 0$, $\alpha \in (0, 1]$, and $\Sigma \in \mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$. Then there exists a constant*

$$C = C(L, R, \alpha, m, n) \geq 1$$

such that for each $\varepsilon \in (0, 1]$ there is an ε-normal map $\Phi^\varepsilon[\Sigma] : \Sigma \to \mathcal{G}$ for Σ satisfying, in addition, $\operatorname{Lip}(\Phi^\varepsilon[\Sigma]) \leq C\varepsilon^{-1/\alpha}$.

Remark. A similar statement for smooth manifolds (including the C^1-case) can be found in [44, Thm. 10A, p.121], but for the convenience of the reader, and to emphasise how the constants depend quantitatively on the parameters determining the class $\mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$ we provide the full argument

here. We are going to construct $\Phi^\varepsilon[\Sigma]$ simply by mollifying the map $x \mapsto (T_x\Sigma)_\natural$. Note that since Σ is embedded we do not need to use the *center of mass* tool known from Riemannian geometry, which was used in [8].

Proof. For $\Phi_0 : \Sigma \to \mathcal{G}$ given by $\Phi_0(x) := (T_x\Sigma^\perp)_\natural$ for $x \in \Sigma$, we first prove a simple Hölder estimate as follows. For $x, y \in \Sigma$ we find

$$\|\Phi_0(x) - \Phi_0(y)\| = \|(T_x\Sigma)^\perp_\natural - (T_y\Sigma)^\perp_\natural\| \leq \frac{2|x-y|^\alpha}{\min\{R^\alpha, (L\sqrt{2})^{-1}\}} \tag{51}$$

if $|x-y|^\alpha \geq \min\{R^\alpha, (L\sqrt{2})^{-1}\}$. If not, then $y \in \mathbb{B}(x,R)$ so that we can use the local graph representation

$$\Sigma \cap \mathbb{B}(x,R) = (x + \operatorname{graph}(f)) \cap \mathbb{B}(x,R)$$

to express the point y as $y = x + \xi + f(\xi)$ for some $\xi \in T_x\Sigma$ and the function $f := f_x \in C^{1,\alpha}(T_x\Sigma, T_x\Sigma^\perp)$ with $f(0) = 0$, $Df(0) = 0$, $\operatorname{Lip}(f) \leq 1$, and the Hölder estimate on Df as in Definition 1. In other words, the mapping $F(\xi) := x + \xi + f(\xi)$ for $\xi \in T_x\Sigma$ parameterises Σ over the tangent plane $T_x\Sigma$ locally near x, so that its differential $DF(\xi) : T_x\Sigma \to T_y\Sigma$ can be used to estimate for an orthonormal basis $\{e_1, \ldots, e_m\}$ of $T_x\Sigma$

$$\begin{aligned}
\operatorname{dist}(e_i, T_y\Sigma) &\leq |e_i - DF(\xi)e_i| = |e_i - (\mathrm{Id} + Df(\xi))e_i| \\
&\leq \|Df(\xi) - Df(0)\| \leq L|\xi|^\alpha \leq L|(x + \xi + f(\xi)) - x|^\alpha \\
&= L|y-x|^\alpha \quad \forall i = 1, \ldots, m,
\end{aligned}$$

where we also used that $f(\xi) \perp \xi$ by definition of f. Since $|y-x|^\alpha < \min\{R^\alpha, (L\sqrt{2})^{-1}\} \leq (L\sqrt{2})^{-1}$ in the present case, we can apply a quantitative linear algebra estimate [25, Prop. 2.5] to find a constant $C = C(m)$ such that

$$\|\Phi_0(x) - \Phi_0(y)\| = \measuredangle(T_x\Sigma, T_y\Sigma) \leq C(m)L|x-y|^\alpha.$$

Combining both cases leads to the desired Hölder estimate for Φ_0 with Hölder constant

$$C_0 = C_0(L, R, \alpha, m) := \max\left\{\frac{2}{\min\{(L\sqrt{2})^{-1}, R^\alpha\}}, C(m)L\right\}.$$

Notice that the constant C_0 does not depend on R or α if $R^\alpha \geq (L\sqrt{2})^{-1}$.

Choosing an orthonormal coordinate system in $\mathbb{R}^n$ we can represent Φ_0 as an $(n \times n)$-matrix of functions $(\Phi_0^{ij})_{i,j=1}^n$ and extend each Φ_0^{ij} to all of $\mathbb{R}^n$ by setting

$$\Phi_1^{ij}(x) = \inf_{z\in\Sigma}\{\Phi_0^{i,j}(z) + C_0|z-x|^\alpha\}$$

preserving the same Hölder exponent α and Hölder constant C_0 for each $i,j = 1,\dots,n$ (cf. for the proof of [15, Theorem 1, p.80] which carries over to all $\alpha \in (0,1]$). The matrix $(\Phi_1^{ij})_{i,j=1}^n$ represents the Hölder continuous mapping $\Phi_1 : \mathbb{R}^n \to \mathrm{Hom}(\mathbb{R}^n,\mathbb{R}^n)$ with $\Phi_1|_\Sigma = \Phi_0$ and the estimate

$$\|\Phi_1(x) - \Phi_1(y)\| \le C_1|x-y|^\alpha \quad \forall x,y \in \mathbb{R}^n\,, \tag{52}$$

where $C_1 = C_1(L,R,\alpha,m,n) := nC_0(L,R,\alpha,m)$. Now let $\phi \in C_0^\infty(\mathbb{B}(0,1))$ with $\phi(x) = 1$ for all $x \in \mathbb{B}(0,1/2))$, $0 \le \phi(x) \le 1$ and $|\nabla\phi(x)| \le 4$ for all $x \in \mathbb{B}(0,1)$, and $\int_{\mathbb{R}^n}\phi(x)\,dx = 1$, and consider for $r > 0$ the usual scaling $\phi_r(x) := r^{-n}\phi(x/r)$ to define the convolution $\Phi_{2,r} : \mathbb{R}^n \to \mathrm{Hom}(\mathbb{R}^n,\mathbb{R}^n)$ as

$$\Phi_{2,r}(x) = \phi_r * \Phi_1(x) = \int_{\mathbb{R}^n}\phi_r(x-z)\Phi_1(z)\,dz\,.$$

Since

$$\begin{aligned}\|\Phi_1(z)\| &\le \|\Phi_1(x)\| + C_1|x-z|^\alpha\\ &= \|\Phi_0(x)\| + C_1|x-z|^\alpha \le 1 + C_1r^\alpha \le 1 + 2C_1\end{aligned}$$

for all $x \in \Sigma$, $z \in \mathbb{B}(x,r)$, $r \in (0,2]$, we find

$$\|\Phi_1(\cdot)\| \le 1 + 2C_1 \quad \text{on } \Sigma + \mathbb{B}(0,2), \tag{53}$$

where the constant on the right-hand side depends on L,R,α,m, and n. Therefore, we can estimate for $x,y \in \Sigma + \mathbb{B}(0,1)$, $e \in \mathbb{S}^{n-1}$, $r \in (0,1)$,

$$\begin{aligned}|\Phi_{2,r}(x)e| &= \left|\int_{\mathbb{R}^n}\phi_r(x-z)\Phi_1(z)e\,dz\right|\\ &\le \int_{\mathbb{B}(x,r)}\|\Phi_1(z)\|\phi_r(x-z)\,dz \le 1 + 2C_1,\end{aligned} \tag{54}$$

because $\mathrm{dist}(z,\Sigma) \le |z-x| + 1 < 2$ for all $z \in \mathbb{B}(x,r)$, whence

$$\|\Phi_{2,r}(x) - \Phi_{2,r}(y)\| \le 2(1+2C_1) \le 2(1+2C_1)\frac{|x-y|}{r} \tag{55}$$

for all $x, y \in \Sigma + \mathbb{B}(0,1)$ with $|x-y| \geq r$, $r \in (0,1)$. On the other hand, for $x, y \in \mathbb{B}(0,1) + \Sigma$ with $|x-y| < r$ and for $e \in \mathbb{S}^{n-1}$ one estimates

$$\begin{aligned} (56) \qquad & \big|\big(\Phi_{2,r}(x) - \Phi_{2,r}(y)\big)e\big| \\ & = \left| \int_0^1 \nabla \Phi_{2,r}(tx + (1-t)y) \cdot (x-y)e \, dt \right| \\ & \leq \frac{1}{r^{n+1}} \int_0^1 \int_{\mathbb{B}(tx+(1-t)y,r)} \big|\nabla\phi\big(\tfrac{tx+(1-t)y-z}{r}\big)\big| \|\Phi_1(z)\| \, dz dt |x-y|. \end{aligned}$$

Since $\operatorname{dist}(tx + (1-t)y, \Sigma) \leq \operatorname{dist}(x, \Sigma) + (1-t)|x-y|$ for all $t \in [1/2, 1]$ and $\operatorname{dist}(tx + (1-t)y, \Sigma) \leq \operatorname{dist}(y, \Sigma) + t|x-y|$ for all $t \in [0, 1/2]$ we find $\operatorname{dist}(tx+(1-t)y, \Sigma) < 1 + r/2$ for all $t \in [0,1]$; hence $\operatorname{dist}(z, \Sigma) < 1 + 3r/2 < 2$ for all $z \in \mathbb{B}(tx + (1-t)y, r)$, $r \in (0, 2/3)$, which implies $\|\Phi_1(z)\| \leq 1 + 2C_1$ for such z by virtue of (53), which inserted in (56) gives

$$\begin{aligned} (57) \qquad \|\Phi_{2,r}(x) - \Phi_{2,r}(y)\| & \leq 4(1 + 2C_1)\frac{\omega_n}{r}|x-y| \leq 24(1+2C_1)\frac{|x-y|}{r} \\ & =: C_2(L, R, \alpha, m, n)\frac{|x-y|}{r} \end{aligned}$$

for all $x, y \in \Sigma + \mathbb{B}(0,1)$ and $r \in (0, 2/3)$. (We have used that the volume ω_n of the n-dimensional unit ball is at most 6 for all $n = 1, 2, \ldots$)

Furthermore, for $x \in \Sigma$,

$$\begin{aligned} \Phi_0(x) - \Phi_{2,r}(x) &= \int_{\mathbb{R}^n} \big(\Phi_0(x) - \Phi_1(z)\big)\phi_r(x-z) \, dz \\ &= \int_{\mathbb{B}(x,r)} \big(\Phi_1(x) - \Phi_1(z)\big)\phi_r(x-z) \, dz, \end{aligned}$$

since $\Phi_1|_\Sigma = \Phi_0$, so that by (52)

$$\begin{aligned} (58) \qquad \|\Phi_0(x) - \Phi_{2,r}(x)\| &\leq C_1 \int_{\mathbb{B}(x,r)} |x-z|^\alpha \phi_r(x-z) \, dz \\ &\leq C_1 r^\alpha < C_2 r^\alpha \quad \forall x \in \Sigma, \, r > 0. \end{aligned}$$

Since $\mathcal{G}$ is a C^∞-submanifold of $\operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^n) \simeq \mathbb{R}^{n^2}$, it has positive reach $r_{\mathcal{G}} = r_{\mathcal{G}}(m,n) > 0$ in the sense of Federer [16, Definition 4.1] such that the nearest point projection $P_{\mathcal{G}} : \mathcal{G} + \mathbb{B}(0, r_{\mathcal{G}}) \to \mathcal{G}$ is C^∞-smooth; see, e.g., [18,

Lemma, p. 153][3]. In addition, for any $\delta \in (0, r_{\mathcal{G}}/2]$ it follows from [16, Theorem 4.8(8)] that $P_{\mathcal{G}}$ has Lipschitz constant $\mathrm{Lip}(P_{\mathcal{G}}) \le 2$ on $\mathcal{G} + \mathbb{B}(0,\delta)$.

According to (58) the map $\Phi_{3,r} := P_{\mathcal{G}} \circ \Phi_{2,r}|_\Sigma$ maps Σ into $\mathcal{G}$ if $C_2 r^\alpha \le r_{\mathcal{G}}/2$. Now choose for a given $\varepsilon \in (0,1]$ first

$$\delta_0 = \delta_0(L,R,\alpha,m,n) := \min\{r_{\mathcal{G}}, (2/3)^\alpha C_2, 1\}, \tag{59}$$

and then $r_\varepsilon := (\delta_0/(2C_2))^{1/\alpha}\varepsilon^{1/\alpha} \in (0,2/3)$. Then $\Phi^\varepsilon[\Sigma] := \Phi_{3,r_\varepsilon} : \Sigma \to \mathcal{G}$ as a composition of C^1-maps is also of class C^1, and according to (57) with Lipschitz constant

$$\begin{aligned}\mathrm{Lip}(\Phi^\varepsilon[\Sigma]) &\le \mathrm{Lip}(P_{\mathcal{G}})C_2/r_\varepsilon \\ &\le (2C_2)^{1+(1/\alpha)}(\delta_0\varepsilon)^{-1/\alpha} =: C(L,R,\alpha,m,n)\varepsilon^{-1/\alpha}\end{aligned} \tag{60}$$

Finally, $\Phi^\varepsilon[\Sigma]$ is an ε-normal map for Σ, since by (58)

$$\begin{aligned}\|\Phi^\varepsilon[\Sigma](x) - (T_x\Sigma)^\perp_{\mathfrak{h}}\| &= \|P_{\mathcal{G}} \circ \Phi_{2,r_\varepsilon}(x) - P_{\mathcal{G}} \circ \Phi_0(x)\| \\ &\le \mathrm{Lip}(P_{\mathcal{G}})\|\Phi_{2,r_\varepsilon}(x) - \Phi_0(x)\| \overset{(58)}{<} 2C_2 r_\varepsilon^\alpha \le \varepsilon.\end{aligned}$$

□

Remark 4.3. An inspection of the proof yields $C_0 \le C(m)L + 2R^{-\alpha} \le c(m,n,l,p)(L+1)$, $\alpha = 1 - p_0(\mathcal{E})/p$, whenever R, L are given by (10) for an energy threshold $E \ge \mathcal{E}(\Sigma)$ for a particular energy $\mathcal{E} \in \{\mathcal{E}_p^l, \mathcal{T}_p, \mathcal{T}_p^G\}$. This gives $C_2 \le c(m,n,l,p)(E^{1/p}+1)$. Assuming w.l.o.g. that $C_2 \ge \frac{3}{2}$, we obtain δ_0 in (59) unrelated to C_2, and finally, for a fixed $\varepsilon \in (0, \frac{1}{100})$ and $\alpha = \alpha(p) = 1 - p_0(\mathcal{E})/p$,

$$\mathrm{Lip}(\Phi^\varepsilon[\Sigma]) \le c(m,n,l,p)(E^{1/p}+1)^{1+(1/\alpha)}\delta_0^{-1/\alpha}, \quad \text{where} \quad \delta_0 = \min\{r_{\mathcal{G}}, 1\}.$$

Definition 4.4. Let $R, L, d > 0$, and $\alpha \in (0,1]$, $\Sigma \in \mathscr{C}^{1,\alpha}_{m,n}(R,L,d)$, and for some $\varepsilon \in (0, 1/100)$ let $\Phi : \Sigma \to \mathcal{G}$ be an ε-normal map for Σ. For $\delta > 0$ define the *δ-normal neighbourhood*

$$\begin{aligned}&N_\delta(\Sigma,\Phi) := \{(x,v) \in \Sigma \times \mathbb{R}^n : \Phi(x)v = v,\ |v| < \delta\} \\ &\text{and the map} \quad \Psi_\delta[\Sigma,\Phi] : N_\delta(\Sigma,\Phi) \to \mathbb{R}^n, \qquad \Psi_\delta[\Sigma,\Phi](x,v) := x + v\,.\end{aligned}$$

[3]Formally, Foote [18, Lemma, p. 153] mentions only *a neighbourhood* of the manifold M. However, this neighbourhood is defined via an application of the inverse function theorem, which — in light of Federer [16, Theorem 4.8(13)] — is possible on the whole $\mathcal{G} + \mathbb{B}(0, r_{\mathcal{G}})$.

Lemma 4.5 (tubular neighbourhoods for $C^{1,\alpha}$ manifolds). *Assume $R, L, d > 0$, $\alpha \in (0,1]$, and let $\Sigma \in \mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$, and for $\varepsilon \in (0, 1/100)$ let $\Phi : \Sigma \to \mathcal{G}$ be an ε-normal map for Σ. Then there is a constant $\delta_{tub} = \delta_{tub}(R, L, \alpha, \varepsilon, \mathrm{Lip}(\Phi)) > 0$ such that for all $\delta \in (0, \delta_{tub}]$*

(i) $\Psi = \Psi_\delta[\Sigma, \Phi]$ *is a C^1-embedding,*

(ii) $(1/4)|(x-y, u-v)| \leq |\Psi(x,u) - \Psi(y,v)| \leq \sqrt{2}|(x-y, u-v)|$ *for all $(x,u), (y,v) \in N_\delta(\Sigma, \Phi)$,*

(iii) $\mathrm{dist}(\Psi(x,v), \Sigma) > \frac{1}{4}|v|$ *for all $(x,v) \in N_\delta(\Sigma, \Phi)$, $v \neq 0$,*

(iv) $\Sigma + \mathbb{B}(0, \delta/2) \subset \Psi_\delta[\Sigma, \Phi](N_\delta(\Sigma, \Phi))$.

Proof. For any $\delta > 0$ the mapping $\Psi = \Psi_\delta[\Sigma, \Phi]$ is the restriction of the smooth function

$$\mathbb{R}^n \times \mathbb{R}^n \ni (x, v) \mapsto x + v \in \mathbb{R}^n$$

to the C^1-submanifold

$$N := N_\delta(\Sigma, \Phi)) = \bigcup_{x \in \Sigma} \big[\{x\} \times \big(\ker(\Phi(x) - \mathrm{Id}) \cap \mathbb{B}(0, \delta)\big)\big];$$

hence Ψ is of class C^1. To show that Ψ is an embedding it suffices to prove that it is bilipschitz, i.e., that (ii) holds, for sufficiently small δ. For any $(x,u), (y,v) \in N$ one has

$$|\Psi(x,u) - \Psi(y,v)| = |(x-y) + (u-v)| \leq \sqrt{2}|(x-y, u-v)|\,, \tag{61}$$

and therefore it is enough to prove the estimate from below in (ii). Set

$$\delta_{tub} := \min\left\{\frac{R}{4}, \frac{1}{4}\Big(\frac{\varepsilon}{L}\Big)^{1/\alpha}, \frac{\varepsilon}{4\,\mathrm{Lip}(\Phi)}, 1\right\}. \tag{62}$$

Assume $0 < \delta \leq \delta_{tub}$. For $(x,u), (y,v) \in N$ define the subspaces $U := \operatorname{im}\Phi(x)$ and $V := \operatorname{im}\Phi(y)$ and observe that if $|x-y| \geq 4\delta$ then, on the one hand,

$$|\Psi(x,u) - \Psi(y,v)| \geq |x-y| - |u| - |v| \geq |x-y| - 2\delta \geq |x-y|/2,$$

and, on the other hand,

$$|(x-y, u-v)| \leq |x-y| + |u-v| \leq |x-y| + 2\delta \leq \tfrac{3}{2}|x-y|,$$

so that

$$|\Psi(x,u)-\Psi(y,v)| \geq \frac{1}{3}|(x-y,u-v)|\,. \tag{63}$$

Thus we have to treat the case $|x-y|<4\delta$. Since $\Sigma \in \mathscr{C}^{1,\alpha}_{m,n}(R,L,d)$ and $|x-y|<4\delta \leq R$, we can use the local graph representation

$$\Sigma \cap \mathbb{B}(x,R) = (x+\operatorname{graph}(f_x)) \cap \mathbb{B}(x,R)$$

for a function $f := f_x \in C^{1,\alpha}(T_x\Sigma, T_x\Sigma^\perp)$ satisfying $f(0)=0,\ Df(0)=0$, $\operatorname{Lip}(f)\leq 1$, and the Hölder estimate for Df as in Definition 1, to find for $y = x+\eta+f(\eta) \in x+\operatorname{graph}(f)$, $\eta \in T_x\Sigma$, by means of (14)

$$\operatorname{dist}(y, x+T_x\Sigma) = |f(\eta)| \overset{(14)}{\leq} L|\eta|^{1+\alpha} \leq L|x-y|^{1+\alpha} \leq L(4\delta)^\alpha |x-y|,$$

so that we obtain by our choice of δ_{tub} in (62)

$$|(T_x\Sigma^\perp)_\natural(y-x)| = \operatorname{dist}(y, x+T_x\Sigma) \overset{(62)}{\leq} \varepsilon|x-y|\,. \tag{64}$$

Using this estimate together with the fact that Φ is an ε-normal map for Σ we can write

$$\begin{aligned} |U^\perp_\natural(x-y)| &\geq |(T_x\Sigma)_\natural(x-y)| - \|U^\perp_\natural - (T_x\Sigma)_\natural\||x-y| \\ &\geq (1-\varepsilon)|x-y| - \|\Phi(x) - (T_x\Sigma)^\perp_\natural\||x-y| \\ &\geq (1-2\varepsilon)|x-y|\,, \end{aligned} \tag{65}$$

which implies by means of $|U_\natural(x-y)|^2 = |x-y|^2 - |U^\perp_\natural(x-y)|^2$ the inequality $|U_\natural(x-y)|^2 \leq \big(1-(1-2\varepsilon)^2\big)|x-y|^2$; hence,

$$|U_\natural(x-y)| \leq 2\sqrt{\varepsilon}|x-y|. \tag{66}$$

Recall our choice of δ_{tub} in (62) to estimate for $u=\Phi(x)u \in U\cap\mathbb{B}(0,\delta)$ and $v=\Phi(y)v \in V\cap\mathbb{B}(0,\delta)$

$$\begin{aligned} |U^\perp_\natural(u-v)| &= |U^\perp_\natural v| = |(U_\natural - V_\natural)v| \\ &\leq |v|\|\Phi(x)-\Phi(y)\| \leq \delta \operatorname{Lip}(\Phi)|x-y| \overset{(62)}{\leq} \varepsilon|x-y|\,, \end{aligned} \tag{67}$$

so that

$$|U_\natural(u-v)| = |(\mathrm{Id}-U^\perp_\natural)(u-v)| \geq |u-v| - \varepsilon|x-y|\,. \tag{68}$$

Combining (65), (66), (67), (68) with the triangle inequality, we arrive at

$$\begin{aligned}
|\Psi(x,u)-\Psi(y,v)| &= |(U_\natural^\perp + U_\natural)\big((x-y)+(u-v)\big)| \\
&\geq |U_\natural^\perp(x-y)+U_\natural(u-v)| - |U_\natural(x-y)| - |U_\natural^\perp(u-v)| \\
&\overset{(66),(67)}{\geq} \frac{1}{\sqrt{2}}\Big(|U_\natural^\perp(x-y)| + |U_\natural(u-v)|\Big) \\
&\qquad - (2\sqrt{\varepsilon}+\varepsilon)|x-y| \\
&\overset{(68),(65)}{\geq} \left(\frac{1}{\sqrt{2}}(1-3\varepsilon) - 2\sqrt{\varepsilon} - \varepsilon\right)|x-y| + \frac{1}{\sqrt{2}}|u-v| \\
&\geq \frac{1}{4}|(x-y,u-v)|\,, \qquad (69)
\end{aligned}$$

since $\varepsilon \in (0, 1/100)$. So, part (ii) of Lemma 4.5 follows from (61), (63), and (69), which — as observed above — implies part (i) as well.

We now turn to the proof of part (iii). For $(x,v) \in N_\delta(\Sigma,\Phi)$ with $v \neq 0$, denote $u := (T_x\Sigma)_\natural^\perp v$ and note that by definition of N and the fact that Φ is an ε-normal map for Σ,

$$|u-v| = \big|\big(T_x\Sigma)_\natural^\perp - \Phi(x)\big)v\big| \leq \|\big(T_x\Sigma)_\natural^\perp - \Phi(x)\||v| < \varepsilon|v|. \qquad (70)$$

Since $\Sigma \in \mathscr{C}^{1,\alpha}_{m,n}(R,L,d)$, we find for any $y \in \Sigma$ with $|y-x| < 4\delta_{tub} \overset{(62)}{\leq} R$ as in (64)

$$\operatorname{dist}(y, x+T_x\Sigma) \leq L|y-x|^{1+\alpha} \leq \varepsilon|y-x|\,. \qquad (71)$$

On the other hand, if $y \in \mathbb{B}(x+u, \frac{1}{2}|v|)$, then

$$|y-x| \leq |y-(x+u)| + |u| \leq \tfrac{1}{2}|v| + \big|\big(T_x\Sigma\big)_\natural^\perp v\big| \leq \tfrac{3}{2}|v|, \qquad (72)$$

and by (70)

$$\begin{aligned}
\operatorname{dist}(y, x+T_x\Sigma) &= \big|\big(T_x\Sigma\big)_\natural^\perp (y-x)\big| \\
&\geq \big|\big(T_x\Sigma\big)_\natural^\perp u\big| - \big|\big(T_x\Sigma\big)_\natural^\perp (y-(x+u))\big| \\
&= |u| - \big|\big(T_x\Sigma\big)_\natural^\perp (y-(x+u))\big| \\
&\overset{(70)}{\geq} |v| - |u-v| - |y-(x+u)| \geq \big(\tfrac{1}{2}-\varepsilon\big)|v|\,. \qquad (73)
\end{aligned}$$

Combining (71), (72) and (73) would yield for $y \in \Sigma \cap \mathbb{B}(x+u, \frac{1}{2}|v|) \cap \mathbb{B}(x, 4\delta_{tub})$

$$\big(\tfrac{1}{2} - \varepsilon\big)|v| \overset{(73)}{\leq} |T_x\Sigma^{\perp}_{\natural}(y-x)| = \operatorname{dist}(y, x+T_x\Sigma) \overset{(71)}{\leq} \varepsilon|y-x| \overset{(72)}{\leq} \tfrac{3}{2}\varepsilon|v|$$

contradicting $\varepsilon \in (0, 1/100)$ because $|v| \neq 0$. Since $\mathbb{B}(x+u, \frac{1}{2}|v|) \subset \mathbb{B}(x, 4\delta_{tub})$ because $|u| \leq |v| < \delta < \delta_{tub}$, this can only mean that $\mathbb{B}(x+u, \frac{1}{2}|v|) \cap \Sigma = \emptyset$, which implies by (70)

$$\begin{aligned} \operatorname{dist}(\Psi(x,v), \Sigma) &\geq \operatorname{dist}(x+u, \Sigma) - |u-v| \\ &\geq \tfrac{1}{2}|v| - |u-v| \overset{(70)}{\geq} \big(\tfrac{1}{2} - \varepsilon\big)|v| > \tfrac{1}{4}|v|\,. \end{aligned}$$

Finally we prove part (iv). For $x \in \Sigma \in \mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$ there exists a function $f = f_x \in C^{1,\alpha}(T_x\Sigma, T_x\Sigma^{\perp})$ with $f(0) = 0$, $Df(0) = 0$, $\operatorname{Lip}(f) \leq 1$, and the Hölder condition for Df in Definition 1, such that by (62)

$$\mathbb{B}(x, 4\delta_{tub}) \cap \Sigma = (x + \operatorname{graph}(f)) \cap \mathbb{B}(x, 4\delta_{tub}).$$

For any $\xi \in T_x\Sigma \cap \mathbb{B}(0, 4\delta_{tub})$ one can use (14) and (62) to estimate

$$|f(\xi)| \overset{(14)}{\leq} L|\xi|^{1+\alpha} \leq L(4\delta_{tub})^{\alpha}|\xi| \leq \varepsilon|\xi|\,. \tag{74}$$

Again by (62) in combination with Definition 4.1 we have for $\zeta \in \mathbb{B}(x, 4\delta_{tub}) \cap \Sigma$

$$\begin{aligned} \|\Phi(\zeta) - \big(T_x\Sigma\big)^{\perp}_{\natural}\| &\leq \|\Phi(\zeta) - \Phi(x)\| + \|\Phi(x) - \big(T_x\Sigma\big)^{\perp}_{\natural}\| \\ &< \operatorname{Lip}(\Phi) 4\delta_{tub} + \varepsilon \leq 2\varepsilon\,. \end{aligned} \tag{75}$$

For fixed $\delta \in (0, \delta_{tub}]$ consider the C^1-functions

$$\begin{aligned} &\psi : \overline{\mathbb{B}}(0,\delta) \to \Sigma \quad \text{given by} \quad \psi(w) := x + (T_x\Sigma)_{\natural} w + f\big((T_x\Sigma)_{\natural} w\big)\,, \\ &F : \overline{\mathbb{B}}(0,\delta) \to N_{\delta}(\Sigma, \Phi) \quad \text{defined by} \quad F(z) := \big(\psi(z), \Phi(\psi(z))\big(T_x\Sigma\big)^{\perp}_{\natural} z\big)\,, \\ &\text{and} \quad G = \big(\Psi_{\delta}[\Sigma, \Phi] \circ F\big) - x : \overline{\mathbb{B}}(0,\delta) \to \mathbb{R}^n\,. \end{aligned}$$

Employing (74) and (75) we obtain for $z \in \mathbb{B}(0,\delta)$

$$\begin{aligned}|G(z)-z| &= |\psi(z)+\Phi(\psi(z))(T_x\Sigma)^\perp_\natural z - x - (T_x\Sigma)_\natural z - (T_x\Sigma)^\perp_\natural z| \\ &\le |\psi(z) - (T_x\Sigma)_\natural z - x| + |\Phi(\psi(z))(T_x\Sigma)^\perp_\natural z - (T_x\Sigma)^\perp_\natural z| \\ &\le |f\big((T_x\Sigma)_\natural z\big)| + \|\Phi(\psi(z)) - (T_x\Sigma)^\perp_\natural\||(T_x\Sigma)^\perp_\natural z| \\ &\overset{(74),(75)}{\le} \varepsilon|(T_x\Sigma)_\natural z| + 2\varepsilon|(T_x\Sigma)^\perp_\natural z| \le 2\sqrt{2}\varepsilon|z| < 2\sqrt{2}\varepsilon\delta,\end{aligned}$$

so that we can apply Proposition 2.5 to get

$$\overline{\mathbb{B}}(0,\delta/2) \subset \mathbb{B}(0,(1-2\sqrt{2}\varepsilon)\delta) \subset G(\overline{\mathbb{B}}(0,\delta)) = \Psi_\delta[\Sigma,\Phi](F(\overline{\mathbb{B}}(0,\delta))) - x;$$

hence

$$x + \mathbb{B}(0,\delta/2) \subset \Psi_\delta[\Sigma,\Phi](F(\overline{\mathbb{B}}(0,\delta))) \subset \Psi_\delta[\Sigma,\Phi](N_\delta(\Sigma,\Phi)). \qquad \square$$

Proposition 4.6. *Let $\theta \in [0,1]$, $\lambda,\gamma \in [0,1)$ and $k \in \{1,\dots,n-1\}$ and suppose that $W,T \in G(n,k)$ and $U,V \in G(n,n-k)$ satisfy $\measuredangle(W,T) \le \theta$, $\measuredangle(T,U^\perp) \le \lambda$, $\measuredangle(T,V^\perp) \le \lambda$, and $\measuredangle(U,V) \le \gamma$. Given any vectors $\mathbf{w} \in W$, $\mathbf{t} \in T$, $\mathbf{u} \in U$ and $\mathbf{v} \in V$ such that $\mathbf{u}+\mathbf{t} = \mathbf{w}+\mathbf{v}$ the following holds:*

$$\Big(|\mathbf{u}| - \tfrac{\theta}{1-\lambda}|\mathbf{w}|\Big)\Big(1 - \tfrac{\gamma}{1-\lambda}\Big) \le |\mathbf{v}| \le \Big(|\mathbf{u}| + \tfrac{\theta}{1-\lambda}|\mathbf{w}|\Big)\Big(1 + \tfrac{\gamma}{1-\lambda}\Big) \tag{76}$$

$$\text{and} \quad |\mathbf{u}-\mathbf{v}| \le \tfrac{\gamma}{1-\lambda}\Big(|\mathbf{u}| + \tfrac{\theta}{1-\lambda}|\mathbf{w}|\Big) + \tfrac{\theta}{1-\lambda}|\mathbf{w}|\,. \tag{77}$$

Proof. Let $P : \mathbb{R}^n \to U$ be the oblique projection onto U with $\ker P = T$ and set $\bar{\mathbf{u}} = P(\mathbf{u}-\mathbf{w}) \in U$, so that

$$\begin{aligned}\mathbf{w} + (\bar{\mathbf{u}} - \mathbf{u}) &= \mathbf{w} + (P(\mathbf{u}-\mathbf{w}) - \mathbf{u}) \\ &= \mathbf{w} - P\mathbf{w} + P\mathbf{u} - \mathbf{u} = \mathbf{w} - P\mathbf{w} \in \ker P = T.\end{aligned}$$

Thus we can apply Proposition 2.4 to $\mathbf{z} := \mathbf{u} - \bar{\mathbf{u}} \in U =: Z$ and $\mathbf{x} := \mathbf{w} \in W =: X$ (with $Y := T$ implying $\measuredangle(X,Y) \le \theta$, $\measuredangle(Y,Z^\perp) \le \lambda$, and $\mathbf{z}-\mathbf{x} \in Y$) to obtain $|\bar{\mathbf{u}} - \mathbf{u}| \le \theta|\mathbf{w}|/(1-\lambda)$ which directly leads to

$$|\mathbf{u}| - \frac{\theta}{1-\lambda}|\mathbf{w}| \le |\mathbf{u}| - |\bar{\mathbf{u}} - \mathbf{u}| \le |\bar{\mathbf{u}}| \le |\mathbf{u}| + |\bar{\mathbf{u}} - \mathbf{u}| \le |\mathbf{u}| + \frac{\theta}{1-\lambda}|\mathbf{w}|\,. \tag{78}$$

Applying Proposition 2.4 now to $\mathbf{x} := \bar{\mathbf{u}} \in U =: X$ and to

$$\mathbf{z} := \bar{\mathbf{u}} - \mathbf{v} = P(\mathbf{u}-\mathbf{w}) - \mathbf{v} = \mathbf{u} - \mathbf{v} - P\mathbf{w} = \mathbf{w} - \mathbf{t} - P\mathbf{w} \in T =: Z$$

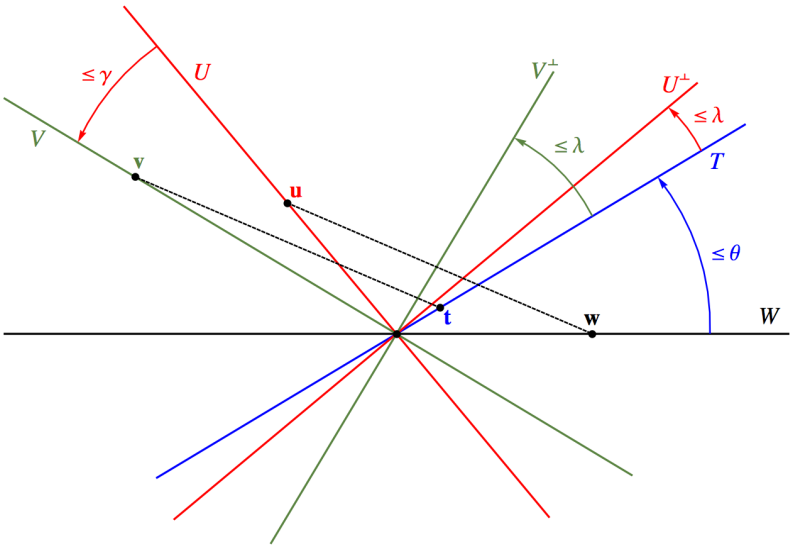

Figure 3: The situation in Proposition 4.6: the vectors $\mathbf{u}-\mathbf{w}$ and $\mathbf{v}-\mathbf{t}$ are equal.

(so that $\mathbf{z}-\mathbf{x}=-\mathbf{v}\in V=:Y$, and hence $\sphericalangle(X,Y)=\sphericalangle(U,V)\le\gamma$ and $\sphericalangle(Y,Z^\perp)=\sphericalangle(V,T^\perp)=\sphericalangle(V^\perp,T)\le\lambda$) to arrive at $|\bar{\mathbf{u}}-\mathbf{v}|\le\gamma|\bar{\mathbf{u}}|/(1-\lambda)$, and in consequence

$$|\bar{\mathbf{u}}|\big(1-\tfrac{\gamma}{1-\lambda}\big)\le|\bar{\mathbf{u}}|-|\bar{\mathbf{u}}-\mathbf{v}|\le|\mathbf{v}|\le|\bar{\mathbf{u}}|+|\mathbf{v}-\bar{\mathbf{u}}|\le|\bar{\mathbf{u}}|\big(1+\tfrac{\gamma}{1-\lambda}\big)\,.$$

This together with (78) gives the first part of the proposition. To get the second we use (78) to write

$$\begin{aligned}|\mathbf{u}-\mathbf{v}|&\le|\bar{\mathbf{u}}-\mathbf{v}|+|\bar{\mathbf{u}}-\mathbf{u}|\\&\le\tfrac{\theta}{1-\lambda}|\mathbf{w}|+\tfrac{\gamma}{1-\lambda}|\bar{\mathbf{u}}|\overset{(78)}{\le}\tfrac{\theta}{1-\lambda}|\mathbf{w}|+\tfrac{\gamma}{1-\lambda}\big(|\mathbf{u}|+\tfrac{\theta}{1-\lambda}|\mathbf{w}|\big)\,.\end{aligned}$$

□

Definition 4.7. For $t\in\mathbb{R}$ we define the continuous map

$$m_t:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^n\times\mathbb{R}^n\,,\quad m_t(x,v):=(x,tv)\,.$$

Lemma 4.8 (bilipschitz diffeomorphisms). *For $R, L, d \in (0, \infty)$, $\alpha \in (0, 1]$, $\varepsilon \in (0, 10^{-2})$ let $\Sigma_1, \Sigma_2 \in \mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$ with $\rho := d_{\mathcal{H}}(\Sigma_1, \Sigma_2) < \delta_{tub}/8$, and with ε-normal map $\Phi_1 : \Sigma_1 \to \mathcal{G}$ for Σ_1, where*

$$\delta_{tub} = \delta_{tub}(R, L, \alpha, \varepsilon, \mathrm{Lip}(\Phi_1))$$

is the radius of the tubular neighbourhood of Σ_1 established in Lemma 4.5. Set $\Psi := \Psi_{\delta_{tub}}[\Sigma_1, \Phi_1]$ and define

$$\begin{aligned} F &: \Sigma_2 \to \mathbb{R}^n \quad \text{by} \quad F := \Psi \circ m_0 \circ \Psi^{-1}|_{\Sigma_2}\,, \\ G &: \Sigma_2 \to \mathbb{R}^n \quad \text{by} \quad G := F - \mathrm{Id}\,. \end{aligned}$$

Then $\Sigma_1 \subset \mathrm{im}(F)$, and there exist $C_l = C_l(L, \mathrm{Lip}(\Phi_1)) \geq 1$ and $\rho_G = \rho_G(R, L, \alpha, \varepsilon, \mathrm{Lip}(\Phi_1)) \in (0, \delta_{tub}/8]$ such that for all $\rho = d_{\mathcal{H}}(\Sigma_1, \Sigma_2) \in (0, \rho_G)$

(i) $\mathrm{Lip}(G) \leq C_l \rho^{\alpha/2}$,

(ii) $|G(x)| \leq 4\,\mathrm{dist}(x, \Sigma_1)$ *for all* $x \in \Sigma_2$,

(iii) F is a bilipschitz diffeomorphism onto its image Σ_1 satisfying

$$\big(1 - C_l \rho^{\alpha/2}\big)|x - y| \leq |F(x) - F(y)| \leq \big(1 + C_l \rho^{\alpha/2}\big)|x - y| \quad \forall x, y \in \Sigma_2\,.$$

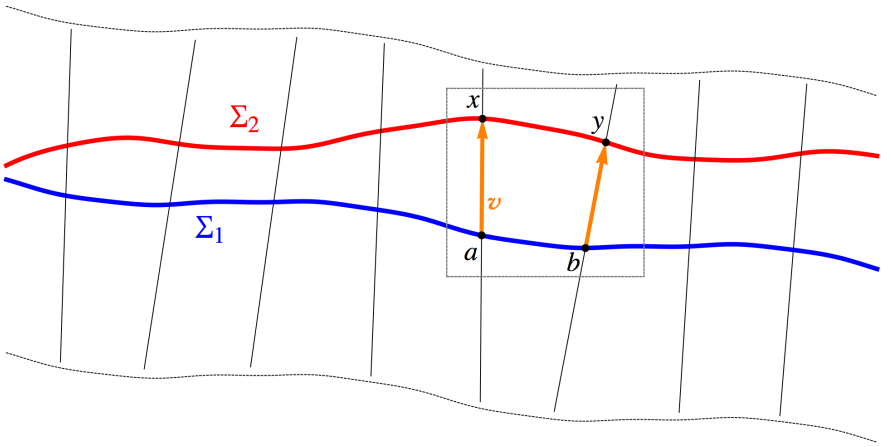

Figure 4: The definition of $F\colon \Sigma_2 \to \Sigma_1$. Thin nearly vertical lines represent ε-normal spaces to Σ_1. We have $x = \Psi(a, v)$, so that $\Psi^{-1}|_{\Sigma_2}(x) = (a, v) \in \mathbb{R}^{2n}$. Next, $m_0(a, v) = (a, 0)$, and $\Psi(a, 0) = a + 0 = a$. This yields $F(x) = a$.

Proof. Notice that Ψ^{-1} is well-defined on a neighbourhood of Σ_2 by virtue of Lemma 4.5 (iv), since $d_{\mathcal{H}}(\Sigma_1, \Sigma_2) < \delta_{tub}/8$, so that for $x \in \Sigma_2$ we find a unique pair $(\xi, v) \in N_{\delta_{tub}}(\Sigma_1, \Phi_1)$ such that $x = \xi + v = \Psi(\xi, v)$.

By definition of the map F it is clear that $\mathrm{im}(F) \subset \Sigma_1$, however, the converse $\Sigma_1 \subset \mathrm{im}(F)$ is not so obvious. To establish that we use a topological argument by means of the degree mod 2 as follows. For the l-plane $P \in G(n, l)$ denote the $l-1$-dimensional sphere

$$\mathbb{S}^{l-1}(\xi, r, P) := \xi + \{v \in P : |v| = r\} \quad \text{for } \xi \in \mathbb{R}^n,\ r > 0,$$

and observe that for $\xi \in \Sigma_1$, $r \in (0, \delta_{tub})$,

$$\mathbb{S}^{n-m-1}(\xi, r, \mathrm{im}(\Phi_1(\xi))) = \Psi(\xi, \mathrm{im}(\Phi_1(\xi))) \cap \partial\mathbb{B}(0, r),$$

so that by virtue of Lemma 4.5 (iii) $\mathbb{S}^{n-m-1}(\xi, r, \mathrm{im}(\Phi_1(\xi))) \cap \Sigma_1 = \emptyset$ and, in addition, $\mathbb{S}^{n-m-1}(\xi, r, \mathrm{im}(\Phi_1(\xi)))$ and Σ_1 are nontrivially linked for all $r \in (0, \delta_{tub})$, that is, the map

$$\Sigma_1 \times \mathbb{S}^{n-m-1}(\xi, r, \mathrm{im}(\Phi_1(\xi))) \ni (w, z) \mapsto \frac{w - z}{|w - z|} \in \mathbb{S}^{n-1}$$

has non-vanishing degree mod 2, for each $\xi \in \Sigma_1$ and $r \in (0, \delta_{tub})$, since Σ_1 is a compact m-dimensional C^1-submanifold without boundary. Since $d_{\mathcal{H}}(\Sigma_1, \Sigma_2) < \delta_{tub}/8$ also Σ_2 and $\mathbb{S}^{n-m-1}(\xi, r, \mathrm{im}(\Phi_1(\xi)))$ are non-trivially linked for each $\xi \in \Sigma_1$ and for all $r \in (\delta_{tub}/2, \delta_{tub})$, because $d_{\mathcal{H}}(\Sigma_1, \Sigma_2) < d_{\mathcal{H}}(\Sigma_1, \mathbb{S}^{n-m-1}(\xi, r, \mathrm{im}(\Phi_1(\xi))))$, again by virtue of Lemma 4.5 (iii). Therefore, each $n-m$-dimensional disk

$$\begin{aligned} \mathbb{D}^{n-m}(\xi, r, \mathrm{im}(\Phi_1(\xi))) &:= \xi + \{v \in \mathrm{im}(\Phi_1(\xi)) : |v| \le r\} \\ &= \mathrm{im}\left(\Psi(\xi, \mathrm{im}(\Phi_1(\xi)) \cap \mathbb{B}(0, r))\right) \end{aligned} \tag{79}$$

for $\xi \in \Sigma_1$ and $r \in (\delta_{tub}/2, \delta_{tub})$ contains at least one point of Σ_2; see [25, Lemma 3.5]. Take for fixed $\xi \in \Sigma_1$ and $r = 3\delta_{tub}/4$ one of those points

$$z \in \Sigma_2 \cap \mathbb{D}^{n-m}(\xi, 3\delta_{tub}/4, \mathrm{im}(\Phi_1(\xi))),$$

and use (79) to express z as $z = \xi + v = \Psi(\xi, v)$ for some $v \in \mathrm{im}(\Phi_1(\xi))$ with $|v| < 3\delta_{tub}/4$ to find

$$F(z) = \Psi \circ m_0 \circ \Psi^{-1}(z) = \Psi \circ m_0(\xi, v) = \Psi(\xi, 0) = \xi,$$

which establishes $\Sigma_1 \subset \mathrm{im}(F)$.

Remark. One can also prove that $\Sigma_1 \subseteq \operatorname{im} F$ later, right after proving that F is bilipschitz (i.e. after proving that G is Lipschitz): once this is established, F is a $C^{1,\alpha}$ diffeomorphism onto its image. Thus, the image of F is a submanifold of $\mathbb{R}^n$ — actually it is a submanifold of Σ_1, of the same dimension as Σ_1. Hence, it is open in Σ_1, which can be seen using a local graph representation; it is also closed in Σ_1 as a continuous image of a compact set. Therefore, $\Sigma_1 \sim \operatorname{im} F$ is a connected component of Σ_1; assuming $\Sigma_1 \sim \operatorname{im} F$ is not empty and using the definition of $\mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$ one sees that $\Sigma_1 \sim \operatorname{im} F$ is at least R away from $\operatorname{im} F$ which contradicts the assumption $d_{\mathcal{H}}(\Sigma_1, \Sigma_2) \leq \delta_{tub}$.

Proof of (i): If $\rho = d_{\mathcal{H}}(\Sigma_1, \Sigma_2) = 0$ then $\Sigma_1 = \Sigma_2$ and $F = \mathrm{Id}$, and there is nothing to prove. Assume $\rho > 0$ from now on. Let $x, y \in \Sigma_2$ and set

$$a := F(x)\,, \quad b := F(y)\,, \quad X = \Phi_1(a)\,, \quad Y = \Phi_1(b)\,.$$

Observe that, by Lemma 4.5(iii), $|x - a| \leq 4 \operatorname{dist}(\Psi(a, x - a), \Sigma_1) \leq 4\rho$ for $\Psi(a, x - a) = x \in \Sigma_1$, and in the same way, $|y - b| \leq 4\rho$, so that we infer immediately

$$\begin{aligned} (80)\quad |G(x) - G(y)| &\leq |a - x| + |b - y| \\ &\leq 8\rho \leq 8\sqrt{\rho}|x - y| \quad \text{for all } x, y \in \Sigma_2 \text{ with } |x - y| \geq \sqrt{\rho}\,. \end{aligned}$$

Assume now that $x, y \in \Sigma_2$ satisfy $|x - y| < \sqrt{\rho}$. Note that by Lemma 4.5(ii)

$$\begin{aligned} (81)\qquad \operatorname{Lip}(F) &\leq \operatorname{Lip}(\Psi) \operatorname{Lip}(m_0) \operatorname{Lip}(\Psi^{-1}) \leq 4\sqrt{2}\,; \\ \text{hence,}\quad |a - b| &= |F(x) - F(y)| \leq 4\sqrt{2}|x - y| < 4\sqrt{2\rho}\,. \end{aligned}$$

Set

$$\rho_0 := \min\left\{2^{-3}\delta_{tub},\, (2L)^{-2/\alpha},\, 2^{-6}R^2,\, 2^{-12/\alpha} C_{\mathrm{ang}}(L, 4)^{-2/\alpha},\, 2^{-9} \operatorname{Lip}(\Phi_1)^{-2}\right\}.$$

As $\delta_{tub} \leq 1$, cf. (62) in the proof of Lemma 4.5, we have $\rho_0 \leq \frac{1}{8}$. If we require

$$(82)\qquad d_{\mathcal{H}}(\Sigma_1, \Sigma_2) = \rho < \rho_0\,,$$

then, as $|a - x| < 4\rho$, we can use Lemma 3.1 with $A := 4$ to write

$$(83)\qquad \measuredangle(T_a\Sigma_1, T_x\Sigma_2) \leq C_{\mathrm{ang}}(L, 4)\rho^{\alpha/2}\,.$$

Moreover, by (81) and the choice of ρ_0 above,

$$\begin{aligned}(84)\qquad \measuredangle(X,Y) &= \|\Phi_1(a)-\Phi_1(b)\| \le \mathrm{Lip}(\Phi_1)|a-b| \\ &\le 4\sqrt{2}\,\mathrm{Lip}(\Phi_1)|x-y| \le 4\sqrt{2\rho}\,\mathrm{Lip}(\Phi_1) < \tfrac{1}{4}\,.\end{aligned}$$

Thus, by (84), (83), and our choice of ρ_0

$$\begin{aligned}(85)\qquad \measuredangle(Y^\perp, T_x\Sigma_2) &\le \measuredangle(Y^\perp, X^\perp) + \measuredangle(X^\perp, T_a\Sigma_1) + \measuredangle(T_a\Sigma_1, T_x\Sigma_2) \\ &\overset{(84)}{\le} 4\sqrt{2\rho}\,\mathrm{Lip}(\Phi_1) + \|\Phi_1(a) - (T_a\Sigma_1)^\perp_\natural\| + \measuredangle(T_a\Sigma_1, T_x\Sigma_2) \\ &\overset{(83)}{\le} 4\sqrt{2\rho}\,\mathrm{Lip}(\Phi_1) + \varepsilon + C_{\mathrm{ang}}(L,4)\rho^{\alpha/2} < \tfrac{1}{2}\,.\end{aligned}$$

Similarly,

$$\begin{aligned}(86)\qquad \measuredangle(Y^\perp, T_a\Sigma_1) &\le \measuredangle(Y^\perp, X^\perp) + \measuredangle(X^\perp, T_a\Sigma_1) \\ &\overset{(84)}{\le} 4\sqrt{2\rho}\,\mathrm{Lip}(\Phi_1) + \varepsilon < \tfrac{1}{2}\,.\end{aligned}$$

These angle conditions imply by Lemma 2.2 that

$$Y \cap T_a\Sigma_1 = Y \cap T_x\Sigma_2 = \{0\},$$

Therefore, there exist points $\bar{b}$, $\bar{y} \in \mathbb{R}^n$ (see Figure 5) such that

$$(87)\qquad (y+Y)\cap(a+T_a\Sigma_1) = \{\bar{b}\}\,, \quad \text{and} \quad (y+Y)\cap(x+T_x\Sigma_2) = \{\bar{y}\}\,.$$

Indeed, the characterisation $\{P(y-a)\} = ((y-a)+Y)\cap T_a\Sigma_1$ of the well-defined oblique projection $P:\mathbb{R}^n \to T_a\Sigma_1$ along Y (see Remark 2.3) immediately gives $\bar{b} := P(y-a)+a$, and similarly one finds $\bar{y}$.

To prove that G is Lipschitz we need to estimate $|G(x)-G(y)| = |(x-a)-(y-b)|$. To this end, we shall first estimate $|(x-a)-(\bar{y}-\bar{b})|$ treating $|(\bar{y}-\bar{b})-(y-b)| \le |\bar{b}-b| + |\bar{y}-y|$ as a small error term. Employing (77) of Proposition 4.6 with

$$\begin{gathered} U := Y, \quad V := X, \quad W := T_x\Sigma_2, \quad T := T_a\Sigma_1, \\ \mathbf{u} := \bar{y}-\bar{b}, \quad \mathbf{v} := x-a, \quad \mathbf{w} := \bar{y}-x, \quad \mathbf{t} := \bar{b}-a, \end{gathered}$$

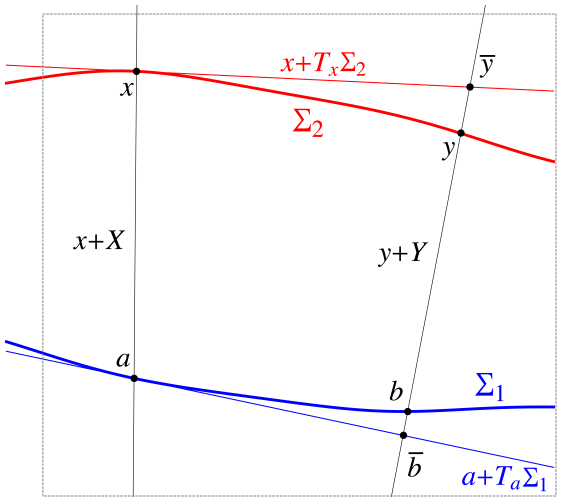

Figure 5: Enlarged fragment of FIGURE 4. We have $G(x) = a - x$, $G(y) = b - y$. However, to prove that G is Lipschitz, we do not deal with $(a - x) - (b-y)$ directly. Instead, we use Proposition 4.6 to estimate $|(a-x)-(\bar{b}-\bar{y})|$, and add an error term $|b - \bar{b}| + |y - \bar{y}|$, which is small as both Σ_1 and Σ_2 are of class $\mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$. The final dependence of $\mathrm{Lip}(G)$ on a power of ρ is due to this error term.

in combination with (83), (84), (85), (86) to estimate the angles by our choice of ρ_0

$$\begin{aligned}
&\measuredangle(W,T) \le \theta := C_{\mathrm{ang}}(L,4)\rho^{\alpha/2}, \qquad \measuredangle(U,V) \le \gamma := 4\sqrt{2}\,\mathrm{Lip}(\Phi_1)|x-y|,\\
&\max\left\{\measuredangle(T,U^\perp), \measuredangle(T,V^\perp)\right\} \le \lambda := \frac{1}{2},
\end{aligned}$$

we obtain

$$\begin{aligned}
(88)\qquad &|(x-a)-(\bar{y}-\bar{b})|\\
&\overset{(77)}{\le} 8\sqrt{2}\,\mathrm{Lip}(\Phi_1)|x-y|\left(|\bar{y}-\bar{b}| + 2C_{\mathrm{ang}}(L,4)\rho^{\alpha/2}|\bar{y}-x|\right)\\
&\quad + 2C_{\mathrm{ang}}(L,4)\rho^{\alpha/2}|\bar{y}-x|\,.
\end{aligned}$$

By (81) and the choice of ρ_0 we have

$$\begin{aligned}
&|a-b| < 4\sqrt{(2\rho)} < 4(2^{-5}R^2)^{1/2} < R\,,\\
&|a-b|^{1+\alpha} \le (4\sqrt{2})^2|x-y|^{1+\alpha} = 2^5|x-y|^{1+\alpha}\,.
\end{aligned}$$

Thus, since $\bar{y} - y \in Y$ and $\bar{b} - b = (\bar{b} - y) + (y - b) \in Y$ (we use (87) and note that $y = \Psi(b, y - b)$ with $y - b \in \Phi_1(b) = Y$), it follows from (81), (85) and (86) that

$$\begin{aligned} &|\bar{y} - y| \le 2\,\mathrm{dist}(y, x + T_x\Sigma_2) \le 2L|x - y|^{1+\alpha} \\ \text{and}\quad &|\bar{b} - b| \le 2\,\mathrm{dist}(b, a + T_a\Sigma_1) \le 2L|a - b|^{1+\alpha} \le 2^6 L|x - y|^{1+\alpha}\,, \end{aligned} \tag{89}$$

where we estimated similarly as in (17). Hence, since we observed $|y - b| \le 4\rho$ earlier, and $|x - y| < \sqrt{\rho} \le \frac{1}{\sqrt{8}} < 1$, we obtain

$$\begin{aligned} |\bar{y} - \bar{b}| &\le |\bar{y} - y| + |y - b| + |\bar{b} - b| \le 2^7 L|x - y|^{1+\alpha} + 4\rho < (2^7 L + 4)\rho^{\alpha/2}\,, \\ |\bar{y} - x| &\le |\bar{y} - y| + |y - x| \le |y - x|\left(1 + 2L\rho^{\alpha/2}\right) \le 2|x - y| < 2\sqrt{\rho} < 1\,. \end{aligned}$$

Therefore, plugging these two estimates into (88), and adding the error $|\bar{y} - y| + |\bar{b} - b|$ which can be estimated by (89), we compute

$$\begin{aligned} &|G(x) - G(y)| \\ &= |(x - a) - (y - b)| \le |(x - a) - (\bar{y} - \bar{b})| + |\bar{y} - y| + |\bar{b} - b| \\ &\overset{(88),(89)}{\le} |x - y|\rho^{\alpha/2}\Big\{8\sqrt{2}\,\mathrm{Lip}(\Phi_1)\big[2^7 L + 4 + 2C_{\mathrm{ang}}(L, 4)\big] \\ &\qquad\qquad + 4C_{\mathrm{ang}}(L, 4) + 2L + 2^6 L\Big\}\,. \end{aligned}$$

As $\rho < 1$, taking into account $C_{\mathrm{ang}}(L, 4) = 257L + 8$ (cf. Lemma 3.1), we finally obtain an estimate of the Lipschitz constant of G,

$$\begin{aligned} &|G(x) - G(y)| \le C_l\rho^{\alpha/2}|x - y|, \\ &C_l = C_l(L, \mathrm{Lip}(\Phi_1)) := 10^4(\mathrm{Lip}(\Phi_1) + 1)(L + 1)\,. \end{aligned} \tag{90}$$

Proof of (ii): Directly from the definition of Ψ we infer $\Psi(F(x), x - F(x)) = F(x) + x - F(x) = x$ for any $x \in \Sigma_2$, so that we obtain from Lemma 4.5(iii)

$$\begin{aligned} \mathrm{dist}(x, \Sigma_1) &= \mathrm{dist}(\Psi(F(x), x - F(x)), \Sigma_1) \\ &\overset{L.4.5(iii)}{\ge} \tfrac{1}{4}|x - F(x)| = \tfrac{1}{4}|G(x)| \quad \forall x \in \Sigma_2\,. \end{aligned}$$

Proof of (iii): Since F is a composition of C^1-smooth functions it is C^1-smooth. We can find $\rho_G = \rho_G(R, L, \alpha, \mathrm{Lip}(\Phi_1)) \in (0, \rho_0)$ so small that

$$C_l\rho^{\alpha/2} < 1 \qquad \text{for all } \rho \in (0, \rho_G), \tag{91}$$

and then

$$\begin{aligned}|F(x)-F(y)| &\le |x-y| + |F(x)-x-(F(y)-y)| \\ &= |x-y| + |G(x)-G(y)| \overset{(90)}{\le} \big(1 + C_l \rho^{\alpha/2}\big)|x-y|.\end{aligned}$$

The lower estimate in (iii) follows in the same manner; hence F is bilipschitz and, in consequence, a diffeomorphism. □

As a corollary we can can establish a bound on the Hausdorff-distance $d_{\mathcal{H}}(\Sigma_1, \Sigma_2)$ under which two submanifolds $\Sigma_1, \Sigma_2 \in \mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$ are actually ambient isotopic. Moreover, in Lemma 4.10 we construct a global diffeomorphism of the ambient space mapping Σ_2 onto Σ_1. Both results will be essential ingredients in the proof of Theorem 4.

Corollary 4.9 (ambient isotopies). *For $R, L, d \in (0, \infty)$, $\alpha \in (0, 1]$, and $\varepsilon \in (0, 1/100)$ let $\Sigma_1, \Sigma_2 \in \mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$ with ε-normal map Φ_1 for Σ_1, such that $d_{\mathcal{H}}(\Sigma_1, \Sigma_2) \in (0, \rho_G)$, where $\rho_G = \rho_G(R, L, \alpha, \varepsilon, \mathrm{Lip}(\Phi_1))$ is the constant of Lemma 4.8. Then Σ_1 and Σ_2 are C^1-ambient isotopic.*

Proof. According to [5, Theorem 1.2] it suffices to come up with a C^1-isotopy $h : \Sigma_2 \times [0, 1] \to \mathbb{R}^n$, i.e., a family of C^1-embeddings $h_t(\cdot) := h(\cdot, t) : \Sigma_2 \to \mathbb{R}^n$, with

$$\Sigma_1 = h(\Sigma_2 \times \{0\}) \quad \text{and} \quad h(\Sigma_2 \times \{1\}) = \Sigma_2. \tag{92}$$

Indeed, the map $h(x, t) := \Psi \circ m_t \circ \Psi^{-1}|_{\Sigma_2}(x)$ for $(x, t) \in \Sigma_2 \times [0, 1]$, and with $m_t(y, v) = (y, tv)$ and $\Psi := \Psi_{\delta_{tub}}[\Sigma_1, \Phi_1]$ for $y, v \in \mathbb{R}^n$ will do. Here δ_{tub} is the constant from Lemma 4.5 defined in (62), and Φ_1 is an ε-normal map for Σ_1.

Observe that part (iv) of Lemma 4.5 implies that $\Sigma_2 \subset \Psi(N_{\delta_{tub}}(\Sigma_1, \Phi_1))$, since we have $d_{\mathcal{H}}(\Sigma_1, \Sigma_2) < \rho_G < \delta_{tub}/8$ (see Lemma 4.8). Therefore, Ψ^{-1} is a well-defined C^1-map in an open neighbourhood of Σ_2, which implies that h itself as a composition of C^1-maps is of class C^1. With Lemma 4.8 (iii) we obtain $h(\Sigma_2 \times \{0\}) = F(\Sigma_2) = \Sigma_1$, and $h(\cdot, 0)$ is a bilipschitz diffeomorphism from Σ_2 onto Σ_1. Moreover, one immediately sees that $h(x, 1) = \Psi \circ m_1 \circ \Psi^{-1}(x) = x$ for all $x \in \Sigma_2$ by the very definition of m_t for $t = 1$, so that $h(\Sigma_2 \times \{1\}) = \Sigma_2$, which proves (92).

So, it remains to be shown that $h(\cdot, t) : \Sigma \to \mathbb{R}^n$ is an embedding for each $t \in (0, 1)$. Note that $\Psi : N_\delta(\Sigma_1, \Phi_1) \to \mathbb{R}^n$ is bilipschitz for all $\delta \in (0, \delta_{tub}]$ by Lemma 4.5 (ii), and hence so is Ψ^{-1} on $\Psi(N_{\delta_{tub}}(\Sigma_1, \Phi_1))$. In addition,

m_t is bilipschitz for $t \in (0,1)$, and $m_t(N_\delta(\Sigma_1,\Phi_1)) \subset N_\delta(\Sigma_1,\Phi_1)$ for all $t \in [0,1]$, $\delta \in (0,\delta_{tub}]$. Recall again from Lemma 4.5 (iv) that $d_{\mathcal{H}}(\Sigma_1,\Sigma_2) < \rho_G < \delta_{tub}/8$ implies that $\Sigma_2 \subset \Psi(N_{2d_{\mathcal{H}}(\Sigma_1,\Sigma_2)}(\Sigma_1,\Phi_1))$, so that $\Psi^{-1}|_{\Sigma_2}$ is just the restriction of a C^1-bilipschitz map, and in consequence $h(\cdot,t)$ is bilipschitz and C^1-smooth, and therefore a C^1-diffeomorphism onto its image $h(\Sigma_2,t)$ for each $t \in [0,1]$. Consequently, $h(\cdot,t) : \Sigma_2 \to \mathbb{R}^n$ is an embedding for each $t \in [0,1]$. □

Lemma 4.10 (diffeomorphisms of the ambient space). *For $R, L, d > 0$, $\alpha \in (0,1]$ there exist constants $\rho_g := \rho_g(R,L,\alpha,n,m)$ and $C_J = C_J(R,L,\alpha,n,m) \geq 0$, such that for any two manifolds $\Sigma_1,\Sigma_2 \in \mathscr{C}^{1,\alpha}_{m,n}(R,L,d)$ with $\rho := d_{\mathcal{H}}(\Sigma_1,\Sigma_2) \in (0,\rho_g]$ there exists a bilipschitz C^1-diffeomorphism $J : \mathbb{R}^n \to \mathbb{R}^n$ satisfying*

1) $J(\Sigma_2) = \Sigma_1$,
2) $J(x) = x$ *for* $x \in \mathbb{R}^n \sim (\Sigma_2 + \mathbb{B}(0,\rho_g))$,
3) $(1 - C_J\rho^{\alpha/2})|z_1 - z_2| \leq |J(z_1) - J(z_2)| \leq (1 + C_J\rho^{\alpha/2})|z_1 - z_2|$ *for all* $z_1, z_2 \in \mathbb{R}^n$.

The constant ρ_G was introduced in Lemma 4.8.

Proof. Set $\varepsilon := 1/200$. Lemma 4.2 guarantees the existence of ε-normal maps $\Phi_i : \Sigma_i \to \mathcal{G}$ for Σ_i, $i = 1,2$. Define $\Psi_2 = \Psi_{\delta_{tub}}[\Sigma_2,\Phi_2]$ as in Definition 4.4.

Choose $\rho_0 = \rho_0(R,L,\alpha,\mathrm{Lip}(\Phi_1),\mathrm{Lip}(\Phi_2)) \in (0,\min\{\delta_{tub}/16,\rho_G/2\})$ so small that

$$4C_l\rho^{\alpha/2} < \varepsilon = \frac{1}{200} \quad \text{for all } \rho \in (0,\rho_0], \tag{93}$$

where we denote by $\delta_{tub} = \delta_{tub}(R,L,\alpha,\mathrm{Lip}(\Phi_2))$ the tubular radius for Σ_2 established in Lemma 4.5 for our fixed $\varepsilon = 1/200$. Moreover, $\rho_G = \rho_G(R,L,\alpha,\mathrm{Lip}(\Phi_1))$ and $C_l = C_l(L,\mathrm{Lip}(\Phi_1))$ are the constants estimating the maps $F, G : \Sigma_2 \to \mathbb{R}^n$ in Lemma 4.8 for $\varepsilon = 1/200$. Consider the projections $\pi_1,\pi_2 : N_2 := N_{\delta_{tub}/2}(\Sigma_2,\Phi_2) \to \mathbb{R}^n$ via $\pi_1(x,v) := x$ and $\pi_2(x,v) := v$ for $(x,v) \in N_2$, define the map $\lambda : N_2 \to \mathbb{R}^n$ by $\lambda(x,v) := F(x) + v$, and finally,

$$\tilde{J} : \Sigma_2 + \mathbb{B}(0,\rho_0) \to \mathbb{R}^n, \quad \tilde{J} = \lambda \circ \Psi_2^{-1}$$
$$\text{and the map} \quad I : \Sigma_2 + \mathbb{B}(0,\rho_0) \to \mathbb{R}^n, \quad I(z) = \tilde{J}(z) - z.$$

measuring the deviation of $\tilde{J}$ from the identity. According to Lemma 4.8(ii) one has for $x \in \Sigma_2$

$$\begin{aligned} |G(x)| = |F(x) - x| &\leq 4\,\mathrm{dist}(x, \Sigma_1) \\ &\leq 4 d_{\mathcal{H}}(\Sigma_2, \Sigma_1) = 4\rho \leq 4\rho_0 < \delta_{tub}/4 \quad \text{for all } \rho \in (0, \rho_0]\,, \end{aligned} \tag{94}$$

and therefore, for $z \in \Sigma_2 + \mathbb{B}(0, \rho_0)$ with $\Psi_2^{-1}(z) = (x, v)$,

$$\begin{aligned} I(z) = \tilde{J}(z) - z &= (F(x) + v) - (x + v) \\ &= F(x) - x = G(x) = G \circ \pi_1 \circ \Psi_2^{-1}(z)\,, \end{aligned} \tag{95}$$

so

$$|I(z)| \leq 4\rho < \delta_{tub}/4, \tag{96}$$

whence $\tilde{J}(\Sigma_2 + \mathbb{B}(0, \rho_0)) \subset \Sigma_2 + \mathbb{B}(0, \delta_{tub}/2)$. The identity (95) together with Lemma 4.5(ii) applied to Σ_2 and Ψ_2 and Lemma 4.8(i) implies

$$\mathrm{Lip}(I) \leq \mathrm{Lip}(G)\,\mathrm{Lip}(\Psi_2^{-1}) \leq 4 C_l \rho^{\alpha/2}\,. \tag{97}$$

Thus, we can estimate the difference $\tilde{J}(z_1) - \tilde{J}(z_2) = I(z_1) - I(z_2) + z_1 - z_2$ using (95) for $z_1, z_2 \in \Sigma_2 + \mathbb{B}(0, \rho_0)$ as

$$(1 - 4C_l\rho^{\alpha/2})|z_1 - z_2| \leq |\tilde{J}(z_1) - \tilde{J}(z_2)| \leq (1 + 4C_l\rho^{\alpha/2})|z_1 - z_2|\,, \tag{98}$$

so that by our choice of ρ_0 in (93), $\tilde{J}$ turns out to be bilipschitz, and since Ψ_2^{-1} is C^1 on $\Psi_2(N_2)$ and

$$\Sigma_2 + \mathbb{B}(0, \rho_0) \subset \Sigma_2 + \mathbb{B}(0, \delta_{tub}/4) \overset{Lem.4.5}{\subset} \Psi_2(N_2),$$

and λ is C^1 on N_2, the map $\tilde{J}$ is a C^1-diffeomorphism from $\Sigma_2 + \mathbb{B}(0, \rho_0)$ onto its image. Note that this image $\tilde{J}(\Sigma_2 + \mathbb{B}(0, \rho_0))$ contains Σ_1, since by Lemma 4.8, F maps Σ_2 diffeomorphically onto Σ_1. In particular, for any $\xi \in \Sigma_1$ there is exactly one $x \in \Sigma_2$ such that $F(x) = \xi$, so that for $z = x + 0$ one has $\tilde{J}(z) = \lambda(x, 0) = F(x) + 0 = \xi$. This also shows that $\tilde{J}(\Sigma_2) = \Sigma_1$.

To construct the global diffeomorphism we smoothly extend $\tilde{J}$ to all of $\mathbb{R}^n$ by the identity in the following way. Let $\phi \in C^1(\mathbb{R})$ be a cut-off function satisfying $0 \leq \phi \leq 1$ on $\mathbb{R}$, $\phi(t) = 0$ for $t \leq \rho_0/8$, $\phi(t) = 1$ for $t \geq \rho_0/4$, and $|\phi'(t)| \leq 16/\rho_0$ for all $t \in \mathbb{R}$. Define $\eta : \Sigma_2 + \mathbb{B}(0, \rho_0) \to \mathbb{R}$ as $\eta(z) := |\pi_2 \circ \Psi_2^{-1}(z)|$, and the transition term $T : \Sigma_2 + \mathbb{B}(0, \rho_0) \to \mathbb{R}^n$ as $T(z) :=$

$\phi(\eta(z))I(z)$, which is of class C^1 since $\phi(\eta(z))$ vanishes for $0 \le \eta(z) \le \rho_0/8$. In addition, we can estimate the Lipschitz constant of the transition term using Lemma 4.5(ii) for Σ_2 and Ψ_2, (96), and (97) as

$$\begin{aligned}(99)\qquad \operatorname{Lip}(T) &\le \operatorname{Lip}(\phi\circ\eta)\|I\|_\infty + \|\phi\circ\eta\|_\infty \operatorname{Lip}(I)\\ &\le \frac{16^2}{\rho_0}\rho + 4C_l\rho^{\alpha/2} \le C_T\rho^{\alpha/2}, \quad \rho\in(0,\rho_0]\end{aligned}$$

for $C_T := 16^2/\rho_0^{\alpha/2} + 4C_l$. The global diffeomorphism $J : \mathbb{R}^n \to \mathbb{R}^n$ can now be defined as

$$J(z) := \begin{cases} \tilde J(z) - T(z) & \text{for } z \in \Sigma_2 + \mathbb{B}(0,\rho_0)\\ z & \text{otherwise}\,,\end{cases}$$

which is of class C^1 since $T(z) = I(z)$ (and hence $\tilde J(z) - T(z) = z$) if z is contained in the transition zone $\Sigma_2 + \mathbb{B}(0,\rho_0) \sim \mathbb{B}(0,\rho_0/2)$. Indeed, then $z = x + v$ for some $(x,v) \in N_2$ satisfying, by Lemma 4.5(ii),

$$\begin{aligned}\frac{\rho_0}{2} \le \operatorname{dist}(z,\Sigma_2) &= \operatorname{dist}(\Psi_2(x,v),\Sigma_2)\\ &\le |\Psi_2(x,v) - \Psi_2(x,0)| \le \sqrt{2}|(x-x, v-0)| = \sqrt{2}|v|,\end{aligned}$$

so that $|v| \ge \rho_0/(2\sqrt{2}) > \rho_0/4$, from which $\eta(z) = |\pi_2 \circ \Psi_2^{-1}(z)| = |\pi_2(x,v)| = |v| > \rho_0/4$ follows, and thus $\phi(\eta(z)) = 1$ for such z in the transition zone. Combining (98) with (99) we arrive at the desired bilipschitz estimate

$$\begin{aligned}\big(1 - (4C_l + C_T)\rho^{\alpha/2}\big)|z_1 - z_2| &\le |J(z_1) - J(z_2)|\\ &\le \big(1 + (4C_l + C_T)\rho^{\alpha/2}\big)|z_1 - z_2|\,,\end{aligned}$$

which establishes Part (3) of Lemma 4.10 if we set $C_J := 4C_l + C_T$, and if we choose $\rho_g = \rho_g(R, L, \alpha, \operatorname{Lip}(\Phi_1), \operatorname{Lip}(\Phi_2)) \in (0,\rho_0)$ so small that $C_J\rho^{\alpha/2} < 1$ for all $\rho \in (0,\rho_g]$. Recall that we have fixed $\varepsilon = 1/200$ and that $\operatorname{Lip}(\Phi_1)$ and $\operatorname{Lip}(\Phi_2)$ depend only on $\varepsilon, R, L, \alpha, m, n$ according to Lemma 4.2, which means that ρ_0 and hence also ρ_g and C_J actually depend on R, L, α, m, n only. □

Remark 4.11. Inspecting the above proof one can see that

$$\operatorname{Lip}(J - \mathrm{Id}) = \operatorname{Lip}(I - T) = \operatorname{Lip}(I(1 - \phi\circ\eta)) \le \frac{16^2 C_l}{\rho_0}\rho^{\alpha/2}\,.$$

Proof of Theorem 4. According to Definition 2 we have $d_{\mathcal{H}}(\Sigma_0, \Sigma_j) \to 0$ as $j \to \infty$, so that we can choose j_0 such that $d_{\mathcal{H}}(\Sigma_0, \Sigma_j) \leq \rho_g < \rho_G/2$ for all $j \geq j_0$, where ρ_G is the constant from Lemma 4.8 for fixed $\varepsilon := 1/200$. Therefore, by Corollary 4.9, Σ_j is ambient isotopic to Σ_0 for all $j \geq j_0$. Moreover, by means of Lemma 4.10 we can find for each $j \geq j_0$ a C^1-diffeomorphism of the ambient space $J_j : \mathbb{R}^n \to \mathbb{R}^n$ such that $\operatorname{biLip}(J_j) \leq 1 + C_J d_{\mathcal{H}}(\Sigma_1, \Sigma_2)^{\alpha/2}$ and $J_j(\Sigma_j) = \Sigma_0$. □

5. Semicontinuity

5.1. Preliminaries

Before passing to the proof of Theorem 2, we set up some notation, and prove two technical lemmata which explain how our discrete curvatures change under small bilipschitz perturbations of the identity map.

Definition 5.1. Let $N \in \mathbb{N}$ and $\mu, \nu \in \mathscr{M}(\mathbb{R}^N)$ be Radon measures. We set

$$D(\nu, \mu, x) = \lim_{r \downarrow 0} \frac{\nu(\mathbb{B}^N(x, r))}{\mu(\mathbb{B}^N(x, r))},$$

where we interpret $0/0 = 0$.

Definition 5.2. For any function $F : X \to Y$ and any $l \in \mathbb{N}$, we define $F^{\times l} : X^l \to Y^l$ by the formula $F^{\times l}(x_1, \ldots, x_l) = (F(x_1), \ldots, F(x_l))$.

Definition 5.3. Let $l \in \{1, 2, \ldots, m+2\}$ and $\Sigma \in \mathscr{C}^{1,\alpha}_{m,n}$. For $l \leq m+1$, we define

$$\mathcal{K}_l[\Sigma](y_0, \ldots, y_{l-1}) = \sup_{y_l, \ldots, y_{m+1} \in \Sigma} \mathcal{K}(y_0, \ldots, y_{m+1}) \quad \text{for } y_0, \ldots, y_{l-1} \in \Sigma$$

for $\mathcal{K}$ given in the introduction by formula (2). We additionally set $\mathcal{K}_{m+2}[\Sigma] \equiv \mathcal{K}$.

Definition 5.4. Let $A \subset \mathbb{R}^n$ and $l \in \mathbb{N}$. We define *the l-diagonal of A*

$$\Delta^l A = \{(x_0, \ldots, x_l) \in A^l : x_0 = x_1 = \cdots = x_l\}.$$

Formally, the integrands $\mathcal{K}$ and R_{tp}^{-1} are defined only off the diagonal $\Delta^l \Sigma$. It does not matter how one defines them on the diagonal: it does not affect the integral, since $\mathcal{H}^{ml}(\Delta^l \Sigma) = 0$. Below, we also freely use the

equivalence of measures $\mathcal{H}^{ml} \llcorner \Sigma^l \simeq (\mathcal{H}^m \llcorner \Sigma)^l$ which holds as long as Σ is an embedded submanifold due to [17, 3.2.23]. (Actually, it holds even if Σ is just a subset of the image of a *single* Lipschitz function.)

Remark 5.5. Let $N \in \mathbb{N}$ and μ, ν be Radon measures on $\mathbb{R}^N$. The Radon-Nikodym theorem (cf. [27, Theorem 2.12]) implies that if ν is absolutely continuous with respect to μ, then for any $f \in L^1(\mathbb{R}^N, \nu)$

$$\int f(x)\ d\nu(x) = \int f(x) D(\nu, \mu, x)\ d\mu(x)\,.$$

Lemma 5.6. *Let $\Sigma_1, \Sigma_2 \in \mathscr{C}^{1,\alpha}_{m,n}$, $F : \mathbb{R}^n \to \mathbb{R}^n$ be a bilipschitz homeomorphism such that $F(\Sigma_1) = \Sigma_2$. Set $\mu = \mathcal{H}^{ml} \llcorner \Sigma_2^l$ and $\nu = (F^{\times l})_*(\mathcal{H}^{ml} \llcorner \Sigma_1^l)$. Then μ and ν are mutually absolutely continuous and*

$$D(\nu, \mu, x) \le \operatorname{Lip}(F^{-1})^{ml} \quad \text{and} \quad D(\mu, \nu, x) \le \operatorname{Lip}(F)^{ml}$$

for all $x \in \mathbb{R}^{nl}$.

Proof. If $x \in \mathbb{R}^{nl} \sim \Sigma_2^l$, then $\operatorname{dist}(x, \Sigma_2^l) > 0$, so for $0 < r < \operatorname{dist}(x, \Sigma_2^l)$ we have $\mu(\mathbb{B}^{ml}(x, r)) = 0 = \nu(\mathbb{B}^{ml}(x, r))$ and, according to Definition 5.1, $D(\mu, \nu, x) = D(\nu, \mu, x) = 0$.

Note that $\operatorname{Lip}(F^{\times l}) = \operatorname{Lip}(F)$ and $(F^{\times l})^{-1} = (F^{-1})^{\times l}$. Furthermore, observe that for $x \in \Sigma_2^l$ and $0 < r < \infty$

$$\begin{aligned} \Sigma_1^l \cap (F^{\times l})^{-1}(\mathbb{B}^{ml}(x, r)) &= \Sigma_1^l \cap (F^{\times l})^{-1}(\Sigma_2^l \cap \mathbb{B}^{ml}(x, r)) \\ &= (F^{\times l})^{-1}(\Sigma_2^l \cap \mathbb{B}^{ml}(x, r))\,; \end{aligned}$$

hence

$$\frac{\nu(\mathbb{B}^{ml}(x, r))}{\mu(\mathbb{B}^{ml}(x, r))} = \frac{\mathcal{H}^{ml}((F^{\times l})^{-1}(\Sigma_2^l \cap \mathbb{B}^{ml}(x, r)))}{\mathcal{H}^{ml}(\Sigma_2^l \cap \mathbb{B}^{ml}(x, r))} \le \operatorname{Lip}(F^{-1})^{ml}$$

and consequently $D(\nu, \mu, x) \le \operatorname{Lip}(F^{-1})^{ml}$. The estimate for $D(\mu, \nu, x)$ is obtained by writing

$$\Sigma_2^l \cap \mathbb{B}^{ml}(x, r) = F^{\times l}((F^{\times l})^{-1}(\Sigma_2^l \cap \mathbb{B}^{ml}(x, r)))\,.$$

□

Lemma 5.7. *Let $\Sigma_1, \Sigma_2 \in \mathscr{C}^{0,1}_{m,n}$, $0 < \varepsilon < 1/2$. Assume $F : \Sigma_1 \to \mathbb{R}^n$ is bilipschitz and satisfies $F(\Sigma_1) = \Sigma_2$, $F(z) = z + G(z)$ for $z \in \Sigma_1$ and some*

$G : \Sigma_1 \to \mathbb{R}^n$ *having* $\operatorname{Lip}(G) \leq \varepsilon$. *Then for any* $T \in \Sigma_1^l \sim \Delta^l \Sigma_1$ *and* $l \in \{2, \ldots, m+2\}$

$$\left| \mathcal{K}_l[\Sigma_2](F^{\times l}(T)) - \mathcal{K}_l[\Sigma_1](T) \right| \leq \varepsilon C_{5.7} \operatorname{diam}(\triangle T)^{-1} ,$$

where $C_{5.7} = C_{5.7}(m) > 0$.

Proof. First we treat the case $l = m+2$. Let $T = (x_0, \ldots, x_{m+1}) \in \Sigma_1^l \sim \Delta^l \Sigma_1$. Set $u_i = x_i - x_0$, $v_i = F(x_i) - F(x_0)$ and $e_i = G(x_i) - G(x_0)$ for $i = 1, 2, \ldots, m+1$. Observe that $v_i = u_i + e_i$ and that $|e_i| \leq \varepsilon |u_i|$. We compute

$$\begin{aligned} &\left| \mathcal{H}^{m+1}(\triangle F^{\times l}(T)) - \mathcal{H}^{m+1}(\triangle T) \right| \\ &= \tfrac{1}{(m+1)!} \big| |v_1 \wedge \cdots \wedge v_{m+1}| - |u_1 \wedge \cdots \wedge u_{m+1}| \big| \\ &\leq \tfrac{1}{(m+1)!} \operatorname{diam}(\triangle T)^{m+1} \sum_{i=1}^{m+1} \tbinom{m+1}{i} \varepsilon^i \leq \tfrac{2^{m+1}}{(m+1)!} \operatorname{diam}(\triangle T)^{m+1} \varepsilon . \end{aligned}$$

Since $(1-\varepsilon) \operatorname{diam}(\triangle T) \leq \operatorname{diam}(\triangle F^{\times l}(T)) \leq (1+\varepsilon) \operatorname{diam}(\triangle T)$ and recalling $\varepsilon < 1/2$, we obtain

$$\begin{aligned} \left| \mathcal{K}(F^{\times l}(T)) - \mathcal{K}(T) \right| &\leq \tfrac{\varepsilon}{1-\varepsilon} \left(\mathcal{K}(T) + \tfrac{2^{m+1}\varepsilon}{(m+1)!(1-\varepsilon)} \tfrac{1}{\operatorname{diam}(\triangle T)} \right) \\ &\leq C \varepsilon \operatorname{diam}(\triangle T)^{-1} , \end{aligned} \tag{100}$$

where $C = C(m) > 0$.

In case $2 \leq l < m+2$ for $T = (x_0, \ldots, x_{l-1}) \in \Sigma_1^l \sim \Delta^l \Sigma_1$ we employ the assumption that F is bilipschitz, so that we can write

$$\begin{aligned} (\mathcal{K}_l[\Sigma_2] \circ F^{\times l})(T) &= \sup_{y_l, \ldots, y_{m+1} \in \Sigma_2} \mathcal{K}(F(x_0), \ldots, F(x_{l-1}), y_l, \ldots, y_{m+1}) \\ &= \sup_{x_l, \ldots, x_{m+1} \in \Sigma_1} \mathcal{K}(F(x_0), \ldots, F(x_{m+1})) \end{aligned}$$

and using (100)

$$\begin{aligned} &\sup_{x_l, \ldots, x_{m+1} \in \Sigma_1} \mathcal{K}(F(x_0), \ldots, F(x_{m+1})) \\ &\leq \sup_{x_l, \ldots, x_{m+1} \in \Sigma_1} \big(\mathcal{K}(x_0, \ldots, x_{m+1}) + C \varepsilon \operatorname{diam}(\{x_0, \ldots, x_{m+1}\})^{-1} \big) \\ &\leq \mathcal{K}_l[\Sigma_1](T) + C \varepsilon \operatorname{diam}(\triangle T)^{-1} . \end{aligned}$$

In the same way we obtain the lower bound

$$\sup_{x_l,\dots,x_{m+1}\in\Sigma_1} \mathcal{K}(F(x_0),\dots,F(x_{m+1})) \geq \mathcal{K}_l[\Sigma_1](T) - C\varepsilon \operatorname{diam}(\triangle T)^{-1}\,.$$

□

A similar lemma does hold for the $R_{\rm tp}$ function.

Lemma 5.8. *Let $\Sigma_1, \Sigma_2 \in \mathscr{C}^{0,1}_{m,n}$, $0 < \varepsilon < 1/2$. Assume $F : \Sigma_1 \to \mathbb{R}^n$ is bilipschitz and satisfies $F(\Sigma_1) = \Sigma_2$, $F(z) = z + G(z)$ for $z \in \Sigma_1$ and some $G : \Sigma_1 \to \mathbb{R}^n$ having $\operatorname{Lip}(G) \leq \varepsilon$. Then for $\mathcal{H}^m$-almost all $x_1, y_1 \in \Sigma_1$, $x_1 \neq y_1$, we have*

$$\left|\frac{1}{R_{\rm tp}[\Sigma_1](x_1,y_1)} - \frac{1}{R_{\rm tp}[\Sigma_2](F(x_1),F(y_1))}\right| \leq \frac{C_{5.8}\,\varepsilon}{|x_1-y_1|}\,, \tag{101}$$

where $C_{5.8} = C_{5.8}(m) > 0$.

Proof. Set $x_2 = F(x_1)$, $y_2 = F(y_1)$. Without loss of generality, by the classic Rademacher theorem, assume that G is differentiable at x_1 and the tangent spaces to both manifolds, $U_i := T_{x_i}\Sigma_i$, are well defined for $i = 1, 2$. Then,

$$\frac{1}{R_{\rm tp}[\Sigma_i](x_i,y_i)} = \frac{2d_i}{|x_i-y_i|^2}\,, \qquad i = 1,2,$$

where $d_i = \operatorname{dist}(y_i - x_i, U_i)$. By the triangle inequality,

$$\begin{aligned}\left|\frac{1}{R_{\rm tp}[\Sigma_1](x_1,y_1)} - \frac{1}{R_{\rm tp}[\Sigma_2](F(x_1),F(y_1))}\right| \\ \leq \frac{2|d_1-d_2|}{|x_1-y_1|^2} + 2d_2\left|\frac{1}{|x_1-y_1|^2} - \frac{1}{|x_2-y_2|^2}\right|\,.\end{aligned} \tag{102}$$

We shall show that each of these two terms is controlled by a constant multiple of $\varepsilon|x_1-y_1|^{-1}$. Indeed, since $d_i \leq |x_i - y_i|$ and

$$(1-\varepsilon)|x_1-y_1| \leq |x_2-y_2| = |F(x_1)-F(y_1)| \leq (1+\varepsilon)|x_1-y_1|\,,$$

we easily estimate the second term on the right hand side of (102),

$$\begin{aligned}2d_2\left|\frac{1}{|x_1-y_1|^2} - \frac{1}{|x_2-y_2|^2}\right| &\leq \frac{2|x_2-y_2|}{|x_1-y_1|^2|x_2-y_2|^2}\left||x_1-y_1|^2 - |x_2-y_2|^2\right| \\ &\leq \frac{2}{(1-\varepsilon)|x_1-y_1|^3}\cdot\varepsilon(2+\varepsilon)|x_1-y_1|^2 \\ &< \frac{10\varepsilon}{|x_1-y_1|} \qquad \text{as } \varepsilon \in (0,\tfrac{1}{2}).\end{aligned} \tag{103}$$

To estimate the first term on the right hand side of (102), it is enough to check that $|d_1 - d_2| \leq C\varepsilon|x_1 - y_1|$. Note that $d_i = \operatorname{dist}(x_i - y_i, U_i) = |(x_i - y_i) - (U_i)_\natural(x_i - y_i)|$, so that

$$\begin{aligned} |d_1 - d_2| &\leq |(x_1 - y_1) - (x_2 - y_2)| + |(U_1)_\natural(x_1 - y_1) - (U_2)_\natural(x_2 - y_2)| \\ &\leq \varepsilon|x_1 - y_1| + \big|(U_1)_\natural\big((x_1 - y_1) - (x_2 - y_2)\big)\big| \\ &\quad + |((U_1)_\natural - (U_2)_\natural)(x_2 - y_2)| \\ &\leq 2\varepsilon|x_1 - y_1| + \|(U_1)_\natural - (U_2)_\natural\| \cdot (1 + \varepsilon)|x_1 - y_1|\,. \end{aligned} \tag{104}$$

By the assumption on F and x_1, we have

$$U_2 = DF(x_1)(U_1) = (\mathrm{Id} + DG(x_1))(U_1)\,, \qquad \|DG(x_1)\| \leq \varepsilon$$

The estimate of the angle between m-planes, see [25, Prop. 2.5], yields $\|(U_1)_\natural - (U_2)_\natural\| \leq C\varepsilon$ for some constant $C = C(m)$, and the lemma follows. □

5.2. Semicontinuity, compactness and existence of minimisers

We are now ready to give the proof of Theorem 2 and Corollary 1. We begin with lower semicontinuity which is crucial for the compactness of sublevel sets of geometric curvature and the existence of energy minimisers in isotopy classes.

Remark. There are several more or less equivalent ways to phrase the argument which yields semicontinuity. If the curvature integrand contains no supremum, then the aim can be achieved by (a) localization, (b) parametrization of all integrals by the same domain, (c) an application of Fatou's lemma which is enabled by the C^1 convergence of the parameterisations. Such an argument is presented in [40, pp. 2297–2298]. However, even in the case of curves the integrands involving a supremum are by no means continuous (see e.g. [39, Section 3]) with respect to C^1 convergence, which requires more care, cf. e.g. the proof of Thm. 3 in [39]. In order to avoid a case by case study of the appropriately understood lower semicontinuity of all the integrands considered, we present here a general argument which is streamlined so that all the cases can be included into the same scheme.

Let us note that the proof is flexible enough to show that all the energies are in fact lower semicontinous w.r.t. the bilipschitz convergence of (not necessarily C^1) submanifolds also in the case $p \leq p_0$, when we have no $C^{1,\alpha}$–regularization effect.

Proof of part (i) of Theorem 2. Fix $l \in \{1, 2, \dots, m+2\}$. For $j \in \mathbb{N} \cup \{0\}$ let $\Sigma_j \in \mathcal{A}^{\mathcal{E}}_{m,n}(E, d)$. Assume that Σ_j converges in Hausdorff distance to Σ and, without loss of generality, such that

$$\lim_{j\to\infty} \mathcal{E}(\Sigma_j) = \liminf_{j\to\infty} \mathcal{E}(\Sigma_j). \tag{105}$$

Hence for some fixed $x \in \Sigma$ we find a sequence of points $x_j \in \Sigma_j$ such that $x_j \to x$ as $j \to \infty$, so that the shifted submanifolds $\tilde{\Sigma}_j := \Sigma_j - x_j$ converge in Hausdorff-distance to $\tilde{\Sigma} := \Sigma - x$. Hence $0 \in \tilde{\Sigma}$ and $0 \in \tilde{\Sigma}_j$, and by translation invariance of the geometric curvature energies, $\mathcal{E}(\tilde{\Sigma}_j) = \mathcal{E}(\Sigma_j) \le E$ for all $j \in \mathbb{N}$. Thus, by the Regularity Theorem, $\tilde{\Sigma}_j \in \mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$, for appropriate R, L given by (10) depending only on the fixed p and on the uniform energy threshold E. Hence, by the compactness result in $\mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$, Theorem 3, we find a subsequence still denoted by $\tilde{\Sigma}_j$ and some submanifold $\tilde{\Sigma}_0 \in \mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$ such that $\tilde{\Sigma}_j \to \tilde{\Sigma}_0$ in $\mathscr{C}^{1,\alpha}_{m,n}$ (see Definition 2), which immediately implies that $\tilde{\Sigma} = \tilde{\Sigma}_0$ is contained in $\mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$, so that we can apply all results of Section 4 to $\tilde{\Sigma}_j$ and $\tilde{\Sigma}$, and, in addition, we may evaluate the energy at $\Sigma = \tilde{\Sigma} + x$ to obtain $\mathcal{E}(\Sigma) = \mathcal{E}(\tilde{\Sigma})$, so it is enough to establish $\mathcal{E}(\tilde{\Sigma}) \le \lim_{j\to\infty} \mathcal{E}(\tilde{\Sigma}_j)$. To simplify notation, we identify $\tilde{\Sigma}$ with Σ, and $\tilde{\Sigma}_j$ with Σ_j from now on.

We may also assume that for each j

$$d_{\mathcal{H}}(\Sigma_j, \Sigma) < \rho_g, \qquad \text{where } \rho_g \text{ is given by Lemma 4.10.}$$

Now, for $j \in \mathbb{N}$ set $\mu_j := \mathcal{H}^{ml} \llcorner \Sigma^l_j$, $\mu := \mathcal{H}^{ml} \llcorner \Sigma^l$, and let $J_j : \mathbb{R}^n \to \mathbb{R}^n$ be the diffeomorphism constructed in Lemma 4.10 such that $J_j(\Sigma) = \Sigma_j$.

Observe that, by Lemma 4.10, $\mathrm{biLip}(J_j) \le 1 + C_J d_{\mathcal{H}}(\Sigma, \Sigma_j)^{\alpha/2}$. Moreover, the restriction $F_j := J_j \,|_{\Sigma}$ satisfies

$$F_j(x) = x + G_j(x) \ \text{ on } \Sigma, \quad \mathrm{Lip}(G_j) =: \varepsilon_j, \quad \varepsilon_j \to 0 \ \text{ as } j \to \infty, \tag{106}$$

since by Lemma 4.8 $\varepsilon_j \le C_l d_{\mathcal{H}}(\Sigma_j, \Sigma)^{\alpha/2}$.

Now, the reader who is not overly keen to follow all the technical details can skim the lines from here to (116), thinking that, roughly, Step 1 below involves the reparameterisations and the control of Jacobians, Step 2 is the place where 'Fatou does the job', and Step 3 is the price to pay for the flexibility of the argument (in the case where the integrand contains too many suprema).

Step 1 (fixing the domain of integration). We shall first check that if $\mathcal{F} \in \{\mathcal{E}^l_p, \mathcal{T}_p, \mathcal{T}^G_p\}$ is one of the energies considered, then — in order to check that

$\mathcal{F}(\Sigma) \leq \liminf \mathcal{F}(\Sigma_j)$ — we can consider the limes inferior of a sequence of integrals over a *fixed* domain Σ, with appropriately perturbed integrands. (In our application we could write actual limits because of (105), but that does not affect any of the following arguments.)

Indeed, in each of the cases considered we have

$$\mathcal{F}(\Sigma) = \int_{\Sigma^l} (K_\Sigma)^p \, d\mathcal{H}^{ml} \quad \text{for an integrand} \quad K_\Sigma \colon \Sigma^l \to [0,\infty] \,. \tag{107}$$

(To fix the ideas, we assume in all cases $K_\Sigma \equiv \infty$ on $\Delta^l\Sigma$; this does not affect the value of $\mathcal{F}$ as $\mathcal{H}^{ml}(\Delta^l\Sigma) = 0$.) Thus, changing the variables and using Lemma 5.6, for a sequence of Σ_j's with $\sup_j \mathcal{F}(\Sigma_j)$ finite we obtain

$$\begin{aligned}
\left| \mathcal{F}(\Sigma_j) - \int \big(K_{\Sigma_j} \circ J_j^{\times l}\big)^p \, d\mu \right| &= \left| \int \big(K_{\Sigma_j}\big)^p \, d\mu_j - \int \big(K_{\Sigma_j} \circ J_j^{\times l}\big)^p \, d\mu \right| \\
&= \left| \int \big(K_{\Sigma_j}\big)^p \, d\mu_j - \int \big(K_{\Sigma_j}\big)^p \, d((J_j^{\times l})_* \mu) \right| \\
&= \left| \int \big(K_{\Sigma_j}\big)^p \, \Big(1 - D\big((J_j^{\times l})_* \mu, \mu_j, \cdot\big)\Big) \, d\mu_j \right| \\
&\leq C \left(\sup_j \mathcal{F}(\Sigma_j) \right) d_{\mathcal{H}}(\Sigma_j, \Sigma)^{\alpha/2} \xrightarrow{j\to\infty} 0 \,.
\end{aligned}$$

The last inequality follows from the fact that by Lemma 5.6 applied to $\nu = (J_j^{\times l})_* \mu$ and $\mu = \mu_j$ we have the density estimate

$$\frac{1}{\big(\operatorname{Lip} J_j\big)^{ml}} \leq D\big((J_j^{\times l})_* \mu, \mu_j, \cdot\big) \leq \big(\operatorname{Lip} J_j^{-1}\big)^{ml}$$

Therefore, as $\operatorname{biLip}(J_j) \leq 1 + C_J d_{\mathcal{H}}(\Sigma, \Sigma_j)^{\alpha/2} \to 1$, we have

$$\Big|1 - D\big((J_j^{\times l})_* \mu, \mu_j, \cdot\big)\Big| \leq C d_{\mathcal{H}}(\Sigma_j, \Sigma)^{\alpha/2} \,. \tag{108}$$

All this yields

$$\liminf_{j\to\infty} \int \big(K_{\Sigma_j} \circ J_j^{\times l}\big)^p \, d\mu = \liminf_{j\to\infty} \mathcal{F}(\Sigma_j) \,. \tag{109}$$

Below, we work with the left hand side of (109).

Step 2 (energies with at least two integrals over the manifold). Suppose now that $l \geq 2$. If $\mathcal{F} = \mathcal{E}^l_p$, $l \geq 2$, with the integrand

$$K_\Sigma = \mathcal{K}_l[\Sigma] \colon \Sigma^l \to [0, \infty)$$

being the discrete curvature from Definition 5.3, then, by Lemma 5.7, we have

$$K_\Sigma(T) \leq K_{\Sigma_j}\big(F_j^{\times l}(T)\big) + \frac{C(m)\,\varepsilon_j}{\operatorname{diam} T}, \qquad T \in \Sigma^l \sim \Delta^l\Sigma\,. \tag{110}$$

If on the other hand $\mathcal{F} = \mathcal{T}_p$ (so that $l = 2$), and

$$K_\Sigma = \frac{1}{R_{\mathrm{tp}}[\Sigma]} \colon \Sigma \times \Sigma \to [0, \infty],$$

with the tangent–point radius R_{tp} defined on $\Sigma^2 \sim \Delta^2\Sigma$ by (4), then inequality (110) holds by Lemma 5.8. Thus, in both cases we can use (110) to write, for a fixed $T \in \Sigma^l \sim \Delta^l\Sigma$,

$$\begin{aligned} K_\Sigma(T)^p &\leq \liminf_{j\to\infty} \left(K_{\Sigma_j}\big(F_j^{\times l}(T)\big) + \frac{C(m)\,\varepsilon_j}{\operatorname{diam} T} \right)^p \\ &= \liminf_{j\to\infty} K_{\Sigma_j}\big(J_j^{\times l}(T)\big)^p \qquad (\text{as } J_j = F_j \text{ on } \Sigma). \end{aligned}$$

Since $\mathcal{H}^{ml}(\Delta^l\Sigma) = 0$, we can now ingrate both sides w.r.t. μ and invoke Fatou's lemma (cf. [17, 2.4.6]) to obtain

$$\begin{aligned} \mathcal{F}(\Sigma) = \int K_\Sigma(T)^p \, d\mu &\leq \int \liminf_{j\to\infty} K_{\Sigma_j}\big(J_j^{\times l}(T)\big)^p \, d\mu \\ &\leq \liminf_{j\to\infty} \int K_{\Sigma_j}\big(J_j^{\times l}(T)\big)^p \, d\mu \\ &= \liminf_{j\to\infty} \mathcal{F}(\Sigma_j) \end{aligned}$$

by (109). This concludes the proof of Theorem 1 for $\mathcal{F} = \mathcal{E}^l_p$ with $l \geq 2$ and for $\mathcal{F} = \mathcal{T}_p$.

Step 3 (energies with a single integral). The case $l = 1$, i.e. when $\mathcal{F} = \mathcal{E}^1_p$, resp. $\mathcal{F} = \mathcal{T}^G_p$, needs a separate treatment. We shall now work with the auxiliary integrands

$$K_\Sigma \colon \Sigma \times \Sigma \to [0, \infty)\,,$$

using $K_\Sigma = \mathcal{K}_2[\Sigma]$ for $\mathcal{F} = \mathcal{E}^1_p$, resp. $K_\Sigma = 1/R_{\mathrm{tp}}[\Sigma]$ for $\mathcal{F} = \mathcal{T}^G_p$.

The argument from Step 2 does not work here, as for $l = 1$ we deal with simplices T that degenerate to one point, and (110) would yield nothing. To avoid this problem, we remove a small neighbourhood of the diagonal, and pass to the limit twice. Here are the details.

For a fixed $s \in \mathbb{N}$, set

$$K_{\Sigma,s}(x) = \sup_{\substack{y\in\Sigma \\ |y-x|\geq 1/s}} K_\Sigma(x,y)\,, \qquad x \in \Sigma\,. \tag{111}$$

Define

$$\mathcal{F}_s(\Sigma) = \int_\Sigma K_{\Sigma,s}^p \, d\mathcal{H}^m\,.$$

Note that $0 \leq K_{\Sigma,1}^p \leq K_{\Sigma,2}^p \leq \ldots$, $K_{\Sigma,s}^p \nearrow K_\Sigma^p$ as $s \to \infty$, so that by the monotone convergence theorem, we have in each of the two cases ($\mathcal{F} = \mathcal{E}_p^1$ or $\mathcal{F} = \mathcal{T}_p^G$) that are being considered

$$\mathcal{F}(\Sigma) = \sup_{s\in\mathbb{N}} \mathcal{F}_s(\Sigma) = \lim_{s\to\infty} \mathcal{F}_s(\Sigma)\,. \tag{112}$$

Repeating Step 1 for each of the $\mathcal{F}_s$, we obtain

$$\liminf_{j\to\infty} \int \big(K_{\Sigma_j,s} \circ J_j\big)^p \, d\mu = \liminf_{j\to\infty} \mathcal{F}_s(\Sigma_j)\,. \tag{113}$$

Rewriting (110) for $l = 2$, $T = (x, y)$, $x \neq y \in \Sigma$, for the auxiliary integrands K_Σ, we obtain

$$K_\Sigma(x,y) \leq K_{\Sigma_j}\big(F_j(x), F_j(y)\big) + \frac{C(m)\,\varepsilon_j}{|x-y|}\,. \tag{114}$$

We shall use this estimate for s fixed and $j > 1$ so large that $\varepsilon_j < \frac{1}{s+1}$ (keep in mind that $\varepsilon_j \to 0$ as $j \to \infty$). Then, for points $x, y \in \Sigma$ with $|x - y| \geq \frac{1}{s}$, we have

$$|F_j(x) - F_j(y)| \geq (1-\varepsilon_j)|x-y| \geq \left(1 - \frac{1}{s+1}\right)\frac{1}{s} = \frac{1}{s+1},$$

and upon taking the suprema of both sides of (114) with respect to $y \in \Sigma$, $|x - y| \geq \frac{1}{s}$, we obtain

$$K_{\Sigma,s}(x) \leq K_{\Sigma_j,s+1}(F_j(x)) + C(m)s \cdot \varepsilon_j\,. \tag{115}$$

Thus, for each $x \in \Sigma$,

$$K_{\Sigma,s}(x)^p \le \liminf_{j\to\infty} K_{\Sigma_j,s+1}(F_j(x))^p \,.$$

Integration and Fatou's lemma yield now

$$\begin{aligned} \mathcal{F}_s(\Sigma) &\le \liminf_{j\to\infty} \int K_{\Sigma_j,s+1}(F_j(x))^p \, d\mu \\ &\overset{(113)}{=} \liminf_{j\to\infty} \mathcal{F}_s(\Sigma_j) \le \liminf_{j\to\infty} \mathcal{F}(\Sigma_j)\,, \end{aligned} \tag{116}$$

as $\mathcal{F}_s \le \mathcal{F}$ for all $s \in \mathbb{N}$. Upon taking the supremum of the left-hand sides with respect to $s \in \mathbb{N}$, in light of (112), we conclude the proof for $\mathcal{F} = \mathcal{E}_p^1$ and for $\mathcal{F} = \mathcal{T}_p^G$. □

Proof of part (ii) of Theorem 2. By the Regularity Theorem, $\Sigma_j \in \mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$ for all $j \in \mathbb{N}$, where the parameters R, L are given by (10) and do not depend on j. Thus, Theorem 3 implies that there is a subsequence (still denoted by Σ_j) and a submanifold $\Sigma \in \mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$, such that $\Sigma_j \to \Sigma$ in $\mathscr{C}^{1,\alpha}_{m,n}$, i.e., in the sense of Definition 2, which implies in particular that $\Sigma_j \to \Sigma$ in Hausdorff-distance, that $\operatorname{diam}\Sigma \le d$, and that $\Sigma \in \mathscr{C}^{0,1}_{m,n}$. Therefore, we may evaluate the energy $\mathcal{E}$ on Σ. Part (i) implies that $\Sigma \in \mathcal{A}^{\mathcal{E}}_{m,n}(E, d)$. □

Proof of Corollary 1. Notice that the class

$$\mathcal{C} := \{\Sigma \in \mathcal{A}^{\mathcal{E}}_{m,n}(E, d) : \Sigma \text{ is ambient isotopic to } M_0\}$$

contains the reference manifold M_0, so that we can find a minimising sequence $(\Sigma_j)_j \subset \mathcal{C}$ with $\mathcal{E}(\Sigma_j) \to \inf_{\mathcal{C}} \mathcal{E}$ as $j \to \infty$. The uniform energy bound E implies by the Regularity Theorem that $\Sigma_j \in \mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$ for all $j \in \mathbb{N}$, where the parameters R, L depend only on the energy bound and on the integrability parameter p, so that we can apply the improved compactness result, Theorem 3, to deduce the existence of a subsequence (still denoted by Σ_j) that converges to a limit submanifold $\Sigma_0 \in \mathscr{C}^{1,\alpha}_{m,n}(R, L, d)$ in $\mathscr{C}^{1,\alpha}_{m,n}$. Then the isotopy result, Theorem 4, implies that Σ_j is ambient isotopic to Σ_0 for j sufficiently large, which implies that $\Sigma_0 \in \mathcal{C}$. Part (i) of Theorem 2 finally leads to

$$\inf_{\mathcal{C}} \mathcal{E} \le \mathcal{E}(\Sigma_0) \le \liminf_{j\to\infty} \mathcal{E}(\Sigma_j) = \inf_{\mathcal{C}} \mathcal{E},$$

which concludes the proof. □

6. Bounds on the number of diffeomorphism and isotopy types

Proof of Theorem 1. Fix an energy $\mathcal{E} \in \{\mathcal{E}_p^l, \mathcal{T}_p, \mathcal{T}_p^G\}$ and a $p > p_0(\mathcal{E})$. Let $\Sigma \in \mathcal{A}_{m,n}^{\mathcal{E}}(E,d)$ be a manifold with controlled energy and diameter, cf. (1). Translating Σ if necessary, we have $\Sigma \in \mathscr{C}_{m,n}^{1,\alpha}(R,L,d)$ by the Regularity Theorem, where the parameters R and L depend only on E and p. Thus, $\Sigma \subset [-d,d]^n$.

Now fix $\varepsilon := 1/200$ and let $\mathfrak{F} = \{Q_1, Q_2, \dots Q_N\}$ be a minimal collection of closed cubes of edge $e := \rho_G/(2\sqrt{n})$ covering $[-d,d]^n$; here $\rho_G > 0$ is the constant of Lemma 4.8 and Corollary 4.9 for $\varepsilon = 1/200$. Notice that the dependence of ρ_G on the Lipschitz constant of an ε-normal map for Σ boils down to $\rho_G = \rho_G(R,L,\alpha,m,n)$ since we have fixed ε; see Lemma 4.2. Clearly, the cardinality of $\mathfrak{F}$ satisfies

$$\mathcal{H}^0(\mathfrak{F}) = N \leq k^n, \quad \text{with} \quad k = \left\lceil \frac{2d\sqrt{n}}{\rho_G} \right\rceil . \tag{117}$$

Following Durumeric [14, Section 5] (see also the remarks in Peters [29, Section 5]), to each $\Sigma \in \mathscr{C}_{m,n}^{1,\alpha}(R,L,d)$ we assign the subset $P(\Sigma) \subset \mathfrak{F}$ which consists of those cubes $Q \in \mathfrak{F}$ that intersect Σ, i.e.

$$Q \in P(\Sigma) \subset \mathfrak{F} \iff Q \cap \Sigma \neq \emptyset .$$

If $P(\Sigma_1) = P(\Sigma_2)$, then obviously $d_{\mathcal{H}}(\Sigma_1, \Sigma_2)$ does not exceed the diameter of all the Q_i which equals $e\sqrt{n} = \rho_G/2$. Hence, by Corollary 4.9, Σ_1 and Σ_2 are ambient isotopic. Therefore, the number of distinct isotopy classes of manifolds $\Sigma \in \mathcal{A}_{m,n}^{\mathcal{E}}(E,d)$ is not larger than $K = 2^N$, the number of all subsets of $\mathfrak{F}$.

Finally, since $\rho_G = \rho_G(R,L,\alpha,m,n)$ depends only on R and L which are given, for a particular energy $\mathcal{E}$ and an upper energy bound E, by (10) in the Regularity Theorem, and on $\alpha = 1 - p_0(\mathcal{E})/p$, it is clear that $K = K(E,d,m,n,p)$. □

Remark 6.1. The estimate $K \leq 2^N$ is obviously not optimal for *connected* manifolds. If Σ is connected, then the union of all cubes in $P(\Sigma)$ is connected, too; thus, one only needs to count those subsets of $\mathfrak{F}$ which have connected unions. (For $n = 1$ there are 2^k subsets of the family of intervals and only $O(k^2)$ connected subsets!) One can prove [32] that the number K_{con} of such

subsets of $\mathfrak{F}$ satisfies

$$(2-a(n))^N \le K_{\mathrm{con}} \le (2-b(n))^N\,, \quad N \le k^n\,, \tag{118}$$

where $a, b\colon \mathbb{N} \to (0,\infty)$ are positive (but go to zero as the dimension $n \to \infty$).

Here is the gist of the argument. Assume for the sake of simplicity that the closed cubes in $\mathfrak{F}$ have disjoint interiors, and that $k = 2d/\varepsilon$ is divisible by 3.

To obtain the upper bound, divide $[-d,d]^n$ into larger cubes $\widetilde{Q}$, each of them consisting of 3^n of the initial Q_j's. In each $\widetilde{Q}_0$, one Q_{j_0} — call it *central* — contains the center of $\widetilde{Q}_0$ and is separated from other $\widetilde{Q}_i$'s by a layer of small Q_j's. Now, if for a connected Σ the subset $P(\Sigma)$ contains one of the small *central* Q_j's, then it must contain at least one small cube from the layer around this Q_j unless the whole $P(\Sigma) = \{Q_j\}$. This limits the number of possible choices of $P(\Sigma)$ and yields the upper bound in (118).

To obtain the lower bound, one constructs a specific family of subsets of $\mathfrak{F}$ with connected unions, e.g. as follows. Let $X \subset \mathfrak{F}$ consist of k^{n-1} little cubes adjacent to a fixed $(n-1)$-dimensional face of $[-d,d]^n$ (think of it as the bottom face) and of $(k/3)^{n-1}$ thin vertical, symmetrically placed 'towers' standing on the bottom face, each of these towers consisting of $k-1$ little cubes and reaching to the top of the whole box $[-d,d]^n$. Thus,

$$\text{the number of cubes in } X = k^{n-1} + \frac{1}{3^{n-1}} k^{n-1}(k-1)\,.$$

Note that adding to X *any* subset of $\mathfrak{F} \sim X$, we obtain a family of cubes with connected union (because each of the cubes in $\mathfrak{F} \sim X$ touches one of the towers in X). From this, one obtains the lower bound for K_{con}.

It is however clear that (118) does not take into account any *global* information on Σ (e.g., it does not exclude those subsets of $\mathfrak{F}$ that are too small or too flat to cover a Σ with $\mathcal{E}(\Sigma) \le E$).

Remark 6.2 (Explicit bounds). One can track an estimate of N (the number of little cubes in $\mathfrak{F}$) in terms of the energy bounds $\mathcal{E}(\Sigma) \le E$ etc. as follows.

(i) Note that R and L given by (10) in the Regularity Theorem satisfy $R^\alpha L = c(m,n,l,p)$.

(ii) Lemma 3.1 with $A = 4$ yields $C_{\mathrm{ang}}(L,4) = 257L + 8$.

(iii) Lemma 4.2 gives the Lipschitz constant of the ε-normal map, cf. (60) and Remark 4.3.

(iv) The number ρ_G for fixed $\varepsilon = 1/200$ emerges in Lemma 4.8 and Corollary 4.9, via the constant C_l; a combination of (90)–(91) with (i) above shows that we can have, e.g.,

$$\begin{aligned}\frac{1}{\rho_G} = C_l^{2/\alpha} &\leq c(m,n,l,p)(L+1)^{2(2\alpha+1)/\alpha^2} \\ &\leq \tilde{c}(m,n,l,p)\big(E^{1/p}+1\big)^{2(2\alpha+1)/\alpha^2}, \quad \alpha = 1 - \frac{p_0(\mathcal{E})}{p}.\end{aligned}$$

Plugging the last estimate into $\log\log K \leq \log N \leq n \log\big(1 + 2d\sqrt{n}/\rho_G\big)$, one obtains the bound (9) stated in the Introduction.

Acknowledgments

Parts of this work were completed between 2012 and 2015, while the authors, in various combinations, were visiting the Albert Einstein Institute in Golm and RWTH Aachen University. They are all grateful for the hospitality and for excellent working conditions.

The first author was partially supported by the Foundation for Polish Science. The first and second author were partially supported by NCN Grant no. 2013/10/M/ST1/00416 *Geometric curvature energies for subsets of the Euclidean space.* The third author's work is partially funded by the Excellence Initiative of the German federal and state governments.

References

[1] *Geometric knot theory*, Oberwolfach Rep. **10** (2013), no. 2, 1313–1358. Abstracts from the workshop held April 28–May 4, 2013, Organized by Dorothy Buck, Jason Cantarella, John M. Sullivan and Heiko von der Mosel.

[2] W. K. Allard, *On the first variation of a varifold*, Ann. of Math. (2) **95** (1972) 417–491.

[3] M. T. Anderson and J. Cheeger, *Diffeomorphism finiteness for manifolds with Ricci curvature and $L^{n/2}$-norm of curvature bounded*, Geom. Funct. Anal. **1** (1991), no. 3, 231–252.

[4] M. T. Anderson and J. Cheeger, *C^α-compactness for manifolds with Ricci curvature and injectivity radius bounded below*, J. Differential Geom. **35** (1992), no. 2, 265–281.

[5] S. Blatt, *Note on Continuously Differentiable Isotopies* (2009). Preprint no. 34, RWTH Aachen.

[6] S. Blatt and S. Kolasiński, *Sharp boundedness and regularizing effects of the integral Menger curvature for submanifolds*, Adv. Math. **230** (2012), no. 3, 839–852.

[7] B. Bojarski, P. Hajłasz, and P. Strzelecki, *Sard's theorem for mappings in Hölder and Sobolev spaces*, Manuscripta Math. **118** (2005), no. 3, 383–397.

[8] P. Breuning, *Compactness of immersions with local Lipschitz representation*, Ann. Inst. H. Poincaré Anal. Non Linéaire **29** (2012), no. 4, 545–572.

[9] G. Burde and H. Zieschang, Knots, Vol. 5 of de Gruyter Studies in Mathematics, Walter de Gruyter & Co., Berlin, second edition (2003). ISBN 3-11-017005-1.

[10] J. Cheeger, *Finiteness theorems for Riemannian manifolds*, Amer. J. Math. **92** (1970) 61–74.

[11] J. Cheeger, *Structure theory and convergence in Riemannian geometry*, Milan J. Math. **78** (2010), no. 1, 221–264.

[12] G. David, *Unrectifiable 1-sets have vanishing analytic capacity*, Rev. Mat. Iberoamericana **14** (1998), no. 2, 369–479.

[13] S. Delladio, *On hypersurfaces in $\mathbf{R}^{n+1}$ with integral bounds on curvature*, J. Geom. Anal. **11** (2001), no. 1, 17–42.

[14] O. C. Durumeric, *Thickness formula and C^1-compactness for $C^{1,1}$ Riemannian submanifolds*, arXiv:math/0204050 [math.DG], (2002).

[15] L. C. Evans and R. F. Gariepy, Measure Theory and Fine Properties of Functions, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL (1992). ISBN 0-8493-7157-0.

[16] H. Federer, *Curvature measures*, Trans. Amer. Math. Soc. **93** (1959), 418–491.

[17] H. Federer, Geometric Measure Theory, Die Grundlehren der Mathematischen Wissenschaften, Band 153, Springer-Verlag New York Inc., New York (1969).

[18] R. L. Foote, *Regularity of the distance function*, Proc. Amer. Math. Soc. **92** (1984), no. 1, 153–155.

[19] M. Gromov, *Almost flat manifolds*, J. Differential Geom. **13** (1978), no. 2, 231–241.

[20] M. Gromov, Structures métriques pour les variétés riemanniennes, Vol. 1 of *Textes Mathématiques [Mathematical Texts]*, CEDIC, Paris (1981). ISBN 2-7124-0714-8. Edited by J. Lafontaine and P. Pansu.

[21] A. Hatcher, Algebraic Topology, Cambridge University Press, Cambridge (2002). ISBN 0-521-79160-X; 0-521-79540-0.

[22] M. W. Hirsch, Differential Topology, Vol. 33 of Graduate Texts in Mathematics, Springer-Verlag, New York (1994). ISBN 0-387-90148-5. Corrected reprint of the 1976 original.

[23] V. Kapovitch, A. Petrunin, and W. Tuschmann, *Non-negative pinching, moduli spaces and bundles with infinitely many souls*, J. Differential Geom. **71** (2005), no. 3, 365–383.

[24] S. Kolasiński, *Geometric Sobolev-like embedding using high-dimensional Menger-like curvature*, Trans. Amer. Math. Soc. **367** (2015), no. 2, 775–811.

[25] S. Kolasiński, P. Strzelecki, and H. von der Mosel, *Characterizing $W^{2,p}$ submanifolds by p-integrability of global curvatures*, Geom. Funct. Anal. **23** (2013), no. 3, 937–984.

[26] J. Langer, *A compactness theorem for surfaces with L_p-bounded second fundamental form*, Math. Ann. **270** (1985), no. 2, 223–234.

[27] P. Mattila, Geometry of Sets and Measures in Euclidean Spaces, Vol. 44 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge (1995). ISBN 0-521-46576-1; 0-521-65595-1. Fractals and rectifiability.

[28] A. Nabutovsky, *Non-recursive functions, knots "with thick ropes", and self-clenching "thick" hyperspheres*, Comm. Pure Appl. Math. **48** (1995), no. 4, 381–428.

[29] S. Peters, *Cheeger's finiteness theorem for diffeomorphism classes of Riemannian manifolds*, J. Reine Angew. Math. **349** (1984), 77–82.

[30] S. Peters, *Convergence of Riemannian manifolds*, Compositio Math. **62** (1987), no. 1, 3–16.

[31] A. Petrunin and W. Tuschmann, *Diffeomorphism finiteness, positive pinching, and second homotopy*, Geom. Funct. Anal. **9** (1999), no. 4, 736–774.

[32] M. Pilipczuk and M. Pilipczuk, *Private communnication* (February 2015).

[33] G. B. Price, *On the completeness of a certain metric space with an application to Blaschke's selection theorem*, Bull. Amer. Math. Soc. **46** (1940), 278–280.

[34] G. Smith, *An Arzela-Ascoli theorem for immersed submanifolds*, Ann. Fac. Sci. Toulouse Math. (6) **16** (2007), no. 4, 817–866.

[35] G. Smith, *Compactness results for immersions of prescribed Gaussian curvature I — analytic aspects*, Adv. Math. **229** (2012), no. 2, 731–769.

[36] P. Strzelecki, M. Szumańska, and H. von der Mosel, *A geometric curvature double integral of Menger type for space curves*, Ann. Acad. Sci. Fenn. Math. **34** (2009), no. 1, 195–214.

[37] P. Strzelecki, M. Szumańska, and H. von der Mosel, *Regularizing and self-avoidance effects of integral Menger curvature*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **9** (2010), no. 1, 145–187.

[38] P. Strzelecki, M. Szumańska, and H. von der Mosel, *On some knot energies involving Menger curvature*, Topology Appl. **160** (2013), no. 13, 1507–1529.

[39] P. Strzelecki and H. von der Mosel, *On rectifiable curves with L^p-bounds on global curvature: self-avoidance, regularity, and minimizing knots*, Math. Z. **257** (2007), no. 1, 107–130.

[40] P. Strzelecki and H. von der Mosel, *Integral Menger curvature for surfaces*, Adv. Math. **226** (2011), no. 3, 2233–2304.

[41] P. Strzelecki and H. von der Mosel, *Tangent-point repulsive potentials for a class of non-smooth m-dimensional sets in $\mathbb{R}^n$. Part I: Smoothing and self-avoidance effects*, J. Geom. Anal. **23** (2013), no. 3, 1085–1139.

[42] X. Tolsa, Analytic Capacity, the Cauchy Transform, and Non-Homogeneous Calderón-Zygmund Theory, Vol. 307 of Progress in Mathematics, Birkhäuser/Springer, Cham (2014). ISBN 978-3-319-00595-9; 978-3-319-00596-6.

[43] W. Tuschmann, *Geometric diffeomorphism finiteness in low dimensions and homotopy group finiteness*, Math. Ann. **322** (2002), no. 2, 413–420.

[44] H. Whitney, Geometric Integration Theory, Princeton University Press, Princeton, N. J. (1957).

Max Planck Institute for Gravitational Physics
(Albert Einstein Institute)
Am Mühlenberg 1, D-14476 Golm, Germany
E-mail address: s.kolasinski@mimuw.edu.pl

Institute of Mathematics, University of Warsaw
Banacha 2, 02-097 Warsaw, Poland
E-mail address: p.strzelecki@mimuw.edu.pl

Institut für Mathematik, RWTH Aachen University
Templergraben 55, D-52062 Aachen, Germany
E-mail address: heiko@instmath.rwth-aachen.de

Received October 2, 2015