HUJI Summer Project - Energy estimates in Elasticity

Author: Florian N. Grün
Advisor: Cy Maor

July 2023

This work was produced under a summer research program at the Einstein Institute of Mathematics at Hebrew University of Jerusalem under guidance of Prof. Cy Maor. We would like to thank Noam Shalev for many fruitful discussions on the topic.

1 Introduction to Gamma convergence

The concept of Gamma (Γ) convergence, originally introduced by de Giorgi in 1975, and finds nowadays wide-ranging applications in analysis and Calculus of Variations.

Often, one encounters a functional F and a sequence of functionals F_n, somehow related to each other. The question is now how minimizers of F_n relate to minimizers of F and whether properties of one can be deduced from properties of the other.

That is, Γ-convergence characterizes this convergence of F_n to F and the *Fundamental theorem of Γ-convergence*(Theorem 1.11) gives convergence of the (approximate) minimizers x_n to a minimizer x of F. For practical purposes, it is then enough to approximately minimize F_n and then pass to the limit in order to obtain a minimizer for F.

Definition 1.1. (sequential) Let $F_n : X \to \bar{\mathbb{R}}$, where X is a metric space. We say it Γ-converges to $F : X \to \bar{\mathbb{R}}$, (denoted $F_n \xrightarrow{\Gamma} F$) if for any x:

(i) For any sequence $x_n \to x$ we have $F(x) \leq \liminf_{n\to\infty} F_n(x_n)$.

(ii) For every $x \in X$, there exists a "recovery" sequence $\tilde{x}_n \to x$ such that such that $F(x) \geq \limsup_n F_n(\tilde{x}_n)$.

Together they imply the existence of $\tilde{x}_n$ such that $F(x) = \lim_{n\to\infty} F_n(\tilde{x}_n)$.

Instead of showing the second condition directly, by the following lemma, it suffices to prove an "approximate" *limsup* inequality.

Lemma 1.2. *If for any $\varepsilon > 0$ there exists a sequence $(x_n^{(\varepsilon)}) \to x$ such that $F(x) \geq \limsup_n F_n(x_n) - \varepsilon$, then there exists some sequence $(\tilde{x}_n) \to x$ such that $F(x) \geq \limsup_n F_n(\tilde{x}_n)$.*

Proof. The proof is based on a double diagonal argument. Take a sequence of sequences $(x_n)^k$ such that (x_n^k denotes the n'th entry of the k'th sequence)

$$\begin{aligned} &F(x) \geq \limsup_n F_n(x_n^k) - \frac{1}{k} \iff \\ &\limsup_n \big(F_n(x_n^k) - F(x)\big) = \limsup_n F_n(x_n^k) - F(x) \leq \frac{1}{k}. \end{aligned} \tag{1.1}$$

Now for all k, there exists N_k, such that whenever $n \geq N_k$ we have

$$F_n(x_n^k) - F(x) \leq \sup_{n \geq N_k} F_n(x_n^k) - F(x) \leq \frac{1}{k}, \tag{1.2}$$

but also $d(x_n^k, x) \leq \frac{1}{k}$. Construct now the sequence

$$(\tilde{x}_n) = (x_{N_1}^1, x_{N_2}^2, x_{N_3}^3, ...).$$

Clearly by construction $d(\tilde{x}_n, x) = d(x_{N_n}^n, x) \leq \frac{1}{n}$, so $\tilde{x}_n \to x$. Moreover,

$$\begin{aligned} \limsup_n F_n(\tilde{x}_n) - F(x) = \limsup_n \left(F_n(x_{N_n}^n) - F(x)\right) &= \lim_{n\to\infty} \sup_{j \geq n} \left(F_n(x_{N_j}^j) - F(x)\right) \\ &\leq \lim_{n\to\infty} \sup_{j\geq n} \frac{1}{j} = \lim_{n\to\infty} \frac{1}{n} = 0, \end{aligned} \tag{1.3}$$

by using (1.2) for the inequality. Hence $\limsup_n F_n(\tilde{x}_n) \leq F(x)$. □

Proposition 1.3. *We have equivalently*

$$F(x) = \inf_{(x_n)} \liminf_{n\to\infty} F_n(x_n) = \inf_{(x_n)} \limsup_{n\to\infty} F_n(x_n) \tag{1.4}$$

where the infimum and minimum are taken over all sequences x_n converging to x.

Proof. We first show that the definition implies (1.4).

$$\begin{aligned} &F(x) \leq \liminf_{n\to\infty} F_n(x_n) \leq \limsup_{n\to\infty} F_n(x_n) \\ &F(x) \leq \inf_{(x_n)} \liminf_{n\to\infty} F_n(x_n) \leq \inf_{(x_n)} \limsup_{n\to\infty} F_n(x_n) \leq \limsup_{n\to\infty} F_n(\tilde{x}_n) \leq F(x), \end{aligned} \tag{1.5}$$

and the infima are attained by taking the recovery sequence $\tilde{x}_n$. For the converse, (i) is trivially satisfied since, $F(x) = \inf_{(x_n)} \liminf_{n\to\infty} F_n(x_n) \leq \liminf_{n\to\infty} F_n(x_n)$ for any sequence x_n. Concerning (ii), by definition of the infimum, for any ε take $(x_n^{(\varepsilon)})$ such that $\limsup F_n(x_n^{(\varepsilon)} - \varepsilon \leq F(x)$, then we apply Lemma 1.2, to get $\tilde{x}_n$ such that $F(x) \geq \limsup F_n(\tilde{x}_n)$. □

Corollary 1.4. *As the infima are actually attained,*

$$F(x) = \min_{(x_n)} \liminf_{n\to\infty} F_n(x_n) = \min_{(x_n)} \limsup_{n\to\infty} F_n(x_n).$$

Definition 1.5. (topological) The general definition of Γ-convergence can be expressed in any topological space X in the following: Let $F_n : X \to \bar{\mathbb{R}}$ and denote by $\mathcal{N}(x)$ the set of all neighborhoods around x. We say F_n Γ-converges to $F : X \to \bar{\mathbb{R}}$, (denoted $F_n \xrightarrow{\Gamma} F$) if for any x:

(i) $F_{inf}(x) = \sup_{U\in\mathcal{N}(x)} \liminf_{n\to\infty} \inf_{y\in U} F_n(y)$

(ii) $F_{sup}(x) = \sup_{U\in\mathcal{N}(x)} \limsup_{n\to\infty} \inf_{y\in U} F_n(y)$

are equal, i.e. $F_{inf}(x) = F_{sup}(x) =: F(x)$.

Proposition 1.6. *The two definitions are equivalent if X is a metric space.*

Proof. Assume that $F_n \xrightarrow{\Gamma} F$, i.e. in both definitions there is equality. In particular for a metric space we can express the topological definition as

$$F(x) = \sup_{U \in \mathcal{N}(x)} \liminf_{n\to\infty} \inf_{y\in U} F_n(y) = \sup_{N\in\mathbb{N}} \liminf_{n\to\infty} \inf_{d(x,y)<\frac{1}{N}} F_n(y) = \lim_{N\to\infty} \liminf_{n\to\infty} \inf_{d(x,y)<\frac{1}{N}} F_n(y), \quad (1.6)$$

using that $\liminf_{n\to\infty} \inf_{d(x,y)<\frac{1}{N}} F_n(y)$ is an increasing function in N as the infimum is taken over a smaller set. We will show that

$$F^1(x) := \inf_{(x_n)\to x} \liminf_n F_n(x_n) = \sup_{N\in\mathbb{N}} \liminf_{n\to\infty} \inf_{d(x,y)<\frac{1}{N}} F_n(y) =: F^2(x), \quad (1.7)$$

similar arguments can be used for the *limsup* inequality.

Suppose that $F^2(x) > F^1(x)$, then also $F^2(x) > F^1(x) + 2\delta$, for some $\delta > 0$. Take a sequence $\bar{x}_n$ such that $\inf_{(x_n)\to x} \liminf_n F_n(x_n) \geq \liminf_n F_n(\bar{x}_n) - \delta$, then

$$\begin{aligned} F^2(x) > F^1(x) + 2\delta &= \inf_{(x_n)\to x} \liminf_n F_n(x_n) + 2\delta \\ &\geq \liminf_n F_n(\bar{x}_n) + \delta \geq \liminf_n \inf_{d(x,y)<\frac{1}{N}} F_n(y) + \delta, \end{aligned} \quad (1.8)$$

for any $N \in \mathbb{N}$, since $\bar{x}_n \to x$, so eventually $\bar{x}_n \in B_{1/N}(x)$ for any N. However passing to the limit in N,

$$\lim_N F^2(x) = \lim_N F^2(x) \geq \lim_N \liminf_n \inf_{d(x,y)<\frac{1}{N}} F_n(y) + \delta = F^2(x) + \delta, \quad (1.9)$$

leads to a contradiction, i.e. $F^2(x) \leq F^1(x)$.

Suppose now that $F^1(x) > F^2(x)$, then also $F^1(x) > F^2(x) + 3\delta$ for some $\delta > 0$. Pick $\bar{N}$ sufficiently large such that whenever $N \geq \bar{N}$ we have,

$$F^2(x) = \lim_N \liminf_n \inf_{d(x,y)<\frac{1}{N}} F_n(y) > \liminf_n \inf_{d(x,y)<\frac{1}{N}} F_n(y) - \delta. \quad (1.10)$$

Enumerate increasingly all $N \geq \bar{N}$ as $\{N_1, N_2, N_3, ...\} =: \mathcal{N}$, then

$$F^1(x) > \lim_N \liminf_n \inf_{d(x,y)<\frac{1}{N}} F_n(y) + 3\delta \geq \liminf_n \inf_{d(x,y)<\frac{1}{N_i}} F_n(y) + 2\delta \qquad \forall N_i \in \mathcal{N}. \quad (1.11)$$

Pick now by definition of the infimum for each N_i and n an element $x_n^{N_i}$ in $B_{1/N_i}(x)$ such that $\inf_{d(x,y)<\frac{1}{N_i}} F_n(y) \geq F_n(x_n^{N_i}) - \delta$, we have (using $a_n \leq b_n \implies \liminf_n a_n \leq \liminf_n b_n$)

$$F^1(x) > \liminf_n \inf_{d(x,y)<\frac{1}{N_i}} F_n(y) + 2\delta \geq \liminf_n \left(F_n(x_n^{N_i}) - \delta\right) + 2\delta = \liminf_n F_n(x_n^{N_i}) + \delta \quad \forall N_i \in \mathcal{N}. \quad (1.12)$$

Take now $(\bar{x}_n)$ as the diagonal sequence $(x_n^{N_n})$, since $N_n \to \infty$ if $n \to \infty$, we have $d(x, \bar{x}_n) \to 0$, i.e. $\bar{x}_n \to x$. Hence

$$F^1(x) > \liminf_n F_n(x_n^{N_n}) + \delta = \liminf_n F_n(\bar{x}_n) + \delta \geq \inf_{(x_n)} \liminf_n F_n(x_n) + \delta = F^1(x) + \delta, \quad (1.13)$$

a contradiction and thus $F^1(x) \leq F^2(x)$, which finishes the proof. □

Lemma 1.7. *If the* Γ*-limit of* F_n *exists, it is unique.*

Proof. Suppose it is not unique, i.e. $F_n \xrightarrow{\Gamma} F$ and $F_n \xrightarrow{\Gamma} F'$ such that $F(x) \neq F'(x)$ for some x. Without loss of generality assume $F(x) < F'(x)$. Let x_n be the recovery sequence for F, namely $F_n(x_n) \to F(x)$. Thus

$$F(x) < F'(x) \leq \liminf_n F_n(x_n) = \lim_n F_n(x_n) = F(x), \tag{1.14}$$

a contradiction, hence the Γ-limit is unique. Note that the recovery sequence is not necessarily unique. □

Proposition 1.8. *Suppose that* $F_n \xrightarrow{\Gamma} F$*, then* F *is lower semicontinuous.* *Proof*

Remark 1.9. We note that for a constant sequence $F_n = F$, the Γ-limit equals F only if F itself is lower semicontinuous. Otherwise the Γ-limit is the lower semicontinous envelope $\tilde{F}$ of F given by

$$\tilde{F}(x) := \min\{\liminf_n F(x_n) : x_n \to x\} = \sup\{\psi : X \to \bar{\mathbb{R}} : \psi \leq F \text{ and } \psi \text{ is l.s.c.}\}. \tag{1.15}$$

Remark 1.10. In contrast to usual limits in analysis,

$$F_n \xrightarrow{\Gamma} F \text{ and } G_n \xrightarrow{\Gamma} G \;\not\Longrightarrow\; (F_n + G_n) \xrightarrow{\Gamma} (F + G). \tag{1.16}$$

However, for G continuous, we have

$$F_n \xrightarrow{\Gamma} F \implies (F_n + G) \xrightarrow{\Gamma} (F + G). \tag{1.17}$$

We can now state the *Fundamental theorem of* Γ*-convergence*, giving a realtion between minimizers of F_n and F.

Theorem 1.11. *Suppose* $F_n \xrightarrow{\Gamma} F$ *in* (X, d) *and that for some compact set* $K \subset X$*,*

$$\inf_X F_n = \inf_K F_n = \min_K F_n \qquad \forall n, \tag{1.18}$$

that is, there is a collection of minimizers $y_n \in X$ *to* F_n *that is localized in* X*. Then* F *has at least one minimizer in X. In particular for any sequence* x_n *with*

$$\lim_n |F_n(x_n) - \inf_X F_n| = 0, \tag{1.19}$$

there is a subsequence that converges to a minimizer of F*.*

Proof. Take any sequence $x_n \in K$ satisfying (1.19). (By taking e.g. $x_n = y_n = \operatorname{argmin}_K F_n$, existence of such a sequence is guaranteed.) Since K is compact, up to a subsequence, $\lim_n x_n = x \in K$. Let $\bar{x} \in X$ and $\bar{x}_n$ its recovery sequence (wrt. $F_n \xrightarrow{\Gamma} F$). Note that

$$\begin{aligned} F_n(\bar{x}_n) &\geq \inf_X F_n \\ \implies F_n(\bar{x}_n) &\geq \inf_{k \geq n} \inf_X F_k \\ \implies \lim_n F_n(\bar{x}_n) &\geq \lim_n \inf_{k \geq n} \inf_X F_k = \liminf_n (\inf_X F_n). \end{aligned} \tag{1.20}$$

Hence

$$F(x) \leq \liminf_n F_n(x_n) \leq \liminf_n (\inf_X F_n) \leq \lim F_n(\bar{x}_n) = F(\bar{x}), \tag{1.21}$$

and so x is a minimizer of F. $\square$

Corollary 1.12. *For any minimizer x of F, there exists a sequence $x_n \to x$, satisfying* (1.19)*.*

Proof. Taking $\bar{x} = x$ in the chain of inequalities, we get $\min_X F = F(x) = \lim_n(\inf_X F_n)$. Let x be a minimizer of F and take its recovery sequence x_n. Then $\lim_n F_n(x_n) = F(x) = \lim_n(\inf_X F_n)$. $\square$

Remark 1.13. The compactness condition $\inf_X F_n = \inf_K F_n$ is satisfied if F_n are equi-coercive and $\sup_n \inf_X F_n \leq C < \infty$. The last condition is equivalent to $F(x) < \infty$, by condition .Let $K = K_{C+1}$ and so for any n, the sublevel set is not empty and

$$\varnothing \neq \{x : F_n(x) \leq C + 1\} \subseteq K_{C+1} = K.$$

Hence it is equivalent whether taking the infimum over K or X.

Example 1.14. The compactness condition in the fundamental theorem of Γ-convergence is necessary. Otherwise, let $F_n(x) = 1 - e^{-(x-n)^2}$. It follows that $F_n \xrightarrow{\Gamma} F \equiv 1$, as for any fixed x and $x_n \to x$,

$$\liminf_n F_n(x_n) = \liminf_n F_n(x) = \lim_n F_n(x_n) = \lim F_n(x) = 1. \tag{1.22}$$

However $\inf_X F_n = 0$ attained at $x = n$ and so for any sequence satisfying (1.19), $x_n = n + o(1)$, i.e. it can not have a converging subsequence. Moreover, $x = 0$, a minimizer of F can not be written as the limit of any such sequence.

2 Quantitative rigidity estimates

The celebrated Korn inequality is a first form of an rigidity estimate. More precisely, given a vector field v, it relates the distance of the gradient ∇v to the linear subspace Skew_N to the distance of ∇v to a specific skew-symmetric matrix A, which is some sort of average of ∇v.

Theorem 2.1. (Korn's inequality) *Let $\Omega \subset \mathbb{R}^N$ be an open, bounded, connected Lipschitz domain. Then for any $v \in H^1(\Omega, \mathbb{R}^N)$, there exists $A = A(v) \in \mathrm{Skew}_N$ such that*

$$\int_\Omega |\nabla v(x) - A|^2 dx \leq C_\Omega \int_\Omega |\operatorname{Sym} \nabla v(x)|^2 dx. \tag{2.1}$$

Lemma 2.2. *We note that for $A = \mathrm{Skew} \fint_\Omega \nabla v dx$.*

$$\mathrm{dist}^2_{L^2}(\nabla v, \mathrm{Skew}_N) = \inf_{M \in \mathrm{Skew}_N} \left(\int_\Omega |\nabla v(x) - A|^2 \right) = \int_\Omega |\nabla v(x) - A|^2. \tag{2.2}$$

Proof. This follows a standard variational argument. Suppose $A \in \mathrm{Skew}_N$ is optimal, i.e.

$$\int_\Omega |\nabla v(x) - A|^2 \leq \int_\Omega |\nabla v(x) - (A + \varepsilon B)|^2 \qquad \forall \varepsilon \in \mathbb{R}, B \in \mathrm{Skew}_N. \tag{2.3}$$

Expanding leads to

$$\int_\Omega |\nabla v(x) - A|^2 \leq \int_\Omega |\nabla v(x) - A|^2 dx + 2\varepsilon \int_\Omega \langle \nabla v(x) - A, B \rangle dx + \varepsilon^2 \int_\Omega |B|^2 dx. \tag{2.4}$$

Divide by ε and then let $\varepsilon \to 0$ gives

$$0 = \int_\Omega \langle \nabla v(x) - A, B \rangle dx = \langle \int_\Omega \nabla v(x) - A dx, B \rangle. \tag{2.5}$$

Hence $\fint_\Omega \nabla v(x) - A dx$ is orthogonal to the subspace Skew_N, i.e. equal to some element S of Sym_N and so

$$0 = \mathrm{Skew}(S) = \mathrm{Skew}\left(\fint_\Omega \nabla v(x) - A dx \right) = \mathrm{Skew}\left(\fint_\Omega \nabla v(x) dx \right) - \mathrm{Skew}(A) = \mathrm{Skew}\left(\fint_\Omega \nabla v(x) dx \right) - A, \tag{2.6}$$

which finishes the proof. ∎

We also have the nonlinear version of Korn's inequality, measuring not only the distance to the tangent space (linear subspace) Skew(N) but to the nonlinear manifold SO(N).

Theorem 2.3. (Friesecke-James-Müller's inequality) *Let $\Omega \subset \mathbb{R}^N$ be an open bounded, connected Lipschitz domain. Then for any $u \in H^1(\Omega, \mathbb{R}^N)$ there exists $R \in \mathrm{SO}(N)$ s.t.*

$$\int_\Omega |\nabla u - R|^2 dx \leq C_\Omega \int_\Omega \mathrm{dist}^2(\nabla u, \mathrm{SO}(N)) dx. \tag{2.7}$$

Similar to the Korn inequality, we characterize the corresponding rotation matrix $R_* \in \mathrm{SO}(N)$, minimizing the LHS, can be given explicitly.

Proposition 2.4. *Decompose $A := \fint_\Omega \nabla u dx$ into its singular value decomposition $A = U\Sigma V^T$ with $U, V \in \mathrm{O}(N)$ and Σ the singular values of A, in decreasing order. We have the following two cases:*

(a) If $\det(A) \geq 0$, then $R_* = UV^T$.

(b) If $\det(A) < 0$, then $R_* = USV^T$, where $S = \begin{pmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & -1 \end{pmatrix}$

Note that for repeated singular values the SVD is not unique, i.e. there are several solutions for R_*.

Proof. We write the LHS as

$$\int_\Omega |\nabla u - R|^2 dx = \int_\Omega \langle \nabla u - R, \nabla u - R\rangle dx = \int_\Omega \langle \nabla u, \nabla u\rangle dx - 2\int_\Omega \langle \nabla u, R\rangle dx + \int_\Omega \langle R, R\rangle dx, \quad (2.8)$$

thus R_* is solution of

$$\min_{R\in \mathrm{SO}(N)} \int_\Omega |\nabla u - R|^2 dx \Leftrightarrow \max_{R\in \mathrm{SO}(N)} \int_\Omega \langle \nabla u, R\rangle dx \ \Leftrightarrow \max_{R\in \mathrm{SO}(N)} \langle A, R\rangle = \max_{R\in \mathrm{SO}(N)} \mathrm{Tr}(A^T R). \quad (2.9)$$

Note that case $\det(A) = 0$ follows by continuity from case where $\det(A) > 0$. We first check that R_* is in $\mathrm{SO}(N)$. In the case (a), $\det(U) = \det(V) = \pm 1$ so $\det(R_*) = 1$ and $R_* R_*^T = UV^T(UV^T) = UV^T VU^T = I$. In the case (b), either U and V have determinant of different sign (since $0 < \det(A) = \det(U)\det(\Sigma)\det(V)$), hence $\det(R_*) = \det(U)\det(S)\det(V) = 1(-1)^2 = 1$ and also $R_* R_*^T = USV^T(USV^T)^T = US^2U^T = I$. It remains to show that $\mathrm{Tr}(A^T R_*) \geq \mathrm{Tr}(A^T R)$ for any $R \in \mathrm{SO}(N)$.

In case (a) this reduces to showing

$$\mathrm{Tr}(\Sigma) = \mathrm{Tr}(A^T R_*) \geq \mathrm{Tr}(A^T R) = \mathrm{Tr}(Q\Sigma) \qquad \forall Q \in \mathrm{SO}(N), \quad (2.10)$$

which holds as

$$(\Sigma Q)_{jj} = \sum_{i=1}^N \Sigma_{j,i} Q_{i,j} = \Sigma_{j,j} Q_{j,j} \leq \Sigma_{jj} = \sigma_j, \quad (2.11)$$

since each entry of the orthonormal column vector Q_j is less than 1.

In the case (b), we need to show

$$-\sigma_n + \sum_{i=1}^{n-1} \sigma_i = \mathrm{Tr}(\Sigma S) = \mathrm{Tr}(V\Sigma U^T USV^T) = \mathrm{Tr}(A^T R_*) \geq \mathrm{Tr}(A^T R) = \mathrm{Tr}(Q\Sigma) \quad \forall Q \in \mathrm{O}(N)\backslash SO(N). \quad (2.12)$$

By von Neumann's trace inequality (7.6),

$$\sum_{i=1}^n \sigma_i + \mathrm{Tr}(Q\Sigma) = \mathrm{Tr}(\Sigma) + \mathrm{Tr}(Q\Sigma) = \mathrm{Tr}(\Sigma(Q+I)) \leq \sum_{i=1}^n \sigma_i^{(\Sigma)} \sigma_i^{(Q+I)} \leq 2\sum_{i=1}^{n-1} \sigma_i. \quad (2.13)$$

using also that $|\sigma_i^{(Q+I)}| = |\lambda_i(Q) + 1| \leq |\lambda_i(Q)| + 1 = 2$ for $i \neq n$. Note that Q has at least one real eigenvalue equal to -1. If they were all complex the complex conjugates are eigenvalues too, and thus the determinant, being the product of eigenvalues is equal to 1. Hence there exists at least one (or at least two if N even) real eigenvalues, whereof one has to be negative to match the condition $\det(Q) = -1$. This in turn gives

$$\mathrm{Tr}(Q\Sigma) \leq 2\sum_{i=1}^{n-1} \sigma_i - \sum_{i=1}^{n} \sigma_i = -\sigma_N + \sum_{i=1}^{n-1} \sigma_i, \tag{2.14}$$

which finishes the proof. $\square$

Proposition 2.5. *Assume that for* $A \in \mathbb{R}^{N\times N}$, , *where* $\sigma_1, ...\sigma_N$ *are the singular values of* A.

(a) *If* $\det A \geq 0$, *then* $\mathrm{dist}^2(A, \mathrm{SO}(N)) = \sum_{i=1}^{N} (\sigma_i - 1)^2$.

(b) *If* $\det A < 0$, *then* $\mathrm{dist}^2(A, \mathrm{SO}(N)) = -\sigma_n + \sum_{i=1}^{n-1} (\sigma_i - 1)^2$.

3 Alternative settings for Korn and FJM inequalities

It is well known that both the Korn inequality and the FJM estimate extend to $p \in (1, \infty)$. For the boundary cases however, there are counterexamples and finer estimates, which will be presented in this chapter.

3.1 The case L^∞

A classical counterexample for the Korn inequality is given in [BD12], we adapt it slightly to use the method from [CFM05] and to arrive at a counterexample for the non-linear FJM case.

The linear case

We look for a sequence of functions $v_k \in W^{1,\infty}$ such that

$$\min_{A \in \mathrm{Skew}(N)} \|\nabla v_k - A\|_{L^\infty(\Omega)} \geq kC \|\nabla v_k + \nabla v_k^T\|_{L^\infty(\Omega)}, \tag{3.1}$$

thus contradicting Korn's inequality for $p = \infty$.
Let $\Omega = B_1(0) \subset \mathbb{R}^2$ and $\mathrm{Skew}(2) \ni Q = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, we define $v_k : \Omega \to \mathbb{R}^N$ as

$$v_k(x) = \begin{cases} Qx \log(|x|) & \text{if } |x| \geq e^{-k} \\ kQx & \text{if } |x| < e^{-k} \end{cases}. \tag{3.2}$$

For the gradient

$$\nabla v_k(x) = \begin{cases} Q \log(|x|) + Qx \otimes \frac{1}{|x|^2} x & \text{if } |x| \geq e^{-k} \\ kQ & \text{if } |x| < e^{-k} \end{cases}, \tag{3.3}$$

we have the estimate $2k \geq \|\nabla v_k\|_{L^\infty} \geq k$, since $|Q| = 1$ and $|Qx \otimes \frac{1}{|x|^2} x| \leq |Q|$. This in turn gives for any $A \in \mathbb{R}^{N \times N}$, $\|\nabla v_k - A\| \geq \frac{k}{3}$ because

$$\begin{aligned} &\text{if } |A| \leq \frac{k}{2}: \quad \|\nabla v_k - A\|_{L^\infty} \geq \|\nabla v_k\|_{L^\infty} - |A| \geq \frac{k}{3}, \\ &\text{if } |A| > \frac{k}{2}: \quad \|\nabla v_k - A\|_{L^\infty} \geq \|Q - A\| \geq |A| - |Q| \geq \frac{k}{3}. \end{aligned} \tag{3.4}$$

However, the symmetric part stays bounded as

$$\|\nabla v_k + \nabla v_k^T\|_{L^\infty(\Omega)} = \|Qx \otimes \frac{1}{|x|^2} x + (Qx \otimes \frac{1}{|x|^2} x)^T\|_{L^\infty(\Omega)} = 2|Q|. \tag{3.5}$$

We conclude

$$\min_{A \in \mathrm{Skew}(N)} \|\nabla v_k - A\|_{L^\infty} \geq \min_{A \in \mathbb{R}^{N \times N}} \|\nabla v_k - A\|_{L^\infty} \geq \frac{k}{3} = \frac{k}{3}|Q| = \frac{k}{6} \|\nabla v_k + \nabla v_k^T\|_{L^\infty}. \tag{3.6}$$

The non-linear case

We are working in a small energy regime where the scaling is chosen appropriately. We look for a sequence of functions $v_k \in W^{1,\infty}$ such that

$$\min_{R\in \mathrm{SO}(N)} \|\nabla v_k - R\|_{L^\infty(\Omega)} \geq kC\| \operatorname{dist}(\nabla v_k, SO(N))\|_{L^\infty(\Omega)}, \tag{3.7}$$

Using the function v_k from the linear case, we set

$$u_k(x) := x + \varepsilon_k + v_k(x), \qquad \nabla u_k(x) = I + \varepsilon_k \nabla v_k(x). \tag{3.8}$$

Now linearizing SO(2) by a Taylor approximation gives for any matrix $R \in R^{N\times N}$

$$\operatorname{dist}(I + R, \mathrm{SO}(2)) \leq \frac{1}{2}|R + R^T| + c|R|^2. \tag{3.9}$$

Thus

$$\begin{aligned}\| \operatorname{dist}(\nabla u_k, \mathrm{SO}(2))\|_{L^\infty} = \| \operatorname{dist}(I + \varepsilon_k \nabla v_k, \mathrm{SO}(2))\|_{L^\infty} &\leq \frac{\varepsilon_k}{2}\|\nabla v_k + \nabla v_k^T\|_{L^\infty} + \varepsilon_k^2 c\|\nabla v_k\|^2_{L^\infty} \\ &\leq \frac{\varepsilon_k}{2}|Q| + \varepsilon_k^2 c(2k)^2 \leq \varepsilon_k|Q| = \varepsilon_k\|\nabla v_k + \nabla v_k^T\|_{L^\infty},\end{aligned} \tag{3.10}$$

if we choose e.g. $\varepsilon_k := \frac{1}{k^3}$ and k sufficiently large. Using the estimate from the linear case (i.e. reverse Korn on ∇v_k),

$$\begin{aligned}\min_{R\in \mathrm{SO}(N)} \|\nabla u_k - R\|_{L^\infty} \geq \min_{R\in \mathbb{R}^{N\times N}} \|\nabla u_k - R\|_{L^\infty} &\geq \min_{A\in \mathbb{R}^{N\times N}} \varepsilon_k\|\nabla v_k - A\|_{L^\infty} \\ &\geq \frac{k}{6}\varepsilon_k\|\nabla v_k + \nabla v_k^T\|_{L^\infty} \geq \frac{k}{6}\| \operatorname{dist}(\nabla u_k, SO(N))\|_{L^\infty}.\end{aligned} \tag{3.11}$$

3.2 The case L^1

A detailed explanation of the counterexample can be found in [CFM05].

3.3 Weak L^p estimates

Despite the counterexamples in L^1 of the previous section, there is some room for improvement.

Definition 3.1. A measurable function f is in *weak*-$L^1(\Omega)$ (denoted also as $w - L^1$) if

$$\sup_{t>0} t|\{x \in \Omega : |f(x)| > t\}| =: \|f\|_{w-L^1} < \infty. \tag{3.12}$$

Lemma 3.2. *It follows directly that $L^1(\Omega) \subsetneq w - L^1(\Omega)$, that is for any f, $\|f\|_{w-L^1} \leq \|f\|_{L^1}$.*

Proof. This follows by Markov's inequality or more explicitly,

$$\int_\Omega |f(x)|dx = \int_{|f(x)|\leq t} |f(x)|dx + \int_{|f(x)|>t} |f(x)|dx \geq \int_{|f(x)|>t} tdx = t|\{|f(x)| > t\}|. \tag{3.13}$$

Then taking the supremum in t gives the inequality. The inclusion is strict, e.g. $f(x) = 1/x$ is not in $L^1[0, \infty]$ but in $w - L^1[0, \infty]$. □

Remark 3.3. Note that $\|\cdot\|_{w-L^1}$ is not a full norm, but only a quasi-norm, i.e. instead of the triangle inequality, $\|f+g\|_{w-L^1} \leq C(\|f\|_{w-L^1} + \|g\|_{w-L^1})$.

It is possible to achieve estimates in weak L^p space ($p \in (1,\infty)$), by Corollary 4.1 in [CDM12]. For the weak L^1 space we have the Korn and the FJM inequality in the following sense, see [CS06]:

Theorem 3.4. (weak-L^1 Korn) *For any $u \in W^{1,1}(\Omega, \mathbb{R}^N)$ we have*

$$\min_{A \in \mathrm{Skew}(N)} \|\nabla u^T - Au\|_{w-L^1(\Omega)} \leq C\|\nabla u^T + \nabla u\|_{L^1(\Omega)} dx. \tag{3.14}$$

Proof. In Conti's papers e.g [CDM12] the result is stated, but no more explanation. □

Theorem 3.5. (weak-L^1 FJM) *For any $u \in W^{1,1}(\Omega, \mathbb{R}^N)$ we have*

$$\min_{R \in \mathrm{SO}(N)} \|\nabla u^T - Ru\|_{w-L^1(\Omega)} \leq C\|\operatorname{dist}(\nabla u, \mathrm{SO}(N))\|_{L^1(\Omega)}. \tag{3.15}$$

Proof. Assume first that $\nabla u \leq M$. If not, replace it by its Lipschitz truncation (7.11), take e.g. $\lambda \geq 2$. Then $\|\bar{u}\| \leq 2C$ and

$$\|\nabla u - \nabla \bar{u}\|_{L^1} \leq C \int_{|\nabla u|>2} |\nabla u| \leq 2C \int_{|\nabla u|>2} \operatorname{dist}(\nabla u, \mathrm{SO}(N)) \leq 2C \int_\Omega \operatorname{dist}(\nabla u, \mathrm{SO}(N)), \tag{3.16}$$

using that $|\nabla u| > 2$ implies that $\operatorname{dist}(\nabla u, \mathrm{SO}(N)) \geq \frac{|\nabla u|}{2}$, since $\mathrm{SO}(N) \subset \mathbb{S}^{N-1}$. If the estimate holds for $\bar{u}$, then

$$\min_{R \in \mathrm{SO}(N)} \|\nabla u - R\|_{L^1} \leq \min_{R \in \mathrm{SO}(N)} \|\nabla \bar{u} - R\|_{L^1} + \|\nabla u - \nabla \bar{u}\|_{L^1} \leq C\|\operatorname{dist}(\nabla u, \mathrm{SO}(N)\|_{L^1}. \tag{3.17}$$

Secondly, assume that $\|\operatorname{dist}(\nabla u, \mathrm{SO}(N)\|_{L^1} < \frac{1}{2}$. If not (i.e. $\|\operatorname{dist}(\nabla u, \mathrm{SO}(N)\|_{L^1} \geq \frac{1}{2}$) the estimate follows directly as

$$\min_{R \in \mathrm{SO}(N)} \|\nabla u - R\|_{L^1} \leq \|\nabla u - I\|_{L^1} \leq \||\nabla u| + N\|_{L^1} \leq |\Omega|(M+N) \leq C\frac{1}{2} \leq C\|\operatorname{dist}(\nabla u, \mathrm{SO}(N)\|_{L^1}. \tag{3.18}$$

By the standard FJM inequality for $p = 2$, for some $R \in \mathrm{SO}(N)$,

$$\|\nabla u - R\|_{L^2} \leq C\left(\int_\Omega \operatorname{dist}^2(\nabla u, \mathrm{SO}(N)\right)^{\frac{1}{2}} \leq C\left(\int_\Omega M \operatorname{dist}(\nabla u, \mathrm{SO}(N)\right)^{\frac{1}{2}} \leq C\|\operatorname{dist}(\nabla u, \mathrm{SO}(N))\|_{L^1}^{\frac{1}{2}}. \tag{3.19}$$

We assume now that $R = I$, otherwise replace u by $R^{-1}u$ and then in the last step multiply by R again. By a Taylor approximation,

$$\left|\frac{\nabla u + \nabla u^T}{2} - I\right| \leq c \operatorname{dist}(\nabla u, \mathrm{SO}(N)) + c|\nabla u - I|^2, \tag{3.20}$$

and combining the previous estimate with the fact that $\|\operatorname{dist}(\nabla u, \mathrm{SO}(N))\| \leq 1/2$, we have

$$\|\nabla u + \nabla u^T - 2I\|_{L^1} \leq C\|\operatorname{dist}(\nabla u, \mathrm{SO}(N))\|_{L^1}. \tag{3.21}$$

Now using the weak-L^1 Korn inequality (3.14) on the function $v = u - id$ ($\nabla v = \nabla u - I$) gives $S \in \mathrm{Skew}(N)$ with

$$\|\nabla u - I - S\|_{w-L^1} \leq C\|\nabla u + \nabla u^T - 2I\|_{L^1} \leq C\| \operatorname{dist}(\nabla u, \mathrm{SO}(N))\|_{L^1}. \tag{3.22}$$

Since $|S| \leq |\nabla u - I - S| + |\nabla u - I|$ and $\|\cdot\|_{w-L^1}$ being a quasi-norm that is upper bounded by $\|\cdot\|_{L^1}$ (and thus by Cauchy-Schwarz also by $\|\cdot\|_{L^2}$),

$$\begin{aligned}|S| = |\Omega|^{-1}\|S\|_{w-L^1} &\leq C\left(\||\nabla u - I - S|\|_{w-L^1} + \||\nabla u - I|\|_{w-L^1}\right)\\ &\leq C\left(\| \operatorname{dist}(\nabla u, \mathrm{SO}(N))\|_{L^1} + \||\nabla u - I|\|_{L^1}\right)\\ &\leq C\left(\| \operatorname{dist}(\nabla u, \mathrm{SO}(N))\|_{L^1} + \||\nabla u - I|\|_{L^2}\right)\\ &\leq C\left(\| \operatorname{dist}(\nabla u, \mathrm{SO}(N))\|_{L^1} + \| \operatorname{dist}(\nabla u, \mathrm{SO}(N))\|_{L^1}^{\frac{1}{2}}\right)\\ &\leq C\| \operatorname{dist}(\nabla u, \mathrm{SO}(N))\|_{L^1}^{\frac{1}{2}},\end{aligned} \tag{3.23}$$

where again we use that $A + A^{1/2} \leq 2A^{1/2}$ for $A \leq 1/2$. Now for $e^S \in SO(N)$,

$$\begin{aligned}\min_{R\in \mathrm{SO}(N)} \|\nabla u^T - Ru\|_{w-L^1(\Omega)} \leq \|\nabla u - e^S\|_{w-L^1} &= \|\nabla u - (I + S + \frac{1}{2}S^2 + h.o.t)\|_{w-L^1}\\ &\leq C\|\nabla u - I - S\|_{w-L^1} + C|S|^2 \leq C\| \operatorname{dist}(\nabla u, \mathrm{SO}(N))\|_{L^1},\end{aligned} \tag{3.24}$$

which finishes the proof. □

Since we do not have the complete L^1 estimate the best one can hope for is something slightly weaker.
I am wondering if we can also derive an estimate of this $x\log(1/x)$ form for L^∞...

Lemma 3.6. *For $f \in L^1(\Omega)$ we have $\int_\Omega |f|dx = \int_0^\infty |\{x \in \Omega : f(x) > t\}|dt$.*

Proof. We assume that f is a simple function, i.e. $|f| = \sum_{i=1}^n a_i\chi_{A_i}$ with the A_i's disjoint, $\Omega = \cup A_i$ and $0 \leq a_1 < a_2 < ...a_n$. Then

$$\int_\Omega |f|dx = \sum_{i=1}^n a_i|A_i| \quad , \tag{3.25}$$

$$\int_0^\infty |\{x : f(x) > t\}|dt = a_1\sum_{i=1}^n |A_i| + ... + (a_k - a_{k-1})\sum_{i=k}^n |A_i| + ... + (a_n - a_{n-1})|A_n| = \sum_{i=1}^n a_i|A_i|, \tag{3.26}$$

and by density of simple functions, the result holds for any $f \in L^1(\Omega)$. □

Corollary 3.7. *For any $u \in W^{1,1}(\Omega, \mathbb{R}^N)$ we have*

$$\min_{R\in \mathrm{SO}(N)} \|\nabla u^T - R\|_{L^1(\Omega)} \leq C\| \operatorname{dist}(\nabla u, \mathrm{SO}(N))\|_{L^1(\Omega)} \max\left(1, \log\left(\frac{1}{\| \operatorname{dist}(\nabla u, \mathrm{SO}(N))\|_{L^1(\Omega)}}\right)\right). \tag{3.27}$$

Proof. Let $R \in \mathrm{SO}(N)$ be optimal in the weak-L^1 estimate. Denote $\varepsilon := \|\operatorname{dist}(\nabla u, \mathrm{SO}(N))\|_{L^1}$, as usual the difficulty is showing the estimate when $\varepsilon < 1$. By the previous lemma,

$$\begin{aligned}
\|\nabla u - R\|_{L^1} &= \int_0^\infty |\{x : |\nabla u - R| > t\}| dt = \int_0^{4N} |\{|\nabla u - R| > t\}| dt + \int_{4N}^\infty |\{|\nabla u - R| > t\}| dt \\
&\leq \int_0^{4N} |\{x : |\nabla u - R| > t\}| dt + C\|\operatorname{dist}(\nabla u, \mathrm{SO}(N))\|_{L^1} \\
&\leq \int_0^{4N} \min\left(|\Omega|, \frac{1}{t}\|\nabla u - R\|_{w-L^1}\right) dt + C\|\operatorname{dist}(\nabla u, \mathrm{SO}(N))\|_{L^1} \\
&\leq \int_0^{4N} \min\left(|\Omega|, \frac{1}{t}\varepsilon\right) dt + C\varepsilon \\
&= \int_0^{\varepsilon|\Omega|^{-1}} |\Omega| dt + \int_{\varepsilon|\Omega|^{-1}}^{4N} \frac{1}{t}\varepsilon dt + C\varepsilon \\
&\leq C\varepsilon + \varepsilon(\log(4N) - \log(\varepsilon|\Omega|^{-1})) \leq C\varepsilon\left(1 + \log\left(\frac{1}{\varepsilon}\right)\right) \leq C\varepsilon \max\left(1, \log\left(\frac{1}{\varepsilon}\right)\right),
\end{aligned}$$

which finishes the proof. □

4 Preliminaries

The setting now changes slightly, instead of deforming a domain $\Omega \subset \mathbb{R}^n$ in $\mathbb{R}^n$ itself, now $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ for $m > n$. That is, the elastic energy involves not only stretching but also bending. The bending of objects is essentially how the normal space of $f(\Omega)$ changes.

Definition 4.1. We define the set (not a group any longer) of orthogonal $(N+k)$-by-N matricies as

$$\mathrm{O}(N+k, N) = \{A \in \mathbb{R}^{(N+k)\times N} : A^T A = I_{N\times N}\}. \tag{4.1}$$

Definition 4.2. A differentiable function between two manifolds $f : \mathcal{M} \to \mathcal{N}$ is an **immersion** if its differential, $D_p f : T_p\mathcal{M} \to T_{f(p)}\mathbb{N}$, is injective at every point $p \in \mathcal{M}$. In the case of $\mathcal{M} = \mathbb{R}^n$ and $\mathcal{N} = \mathbb{R}^{n+k}$ we require ∇f to be of rank n.

Definition 4.3. We define for an n-dimensional manifold $\mathcal{M}$ the set of immersions as

$$\mathrm{Imm}_p(\mathcal{M}, \mathbb{R}^m) := \{f \in W^{1,p}(\mathbb{M}, \mathbb{R}^m) : rk\nabla f = n \text{ a.e. }\}. \tag{4.2}$$

Remark 4.4. An immersion that is also a homeomorphism onto its image is called an embedding. Note that e.g. the lemniscate $\beta : (-\pi, \pi) \to \mathbb{R}^2$ given by $\beta(t) = (\sin 2t, \sin t)$ is an immersion, but not an embedding (the domain is open, while the image is closed). However each immersion is locally an embedding. In FJM-type estimates only local quantities are integrated, hence without loss of generality (apply partition of unity) we can assume that $f(\mathcal{M})$ is a Riemannian submanifold of $\mathbb{R}^m$.

Proposition 4.5. *Let $A \in \mathbb{R}^{m\times n}$ and $N = (n_1, ...n_{m-n}) \in \mathbb{R}^{m\times(m-n)}$ with $|n_i| = 1$ and $n_i \perp n_j$ for $i \neq j$. Suppose that $A \perp N$ (that is each column in A is orthogonal to each column in N) and* $\det(A|N) > 0$*, then*

$$\mathrm{dist}(A, \mathrm{O}(m, n)) = \mathrm{dist}((A|N), \mathrm{SO}(m)). \tag{4.3}$$

Proof. It suffices to show the equality for dist^2, that is

$$\min_{B\in \mathrm{O}(m,n)} \mathrm{Tr}\left((A-B)^T(A-B)\right) = \min_{R\in \mathrm{SO}(m)} \mathrm{Tr}\left(((A|N)-R)^T((A|N)-R)\right) \tag{4.4}$$

Suppose now that B and R are optimal, then

$$\begin{aligned} \mathrm{Tr}((A-B)^T(A-B)) &= \mathrm{Tr}(((A|N)-R)^T((A|N)-R)) \\ \mathrm{Tr}(A^TA) - 2\,\mathrm{Tr}(A^TB) + \mathrm{Tr}(B^TB) &= \mathrm{Tr}((A|N)^T(A|N)) - 2\,\mathrm{Tr}((A|N)^TR) + \mathrm{Tr}(R^TR) \\ \mathrm{Tr}(A^TA) - 2\,\mathrm{Tr}(A^TB) + n &= \mathrm{Tr}(A^TA) + (m-n) - 2\,\mathrm{Tr}((A|N)^TR) + m \\ -\mathrm{Tr}(A^TB) &= (m-n) - \mathrm{Tr}((A|N)^TR) \end{aligned} \tag{4.5}$$

Consider the SVD decompositions $A = U\Sigma V^T$ and $(A|N) = U'\Sigma' V'^T$. Since $(A|N)^T(A|N) = \begin{pmatrix} A^TA & 0 \\ 0 & I_{(m-n)\times(m-n)} \end{pmatrix}$, we get $\sigma_1 = ... = \sigma_{m-n} = 1$ and $\sigma_i = \sigma'_{i-(m-n)}$ for $i = m-n+1, , , m$. Here it is used that $A \perp N$ to obtain a block diagonal matrix with the exact same eigenvalues as A^TA, plus the additional eigenvalues of 1.

By Proposition 2.4, R is given as $R = UV^T$, so the RHS evaluates to

$$(m-n) - \sum_{i=1}^{m} \sigma_i' = \sum_{i=1}^{N} \sigma_i.$$

Thus it remains to show that $\max_{B\in \mathrm{O}(N+1,N)} \mathrm{Tr}(A^T B) = \sum_{i=1}^{N} \sigma_i$. We rewrite $B = UDV^T$, with $D \in \mathrm{O}(N+1,1)$, then

$$\mathrm{Tr}(A^T B) = \mathrm{Tr}(V\Sigma^T U^T UDV^T) = \mathrm{Tr}(\Sigma^T D) = \mathrm{Tr}\left(\begin{pmatrix} \sigma_1 & 0 & \dots & 0 & 0 \\ 0 & \sigma_2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sigma_N & 0 \end{pmatrix} D\right) = \sum_{i=1}^{N} \sigma_i, \quad (4.6)$$

since the optimal is achieved by choosing $D_{i,j} = \delta_{i,j}$. $\square$

Lemma 4.6. *Let $A \in \mathbb{R}^{m\times n}$ be of full rank ($\mathrm{rk}(A) = n$) and $N \in \mathrm{O}(m, m-n) \subset \mathbb{R}^{m\times(m-n)}$ such that $A \perp N$. Assume also that $\det(A|N) > 0$, then for the projection $\mathrm{O}(A)$ of A onto the manifold $\mathrm{O}(m,n)$ it also holds $\mathrm{O}(A) \perp N$. Moreover, $\det(A|N) \cdot \det(\mathrm{O}(A)|N) > 0$.*

Proof. Let v be a column of N. Using the singular value decomposition ($A = U\Sigma V^T$ with $U \in \mathrm{O}(m)$, $V \in \mathrm{O}(n)$ and $\Sigma \in \mathbb{R}^{m\times n}$ diagonal),

$$0 = A^T v = V\Sigma^T U^T v. \quad (4.7)$$

Since A is of full rank ($\sigma_i > 0$ for all i) and V invertible

$$0 = V^T 0 = \Sigma^T U^T v = \begin{pmatrix} \sigma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sigma_n & 0 & \dots & 0 \end{pmatrix} U^T v \implies (U^T v)_i = 0 \quad \text{for } i = 1, \dots, n. \quad (4.8)$$

Thereby also for $\mathrm{O}(A) = UI_{m\times n}V^T$, the projection of A onto $\mathrm{O}(m,n)$, we have $\mathrm{O}(A)^T v = 0$ and since v was arbitrary $\mathrm{O}(A) \perp N$. For the second part,

$$\begin{aligned} \det(A|N) \cdot \det(\mathrm{O}(A)|N) &= \det((A|N)^T (\mathrm{O}(A)|N)) \\ &= \det\left(\begin{pmatrix} A^T \\ N^T \end{pmatrix} (\mathrm{O}(A)|N)\right) \\ &= \det\left(\begin{pmatrix} A^T\,\mathrm{O}(A) & 0 \\ 0 & N^T N \end{pmatrix}\right) \\ &= \det(A^T\,\mathrm{O}(A)) \\ &= \det(V\Sigma^T U^T U I_{m\times n} V^T) \\ &= \det(\Sigma^T I_{m\times n}) > 0. \end{aligned} \quad (4.9)$$

That is the orientation of the frame does not change if part of it is projected. This should also hold for a general $O(g,e)$, maybe can use coordinates and the proof for the Euclidean case or maybe need some more geometrical argument. $\square$

5 The Euclidean case

We now prove that under the assumption that the initial and target domain are Euclidean (i.e. zero curvature) the distance of the transformation f to a rigid motion (rotation and translation) cis comparable to the elastic energy.

Theorem 5.1. *For $\Omega \subset \mathbb{R}^n$, let $f : \Omega \to \mathbb{R}^m$ and any unit vector field $v = (v_1, ..., v_{m-n}) : \Omega \to \mathbb{R}^{m\times(m-n)}$. Then there exists $R \in \mathrm{SO}(m)$ such that*

$$\int_\Omega |(\nabla f|v) - R|^2 \leq C \left(\int_\Omega \operatorname{dist}^2((\nabla f|v), \mathrm{SO}(m)) + \int_\Omega |\nabla v|^2 \right) =: CE_R(f, v), \tag{5.1}$$

where E_R denotes the relaxed elastic energy.

Proof. Let $\Omega_1 = \Omega \times (-1, 1)^{m-n}$ and define

$$F(q, t) = f(q) + \sum_{i=1}^{m-n} t_i v_i = f(q) + v \cdot t \qquad \text{for } (q_1, ..., q_n, t_1, ...t_{m-n}) \in \Omega \times (-1, 1)^{m-n}. \tag{5.2}$$

Use now the FJM-estimate on ∇F, i.e. there exists $R \in \mathrm{SO}(m)$ such that

$$\|\nabla F - R\|_{L^2(\Omega_1)} \leq C\| \operatorname{dist}(\nabla F, \mathrm{SO}(m))\|_{L^2(\Omega_1)}. \tag{5.3}$$

We now lowerbound the LHS and upper bound the RHS by the desired quantities from the theorem. First compute

$$\nabla F(q, t) = \nabla f(q) \oplus v(q) + \sum_{i=1}^{m-n} t_i \nabla v_i(q) \oplus 0 = \nabla f(q) \oplus v(q) + \nabla v(q) \cdot t \oplus 0. \tag{5.4}$$

Then,

$$\begin{aligned} \operatorname{dist}^2(\nabla F, \mathrm{SO}(m)) &= \min_{R\in \mathrm{SO}(m)} |\nabla F - R|^2 \\ &\leq \min_{R\in \mathrm{SO}(m)} |\nabla f \oplus v + \nabla v \cdot t \oplus 0 - R|^2 \\ &\leq \min_{R\in \mathrm{SO}(m)} 2|\nabla f \oplus v - R|^2 + 2|\nabla v \cdot t|^2 \\ &\leq 2 \min_{R\in \mathrm{SO}(m)} |\nabla f \oplus v - R|^2 + 2\sqrt{m}|\nabla v|^2 \\ &\leq c \left(\operatorname{dist}^2((\nabla f|v), \mathrm{SO}(m)) + |\nabla v|^2\right), \end{aligned} \tag{5.5}$$

thus we upper bound the RHS as

$$\begin{aligned} \|\nabla F - R\|_{L^2(\Omega_1)} &\leq C\| \operatorname{dist}(\nabla F, \mathrm{SO}(m))\|_{L^2(\Omega_1)} \\ &\leq C \left(\int_{\Omega_1} c \operatorname{dist}^2((\nabla f(q)|v), \mathrm{SO}(m)) + c|\nabla v(q)|^2 \right)^{\frac{1}{2}} \\ &= C \left(\int_\Omega \int_{(-1,1)^{m-n}} \operatorname{dist}^2 ((\nabla f(q)|v(q)), \mathrm{SO}(m)) + |\nabla v(q)|^2 dt dq \right)^{\frac{1}{2}} \\ &\leq C \left(\int_\Omega \operatorname{dist}^2((\nabla f(q)|v(q)), \mathrm{SO}(m)) + |\nabla v(q)|^2 dq \right)^{\frac{1}{2}} \\ &\leq C\| \operatorname{dist}((\nabla f|v), \mathrm{SO}(m))\|_{L^2(\Omega)} + C\|\nabla v\|_{L^2(\Omega)}. \end{aligned} \tag{5.6}$$

We now lower bound the LHS using $\|a+b\| \geq \|a\| - \|b\|$,

$$
\begin{aligned}
\|\nabla F - R\|_{L^2(\Omega_1)} &= \|\nabla f \oplus v - R + \nabla v \cdot t \oplus 0\|_{L^2(\Omega_1)} \\
&\geq \|\nabla f \oplus v - R\|_{L^2(\Omega_1)} - \|\nabla v \cdot t \oplus 0\|_{L^2(\Omega_1)} \\
&= \left(\int_{(-1,1)^{m-n}} |t|^2 dt + 1 \right)^{1/2} \left(\|\nabla f \oplus v - R\|_{L^2(\Omega)} - \|\nabla v\|_{L^2(\Omega)} \right) \\
&\geq C' \left(\|\nabla f \oplus v - R\|_{L^2(\Omega)} - \|\nabla v\|_{L^2(\Omega)} \right).
\end{aligned}
\tag{5.7}
$$

Combining the two estimates and moving terms,

$$
\begin{aligned}
C' \left(\|\nabla f \oplus v - R\|_{L^2(\Omega)} - \|\nabla v\|_{L^2(\Omega)} \right) &\leq C\| \operatorname{dist}(\nabla f, \mathrm{SO}(m))\|_{L^2(\Omega)} + C\|\nabla v\|_{L^2(\Omega)} \\
\implies \|(\nabla f|v) - R\|_{L^2(\Omega)} &\leq C \left(\| \operatorname{dist}(\nabla f, \mathrm{SO}(m))\|_{L^2(\Omega)} + \|\nabla v\|_{L^2(\Omega)} \right).
\end{aligned}
\tag{5.8}
$$

Upon squaring and using $(a+b)^2 \leq 2a^2 + 2b^2$, the proof is complete. □

Corollary 5.2. *Assume that* $v = N = (N_1, ..., N_{m-n})$*, a normal vector field, i.e.* $N \perp \nabla f$, $\|N_i\| = 1$ *for* $i = 1, ...m-n$ *and* $\det(\nabla f|n) > 0$*. Then*

$$
\min_{Q \in \mathrm{O}(m,n)} \int_\Omega |\nabla f - Q|^2 \leq C \left(\int_\Omega \operatorname{dist}^2(\nabla f, \mathrm{O}(m,n)) + \int_\Omega |\nabla N|^2 \right) =: C\tilde{E}_R(f, N), \tag{5.9}
$$

with the normal vector field N *only appearing in the bending energy.*

Proof. By Proposition 4.5 (giving $\operatorname{dist}((\nabla f|N), \mathrm{SO}(m)) = \operatorname{dist}(\nabla f, \mathrm{O}(m,n))$ pointwise), for the RHS

$$
\int_\Omega \operatorname{dist}^2((\nabla f|v), \mathrm{SO}(m)) + \int_\Omega |\nabla N|^2 = \int_\Omega \operatorname{dist}^2(\nabla f, \mathrm{O}(m,n)) + \int_\Omega |\nabla N|^2. \tag{5.10}
$$

By the theorem, we get an optimal R, since $R_{m\times n} \in \mathrm{O}(m,n)$,

$$
\min_{Q \in \mathrm{O}(m,n)} \int_\Omega |\nabla f - Q|^2 \leq \int_\Omega |\nabla f - R_{m\times n}|^2 \leq \int_\Omega |\nabla f - R|^2, \tag{5.11}
$$

which finishes the proof. □

Corollary 5.3. *Let* $v = (v_1, ..., v_{m-n}) : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ *be a vector field with* $\det(\nabla f|v) \geq c' > 0$ *in* Ω*. Then*

$$
\min_{Q \in \mathrm{O}(m,n)} \int_\Omega |\nabla f - Q|^2 \leq C \left(\int_\Omega \operatorname{dist}^2(\nabla f, \mathrm{O}(m,n)) + \int_\Omega |\nabla v|^2 \right) =: C\tilde{E}(f, v). \tag{5.12}
$$

Proof. We decompose $v = w + v_P$ with $w \perp \nabla f$ and $v_P \in \operatorname{span}(\nabla f)$. Since $\det(\nabla f|v) \geq c'$, this essentially means $|w| \geq c$ for some $c > 0$. We now define the normal vector field $N := \frac{w}{|w|}$, by the previous corollary and get

$$
\min_{Q \in \mathrm{O}(m,n)} \int_\Omega |\nabla f - Q|^2 \leq C \int_\Omega \operatorname{dist}^2(\nabla f, \mathrm{O}(m,n)) + \int_\Omega |\nabla N|^2. \tag{5.13}
$$

We now estimate $\nabla N = \nabla\left(\frac{w}{|w|}\right)$, a $m\times(m-n)\times n$ tensor. For simplicity, we denote $\frac{\partial}{\partial x_k}f = f_k$.

$$\begin{aligned}
\frac{\partial}{\partial x_k}\left(\frac{w^{ij}}{|w|}\right) &= \frac{\partial}{\partial x_k}\left(\frac{w^{ij}}{\sqrt{\sum_{a=1}^{m}\sum_{b=1}^{m-n}(w^{ab})^2}}\right) \\
&= w_k^{ij}\frac{1}{|w|} + w^{ij}\frac{\partial}{\partial x_k}\left(\sum_{a=1}^{m}\sum_{b=1}^{m-n}(w^{ab})^2\right)^{-1/2} \\
&= w_k^{ij}\frac{1}{|w|} - w^{ij}\left(\sum_{a=1}^{m}\sum_{b=1}^{m-n}(w^{ab})^2\right)^{-3/2}\sum_{a=1}^{m}\sum_{b=1}^{m-n} w^{ab}w_k^{ab} \\
&= \frac{1}{|w|}\left(w_k^{ij} - \frac{1}{|w|^2}w^{ij}\sum_{a=1}^{m}\sum_{b=1}^{m-n} w^{ab}w_k^{ab}\right),
\end{aligned} \tag{5.14}$$

$$\begin{aligned}
|\nabla N|^2 &= \sum_{i=1}^{m}\sum_{j=1}^{m-n}\left(\frac{\partial}{\partial x_k}\left(\frac{w^{ij}}{|w|}\right)\right)^2 \\
&= \frac{1}{|w|^2}\sum_{i=1}^{m}\sum_{j=1}^{m-n}\left(w_k^{ij} - \frac{1}{|w|^2}w^{ij}\sum_{a=1}^{m}\sum_{b=1}^{m-n} w^{ab}w_k^{ab}\right)^2 \\
&\le \frac{2}{|w|^2}\sum_{i=1}^{m}\sum_{j=1}^{m-n}\left((w_k^{ij})^2 + \frac{1}{|w|^4}(w^{ij})^2\left(\sum_{a=1}^{m}\sum_{b=1}^{m-n} w^{ab}w_k^{ab}\right)^2\right) \\
&\le C\frac{1}{|w|^2}\sum_{i=1}^{m}\sum_{j=1}^{m-n}\left((w_k^{ij})^2 + \frac{1}{|w|^4}|w|^2\sum_{a=1}^{m}\sum_{b=1}^{m-n} |w|^2(w_k^{ab})^2\right) \\
&\le C\frac{1}{|w|^2}\sum_{i=1}^{m}\sum_{j=1}^{m-n}\left((w_k^{ij})^2 + |\nabla w|^2\right) \\
&\le C\frac{1}{|w|^2}|\nabla w|^2.
\end{aligned} \tag{5.15}$$

Since $|\nabla v|^2 = |\nabla w|^2 + |\nabla v_P|^2$ and $|w|\ge c$,

$$\int_\Omega |\nabla N|^2 \le C\int_\Omega \frac{1}{|w|^2}|\nabla w|^2 \le \frac{C}{c}\int_\Omega |\nabla w|^2 \le C\int_\Omega |\nabla v|^2, \tag{5.16}$$

in (5.13) gives the desired bound. □

Corollary 5.4. (Reshetnyak stability) *Let* $(f_k, N_k)\in W^{1,2}(\Omega;\mathbb{R}^m)$ *such that* $\det(\nabla f_k|N_k) = 1$ *and* $\nabla f_k \perp N_k$. *Assume that* $\tilde{E}_R(f_k,N_k)\to 0$, *then there exists an isometry* f *(up to translation) and a normal frame* N *with* $(\nabla f|N)\in \mathrm{SO}(m)$ *and up to a subsequence,* $f_k \xrightarrow{W^{1,2}(\Omega)} f$ *and* $N_k \xrightarrow{W^{1,2}(\Omega)} N$.

Proof. For each k, let $R_k\in\mathrm{SO}(m)$ be the corresponding matrix from the theorem. Since $\mathrm{SO}(m)$ is compact, up to a subsequence R_k converges to some $R\in\mathrm{SO}(m)$. Set now $f(x) = R_{m\times(m-n)}x + b$ and $N = R_{m\times n}$, the last n components of R. □

6 The non-Euclidean case

We try to adapt the result for general codimension to the case where the domain in not only $\mathbb{R}^n$, but a general curved n-dimensional manifold $\mathcal{M}$ embedded in $\mathbb{R}^m$ for $m > n$. This encapsulates the setting where there is not only the deformation (stretching and bending) energy from a flat reference configuration, but also from an already deformed one. As an analogy (for codimension 1), imagine a part of a tire, its zero energy configuration is already curved (manufactured in that way), yet when trying to make it flat, it stretches and bends. Note that to avoid confusion with Riemannian connections, from now on df denotes ∇f, the usual Euclidean derivative in the ambient space $\mathbb{R}^m$. We also denote the standard Euclidean connection on $\mathbb{R}^m$ by $\bar{\nabla}$.

First we introduce O and SO on manifolds and recall some definitions from differential geometry of surfaces.

Definition 6.1. Given an n-dimensional Riemannian manifold $(\mathcal{M}, g)$, define the set

$$\mathrm{O}(g,e) = \mathrm{O}(g_q, e) = \{h : T_q\,\mathcal{M}(\simeq \mathbb{R}^n) \to \mathbb{R}^m : \quad h \text{ is an isometric linear map w.r.t } g_q\}. \tag{6.1}$$

Note that we can still think of h being a matrix in $\mathbb{R}^{m\times n}$, but now

$$u^T(g_{ij})v = g_{ij}u^iv^j = \langle u, v\rangle_g = \langle hu, hv\rangle_{\mathbb{R}^m} = (hu)^T(hv) = u^Th^Thv \implies (g_{ij}) = h^Th. \tag{6.2}$$

Since the metric g_q depends on the point $q \in \mathcal{M}$, $\mathrm{O}(g, e)$ also depends on $q \in \mathcal{M}$.

Definition 6.2. In the same spirit we define for a Riemannian manifold $(\mathcal{M}, g)$ of dimension m, the set

$$\mathrm{SO}(g,e) = \{h : T_q\,\mathcal{M}(\simeq \mathbb{R}^m) \to \mathbb{R}^m : \; h \text{ is an orientation preserving isometric linear map w.r.t } g_q\}. \tag{6.3}$$

Definition 6.3. Let $(\mathcal{M}, g)$ be an n-dimensional smooth Riemannian submanifold of $\mathbb{R}^m$. Let N be a normal vector field to $(\mathcal{M}, g)$, that is $N \in \Gamma(N\,\mathcal{M}) \subset \Gamma(T\,\mathbb{R}^m)$ and $N(p) \perp T_p\,\mathcal{M}$ (perpendicular w.r.t the Euclidean inner product in $\mathbb{R}^m$) for all $p \in \mathcal{M}$. The **shape operator in the direction** N, $S_N : \Gamma(\mathcal{M}) \to \Gamma(\mathcal{M})$ is defined as the self-adjoint linear map associated to

$$\langle S_N(X), Y\rangle = \langle N, II(X,Y)\rangle. \tag{6.4}$$

It is a $(1,1)$ tensor field, since for each p it is a linear map between $T_p\,\mathcal{M}$ and $T_p\,\mathcal{M}$.

Proposition 6.4. (Weingarten equations) *Let $(\mathcal{M}, g)$ be a Riemannian submanifold of $\mathbb{R}^m$ ($\mathbb{R}^m$ has standard connection $\bar{\nabla}$) and $X \in \Gamma(\mathcal{M})$ and $N \in \Gamma(N\,\mathcal{M})$, then*

$$S_N(X) = -(\bar{\nabla}_X N)^\top. \tag{6.5}$$

This formula is independent of the extensions of X and N to $\Gamma(\mathbb{R}^m)$.

Definition 6.5. Let $(\mathcal{M}, g)$ be an n-dimensional Riemannian manifold (with Levi-Civita connection ∇) and E be a k-plane bundle over $\mathcal{M}$, that is for each $p \in \mathcal{M}$, E_p is a k-dimensional vector space (isomorphic to $\mathbb{R}^k$) equipped with a metric g^E and a connection $\nabla^E : T\,\mathcal{M} \times E \to E$ compatible with g^E. We define a **shape operator on** E as $S : T\,\mathcal{M} \times E \to T\,\mathcal{M}$ being a homomorphism (that is a bilinear map) satisfying

$$\langle S(X,N), Y\rangle - \langle X, S(Y,N)\rangle = 0 \qquad \forall N \in \Gamma(\mathcal{M}, E), \quad \forall X, Y \in \mathfrak{X}(\mathcal{M}) = \Gamma(\mathcal{M}, T\,\mathcal{M}) \tag{6.6}$$

How can we make sure it exists?
By Riesz representation theorem, there exists a bilinear form (the **second fundamental form**) $B : T\mathcal{M} \times T\mathcal{M} \to E$ given by the correspondence

$$\langle B(X,Y), N\rangle = \langle S(X,N), Y\rangle \qquad \forall N \in \Gamma(\mathcal{M}, E), \quad \forall X, Y \in \mathfrak{X}(\mathcal{M}) = \Gamma(\mathcal{M}, T\mathcal{M}) \tag{6.7}$$

Definition 6.6. We say that S is **compatible** (with g, g_E), if for the Riemann curvature tensors

$$\begin{aligned} &R(X,Y)Z := \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z \quad \forall X, Y, Z \in \mathfrak{X}(\mathcal{M}) \\ &R^E(X,Y)N := \nabla^E_X \nabla^E_Y N - \nabla^E_Y \nabla^E_X N - \nabla^E_{[X,Y]} N \quad \forall X, Y \in \mathfrak{X}(\mathcal{M}), \quad \forall N \in \Gamma(\mathcal{M}, E) \end{aligned} \tag{6.8}$$

the Gauss equations

$$\begin{aligned} R(X,Y)Z &= S(X, B(Y,Z)) - S(Y, B(X,Z)), \\ R^E(X,Y)N &= B(X, S(Y,N)) - B(S(X,N), Y), \end{aligned} \tag{6.9}$$

and the Codazzi-Mainard equation

$$\nabla_X S(Y,N) - \nabla_Y S(X,N) - S([X,Y],N) = S(Y, \nabla^E_X N) - S(X, \nabla^E_Y N) \tag{6.10}$$

are satisfied.

Now we introduce a local variant of the shape operator, that is for a chosen basis of E, we get a more explicit description of S which then leads to a direct definition of the elastic energy.

Remark 6.7. Given a basis $e_1, ..e_k$ for E and $S^1, ..., S^k$ symmetric $(1,1)$ tensor fields on $\mathcal{M}$ (i.e. $S^i : T\mathcal{M} \to T\mathcal{M}$), we can define a shape operator $S : T\mathcal{M} \times E \to T\mathcal{M}$ as

$$\forall p \in \mathcal{M} : \quad S_p(v,e) = S_p(v, a^1 e_1 + ... + a^k e_k) := a^1 S^1_p(v) + ... + a^k S^k_p(v). \tag{6.11}$$

The symmetry and bilinearity follows directly from the symmetry and linearity of $S^1, ..., S^k$. We say that $S^1, ..., S^k$ are compatible with g if S is compatible.

Definition 6.8. Let $(\mathcal{M}, g)$ be an n-dimensional Riemannian manifold and E an $(m-n)$-plane bundle over $\mathcal{M}$ with a bundle metric $h : E \times E \to \mathbb{R}$ and a compatible connection $\nabla^E : T\mathcal{M} \times E \to E$. Let $S^1, ..., S^{m-n}$ be $(1,1)$ tensor fields on $\mathcal{M}$, (i.e. $S^i(p) : T_p\mathcal{M} \to T_p\mathcal{M}$ is a linear map) such that for an orthonormal basis $e_1, ..., e_{m-n}$ for E and S, the shape operator $S : T\mathcal{M} \times E \to T\mathcal{M}$ defined by Remark 6.7 is compatible. Let $f \in \mathrm{Imm}_2(\mathcal{M}, \mathbb{R}^m)$ and a map $N_f : E \to \mathbb{R}^m$ such that

$$\forall p \in \mathcal{M} \quad \forall v \in T_p\mathcal{M} : \quad N_f|_p(v) \perp T_p f(\mathcal{M}). \tag{6.12}$$

Define now $N^i_f := N_f(e_i) : \mathcal{M} \to \mathbb{R}^m$, note that

$$N_f(a^1 e_1 + ... + a^{m-n} e_{m-n}) = a^1 N^1_f + ... + a^{m-n} N^{m-n}_f. \tag{6.13}$$

Then for $X \in \mathfrak{X}(\mathcal{M})$ and the Euclidean connection $\bar{\nabla}$ on $\mathbb{R}^m$,

$$\nabla_X N^i_f = \bar{\nabla}_X (N_f(e_i)) = (\nabla_X N_f)(e_i) + N_f(\nabla^E_X e_1) = (\nabla_X N_f)(e_i) + (\nabla^E)^{ij}_X N^j_f, \tag{6.14}$$

using $(\nabla^E)^{ij}_X := h(\nabla^E_X e_i, e_j)$. We define the **elastic energy** of the couple (f, N_f) as

$$E(f, N_f) = E_s + E_b = \int_{\mathcal{M}} \operatorname{dist}^2(df, \mathrm{O}(g,e)) + E_b \tag{6.15}$$

The first term is known as stretching energy and the second as bending energy.

The bending energy is defined as

$$E_b(f, N_f) := \sum_{i=1}^{m-n} \int_{\mathcal{M}} |dN^i_f + df \circ S^i - (\nabla^E)_{ij} N^j_f|^2 = \sum_{i=1}^{m-n} \int_{\mathcal{M}} |(dN^i_f)^\top + df \circ S^i|^2 + |(dN^i_f)^\perp - (\nabla^E)_{ij} N^j_f|^2, \tag{6.16}$$

where for each $p \in \mathcal{M}$, the integrand is the norm of a linear map from $T_p\mathcal{M}$ to $\mathbb{R}^m$. Then it is also possible to consider the part in $T_p f(\mathcal{M})$ and $N_p f(\mathcal{M})$ separately, using the corresponding projections $\top$ and $\perp$.

We now look for a generalization of Theorem A.1 in [AKM23], that is if an immersion f is locally close to being isometric and locally close to having a prescribed shape operator then it is close to a particular isometric immersion f_0 with this prescribed shape operator. We make use of the well established fundamental result.

Theorem 6.9. (Generalization of fundamental theorem of surface theory [Ten71])
Let $(\mathcal{M}, g)$ be an n-dimensional manifold and E a k-plane bundle with a bundle metric and a compatible (with the bundle metric) connection $\nabla^E : T\mathcal{M} \times E \to E$. Let S be a compatible shape operator on E.

Then there exists a local isometric immersion $\iota : U \subset \mathcal{M} \to \mathbb{R}^m$ such that $N\iota(\mathcal{M}) \simeq E$. That is, we have an associated linear isometry $j : E \to N\iota(\mathcal{M})$ ($j_p : E_p \to N_p\iota(\mathcal{M})$ at each point). Moreover, the shape operator of $\iota(\mathcal{M})$ agrees with S and the connection on $\iota(\mathcal{M})$ agrees with ∇^E, that is the connection on $T\iota(\mathcal{M})$ and $N\iota(\mathcal{M})$ are just the projections from the standard Euclidean connection on $\mathbb{R}^m$. The immersion is unique up to a rigid motion.

In the original paper they talk about second fundamental forms, which are the same as shape operators defined here. They are equivalent anyway, through the Riesz representation theorem.

Theorem 6.10. (Version 2)
Let $(\mathcal{M}, g)$ be an oriented, connected, simply connected, compact n-dimensional manifold with Lipschitz boundary and E an $(m-n)$-plane bundle over $\mathcal{M}$ with a bundle metric h and a compatible connection ∇^E. Let $S^1, ..., S^{m-n}$ be $(1,1)$ tensor fields such that for a basis $e_1, ..., e_{m-n}$ the shape operator $S : T\mathcal{M} \times E \to T\mathcal{M}$, given by Remark 6.7 is compatible.

For every $f \in \mathrm{Imm}_2(\mathcal{M}, \mathbb{R}^m)$ and $N_f : E \to \mathbb{R}^m$ with $N_f(e) \perp Tf(\mathcal{M})$, set $N^i_f : \mathcal{M} \to \mathbb{R}^m$ as $N^i_f(p) = N_f(p, e_i(p))$. Then there exist a smooth isometric immersion $f_0 : \mathcal{M} \to \mathbb{R}^m$ and $N^1_{f_0}, ..., N^{m-n}_{f_0} : \mathcal{M} \to \mathbb{R}^m$ with $S^i_{f_0} = S^i$ for $i = 1, ..., m-n$, satisfying

$$\begin{aligned}
\|f - f_0\|_{W^{1,2}(\mathcal{M},\mathbb{R}^m)} &+ \sum_{i=1}^{m-n} \left\|N^i_f - N^i_{f_0}\right\|_{W^{1,2}(\mathcal{M},\mathbb{R}^m)} \\
&\le C E^{1/2}(f, N_f) \\
&= C \left(\|\operatorname{dist}(\nabla f, \mathrm{O}(g,e))\|_{L^2(\mathcal{M})} + \sum_{i=1}^{m-n} \left\| dN^i_f + df \circ S^i - (\nabla^E)_{ij} N^j_f \right\|_{L^2(\mathcal{M},\mathbb{R}^m)} \right).
\end{aligned} \tag{6.17}$$

The constant C is universal and depends on $\mathcal{M}, g, S, \nabla^E$ and h.

Proof. We follow the strategy of the proof of [AKM23, Theorem A.1].
Step 1: Since S is compatible, by the generalization of the fundamental theorem of surface theory (Theorem 6.9) there exists a unique (up to rigid motions) isometric immersion ι such that the shape operator of $\iota(\mathcal{M})$, $\bar{S}$ coincides with S.

Furthermore there exists an isometric bijective linear map $N_0 : E \to \mathbb{R}^m$ such that at each p, $N_0(E_p) \perp T_p\iota(\mathcal{M})$. Thus for $\bar{S} : T_{\iota(p)}\iota(\mathcal{M}) \times N_{\iota(p)}\iota(\mathcal{M}) \to T_{\iota(p)}\iota(\mathcal{M})$, we have

$$\bar{S}(v, e) = d\iota \circ S(d\iota^{-1}(v), N^{-1}(e)) \qquad \forall (v, n) \in T_{\iota(p)}\iota(\mathcal{M}) \times N_{\iota(p)}\iota(\mathcal{M}). \tag{6.18}$$

Define $N_0^i : \mathcal{M} \to \mathbb{R}^m$ as $N_0^i = N_0(e_i)$ for $i = 1, ..., m-n$. This induces the maps $\bar{S}^i : T_p\iota(\mathcal{M}) \to \mathbb{R}^m$, note that $S_X N_0^i = S^i$. We now calculate $dN_0^i|_p : T_p\mathcal{M} \to \mathbb{R}^m$ as ($\bar{\nabla}$ denotes the Euclidean connection on $\mathbb{R}^m$),

$$\begin{aligned} dN_0^i|_p(X) &= \bar{\nabla}_X N_0^i = \bar{\nabla}_X(N_0(e_i)) = (\bar{\nabla}_X(N_0(e_i)))^\top + (\bar{\nabla}_X(N_0(e_i)))^\perp \\ &= -\bar{S}^1(X) + (\bar{\nabla}_X(N_0(e_i)))^\perp \\ &= -d\iota \circ S^1(X) + (\nabla^E)_{ij} N_0^j(X). \end{aligned} \tag{6.19}$$

Step 2: Let $\mathcal{M}_h = \mathcal{M} \times (-h, h)^{m-n}$ with h sufficiently small and define $\Phi : \mathcal{M}_h \to \mathbb{R}^m$ as

$$\Phi(q, t_1, ..., t_{m-n}) = \iota(q) + t_1 N_0^1(q) + ... + t_{m-n} N_0^{m-n}(q). \tag{6.20}$$

The metric G on $\mathcal{M}_h$ is defined as

$$\begin{aligned} &\langle (v, s_1, ..., s_{m-n}), (w, r_1, ..., r_{m-n}) \rangle_G \big|_{(q, t_1, ..., t_{m-n})} \\ &:= \left\langle v - \sum_{i=1}^{m-n} t_i S_q^i(v), w - \sum_{i=1}^{m-n} t_i S_q^i(w) \right\rangle_g \\ &\quad + \left\langle \sum_{i=1}^{m-n} s_i N_0^i + \left(\sum_{i=1}^{m-n} s_i t_i (\nabla_{ij}^E N_0^j)(v) N_0^i \right), \sum_{i=1}^{m-n} r_i N_0^i + \left(\sum_{i=1}^{m-n} r_i t_i (\nabla_{ij}^E N_0^j)(w) N_0^i \right) \right\rangle \\ &= \left\langle v - \sum_{i=1}^{m-n} t_i S_q^i(v), w - \sum_{i=1}^{m-n} t_i S_q^i(w) \right\rangle_g \\ &\quad + \left\langle \sum_{i=1}^{m-n} s_i \left(1 + t_i (\nabla_{ij}^E N_0^j)(v)\right) N_0^i, \sum_{i=1}^{m-n} r_i \left(1 + t_i (\nabla_{ij}^E N_0^j)(w)\right) N_0^i \right\rangle \\ &= \left\langle v - \sum_{i=1}^{m-n} t_i S_q^i(v), w - \sum_{i=1}^{m-n} t_i S_q^i(w) \right\rangle_g + \sum_{i=1}^{m-n} s_i r_i \left(1 + t_i (\nabla_{ij}^E N_0^j)(v)\right) \left(1 + t_i (\nabla_{ij}^E N_0^j)(w)\right). \end{aligned} \tag{6.21}$$

Since $t \le h := \min\{2/\|S\|_\infty, 1/\|(\nabla^E)_{ij} N_0\|_\infty\}$, G is nondegenerate, in other words h, depends only on the information $\mathcal{M}, g, h, \nabla^E, S$ since they determine the immersion ι and N_0.
Interestingly in [AKM23], h depended only on S but not on the manifold $\mathcal{M}, g$.

Linearity and symmetry follows from the metrics g and h. Moreover, G is equivalent to the product metric $G_1 = g + h$ and $\tilde{G} = g + dt_1 \otimes dt_1 + ... + dt_{m-n} \otimes dt_{m-n}$, that is

$$c(|v|_g^2 + s_1^2 + ... + s_{m-n}^2) \leq |(v, s_1, ..., s_{m-n})|_G^2 \leq C(|v|_g^2 + s_1^2 + ... + s_{m-n}^2). \tag{6.22}$$

Also, Φ is an isometric immersion from $\mathcal{M}_h$ to $\mathbb{R}^m$.
Actually, that it is an isometric immersion is not used, if lower and upper bounded, then also okay.

Step 3: We extend $f \in \mathrm{Imm}_2(\mathcal{M}, \mathbb{R}^m)$ to $F : \mathcal{M}_h \to \mathbb{R}^m$ as

$$F(q, t_1, ..., t_{m-n}) = f(q) + t_1 N_f^1(q) + ... + t_{m-n} N_f^{m-n}(q). \tag{6.23}$$

Now to the map $F \circ \Phi^{-1} : \Phi(\mathcal{M}_h) \subset \mathbb{R}^m \to \mathbb{R}^m$ we apply the FJM inequality to get an isometry $Q \in \mathrm{SO}(m)$ such that for $C = C(\mathcal{M}, g, h)$

$$\|d(F \circ \Phi^{-1}) - Q\|_{L^p(\Phi(\mathcal{M}_h);\mathbb{R}^m)} \leq C \| \operatorname{dist}(d(F \circ \Phi^{-1}), \mathrm{SO}(m)) \|_{L^p(\Phi(\mathcal{M}_h);\mathbb{R}^m)}. \tag{6.24}$$

Since Φ is an isometric (equivalence of metric would be sufficient) immersion, by a change of variables

$$\|dF - Q \circ d\Phi\|_{L^p(\mathcal{M}_h;\mathbb{R}^m)} \leq C \| \operatorname{dist}(dF, \mathrm{SO}(G, e) \|_{L^p(\mathcal{M}_h;\mathbb{R}^m)}. \tag{6.25}$$

Step 4: As in the Euclidean case, the idea is to lower bound the LHS and upper bound the RHS separately. We start with the upper bound,

$$\begin{aligned} dF &= df \oplus N_f^1 \oplus ... \oplus N_f^{m-n} + \sum_{i=1}^{m-n} t_i dN_f^i \\ &= df \oplus N_f^1 \oplus ... \oplus N_f^{m-n} + \left(\sum_{i=1}^{m-n} -t_i df \circ S_f^i \right) \oplus \left(\sum_{i=1}^{m-n} t_i (\bar{\nabla} N_f^i)^\perp \right). \end{aligned} \tag{6.26}$$

For a candidate in $\mathrm{SO}(G, e)$, set

$$A := \mathrm{O}(df) \oplus N_f^1 \oplus ... \oplus N_f^{m-n} + \left(\sum_{i=1}^{m-n} -t_i \, \mathrm{O}(df) \circ S^i \right) \oplus \left(\sum_{i=1}^{m-n} t_i \nabla_{ij}^E(\bullet) N_f^j \right), \tag{6.27}$$

that is for $(v, s_1, ..., s_{m-n}) \in T_{(q,t_1,...,t_{m-n})} \mathcal{M}_h$, A is a linear map (linearity follows from linearity of the $\mathrm{O}(df)$, ∇^E and S^i) given by,

$$\begin{aligned} A(v, s_1, ..., s_{m-n}) &= \mathrm{O}(df)(v) + s_1 N_f^1 + ... + s_{m-n} N_f^{m-n} \\ &\quad + \left(\sum_{i=1}^{m-n} -t_i \, \mathrm{O}(df) \circ S^i(v) \right) + \left(\sum_{i=1}^{m-n} s_i t_i \nabla_{ij}^E(v) N_f^j \right). \end{aligned} \tag{6.28}$$

By Lemma 4.6, $\mathrm{O}(df)(v) \perp N_f^i$ for all $i = 1, ..., m-n$, and the N_f^i's are orthogonal to each other, thus

$$
\begin{aligned}
|A(v,s_1,...,s_{m-n})|_e^2 &= \left|\mathrm{O}(df)(v) - \sum_{i=1}^{m-n} t_i\,\mathrm{O}(df)\circ S^i(v)\right|^2 \\
&\quad + \left|s_1 N_f^1 + ... + s_{m-n}N_f^{m-n} + \left(\sum_{i=1}^{m-n} s_i t_i \nabla_{ij}^E(v)N_f^j\right)\right|^2 \quad OK \\
&= \left|v - \sum_{i=1}^{m-n} t_i S^i(v)\right|_g^2 + \left|\sum_{i=1}^{m-n} s_i\left(1 + t_i\nabla_{ij}^E(v)N_f^j\right)\right|^2 \\
&= \langle (v,s_1,...,s_{m-n}),(v,s_1,...,s_{m-n})\rangle_G\big|_{(q,t_1,...,t_{m-n})},
\end{aligned}
\tag{6.29}
$$

where in the first equality we use the fact that all N_f^i are orthogonal to $T_q\mathcal{M}$ and in the last the definition of G. Hence $A\in\mathrm{O}(G,e)$ and since the orientation holds by definition of N_f^i, $A\in\mathrm{SO}(G,e)$. Actually, so far I have never specified any orientation on the N_f, maybe have to be a bit more accurate here...
Then

$$
\begin{aligned}
\operatorname{dist}(dF,\mathrm{SO}(G,e)) &\le |dF - A|_e \\
&\le |(df\oplus N_f^1\oplus ...\oplus N_f^{m-n} - \mathrm{O}(df)\oplus N_f^1\oplus ...\oplus N_f^{m-n})|_e \\
&\quad + \left|\sum_{i=1}^{m-n} t_i dN_f^i - \left(\sum_{i=1}^{m-n} -t_i\,\mathrm{O}(df)\circ S^i\right)\oplus\left(\sum_{i=1}^{m-n} t_i\nabla_{ij}^E(\bullet)N_f^j\right)\right|_e \\
&\lesssim |df - \mathrm{O}(df)| + \sum_{i=1}^{m-n}\left|df\circ S_f^i + \mathrm{O}(df)\circ S^i - \nabla_{ij}^E(\bullet)N_f^j\right| \\
&\le |df - \mathrm{O}(df)| + \sum_{i=1}^{m-n}|df\circ S_f^i + df\circ S^i - \nabla_{ij}^E(\bullet)N_f^j| + |df\circ S^i - \mathrm{O}(df)\circ S^i| \\
&\le C\left(|df - \mathrm{O}(df)| + \sum_{i=1}^{m-n}|df\circ S_f^i + df\circ S^i - \nabla_{ij}^E(\bullet)N_f^j|\right) \\
\implies \|\operatorname{dist}(dF,\mathrm{SO}(G,e)\|_{L^2(\mathcal{M}_h)} &\le CE^{1/2}(f,N_f),
\end{aligned}
\tag{6.30}
$$

with C depending on g,S,h, giving the desired bound on the RHS.

Step 5: To lower bound the LHS,

$$
\begin{aligned}
Q\circ d\Phi &= Q\circ\left(d\iota\oplus N_0^1\oplus ...\oplus N_0^{m-n} + \sum_{i=1}^{m-n} t_i dN_0^i\oplus 0\oplus ...\oplus 0\right) \\
&= Q\circ\left(d\iota\oplus N_0^1\oplus ...\oplus N_0^{m-n} - \sum_{i=1}^{m-n} t_1(d\iota\circ S^i - (\bar\nabla_X N_0^i)^\perp)\oplus 0\oplus ...\oplus 0\right).
\end{aligned}
\tag{6.31}
$$

We define now $f_0 := Q \circ \iota$, then $Q \circ d\iota = df_0$, $Q \circ N_0^1 = N_{f_0}^1$ and $Q \circ (\bar{\nabla}_X N_0^i)^{\perp} = (\bar{\nabla}_X N_{f_0}^i)^{\perp}$, hence

$$Q \circ d\Phi = df_0 \oplus N_{f_0}^1 \oplus \ldots \oplus N_{f_0}^{m-n} + \sum_{i=1}^{m-n} t_i dN_{f_0}^i. \tag{6.32}$$

In other words, f_0 is an isometric immersion with shape operator $S_{f_0} = S$, since the isometric immersion ι from Theorem 6.9 is invariant under rigid motions. This in turn leads to

$$dF - Q \circ d\Phi = (df - df_0) \oplus (N_f^1 - N_{f_0}^1) \oplus \ldots \oplus (N_f^{m-n} - N_{f_0}^{m-n}) - \left(\sum_{i=1}^{m-n} t_i (dN_f^i - dN_{f_0}^i) \right). \tag{6.33}$$

We now calculate

$$\begin{aligned}
|dF - Q \circ d\Phi|^2 &= \langle dF - Q \circ d\Phi, dF - Q \circ d\Phi \rangle \\
&= \Bigg\langle (df - df_0) \oplus (N_f^1 - N_{f_0}^1) \oplus \ldots \oplus (N_f^{m-n} - N_{f_0}^{m-n}) - \left(\sum_{i=1}^{m-n} t_i (dN_f^i - dN_{f_0}^i) \right), \\
&\qquad (df - df_0) \oplus (N_f^1 - N_{f_0}^1) \oplus \ldots \oplus (N_f^{m-n} - N_{f_0}^{m-n}) - \left(\sum_{i=1}^{m-n} t_i (dN_f^i - dN_{f_0}^i) \right) \Bigg\rangle \\
&\gtrsim |(df - df_0) \oplus (N_f^1 - N_{f_0}^1) \oplus \ldots \oplus (N_f^{m-n} - N_{f_0}^{m-n})|^2 \\
&\quad + \left\langle \sum_{i=1}^{m-n} t_i (dN_f^i - dN_{f_0}^i), \sum_{i=1}^{m-n} t_i (dN_f^i - dN_{f_0}^i) \right\rangle \\
&\quad + 2 \left\langle (df - df_0) \oplus (N_f^1 - N_{f_0}^1) \oplus \ldots \oplus (N_f^{m-n} - N_{f_0}^{m-n}), \sum_{i=1}^{m-n} t_i (dN_f^i - dN_{f_0}^i) \right\rangle \\
&\gtrsim |df - df_0|^2 + \sum_{i=1}^{m-n} |N_f^i - N_{f_0}^i|^2 + \left| \sum_{i=1}^{m-n} t_i (dN_f^i - dN_{f_0}^i) \right|^2 \\
&\quad + \underbrace{2 \sum_{i=1}^{m-n} t_i \left\langle (df - df_0) \oplus (N_f^1 - N_{f_0}^1) \oplus \ldots \oplus (N_f^{m-n} - N_{f_0}^{m-n}), (dN_f^i - dN_{f_0}^i) \right\rangle}_{B}.
\end{aligned} \tag{6.34}$$

Step 6: The idea is to ensure that the last term B is non-negative and then lower bound $|dF - Q \circ d\Phi|^2$ by the first three terms. Define for $i = 1, \ldots, m-n$ the sets

$$A_+^i = \{q \in \mathcal{M} : \quad \left\langle (df - df_0) \oplus (N_f^1 - N_{f_0}^1) \oplus \ldots \oplus (N_f^{m-n} - N_{f_0}^{m-n}), (dN_f^i - dN_{f_0}^i) \right\rangle \geq 0\} \subset \mathcal{M}. \tag{6.35}$$

Then for $I \in \mathcal{P} = \mathbb{P}(\{1, \ldots, m-n\})$, let

$$\begin{aligned}
D_I = &\{(q, t_1, \ldots, t_{m-n}) \in \mathcal{M}_h : \forall i \in I \quad q \in A_+^i, t_i \geq 0\} \\
&\cup \{(q, t_1, \ldots, t_{m-n}) \in \mathcal{M}_h : \forall j \notin I \quad q \notin A_+^j, t_j \leq 0\},
\end{aligned} \tag{6.36}$$

and $D = \bigcup_{I\in\mathcal{P}} D_I$. Hence on D, the term B is non-negative and

$$|dF - Q\circ d\Phi|^2 \gtrsim |df - df_0|^2 + \sum_{i=1}^{m-n} |N_f^i - N_{f_0}^i|^2 + \left|\sum_{i=1}^{m-n} t_i(dN_f^i - dN_{f_0}^i)\right|^2 \geq |df - df_0|^2 + \sum_{i=1}^{m-n} |N_f^i - N_{f_0}^i|^2. \tag{6.37}$$

Note now that $\mathcal{M} \subset D$ and $|D| = \frac{1}{2^{m-n}}|\mathcal{M}_h|$, D is essentially a union of subblocks of $\mathcal{M}_h$. Hence

$$\begin{aligned}
\int_{\mathcal{M}_h} |dF - Q\circ d\Phi|^2 &\geq \int_D |dF - Q\circ d\Phi|^2 \\
&\gtrsim \int_D |df - df_0|^2 + \sum_{i=1}^{m-n} |N_f^i - N_{f_0}^i|^2 \\
&\gtrsim \int_{\mathcal{M}} |df - df_0|^2 + \sum_{i=1}^{m-n} |N_f^i - N_{f_0}^i|^2 \\
&= \|df - df_0\|^2_{L^2(\mathcal{M},\mathbb{R}^m)} + \sum_{i=1}^{m-n} \|N_f^i - N_{f_0}^i\|^2_{L^2(\mathcal{M},\mathbb{R}^m)}.
\end{aligned} \tag{6.38}$$

Using Poincaré inequality and possibly translating f_0,

$$\|dF - Q\circ d\Phi\|_{L^2(\mathcal{M}_h;\mathbb{R}^m)} \gtrsim \|f - f_0\|_{W^{1,2}(\mathcal{M},\mathbb{R}^m)} + \sum_{i=1}^{m-n} \|N_f^i - N_{f_0}^i\|_{L^2(\mathcal{M},\mathbb{R}^m)}. \tag{6.39}$$

Step 7: We now bootstrap the L^2 estimate on the normals to $W^{1,2}$. First

$$\begin{aligned}
|dN_f^i - dN_0^i|^2 &\leq |(dN_f^i)^\top + (dN_f^i)^\perp + d\iota\circ S^i - (dN_0^i)^\perp|^2 \\
&= \left|\left((dN_f^i)^\top + d\iota\circ S^i + df\circ S^i - df\circ S^i\right) + \left((dN_f^i)^\perp - (dN_0^i)^\perp + (\nabla^E)_{ij}N_f^j - (\nabla^E)_{ij}N_f^j\right)\right|^2 \\
&\lesssim \left|(dN_f^i)^\top + d\iota\circ S^i + df\circ S^i - df\circ S^i\right|^2 + \left|(dN_f^i)^\perp - (dN_0^i)^\perp + (\nabla^E)_{ij}N_f^j - (\nabla^E)_{ij}N_f^j\right|^2 \\
&\lesssim \left|(dN_f^i)^\top + df\circ S^i\right|^2 + \left|(df - d\iota)\circ S^i\right|^2 + \left|(dN_f^i)^\perp - (\nabla^E)_{ij}N_f^j\right|^2 + \left|(\nabla^E)_{ij}\left(N_f^j - N_{f_0}^j\right)\right|^2 \\
&\lesssim \left|(dN_f^i)^\top + df\circ S^i\right|^2 + \left|(dN_f^i)^\perp - (\nabla^E)_{ij}N_f^j\right|^2 \\
&\qquad + |df - d\iota|^2|S^i|^2 + \sum_{j=1}^{m-n} |N_f^j - N_{f_0}^j|^2 |(\nabla^E)_{ij}|^2.
\end{aligned} \tag{6.40}$$

The first term corresponds to the bending energy and the second term is already on the LHS, i.e.

$$\sum_{i=1}^{m-n} \|dN_f^i - dN_0^i\|_{L^2} \lesssim E(f, N_f)^{1/2} + \|S\|_{Op}\|f - f_0\|_{L^2(\mathcal{M},\mathbb{R}^m)} + \|\nabla^E\|_{Op} \sum_{j=1}^{m-n} \|N_f^j - N_{f_0}^j\|_{L^2(\mathcal{M},\mathbb{R}^m)}, \tag{6.41}$$

where $\|\cdot\|_{Op}$ denotes the corresponding operator norms of S and ∇^E, which are known a priori and can be absorbed into the constant on the RHS.

Combining the bound for the LHS (6.39) with the bound for the RHS (6.30), the FJM inequality (6.25) and then finally (6.41) finishes the proof. $\square$

Remark 6.11. In the case $m-n=2$, the term $\sum_{i=1}^{m-n}\|N_f^i-N_{f_0}^i\|^2_{L^2(\mathcal{M},\mathbb{R}^m)}$ can be replaced directly by $\sum_{i=1}^{m-n}\|N_f^i-N_{f_0}^i\|^2_{W^{1,2}(\mathcal{M},\mathbb{R}^m)}$, i.e. the norm on the LHS is the full $W^{1,2}(\mathcal{M},\mathbb{R}^m)$ norm.

Proof. Start from (6.37), instead of using just its non-negativity, by Cauchy-Schwarz,

$$
\begin{aligned}
&\left|\sum_{i=1}^{m-n} t_i(dN_f^i-dN_{f_0}^i)\right|^2\\
&\quad\geq |t_1|^2|dN_f^1-dN_{f_0}^1|^2+|t_2|^2|dN_f^2-dN_{f_0}^2|^2-2|t_1||t_2||dN_f^1-dN_{f_0}^1||dN_f^2-dN_{f_0}^2|\\
&\quad\geq |t_1|^2|dN_f^1-dN_{f_0}^1|^2+|t_2|^2|dN_f^2-dN_{f_0}^2|^2-|t_1||t_2|(|dN_f^1-dN_{f_0}^1|^2+|dN_f^2-dN_{f_0}^2|^2).
\end{aligned}
\tag{6.42}
$$

Since for the projection of D onto $\mathcal{M}$, $\pi(D)=\mathcal{M}$, and the integrand does not depend on the sign of t_1,t_2,

$$
\begin{aligned}
&\int_D |t_1|^2|dN_f^1-dN_{f_0}^1|^2+|t_2|^2|dN_f^2-dN_{f_0}^2|^2-|t_1||t_2|(|dN_f^1-dN_{f_0}^1|^2+|dN_f^2-dN_{f_0}^2|^2)\\
&=\int_{\mathcal{M}}\int_0^h\int_0^h |t_1|^2|dN_f^1-dN_{f_0}^1|^2+|t_2|^2|dN_f^2-dN_{f_0}^2|^2-|t_1||t_2|(|dN_f^1-dN_{f_0}^1|^2+|dN_f^2-dN_{f_0}^2|^2)\\
&=\int_{\mathcal{M}}|dN_f^1-dN_{f_0}^1|^2\int_0^h\int_0^h(|t_1|^2-|t_1||t_2|)+\int_{\mathcal{M}}|dN_f^2-dN_{f_0}^2|^2\int_0^h\int_0^h(|t_2|^2-|t_1||t_2|)\\
&=\frac{h^4}{12}\int_{\mathcal{M}}(|dN_f^1-dN_{f_0}^1|^2+|dN_f^2-dN_{f_0}^2|^2)
\end{aligned}
\tag{6.43}
$$

Hence,

$$
\int_{\mathcal{M}_h}|dF-Q\circ d\Phi|^2\gtrsim\|df-df_0\|^2_{L^2(\mathcal{M},\mathbb{R}^m)}+\sum_{i=1}^{2}\|N_f^i-N_{f_0}^i\|^2_{L^2(\mathcal{M},\mathbb{R}^m)}+\sum_{i=1}^{2}\|dN_f^i-dN_{f_0}^i\|^2_{L^2(\mathcal{M},\mathbb{R}^m)},
\tag{6.44}
$$

and upon translating and using Poincaré inequality, we obtain the result. Note that this calculation only works in the codimension-2 case, for higher order terms, there are multiple mixed terms and it is not guaranteed that the integral over $(0,h)^{m-n}$ is positive. $\square$

Definition 6.12. Let $(\mathcal{M}, g)$ be an n-dimensional Riemannian manifold and E an $(m-n)$-plane bundle over $\mathcal{M}$ with a bundle metric h and a compatible connection ∇^E. Let S be a shape operator $S : T\mathcal{M} \times E \to T\mathcal{M}$. Let $f \in \mathrm{Imm}_2(\mathcal{M}, \mathbb{R}^m)$ and a map $N_f : E \to \mathbb{R}^m$ such that

$$\forall p \in \mathcal{M} \quad \forall e \in E_p : \quad N_f|_p(e) \perp T_p f(\mathcal{M}). \tag{6.45}$$

We have for $X \in \mathfrak{X}(M)$ the map $\nabla_X N_f : E \to \mathbb{R}^m$ given as

$$(\nabla_X N_f)(e) = (\bar{\nabla}_X N_f)^{\perp}(e) = \langle (\bar{\nabla}_X N_f)^{\perp}, N_f(e) \rangle. \tag{6.46}$$

Set now $\nabla N_f : T\mathcal{M} \times E \to \mathbb{R}^m$ as $(\nabla N_f)(X, e) = (\nabla_X N_f)(e)$.
We define the **elastic energy** of the couple (f, N_f) as

$$E(f, N_f) = E_s + E_b := \int_{\mathcal{M}} \mathrm{dist}^2(df, \mathrm{O}(g, e)) + \int_{\mathcal{M}} |dN_f + df \circ S|^2. \tag{6.47}$$

The first term is known as stretching energy and the second as bending energy. Since at a point p, $(\nabla N)_p : T_p\mathcal{M} \times E_p \to \mathbb{R}^m$ and $df_p : T_p\mathcal{M} \to \mathbb{R}^m$, $S_p : T_p\mathcal{M} \times E_p \to T_p\mathcal{M}$, the integrand of the bending energy is the norm of a bilinear map from $T_p\mathcal{M} \times E_p \to \mathbb{R}^m$.

Theorem 6.13. (Version 1)
Let $(\mathcal{M}, g)$ be an oriented, connected, simply connected, compact n-dimensional manifold with Lipschitz boundary and E an $(m-n)$-plane bundle over $\mathcal{M}$ with a bundle metric h and a compatible connection ∇^E. Let S be a compatible shape operator $S : T\mathcal{M} \times E \to T\mathcal{M}$. Then for every $f \in \mathrm{Imm}_2(\mathcal{M}, \mathbb{R}^m)$ and $N_f : E \to \mathbb{R}^m$ with $N_f(e) \perp Tf(\mathcal{M})$, there exist a smooth isometric immersion $f_0 : \mathcal{M} \to \mathbb{R}^m$ and $N_0 : E \to \mathbb{R}^m$ ($N_0(e) \perp T\iota(\mathcal{M})$) with shape operator $S_{f_0} = S$, satisfying

$$\begin{aligned} \|f - f_0\|_{W^{1,2}(\mathcal{M})} &+ \big\| \|N_f - N_0\|_{E_p \to \mathbb{R}^m} \big\|_{W^{1,2}(\mathcal{M})} \\ &\le C \left(\|\mathrm{dist}(\nabla f, \mathrm{O}(g, e))\|_{L^2(\mathcal{M})} + \| \|\nabla N_f + df \circ S\|_{T_p\mathcal{M} \times E_p \to \mathbb{R}^m} \|_{L^2(\mathcal{M})} \right) \\ &\simeq C \tilde{E}^{1/2}(f, N_f). \end{aligned} \tag{6.48}$$

The constant C is universal and depends on $\mathcal{M}, g, S, \nabla^E$ and h.

I did not manage to prove it, somehow I still had a normal term in the dN term, which did not appear in the energy, as we defined it. But according to our discussion, the normal term is in the energy implicitly... .However is it not possible to prove it as corollary of the case where we choose a basis?

7 Appendix: Useful tools and facts

Lemma 7.1. *Useful vector calculus identities:*

(a) *For* $u, v : \mathbb{R}^N \to \mathbb{R}^N$, *we have* $\Delta\langle u, v\rangle = \nabla^2\langle u, v\rangle = \langle u, \nabla^2 v\rangle + \langle \nabla u, \nabla v\rangle + \langle \nabla^2 u, v\rangle$.

(b) *The product rule for the divergence of vector fields yields the following:*

$$\mathrm{div}(Vf) = f \cdot \mathrm{div}(V) + \langle \nabla f, V\rangle \iff f \cdot \mathrm{div}\, V = \mathrm{div}(Vf) - \langle \nabla f, V\rangle. \tag{7.1}$$

We can consider $f : \mathbb{R}^N \to \mathbb{R}$ *and* $V : \mathbb{R}^N \to R^N$, *but also* $f : \mathbb{R}^N \to \mathbb{R}^N$ *and* $V : \mathbb{R}^N \to \mathbb{R}^{N\times N}$. *In particular for a vector field* v, *we have*

$$v \cdot \Delta v = v \cdot \mathrm{div}(\nabla v) = \mathrm{div}(\nabla v\ v) - \langle \nabla v, \nabla v\rangle. \tag{7.2}$$

(c) $\Delta v = \Delta v + \mathrm{div}(\nabla v)^T - \nabla(\mathrm{Tr}(\nabla v)) = 2\,\mathrm{div}\left(\mathrm{Sym}\,\nabla v - \frac{1}{2}\,\mathrm{Tr}(\mathrm{Sym}\,\nabla v)Id\right)$

(d) $\Delta(|\nabla u|^2) = \Delta\langle \nabla u, \nabla u\rangle = 2\langle \Delta\nabla u, \nabla u\rangle + \langle \nabla^2 u, \nabla^2 u\rangle$ *by using* $\Delta(\psi\phi) = \psi\Delta\phi + 2\Delta\psi\cdot\phi + \phi\Delta\psi$. *We denote by* $\nabla^2 u$ *the Hessian of* u, *a* $(0, 3)$*-tensor and not the Laplacian.*

Lemma 7.2. *For* $w : \mathbb{R}^N \to \mathbb{R}^N$, *we have*

$$[\nabla^2 w^i]_{jk} = \partial_j[\mathrm{Sym}\,\nabla w]_{ik} + \partial_k[\mathrm{Sym}\,\nabla w]_{ij} + \partial_i[\mathrm{Sym}\,\nabla w]_{jk}. \tag{7.3}$$

Lemma 7.3. *Let* $x_n \rightharpoonup x$ *in a Hilbert space* H *and* $\|x_n\| \to \|x\|$ *in* $\mathbb{R}$. *Then* $x_n \to x$ *in* H.

Definition 7.4. (Coercivity)

- A scalar function $f : \mathbb{R}^N \to \mathbb{R}$ is coercive if $\lim f(x) = \infty$ as $\|x\| \to \infty$.
- A vector field $v : \mathbb{R}^N \to \mathbb{R}^N$ is coercive if $\lim \frac{f(x)\cdot x}{\|x\|} = \infty$ as $\|x\| \to \infty$.
- For a topological space X, a function $F : X \to \bar{\mathbb{R}}$ is coercive if the closure of the sublevel sets $\{x \in X : F(x) \leq \alpha\}$ is compact for all $\alpha \in \mathbb{R}$. In the Euclidean case (or when X is a normed space), this matches the above definitions involving limits.
- A sequence of functions $F_n : X \to \bar{\mathbb{R}}$ is equi-coercive if

$$\forall \alpha \in \mathbb{R} : \quad \exists K_\alpha \text{ compact such that } \forall n : \quad \{x \in X : F_n(x) \leq \alpha\} \subset K_\alpha. \tag{7.4}$$

In particular for a normed space, this means

$$\forall \alpha \in \mathbb{R} : \quad \exists \delta \in \mathbb{R}^+ \text{ such that } \forall n : \|x\| \geq \delta \implies F_n(x) \geq \alpha \tag{7.5}$$

Lemma 7.5. *From linear algebra:*

(i) *For any* $P \in \mathbb{R}^{N\times N}$, *we have* $(\det P)Id = P^T(\mathrm{cof}\, P)$.

(ii) *For any* $u \in C^2(\Omega, \mathbb{R}^N)$ *(and by density for any* $u \in H^1(\Omega, \mathbb{R}^N)$*), we have Piola's identity* $\mathrm{div}(\mathrm{cof}\,\nabla u) = 0$.

(iii) (von Neumann Trace inequality) *For matricies A and B in $\mathbb{R}^{N\times N}$ with singular values $a_1 \geq ... \geq a_N \geq 0$ and $b_1 \geq ... \geq b_N \geq 0$ respectively, we have*

$$\operatorname{Tr}(AB) \leq |\operatorname{Tr}(AB)| \leq \sum_{i=1}^{N} a_i b_i. \tag{7.6}$$

Lemma 7.6. (Embedding theorems)

- *For Ω bounded we have for any $1 \leq p < q \leq \infty$ that $\|f\|_{L^p(\Omega)} \leq C\|f\|_{L^q(\Omega)}$, i.e. the inclusion $L^q(\Omega) \subset L^p(\Omega)$.*
- *For Ω bounded we have for any $1 \leq p < q \leq \infty$ that $\|f\|_{W^{1,p}(\Omega)} \leq C\|f\|_{W^{1,q}(\Omega)}$, i.e. the inclusion $W^{1,q}(\Omega) \subset W^{1,p}(\Omega)$.*
- *For Ω bounded we have for any $1 \leq p < q \leq \infty$ that $\|f\|_{L^{p,w}(\Omega)} \leq C\|f\|_{L^{q,w}(\Omega)}$, i.e. the inclusion $L^{q,w}(\Omega) \subset L^{p,w}(\Omega)$.*

Lemma 7.7. *In a topological space (X, τ), the convergence $x_n \to x$ with respect to τ is characterized by*

$$\forall x_{n_k},\ \exists x_{n_{k_l}} \to x \quad \Longleftrightarrow \quad x_n \to x. \tag{7.7}$$

Proof. ($\Longleftarrow$): If $x_n \to x$, then for any neighborhood $U(x)$, there exists N such that $n \geq N$ implies that $x_n \in U$. Thus also any subsequence x_{n_k} of x_n lies eventually within U, i.e. take $x_{n_{k_l}} = x_{n_k}$.

($\Longrightarrow$): We show the contrapositive, that is if $x_n \nrightarrow x$, then there exists a subsequence x_{n_k} without a converging subsubsequence. Suppose $x_n \nrightarrow x$, then there exists a neighborhood $U(x)$ such that for any n, there exists $m_n > n$ with $x_m \notin U(x)$. Take now $(x_{n_k}) = (x_{m_1}, x_{m_2}, ...)$, which lies entirely outside $U(x)$ and so can not have any convergent (to x) subsubsequence. □

Example 7.8. Convergence almost everywhere does not come from a topology. Take $X = L^\infty([0,1])$. It suffices to find a sequence f_n that does not satisfy (7.7), i.e. for any subsequence, there exists an almost everywhere converging subsubsequence, without the original sequence f_n converging. Denote first $n = 2^p + q$ with $p \in \mathbb{N}$ and $0 \leq q < 2^p$, and define

$$f_n(x) = f_{2^p+q}(x) = \chi_{[\frac{q}{2(p+1)}, \frac{q+1}{2(p+1)})}. \tag{7.8}$$

Clearly $f_n(x)$ does not converge almost everywhere to 0. However for any subsequence $f_{n_k} = f_{2^{p_k}+q_k}$ we note that $p_k \to \infty$, so $|\operatorname{spt} f_{2^{p_k}+q_k}| = \frac{1}{2p_k} \to 0$. Let now $\varepsilon > 0$ be fixed and partition $[0,1]$ into $\lceil 1/\varepsilon \rceil$ equi-sized intervals of size $\leq \varepsilon$. By the pigeonhole principle, there exists at least one interval I overlapping the supports of infinitely elements of f_{n_k}, i.e. $I \cap \operatorname{spt} f_{n_{k_l}} \neq \varnothing$ for some subsubsequence $f_{n_{k_l}}$. Since $|\operatorname{spt} f_{n_{k_l}}| \to 0$, the sequence $f_{n_{k_l}}$ is eventually supported only in a fixed interval of size $\leq 2\varepsilon$ (centered around I) and vanishes outside. However as ε was arbitrary, this implies convergence almost everywhere.

Theorem 7.9. (Poincaré-Wirtinger) *Let $1 \leq p \leq \infty$ and $\Omega \subset \mathbb{R}^N$ an open, bounded, connected Lipschitz domain. Then there exists $C = C(\Omega, p)$ such that for all $u \in W^{1,p}(\Omega, \mathbb{R})$,*

$$\|u - u_\Omega\|_{L^p(\Omega)} = \left\| u - \fint_\Omega u(y)dy \right\|_{L^p(\Omega)} \leq C\|\nabla u\|_{L^p(\Omega)}. \tag{7.9}$$

The result holds equally for vector fields $v \in W^{1,p}(\Omega, \mathbb{R}^N)$.

Lemma 7.10. *Let $\Omega \subset \mathbb{R}^N$ be an open, bounded Lipschitz domain. Then it can be written as a finite union $\Omega = \bigcup_{i=1}^n \Omega_i$ where each Ω_i is open, bounded and star shaped with respect to some interior ball.*

Theorem 7.11. Lipschitz truncation *Let Ω be an open bounded Lipschitz domain of $\mathbb{R}^N$. Then for every $g \in H^1(\Omega, \mathbb{R})$ and every $\lambda > 0$, there exists $\bar{g} \in W^{1,\infty}(\Omega, \mathbb{R})$ such that:*

(i) $\|\nabla \bar{g}\|_{L^\infty} \leq C\lambda$

(ii) $|\{g(x) \neq \bar{g}(x)\}| \leq \frac{C}{\lambda} \int_{\{|\nabla g| > \lambda\}} |\nabla g|$

(iii) $\int_\Omega |\nabla g - \nabla \bar{g}|^p \leq C \int_{\{|\nabla g| > \lambda\}} |\nabla g|^p \qquad \forall p \in [1, \infty)$

Theorem 7.12. (Kirszbraun Lipschitz extension theorem) *Let X, Y be Hilbert spaces and $E \subset X$. For any Lipschitz map $f : E \to Y$ with Lipschitz constant L, there exists a Lipschitz map $\tilde{f} : X \to Y$ with $\tilde{f}\big|_E = f$ and the same Lipschitz constant L.*

References

[AKM23] Itai Alpern, Raz Kupferman, and Cy Maor. Stability of isometric immersions of hypersurfaces, 2023.

[BD12] Dominic Breit and Lars Diening. Sharp Conditions for Korn Inequalities in Orlicz Spaces. *Journal of Mathematical Fluid Mechanics*, 14(3):565–573, 2012.

[CDM12] Sergio Conti, Georg Dolzmann, and Stefan Müller. Korn's second inequality and geometric rigidity with mixed growth conditions. *Calculus of Variations and Partial Differential Equations*, 50, 03 2012.

[CFM05] Sergio Conti, Daniel Faraco, and Francesco Maggi. A New Approach to Counterexamples to L1 Estimates: Korn's Inequality, Geometric Rigidity, and Regularity for Gradients of Separately Convex Functions. *Archive for Rational Mechanics and Analysis*, 175:287–300, 02 2005.

[CS06] Sergio Conti and Ben Schweizer. Rigidity and gamma convergence for solid-solid phase transitions with SO(2) invariance. *Communications on Pure and Applied Mathematics*, 59, 2006.

[Ten71] Keti Tenenblat. On isometric immersions of riemannian manifolds. *Boletim da Sociedade Brasileira de Matemática*, 2:23–36, 1971.