Quantitative stratification for singular sets and Reifenberg theorem

valto

August 5, 2017

1 SNAP summer school

This is a DRAFT of lecture notes for the SNAP course "Singularities of harmonic maps".

These notes are intended for the Reifenberg part, and contain A LOT MORE than what will be covered in the lectures. The (tentative) program of the lectures (Reifenberg part) is the following:

1.1 Lecture 1: intro and statements

1. General overview of the subject
2. Brief introduction on rectifiable sets and Hausdorff measure (references: [DL08, Mat95, Fed69]), area formula (section 4.1 of the notes)
3. Definition of Reifenberg flat sets with examples (Lip graph with small norm, snowflake) (section 4.3 of the notes)
4. Statement of classical Reifenberg theorem (section 4.2 of the notes)
5. Definition of β numbers with examples (C^2 graph example in the exercise session) (section 5.1 of the notes)
6. Statement of the $W^{1,p}$ Reifenberg, discrete Reifenberg and rectifiable Reifenberg (section 5.3 of the notes)

1.2 Lecture 2: technical constructions

1. squash lemma (improved bi-Lipschitz version) with partial proof (section 4.7 of the notes)
2. distance between L^2 subspaces (with proof) (section 5.4 of the notes)

1.3 Lecture 3: Proof of discrete Reifenberg

1. proof of the discrete Reifenberg theorem (with as many details as possible) (section 5.7 of the notes)

Aaron will use the Reifenberg theorem to prove estimates on the singular set of harmonic maps. He will use different notes for that part. In some sense, this part is a stand-alone part of the course. It is needed to study singularities of harmonic maps, but it is not about harmonic maps.

2 Introduction and notes

In this course, we will prove sharp estimates on the singular sets of harmonic maps between Riemannian manifolds. In order to do so, we will study a generalized version of Reifenberg's theorem in generic dimension. This course is mainly based on the works [NVa] and [NVc].

The main theorem is:

Theorem 2.1. *Let $u : B_2(0) \subseteq \mathbb{R}^m \to N$ be a minimizing harmonic map with $\theta(0,2) \leq \Lambda$. Its singular set $\mathcal{S}(u)$ satisfies*

$$\operatorname{Vol}(B_r(\mathcal{S}(u) \cap B_1(0))) \leq C(m,\Lambda) r^3 , \tag{2.1}$$

and $\mathcal{S}(u)$ is $m-3$-rectifiable.

As a first remark, note that this statement is just the simplest and least technical statement that can be proved with this techniques. The same ideas and an ϵ of technical work lead to the proof of more powerful theorems about effective singularities of minimizing and stationary harmonic maps. We refer the reader to [NVa] for more details on this. En passant, we mention that a corollary of the proof is that, under the hypothesis of the previous theorem, ∇u belongs to weak-$L^3(B_1(0))$ with weak L^3 norm bounded by $C(m,\Lambda)$.

For the moment, these lecture notes are just a sketch, they are a work in progress.

TO DO BEFORE THE END: add an intro, add all the relevant references, add citations to various generalizations and other results like obstacle problem or other stuff.

2.1 Harmonic functions in $\mathbb{R}$

Energy-minimizers, Laplace equation, mean value with

$$\frac{d}{dr} \fint_{\partial B_r(x)} u = \frac{d}{dr} \fint_{\partial B_r(x)} u \left\langle \nabla r, \frac{\nabla r}{|\nabla r|} \right\rangle = \ldots \tag{2.2}$$

Radial mollifiers and regularity. Observation

$$\frac{d}{dr} \fint_{\partial B_r(x)} |u|^2 = \frac{1}{\omega_n r^{n-1}} \int_{B_r(x)} \Delta(u^2) \geq 0 . \tag{2.3}$$

Maybe observe also that

$$\frac{d}{dr} \fint_{\partial B_r(x)} |\nabla u|^2 \geq 0 . \tag{2.4}$$

2.2 Harmonic maps

For this case, I would use Moser's book [Mos05] (which is complete and nice). We need to see that

$$\Delta(u) = A(u)(\nabla u, \nabla u) . \tag{2.5}$$

Using the nearest point projection.

2.3 Stationary harmonic maps

If u is smooth, than it is also stationary because inner variations are a special case of outer variations for smooth u.

In general, this is not true. For minimizing harmonic maps, we have the following stationary equation coming from inner variations.

Lemma 2.2. *Let u be a stationary harmonic map. Then for all smooth compactly supported vector fields* $\xi : \mathbb{R}^n \to \mathbb{R}^n$ *we have*

$$\int_{\mathbb{R}^n} \left(|\nabla u|^2 \, \delta_{ij} - 2\nabla_i u \nabla_j u \right) \partial_i \xi_j = 0 \,. \tag{2.6}$$

In other words, we have in the distributional sense:

$$\partial_i \left(|\nabla u|^2 \, \delta_{ij} - 2\nabla_i u \nabla_j u \right) = 0 \,. \tag{2.7}$$

2.4 Normalized energy

Definition 2.3. Let u be a minimizing harmonic map $u : B_2(0) \subseteq \mathbb{R}^m \to N \subseteq \mathbb{R}^n$. For $x \in B_1(0)$ and $r \leq 1$, define

$$\theta(x, r) = r^{2-m} \int_{B_r(x)} |\nabla u(y)|^2 \, dy \,. \tag{2.8}$$

Remark 2.1. θ is scale invariant. In particular, set $w(y) \equiv u(x + ry)$. Then we have

$$\nabla w(y) = \nabla \left(u(x + ry)\right) = r \nabla u|_{x+ry} \,. \tag{2.9}$$

And thus

$$\int_{B_1(0)} |\nabla w|^2 = \int_{B_1(0)} |\nabla \left(u(x + ry)\right)|^2 \, dy = r^2 \int_{B_1(0)} \left|\nabla u|_{x+ry}\right|^2 dy = r^{2-m} \int_{B_r(x)} |\nabla u|_z|^2 \, dz \,. \tag{2.10}$$

Thus

$$\theta_w(0, 1) = \theta_u(x, r) \,. \tag{2.11}$$

More in general, we have

$$\theta_w(z, s) = \theta_u(x + rz, rs) \,. \tag{2.12}$$

The quantity $\theta(x, r)$ is monotone in r with a nice formula for its derivative.

Lemma 2.4. *We have for all $s \leq r$ for which the following makes sense:*

$$\theta(x, r) - \theta(x, s) = 2 \int_{B_r(x) \setminus B_s(x)} \frac{|\langle \nabla u, x - y \rangle|^2}{|x - y|^m} \,. \tag{2.13}$$

In other words, $\theta(x, r)$ for fixed x is monotone in r and absolutely continuous, and for a.e. r we have

$$\theta(x, r)' = 2 r^{2-m} \int_{\partial B_r(x)} \left|\frac{\partial u}{\partial n}\right|^2 = 2 r^{-m} \int_{\partial B_r(x)} |\langle \nabla u, x - y \rangle|^2 \,, \tag{2.14}$$

where $n = n(y) = \frac{y-x}{|y-x|}$ is the radial vector field from x.

Proof. Set for simplicity $x = 0$. The proof is an application of the stationary equation. Let ψ be any Lipschitz function $\psi : B_2(0) \to \mathbb{R}$ with compact support in $B_2(0)$ and set $\phi : B_2(0) \to \mathbb{R}^n$ to be the vector field

$$\phi(y)_j = \psi(|y|) y_j \quad \Longrightarrow \quad \partial_i \phi_j = \psi(|y|) \delta_{ij} + \psi'(|y|) \frac{y_i y_j}{|y|} \,. \tag{2.15}$$

By (2.6) we get

$$(m - 2) \int |\nabla u|^2 \, \psi(|y|) = - \int \psi'(|y|) \, |y| \left[|\nabla u|^2 - 2 \left|\frac{\partial u}{\partial n}\right|^2 \right] . \tag{2.16}$$

Consider the following sequence ψ_α for $\alpha < r/10$.

$$\psi_\alpha(|y|) = \begin{cases} 1 & \text{for} \quad |y| \leq r - \alpha\,, \\ \frac{r+\alpha-|y|}{2\alpha} & \text{for} \quad |y| \in [r-\alpha, r+\alpha]\,, \\ 0 & \text{for} \quad |y| \geq r + \alpha\,. \end{cases} \tag{2.17}$$

I.e., ψ_α is piecewise linear. This gives

$$(m-2) \int |\nabla u|^2 \, \psi_\alpha(|y|) = \frac{1}{2\alpha} \int_{|y| \in [r-\alpha, r+\alpha]} |y| \left[|\nabla u|^2 - 2 \left| \frac{\partial u}{\partial n} \right|^2 \right] . \tag{2.18}$$

As $\alpha \to 0$, we get by dominated convergence that the lhs converges to

$$\lim_{\alpha \to 0} (m-2) \int |\nabla u|^2 \, \psi_\alpha(|y|) = (m-2) \int_{B_r(0)} |\nabla u|^2 \ . \tag{2.19}$$

In order to study the rhs, consider the function

$$f(s) = \int_{\partial B_s(0)} \left(|\nabla u|^2 - 2 \left| \frac{\partial u}{\partial n} \right|^2 \right) . \tag{2.20}$$

By Fubini's theorem, we have $f \in L^1([0,2])$ with

$$\int_0^2 |f| \leq \int_{B_2(0)} |\nabla u|^2 \ . \tag{2.21}$$

Thus, for almost all $r \in (0,2)$, r is a Lebesgue point for f, meaning that

$$f(r) = \lim_{\alpha \to 0} \frac{1}{2\alpha} \int_{r-\alpha}^{r+\alpha} f(s) ds\,. \tag{2.22}$$

Thus we obtain for these values of r that

$$\lim_{\alpha \to 0} \frac{1}{2\alpha} \int_{|y| \in [r-\alpha, r+\alpha]} |y| \left[|\nabla u|^2 - 2 \left| \frac{\partial u}{\partial n} \right|^2 \right] = \tag{2.23}$$

$$= \lim_{\alpha \to 0} \frac{1}{2\alpha} \int_{|y| \in [r-\alpha, r+\alpha]} r \left[|\nabla u|^2 - 2 \left| \frac{\partial u}{\partial n} \right|^2 \right] + \lim_{\alpha \to 0} \frac{1}{2\alpha} \int_{|y| \in [r-\alpha, r+\alpha]} (|y| - r) \left[|\nabla u|^2 - 2 \left| \frac{\partial u}{\partial n} \right|^2 \right] . \tag{2.24}$$

The second limit is zero by dominated convergence, since

$$\frac{1}{2\alpha} \chi_{\{|y| \in [r-\alpha, r+\alpha]\}} \ \ ||y| - r| \left[|\nabla u|^2 - 2 \left| \frac{\partial u}{\partial n} \right|^2 \right] \leq \frac{3}{2} |\nabla u|^2 \ . \tag{2.25}$$

And the first limit is exactly

$$\lim_{\alpha \to 0} \frac{1}{2\alpha} \int_{|y| \in [r-\alpha, r+\alpha]} r \left[|\nabla u|^2 - 2 \left| \frac{\partial u}{\partial n} \right|^2 \right] = r \lim_{\alpha \to 0} \frac{1}{2\alpha} \int_{r_\alpha}^{r+\alpha} f(s) ds = r f(r)\,. \tag{2.26}$$

Thus we obtain that for almost all r:

$$(2-m)\int_{B_r(0)} |\nabla u|^2 + r\int_{\partial B_r(0)} |\nabla u|^2 = 2r\int_{\partial B_r(0)} \left|\frac{\partial u}{\partial n}\right|^2 \geq 0\,. \tag{2.27}$$

Now consider that θ is absolutely continuous wrt r and, using the previous equation, for almost all r we have

$$\theta(0,r)' = \frac{2-m}{r}\theta(0,r) + r^{2-m}\int_{\partial B_r(0)} |\nabla u|^2 = 2r^{2-m}\int_{\partial B_r(0)} \left|\frac{\partial u}{\partial n}\right|^2 \geq 0\,. \tag{2.28}$$

This proves that θ is monotone in r. Integrating this relation between s and r we obtain the desired result. □

Definition 2.5. Since θ is monotone in r, for all $x \in B_1(0)$, we can define

$$\theta(x) = \theta(x,0) = \lim_{r\to 0}\theta(x,r)\,. \tag{2.29}$$

2.5 TO DO: STUDY THE MOLLIFIED ENERGY INSTEAD, WITH ALL THE NECESSARY FORMULAS

2.6 ϵ-regularity theorem and strong convergence

Here we simply quote two results about regularity without proof. These results are deep and it would require too long to study them.

Theorem 2.6. *Let u be a minimizing (or stationary) harmonic map. There exists $\epsilon_0(n,N) > 0$ such that*

$$\theta(x,2r) < \epsilon \quad \Longrightarrow \quad u \in C^\infty(B_r(x))\,. \tag{2.30}$$

Moreover, we have for all k:

$$\sup_{y\in B_r(x)} r^k \left|\nabla^k u(y)\right| \leq C_k\theta(x,r)^{1/2}\,. \tag{2.31}$$

This immediately tell us two things: first of all, $\theta(x) = 0$ or $\theta(x) \geq \epsilon_0$. There's a "gap" here. Indeed, if $\theta(x) < \epsilon_0$, then there exists $r = r(x) > 0$ such that $\theta(x,2r) < \epsilon_0$. Thus $\|\nabla u\|_{L^\infty(B_r(x))} \leq Cr_x^{-1}$. But then for $s \leq r$:

$$\theta(x,s) = s^{2-m}\int_{B_s(x)} |\nabla u|^2 \leq Cs^2 r_x^{-2} \to_{s\to 0} 0\,. \tag{2.32}$$

We also have a characterization of the singular set of u. Indeed

$$\mathcal{S}(u) = \{x \in B_1(0) \;\; s.t. \;\; u \text{ is not continuous at } x\} = \{x \;\; s.t. \;\; \theta(x,r) \geq \epsilon_0\}\,. \tag{2.33}$$

This gives us a first crude estimate on the singular set.

Theorem 2.7. *Let u be a minimizing harmonic map (actually stationary is enough). Let $\Lambda = \theta(0,2)$. Then we have*

$$\mathrm{Vol}\left\{B_r\left(\mathcal{S}(u)\right) \cap B_1\left(0\right)\right\} \leq C(m)\frac{\Lambda}{\epsilon_0} r^2 . \tag{2.34}$$

Moreover,

$$\lim_{r\to 0} r^{-2}\,\mathrm{Vol}\left\{B_r\left(\mathcal{S}(u)\right) \cap B_1\left(0\right)\right\} = 0 . \tag{2.35}$$

As a corollary, $\mathcal{H}^{n-2}(\mathcal{S}(u) \cap B_1(0)) = 0$.

Remark 2.2. Note that (2.34) depends only on Λ. While the second limit, although it is always zero for every function, "depends" on the function u. In the sense that in general we might have that given a sequence of maps u_i with bounded energy:

$$\lim_{r\to 0} \sup_i r^{-2}\,\mathrm{Vol}\left\{B_r\left(\mathcal{S}(u_i)\right) \cap B_1\left(0\right)\right\} > 0 . \tag{2.36}$$

This is not the case for minimizing harmonic maps, but it is the case for stationary harmonic maps. For the experts, the first limit is valid also for defect measures, the second limit is not valid for defect measures.

Proof. Fix any $r > 0$. Consider the trivial covering of $B_r(\mathcal{S}(u)) \cap B_1(0)$ given by

$$B_r\left(\mathcal{S}(u)\right) \cap B_1\left(0\right) \subseteq \bigcup_{x\in\mathcal{S}(u)\cap B_1(0)} B_r\left(x\right) . \tag{2.37}$$

Let $\{B_{5r}(x_i)\}_{i\in I}$ be a Vitali-subcovering (see Vitali covering lemma on Wikipedia) of this open cover, where $\{x_i\}_{i\in I}$ is a finite subset of $\mathcal{S}(u) \cap B_1(0)$. Then

$$B_r\left(\mathcal{S}(u)\right) \cap B_1\left(0\right) \subset \bigcup_{i\in I} B_{5r}\left(x_i\right) . \tag{2.38}$$

Since $B_r(x_i)$ are disjoint, we have

$$\Lambda \geq \int_{B_2(0)} |\nabla u|^2 \geq \int_{\cup_{i\in I} B_r(x_i)} |\nabla u|^2 = \sum_{i\in I} \int_{B_r(x_i)} |\nabla u|^2 \geq \#\left\{I\right\} \epsilon_0 r^{m-2} . \tag{2.39}$$

Here $\#\{I\}$ is just the cardinality (number of elements) of I. Thus we obtain

$$\mathrm{Vol}\left(B_r\left(\mathcal{S}(u)\right) \cap B_1\left(0\right)\right) \leq C(m)\#\left\{I\right\} r^m \leq C(m)\frac{\Lambda}{\epsilon_0} r^2 . \tag{2.40}$$

Moreover, we also have

$$r^{-2}\,\mathrm{Vol}\left(B_r\left(\mathcal{S}(u)\right) \cap B_1\left(0\right)\right) \leq C(m)\#\left\{I\right\} r^{m-2} \leq C(m)\epsilon_0^{-1} \int_{\cup_{i\in I} B_r(x_i)} |\nabla u|^2 . \tag{2.41}$$

Now consider the set $X(r) \equiv \cup_{i\in I} B_r(x_i)$. We have seen that the volume of this set converges to 0 as $r \to 0$. Thus (since $\nabla u \in L^2$) we can apply dominated convergence theorem and obtain

$$\lim_{r\to 0} \int_{X(r)} |\nabla u|^2 = 0 . \tag{2.42}$$

Note that here (dominated convergence theorem) is where we loose the "uniformity" wrt the Λ, and we "gain" the dependence on the function u. This concludes the proof. □

2.7 Strong convergence

Given a sequence of maps (any maps) $u_i \in W^{1,2}$ with $\|u_i\|_{L^2} + \|\nabla u_i\|_{L^2} \leq C$ (i.e., bounded norm and bounded energy), it is well-known that there exists a subsequence u_{i_k} such that u_{i_k} converges to some u strongly in L^2 and weakly in $W^{1,2}$.

If we add the requirement that u_i are minimizing harmonic maps, the limit is also strong in $W^{1,2}$ and u is a minimizing harmonic map.

2.8 Example

Consider the map

$$u(x) = x/\left|x\right| \, . \tag{2.43}$$

This map goes from $B_1(0) \subseteq \mathbb{R}^n$ to $S^{n-1} \subseteq \mathbb{R}^n$. Its gradient is

$$\nabla_i u^k = \frac{1}{|x|}\left(\delta_i^k - \frac{x_i x^k}{|x|^2}\right) , \quad |\nabla u|^2 = \frac{n-1}{|x|^2} \tag{2.44}$$

its Laplacian is

$$\Delta u^k = -\frac{1}{x^3} x_i \left(\delta_i^k - \frac{x_i x^k}{|x|^2}\right) + \frac{1}{|x|}\left(-\frac{n x_k}{|x|^2} - \frac{x_i \delta_{ik}}{|x|}^2 + 2\frac{|x|^2 x^k}{|x|^4}\right) = 0 + \frac{1-n}{|x|^3} x_k \, . \tag{2.45}$$

Note that

$$\Delta u^k = -\left|\nabla u\right|^2 \frac{x^k}{|x|} = -\left|\nabla u\right|^2 u(x) \, . \tag{2.46}$$

Since the second fundamental form over a sphere is

$$A(p)(X,Y) = \langle X, Y\rangle \, p \, , \tag{2.47}$$

this is consistent with (2.5).

Moreover, u is a homogeneous map, and its normalized energy is

$$\int_{B_r(0)} |\nabla u|^2 = \frac{n-1}{n-2}\omega_n r^{n-2} \, . \tag{2.48}$$

The energy is infinite for $n = 1, 2$, and so this map is not harmonic for $n \leq 2$. But it is harmonic for $n \geq 3$ and its normalized energy θ is constant. This is consistent with the monotonicity properties of θ.

The proof that this map is minimizing can be found in [Lin87, CG89].

2.9 Model maps

Here we define homogeneous and k-symmetric minimizing maps and study their basic properties.

Definition 2.8. A minimizing harmonic map u is said to be homogeneous wrt x if equivalently

- for almost all $y \in S^{m-1}$ and almost all $r > 0$:

$$u(x + ry) = u(y) \tag{2.49}$$

- for almost all $y \in \mathbb{R}^m$

$$\left\langle \nabla u|_y, y - x \right\rangle = 0\,. \tag{2.50}$$

Definition 2.9. We say that a minimizing harmonic map u is k-symmetric if ALL of the following holds

1. u is homogeneous wrt the origin
2. there exists a k-dimensional invariant subspace, meaning that for some linear V of dimension k we have

$$\forall x \in \mathbb{R}^m\,, \quad \forall y \in V\,, \quad u(x + y) = u(x)\,. \tag{2.51}$$

Note that we can state the second condition as

$$\forall x \in \mathbb{R}^m\,, \quad \nabla_V u(x) = 0\,. \tag{2.52}$$

These properties interact with each other. Indeed, we have

Lemma 2.10. *Let U be the set of homogeneous points of u, i.e.*

$$U = \{x \in \mathbb{R}^m \ \text{ s.t. } \ u \text{ is homogeneous wrt } x\}\,. \tag{2.53}$$

If $0 \in U$ and U spans a k-dimensional subspace V, then u is k-symmetric wrt V. As a consequence $U = V$.

Proof. The proof is very geometrical. Let v be any vector in V. Then by definition there exists k points $\{x_i\}$ of U with $x_i \neq 0$ such that

$$v = \sum_i \alpha_i x_i\,. \tag{2.54}$$

For all $x \in \mathbb{R}^n$, we have

1. $\langle \nabla u(x), x \rangle = 0$ because u is homogeneous wrt 0
2. $\langle \nabla u(x), x - x_i \rangle = 0$ because u is homogeneous wrt x_i.

Then for almost all $x \in R^m$:

$$\langle \nabla u, v \rangle = \sum_i \left\langle \nabla u(x), \sum_i \alpha_i x_i \right\rangle = \sum_i \left\langle \nabla u(x), \sum_i \alpha_i (x_i - x) \right\rangle = 0\,. \tag{2.55}$$

This concludes the proof. □

Note that $u(x) = x/|x|$ is obviously homogeneous wrt the origin. Note also that if u is continuous and homogeneous, it must be constant. From the fact that $\mathcal{H}^{m-2}(\mathcal{S}(u)) = 0$, we also know that

Lemma 2.11. *Let u be a minimizing harmonic map which is also $m-2$-symmetric. Then u is constant, and thus it has no singular points.*

NOTE: TO BE ADDED: $\theta(x)$ is upper-semicontinuous, and thus $\mathcal{S}(u)$ is a closed set.

MOREOVER: ADD (AND PROVE) UNIQUE CONTINUATION FOR STATIONARY MAPS, assuming unique continuation for elliptic systems in $\mathbb{R}^m$.

3 Quantitative stratification by Cheeger and Naber

In this lecture, we cover the quantitative stratification technique introduced by Cheeger and Naber in [CN13]. In particular, our objective today is to prove the following theorem.

Theorem 3.1. *Let u be a minimizing harmonic map on $B_2(0) \subseteq \mathbb{R}^m$ with $\theta(0,2) \leq \Lambda$. Then for all $\eta > 0$*

$$\operatorname{Vol}\{B_r(\mathcal{S}(u)) \cap B_1(0)\} \leq C(m, \Lambda, N, \eta) r^{3-\eta} . \tag{3.1}$$

Remark 3.1. Note that this theorem is "just an η" away from the final objective of the course. However, this η is really a big deal. Roughly speaking, it is the difference between fractal behaviour and non-fractal behaviour.

We start by showing that the properties of model maps are stable under suitable perturbation.

Remark 3.2. Throughout this section, we will always assume that u satisfies the hypotheses of Theorem 3.1.

The first property that we notice is that, with a simple compactness argument, we can make the statement in Lemma 2.11 effective in the following sense.

Lemma 3.2. *There exists an $\epsilon(m, N, \Lambda) > 0$ such that if for some constant c we have*

$$\|u - c\|_{L^2(B_2(0))} \leq \epsilon , \tag{3.2}$$

then $u \in C^\infty(B_1(0))$.

Proof. The proof is just based on a compactness argument. Assume by contradiction that this is false, and consider a sequence u_i of minimizing harmonic maps with $\|u_i - c_i\|_{L^2(B_2(0))} \leq i^{-1}$. Since N is compact, up to passing to a subsequence we have $c_i \to c$, and thus (again up to passing to a subsequence) $\|u_i - c\|_{L^2(B_2(0))} \leq 2i^{-1}$.

Since $\|\nabla u_i\|_{L^2(B_2(0))} \leq \Lambda$, and by the strong convergence result for minimizing harmonic maps, we have that (up to a subsequence) u_i converges strongly to some u in $W^{1,2}$. By uniqueness of the limit, $u = c$ and so

$$\lim_{i\to\infty} \int_{B_2(0)} |\nabla u_i|^2 = 0 . \tag{3.3}$$

Now the statement is a corollary of the ϵ-regularity theorem. □

In order to make this statement even more effective, we introduce the following quantity.

Definition 3.3. The regularity scale $r_u(x)$ is defined as

$$r_u(x) = \sup\{s \geq 0 \ \ s.t. \ \ \forall y \in B_s(x) \ \ s|\nabla u(y)| \leq 1\} . \tag{3.4}$$

By the regularity theory, we know that

$$\mathcal{S}(u) = \{x \ \ s.t. \ \ r_u(x) = 0\} . \tag{3.5}$$

Note that this regularity scale is a scale-invariant quantity in the sense of Remark 2.1. We can rephrase the ϵ-regularity theorem in the following way.

Theorem 3.4. *There exist an* $\epsilon_0(m, N)$ *such that if* $\theta(x, 2r) \le \epsilon_0$*, then* $r_u(x) \ge r$*.*

Using the same proof as above, we can turn the previous lemma into the following more effective and scale-invariant version.

Lemma 3.5. *There exists an* $\epsilon(m, N, \Lambda) > 0$ *such that if for some constant* c *we have*

$$\fint_{B_{2r}(0)} |u - c|^2 \le \epsilon\,, \tag{3.6}$$

then $r_u(x) \ge r$*.*

Note that, in light of Lemma 2.11, we can rephrase this lemma in the following way.

Lemma 3.6. *There exists an* $\epsilon(m, N, \Lambda) > 0$ *such that if for some* $m - 2$*-symmetric map* h *we have*

$$\fint_{B_{2r}(x)} |u - h|^2 \le \epsilon\,, \tag{3.7}$$

then $r_u(x) \ge r$*. In particular,* $\mathcal{S}(u) \cap B_r(x) = \emptyset$*.*

3.1 Unique continuation for minimizing harmonic maps

Here we recall another important property of harmonic maps: unique continuation.

Theorem 3.7. *Let* $u : \Omega_1 \subseteq \mathbb{R}^m \to N$ *and* $w : \Omega_2 \subseteq \mathbb{R}^m \to N$ *be two minimizing harmonic maps, where* Ω_1 *and* Ω_2 *are open connected domains. If there exists an open set* B *such that*

$$u|_B = w|_B\,, \tag{3.8}$$

then $u = w$ *on* $\Omega_1 \cap \Omega_2$*. In particular, both maps can be extended to* $\Omega_1 \cup \Omega_2$*.*

Proof. With the basic estimates on the singular set given in Theorem 2.7, we know that $\mathcal{S}(u)$ is a non-disconnecting set. Thus, this theorem is a corollary of unique continuation for elliptic equations in $\mathbb{R}^m$. □

3.2 Almost homogeneity and symmetry

Now the question is: under which circumstances do we have that a generic map u is close to one of our models, and in particular to one of our $m - 2$-symmetric models? The answer is given by the monotonicity formula.

First of all, a characterization of the models.

Lemma 3.8. u *is homogeneous wrt* x *if and only if* $\theta(x, r)$ *is constant in* r*.*

Proof. This is a corollary of the formula (2.13). Indeed, if u is homogeneous wrt x, then by definition we have that for almost $y \in B_2(0)$:

$$\langle \nabla u, y - x\rangle = 0\,. \tag{3.9}$$

Thus, we obtain that for all $s \leq r \leq d(x, \partial B_2(0))$:

$$\theta(x,r) - \theta(x,s) = 0\,. \tag{3.10}$$

The converse direction follows from the same reasoning. □

By the unique continuation theorem, we can obtain a less stringent characterization of these maps. Indeed, we don't need to ask that $\theta(x,r)$ is constant *for all* r, but just that it is constant for an interval with nonempty interior.

Lemma 3.9. *u is homogeneous wrt x if and only if there exists some* $0 \leq s < r$ *such that* $\theta(x,r) - \theta(x,s) = 0$.

Proof. □

3.3 Dimension 3

In order to fix some ideas, we consider the 3-dimensional case before moving to the general case. Here, we already know that $\mathcal{H}^{n-2}(\mathcal{S}(u)) = 0$, but we would like to say some more on the singular set.

Lemma 3.10. *Let* $m = 3$. *Then there exists an* $\epsilon(m, N, \Lambda) > 0$ *such that if* $\theta(0, 3/2) - \theta(0, 1/4) \leq \epsilon$, *then* $\mathcal{S}(u) \cap B_1(0) \setminus B_{1/2}(0) = \emptyset$.

Proof. This is another contradiction argument. First, suppose that $\epsilon = 0$. In this case, we know that u is 0-symmetric and thus it cannot have any singular points away from the origin, otherwise this singular point would propagate in $m - 2 = 1$ dimension, and we know that this is not possible.

Now assume by contradiction that there is a sequence u_i with $\theta_{u_i}(0,2) \leq \Lambda$, $\theta_{u_i}(0,3/2) - \theta_{u_i}(0,1/4) \leq i^{-1}$ and with some $x_i \in B_1(0) \setminus B_{1/2}(0)$ which is a singular point. Thus, $\theta_{u_i}(x_i) \geq \epsilon_0$.

By compactness, we can pass to a subsequence and obtain that

1. $u_i \to u$ in $W^{1,2}$, where u is a minimizing harmonic map with $\theta(0,1) - \theta(0,1/2) = 0$, thus u is 0-symmetric,

2. $x_i \to x \in \overline{B_1(0)} \setminus B_{1/2}(0)$.

We show that $x \in \mathcal{S}(u)$, and this is a contradiction.

Since $u_i \to u$ in $W^{1,2}$, for all fixed $r > 0$ we have

$$\theta_u(x,r) = \lim_i \theta_{u_i}(x,r)\,. \tag{3.11}$$

We claim that this limit is $\geq \epsilon_0/2$, and this proves that $x \in \mathcal{S}(u)$.

Since $x_i \in \mathcal{S}(u_i)$, then $\theta_{u_i}(x_i) \geq \epsilon_0$. Take i sufficiently large in order to have $|x_i - x| \leq r/100$. Then

$$\int_{B_r(x)} |\nabla u|^2 = \lim_i \int_{B_r(x)} |\nabla u_i|^2 \geq \liminf_i \int_{B_{99r/100}(x_i)} |\nabla u_i|^2 \geq \frac{99r}{100}\epsilon_0\,. \tag{3.12}$$

Thus

$$\theta_u(x,r) \geq \frac{99}{100}\epsilon_0 \tag{3.13}$$

as claimed. This finishes the proof. □

Thus, with a simple compactness argument, we have shown the following.

Corollary 3.11. *Let u be a minimizing harmonic map in $\mathbb{R}^3$. Then its singular points are isolated.*

Proof. Let x be a singular point for u, and let $r_x > 0$ such that $\theta(x, 2r_x) - \theta(x, 0) \leq \epsilon$, where ϵ is the constant in the previous lemma. Then the previous lemma applied to all the annuli

$$B_s(x) \setminus B_{s/2}(x) = \emptyset \tag{3.14}$$

for all $0 < s \leq r_x$ tells us that in all of these annuli there is no singular point, and thus $\mathcal{S}(u) \cap B_{r_x}(x) = \{x\}$, as desired. $\square$

ADD IN THE NOTES: EFFECTIVE ARGUMENT IN THIS CASE, i.e.

Theorem 3.12. *Let u be a minimizing harmonic map in $B_2(0) \subseteq \mathbb{R}^3$. Then*

$$\#\{\mathcal{S}(u) \cap B_1(0)\} \leq C(m, N, \Lambda). \tag{3.15}$$

See [NVV,].

3.4 General dimension and rigid cone-splitting

For general dimensions, even model maps can have non-isolated singularities, and this makes bounding the singular set much more complicated. First of all, we want to obtain a suitable generalization of Lemma 3.10.

A "complicated" way to restate Lemma 2.11 is the following.

Lemma 3.13. *Let u be a minimizing harmonic map. If there exists points $\{x_i\}_{i=0}^{m-2}$ that span an $m-2$ dimensional affine subspace $L \subset \mathbb{R}^m$ such that for all i and some $r > 0$*

$$\theta(x_i, 3r/2) - \theta(x_i, r/4) = 0, \tag{3.16}$$

then u is constant.

Proof. We are just saying that the points wrt which u is homogeneous span an $m-2$ dimensional space, which in turn implies that u is $m-2$-symmetric, and thus constant. $\square$

Now we want to prove a rigid version of this statement via a compactness argument. In order to do that, we need an effective notion of linear independence, which is stable under limits.

Definition 3.14. We say that a set of points $\{x_i\}_{i=0}^{k}$ is in τ-general position (or τ-linear independent) if $\{x_i\}_{i=0}^{m-2} \subset B_2(0)$, and moreover for every $i = 1, \cdots, k$ we have

$$d(x_i, x_0 + \operatorname{span}\{x_1 - x_0, \cdots, x_{i-1} - x_0\}) \geq \tau. \tag{3.17}$$

As a corollary, we obtain that

Corollary 3.15. *Let $\{x_i\}$ be points in τ-general position, then for all $p \in V = x_0 + \operatorname{span}\{x_i - x_0\}$ there exists a unique set of numbers $\alpha_i(p)$ such that*

$$p = x_0 + \sum_i \alpha_i (x_i - x_0) , \quad |\alpha_i| \leq c(k, \tau)|p - x_0| . \tag{3.18}$$

Proof. This proof can be obtained as a simple application of Gram-Schmidt orthonormalization procedure. □

Remark 3.3. One can see with simple examples that the notion of being τ-linear independent is not invariant under permutations of $\{x_i\}$. However, using the previous lemma, one can show that if $\{x_i\}_{i=0}^k$ are τ-linearly independent, then for all permutations σ of k-indexes, $\{x_{\sigma(i)}\}_{i=0}^k$ are still $c(k, \tau)$-linearly independent.

Proof. □

It is easy to see that if $\left\{x_{i,j}\right\}_{i=0}^k$ are linearly independent for all j, it is not the case that $x_i = \lim_{j\to\infty} x_{ij}$ are still linearly independent. However, the notion of being τ-linearly independent passes to the limit.

Lemma 3.16. *The notion of τ-linear independence passes to the limit.*

Proof. □

Now we state the main lemma of this section.

Lemma 3.17. *For all $\tau > 0$, there exists an $\epsilon = \epsilon(m, N, \Lambda, \tau)$ such that if there exists points $\{x_i\}_{i=0}^{m-2}$ that are in τ-general position and for all i:*

$$\theta(x_i, 3r/2) - \theta(x_i, r/4) \leq \epsilon , \tag{3.19}$$

then

$$\theta(0, 1) \leq \epsilon_0 \quad \textit{and thus} \qquad r_u(0) \geq 1/2 . \tag{3.20}$$

Proof. The proof is based on a simple compactness argument as before, HERE IT'S JUST SKETCHED.

Take a contradicting sequence u_j with points $\left\{x_{ij}\right\}$. In the limit, we get a map u with points x_i still in τ-general position such that for all i

$$\theta(x_i, 3r/2) - \theta(x_i, r/4) = 0 . \tag{3.21}$$

By the previous lemma, u is constant, and the ϵ-regularity theorem concludes the proof. □

We can restate the previous lemma in the following way.

Lemma 3.18. *For all $\rho > 0$ there exists a $\bar{\epsilon} = \epsilon(m, N, \Lambda, \rho) > 0$ such that*

$$\mathcal{A} = \{y \cap B_r(x) \;\; s.t. \;\; \theta(y, 3r/2) - \theta(y, r/4) \leq \epsilon\} \subseteq B_{\rho r}(V) , \tag{3.22}$$

where V depends on u (and on the ball we are considering) and it is an affine $m-3$-dimensional subspace of $\mathbb{R}^m$.

Proof. Corollary of the previous one. □

3.5 Two easy covering lemmas

Before we proceed with the final proof, we recall two easy covering lemmas. First of all, in $\mathbb{R}^m$ we can always cover a ball of radius r with $c(m)\rho^{-m}$ balls of radius ρr

Lemma 3.19. *Let $B_r(x) \subseteq \mathbb{R}^m$. If $\rho \leq 1/5$, there exists a covering of this ball by $c(m)\rho^{-m}$ balls of radius ρr centered on $B_r(x)$.*

Proof. Standard covering. □

Second, if we want to cover a set $S \subseteq B_r(x)$ such that $S \subseteq B_{\rho r}(V)$, where V is a k-dimensional space, then we can do it with $c(m)\rho^{-k}$ balls instead of $c(m)\rho^{-m}$ as in the previous covering.

Lemma 3.20. *Let $S \subseteq B_r(x) \subseteq \mathbb{R}^m$ be a set such that $S \subseteq B_{\rho r}(V)$, where V is an affine k-dimensional subspace. If $\rho \leq 1/5$, there exists a covering of this ball by $c_0(m)\rho^{-k}$ balls of radius ρr centered on $B_r(x)$.*

Proof. We can assume wlog that $x = 0$ and $r = 1$.

First of all, we cover $B_{\rho r}(V)$ by $c(k)\rho^{-k}$ balls of radius 5ρ. Indeed, consider the open covering

$$B_\rho(V) \subseteq \bigcup_{x \in V \cap B_1(0)} B_\rho(x) \ . \tag{3.23}$$

Take a Vitali subcovering of this covering, so that

$$B_\rho(V) \subseteq \bigcup_{x_i \in I} B_{5\rho}(x_i) \ , \tag{3.24}$$

where $B_\rho(x_i)$ are pairwise disjoint. Thus, we get that

$$\sum_{i \in I} \mathcal{H}^k(B_{5\rho}(x_i)) = c(k)\rho^k \#(I) \, . \tag{3.25}$$

Moreover, $\cup_{i \in I} B_\rho(x_i) \cap V \subseteq B_2(0) \cap V$, thus

$$\sum_{i \in I} \mathcal{H}^k(B_{5\rho}(x_i) \leq \mathcal{H}^k(B_2(0) \cap V) \leq c(k) \, . \tag{3.26}$$

Now apply the previous lemma to each of the balls $B_{5\rho}(x_i)$ in order to obtain a covering of these balls by

$$B_{5\rho}(x_i) \subseteq \bigcup_{j \in J_i} B_\rho\left(y_j\right) \ , \tag{3.27}$$

with $\#(J_i) \leq c(m)$. Then we obtain that

$$S \subseteq \bigcup_{j \in \cup_i J_i} B_\rho\left(y_j\right) \ , \tag{3.28}$$

where

$$\sum_i \#(J_i) \leq c(m)\#I \leq c(m) \, . \tag{3.29}$$

□

3.6 First attempt to prove the theorem

Here we describe a proof of the Theorem 3.1 with a "cheat". Basically, we will assume that for all $x \in \mathcal{S}$ and for all $r \leq 1$,

$$\theta(x, 3r/2) - \theta(x, r/4) \leq \bar{\epsilon} \, . \tag{3.30}$$

Of course, this is a big cheat, and we will need to fix this later. However, this may be helpful in separating the two parts of the argument for the proof of the main theorem.

The proof *with a cheat* is basically a corollary of the following covering lemma.

Lemma 3.21. *For all $\delta > 0$, there exists a parameter $\rho(\delta, m) > 0$ such that the following is true. Let $\mathcal{S} \subset B_1(0)$ be a set such that for all $x \in B_1(0)$ and all $0 \leq r \leq 1$ there exits a k-dimensional affine space $V = V(x, r) \subset \mathbb{R}^m$ such that*

$$\mathcal{S} \cap B_r(x) \subseteq B_{\rho r/5}(V) \, . \tag{3.31}$$

Then

$$\operatorname{Vol}(B_r(\mathcal{S})) \leq C(m) r^{m-k+\delta} \, . \tag{3.32}$$

Proof. The proof is easy, it is carried out with an inductive covering.

We will prove that for all $s \in \mathbb{N}$, there exists a covering of $\mathcal{S}$ made by balls of radius ρ^s such that the number of these balls is bounded by

$$\left(c(m)\rho^{-k}\right)^s \, . \tag{3.33}$$

With this, we would have

$$\operatorname{Vol}(B_r(\mathcal{S})) \leq c(m) \left(c(m)\rho^{-k}\right)^s (\rho^s)^m \, . \tag{3.34}$$

If we pick $\rho = \rho(m, \delta)$ such that

$$c(m)\rho^{\delta} \leq 1 \, , \tag{3.35}$$

we obtain the thesis.

In order to prove the claim, we prove inductively that there exists a family of balls $\mathcal{C}(s)$ of radius ρ^s such that they cover the set $\mathcal{S}$ and whose number is bounded by (3.33) as claimed.

Let $\mathcal{C}(0) = \{B_1(0)\}$. We procede by induction. If we have $\mathcal{C}(s)$, pick any ball inside this family, and apply the covering lemma 3.20 to this ball.

The family $\mathcal{C}(s+1)$ will simply be the union of all these covering for each of the balls in $\mathcal{C}(s)$. This concludes the proof.

□

3.7 A less easy covering lemma

Now we are ready for the proof of the main estimate in this section *without any cheat*. As already stated, this result is in [CN13].

First of all, we reduce the main theorem to a covering lemma. Recall that the theorem we want to prove is:

Theorem 3.22. *Let u be a minimizing harmonic map on $B_2(0) \subseteq \mathbb{R}^m$ with $\theta(0,2) \leq \Lambda$. Then for all $\eta > 0$*

$$\operatorname{Vol}\{B_r(\mathcal{S}(u)) \cap B_1(0)\} \leq C(m,\Lambda,N,\eta) r^{3-\eta}\,. \tag{3.36}$$

For convenience, from now on we set $\mathcal{S} \equiv \mathcal{S}(u) \cap B_1(0)$.

Basically we will fix a scale ρ and partition the set $\mathcal{S}(u)$ into subsets according to wether x is "good" at scale ρ^j or not. In particular, we say that $x \in \mathcal{G}_j \subseteq B_1(0)$ if

$$\theta(x, 2\rho^j) - \theta(x, \rho^j/4) \leq \bar{\epsilon}\,, \tag{3.37}$$

where $\bar{\epsilon} = \epsilon(m,N,\Lambda,\rho) > 0$ is the parameter of Lemma 3.18.

For each x, let $T(x)$ be an infinite dimensional vector of zeros and ones build in this way:

$$\text{if} \quad x \in \mathcal{G}_j\,, \quad T(x)[j] = 0\,, \quad \text{otherwise} \;\; T(x)[j] = 1\,. \tag{3.38}$$

We also denote by $T_s(x)$ the s-dimensional vector obtained with the first s-components of $T(x)$.

A first observation is the following.

Lemma 3.23. *For all $x \in B_1(0)$,*

$$|T(x)| = \sum_{j=0}^{\infty} T(x)[j] \leq c(m)\Lambda/\bar{\epsilon} < \infty\,. \tag{3.39}$$

Proof. The proof is a straighforward application of the monotonicity of θ. Indeed we have

$$c(m)\Lambda \geq \theta(x,1) - \theta(x,0) \geq \sum_{j=0}^{\infty} \theta(x, 2\rho^j) - \theta(x, \rho^j/4) \geq \bar{\epsilon}\,|T(x)|\,. \tag{3.40}$$

□

Definition 3.24. For each finite s, we denote by $\mathcal{F}(T_s)$ the set of $x \in \mathcal{S}$ such that $T_s(x) = T_s$. In other words, it is the subset of $\mathcal{S}$ such that

$$x \in \mathcal{G}_j \;\; \forall j \;\; s.t. \;\; T_s[j] = 0 \quad \text{and} \quad x \notin \mathcal{G}_j \;\; \forall j \;\; s.t. \;\; T_s[j] = 1\,. \tag{3.41}$$

The objective now is to build inductively for all $s \in \mathbb{N}$ a covering of $\mathcal{S}$ such that

$$\mathcal{S} \subseteq \cup_{T_s} \mathcal{C}(T_s)\,, \tag{3.42}$$

where $\mathcal{C}(T_s)$ is a union of a controlled number of balls, and the number of "nontrivial" s-tuples T_s is controlled as well.

Proposition 3.25. *For all $\rho > 0$, we can build a covering of $\mathcal{S}$ with*

$$\mathcal{S} \subseteq \cup_{T_s} \mathcal{C}(T_s)\,, \tag{3.43}$$

where

1. *if $|T_s| \geq K$, then $\mathcal{C}(T_s) = \emptyset$, where $K = K(m, N, \Lambda, \rho)$*
2. *$\mathcal{F}(T_s)$ is covered by $\mathcal{C}(T_s)$, so if $\mathcal{F}(T_s) = \emptyset$, also $\mathcal{C}(T_s)$ will be an empty covering*
3. *each $\mathcal{C}(T_s)$ is a union of at most*

$$(c(m)\rho^{-m})^K (c_0(m)\rho^{-m+3})^{s-K} \tag{3.44}$$

balls of radius ρ^{s+1}.

Before proving this proposition, we use it to prove our main theorem.

Proof of Theorem 3.1. Now we choose the parameter ρ as

$$\rho(m, \eta) = c_0(m)^{-2/\eta}\,. \tag{3.45}$$

Note that as long as $\eta > 0$, then $0 < \rho << 1$.

It is sufficient to prove the theorem for $r = \rho^s$ and $s \in \mathbb{N}$ arbitrary. With a simple reasoning, the result will extend to all r, up to losing another constant depending on (m, η).

Fix some $s \in \mathbb{N}$. Recall that for any set A, if we know that

$$A \subseteq \bigcup_{s \in S} B_r(x_i)\ , \tag{3.46}$$

then

$$\mathrm{Vol}(B_r(A)) \leq c(m)\#(S) r^m\,. \tag{3.47}$$

Thus, we just need to check how many balls we have in our covering to obtain the final estimate.

Given condition (1), we know that T_s is "meaningful" (in the sense that $\mathcal{F}(T_s) \neq \emptyset$) only if $|T_s| \leq K$. Thus, we can have at most

$$\binom{s}{0} + \binom{s}{1} + \binom{s}{2} + \cdots + \binom{s}{\min s, K} \tag{3.48}$$

of these s-tuples. We can give a rough bound to this and say that

$$\binom{s}{0} + \binom{s}{1} + \binom{s}{2} + \cdots + \binom{s}{\min\{s, K\}} \leq K s^K\,. \tag{3.49}$$

Note that, since $K = K(m, N, \Lambda, \rho)$ and since we have fixed ρ according to (3.45), we can bound extremely roughly this quantity by

$$K s^K \leq c(m, N, \Lambda, \eta)\,(\rho^s)^{-\eta/2}\ . \tag{3.50}$$

Thus, by (3.44), we have that $\mathcal{S}$ is covered by a number of balls of radius ρ^s which is controlled by

$$c(m,N,\Lambda,\eta)\,(\rho^s)^{-\eta/2}\,(\rho^s)^{-\eta/2}\left(\rho^{-m+3}\right)^{-\eta/2}(\rho^s)^{-m}\ . \tag{3.51}$$

This concludes the proof. □

Now we turn to the proof of Proposition 3.25.

Proof of Proposition 3.25. We prove this proposition by induction. At the very first step, our covering of $\mathcal{S}$ is just given by $B_1(0)$.

Now there are two admissible 1-tuples T_0, which are given by $T_0 = [0]$ and $T_0 = [1]$. The family $\mathcal{F}([1])$ is given by the points which have no small pinching at scale 1, while the family $\mathcal{F}([0])$ are the points in $\mathcal{S}$ that have the small pinching condition on at scale 1.

The family $\mathcal{C}([1])$ is obtained by covering $B_1(0)$ in a “stupid” way, so by using the covering of Lemma 3.19. Thus $\mathcal{C}([1])$ is made up of $c(m)\rho^{-m}$ balls of radius ρ covering $B_1(0)$.

The family $\mathcal{C}([1])$ is obtained by covering $B_1(0)$ in a “more clever” way. Since we have the pinching condition for these points, we can apply Lemma 3.18 and in turn Lemma 3.20 to cover all the points in $\mathcal{F}([0])$ by $c(m)\rho^{3-m}$ balls of radius ρ^m.

Induction step Assuming that we have build all the families $\mathcal{C}(T_i)$ up to the index $i \leq s$, and that all these families cover the sets $\mathcal{F}(T_i)$ respectively, we want to build $\mathcal{C}(T_{i+1})$ for all the possible choices of T_{i+1}.

Fix an $s+1$-tuple T_{s+1}. As noted before in Lemma 3.23, if $|T| > c(m)\Lambda/\bar{\epsilon}$, then $\mathcal{F}(T_{s+1})$ is empty and there’s nothing to cover. So in this case $\mathcal{C}(T_{s+1})$ is empty as well.

Now, if $T_{s+1}[s+1] = 1$, this means that $\mathcal{F}(T_{s+1})$ is the set of points inside $\mathcal{F}(T_s)$ that have no pinching condition at scale $s+1$. Thus take all the balls in $\mathcal{C}(T_s)$, and apply Lemma 3.19 to all of these balls separarely. We obtain a covering of $\mathcal{F}(T_{s+1})$ of balls of radius ρ^{s+1}, and the number of these balls is equal to the number of balls in $\mathcal{C}(T_s)$ multiplied by $c(m)\rho^{-m}$.

If instead $T_{s+1}[s+1] = 0$, this means that $\mathcal{F}(T_{s+1})$ is the set of points inside $\mathcal{F}(T_s)$ that have a good pinching condition at scale $s+1$. Consider any of the balls in the family $\mathcal{C}(T_s)$. The points in $\mathcal{F}(T_{s+1})$ inside one of these balls satisfy Lemma 3.18, and so we can cover them using Lemma 3.20. We let the union of all these balls be $\mathcal{C}(T_{s+1})$. The number of the balls in this family is equal to the number of balls in the family $\mathcal{C}(T_s)$ multiplied by $c_0(m)\rho^{3-m}$.

This concludes the induction step.

□

TO DO: EXTEND THE RESULT FOR THE QUANTITATIVE STRATIFICATION, NOT JUST THE TOP STRATUM.

4 Classical Reifenberg theorem

In this lecture, we introduce the classical Reifenberg theorem, and prove it. We will prove some technical lemmas with more generality than needed, in such a way that in the coming lecture we will be able to use them again for more general versions of the theorem.

4.1 Preliminaries

First of all, we briefly recall some definitions and properties of Hausdorff measure and rectifiability. Since this subject has been studied extensively, there are many books that can be used to learn this subject in more details. In particular, we refer the reader to [DL08, Mat95, Fed69].

Definition 4.1. Given a generic set $E \subseteq \mathbb{R}^n$, we can set

$$\text{for all}\delta > 0\,, \quad \mathcal{H}^k_\delta(E) = \frac{\omega_k}{2^k} \inf \left\{ \sum_{i\in I} \operatorname{diam}(E_i)^k \;\; s.t. \;\; E \subseteq \bigcup_{i\in I} E_i \;\text{ and }\; \operatorname{diam}(E_i) \leq \delta \right\}\,, \tag{4.1}$$

where E_i are generic subsets of $\mathbb{R}^n$, and ω_k is the volume of the unit sphere in $\mathbb{R}^k$.

Taking the limit as $\delta \to 0$, we obtain the (outer) Hausdorff measure of E. In particular

$$\mathcal{H}^k(E) = \lim_{\delta\to 0} \mathcal{H}^k_\delta(E) = \sup_{\delta>0} \mathcal{H}^k_\delta(E)\,. \tag{4.2}$$

Note that $\mathcal{H}^k$ is a measure on Borel sets.

We recall the main properties of $\mathcal{H}^k$:

Proposition 4.2. *The following properties hold:*

1. *for $k = n$, $\mathcal{H}^n$ is the Lebesgue measure of $\mathbb{R}^n$,*
2. *given an embedded k-dimensional Riemannian submanifold $M \subseteq \mathbb{R}^n$, $\mathcal{H}^k \llcorner M$ coincides with the usual Riemannian measure on M,*
3. *blow-up properties: given $E \subseteq \mathbb{R}^n$, let $E_{x,r}$ be the set*

   $$E_{x,r} = \{y \in \mathbb{R}^n \;\; s.t. \;\; \exists z \in E \;\; : \;\; y = x + rz\} \equiv x + rE\,. \tag{4.3}$$

   We have that

   $$\mathcal{H}^k(E_{x,r}) = r^k \mathcal{H}^k(E)\,. \tag{4.4}$$

4. *Lipschitz bounds: give a Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}^m$ with Lipschitz constant L, for all k and all (Borel) subsets $E \subseteq \mathbb{R}^n$, we have*

   $$\mathcal{H}^k(f(E)) \leq L^k \mathcal{H}^k(E)\,. \tag{4.5}$$

Definition 4.3. Given a Borel set $S \subseteq \mathbb{R}^n$, we say that S is k-rectifiable if there exist a countable collection of Lipschitz maps $f_i : \mathbb{R}^k \to \mathbb{R}^n$ such that

$$\mathcal{H}^k\left(S \setminus \bigcup_i f_i(\mathbb{R}^k)\right) = 0\,. \tag{4.6}$$

A measure μ is called k-rectifiable if its support S is k-rectifiable and there exists a measurable function θ such that

$$\mu = \theta \cdot \mathcal{H}^k \llcorner S\,. \tag{4.7}$$

An important consequence of rectifiability is the existence of tangent cones/measures at almost all points of S. In particular

Proposition 4.4. *Given a measure μ, we define the blow-up measure $\mu_{x,r}$ as*

$$\mu_{y,r}(A) = \mu(y + rA)\,, \tag{4.8}$$

where A is a generic Borel set, and

$$y + rA = \{z \in \mathbb{R}^n \;\; s.t. \;\; \exists x \in A \; : \; z = y + rz\}\,. \tag{4.9}$$

If μ is k-rectifiable, then for μ-almost all $x \in \mathbb{R}^n$,

$$r^{-k}\mu_{x,r} \rightharpoonup^* \theta(x) \cdot \mathcal{H}^k \llcorner L_x\,, \tag{4.10}$$

where L_x is a k-dimensional plane. For a set S, this means that S has a unique k-dimensional tangent plane for almost all x.

4.2 Classical Reifenberg statement

Now we are ready to define the notion of "Reifenberg flat set".

Definition 4.5. The Hausdorff distance between two sets A, B is defined by

$$d_{\mathcal{H}}(A, B) = \inf\{t > 0 \;\; s.t. \;\; A \subseteq B_t(B) \;\; \text{and} \;\; B \subseteq B_t(A)\}\,. \tag{4.11}$$

Definition 4.6. Let $S \subseteq \mathbb{R}^m$ be a set. Assume that $0 \in S \subset B_4(0)$. We say that S is (k, δ)-Reifenberg flat in $B_1(0)$, or that S has the (k, δ)-Reifenberg approximation property $B_1(0)$ if for all $y \in S \cap B_2(0)$, and for each $0 < r \leq 2$ there exists a k-dimensional affine subspace $L(y, r)$ such that

$$d_{\mathcal{H}}(S \cap B_r(y)\,, L(y, r) \cap B_r(y)) \leq \delta r\,. \tag{4.12}$$

Remark 4.1. Up to a factor of 10 in front of the δ, we can assume wlog that $L(y, r) = y + L'(y, r)$, where L' is a linear subspace. In other words, we can assume that $y \in L(y, r)$.

Example 4.1. If a set S can be written as the graph of a Lipschitz function f, in particular if for some

1. k-dimensional subspace $V_k \subseteq \mathbb{R}^n$
2. δ-Lipschitz function $f : V_k \to V_k^{\perp}$, and $f(0) = 0$
3. we have

$$S = \{x + f(x) \ \ x \in V_k\} = \{(x, f(x)), \ \ x \in V_k\} \tag{4.13}$$

then it is easy to see that S is $c(k)\epsilon$-Reifenberg flat. Indeed, for all $p \in S$, we have $p = (x, f(x))$. Consider the affine plane $L(p) = (x, f(x)) + V_k$, it is easy to see that

$$d_{\mathcal{H}}(S \cap B_r(p), L(p) \cap B_r(p)) \le c(k)\delta r\,. \tag{4.14}$$

Theorem 4.7 (Classical Reifenberg Theorem [Rei60, Sim, Mor66]). *For each $0 \le \alpha < 1$ and $\epsilon > 0$ there exists $\delta(m, \alpha, \epsilon) > 0$ such that the following holds. Assume $0^n \in S \subseteq B_2 \subseteq \mathbb{R}^n$ is a closed subset, and that it is (k, δ)-Reifenberg flat. Then there exists $\phi : \mathbb{R}^k \to \mathbb{R}^m$ which is a C^α bi-Hölder homeomorphism onto its image*

1. $S \cap B_1(0) \subseteq \phi(\mathbb{R}^k)$

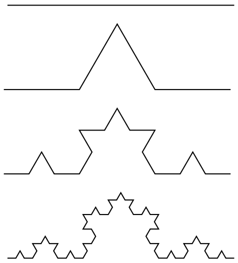
2. $[\phi]_{C^\alpha}, [\phi^{-1}]_{C^\alpha} < 1 + \epsilon$

First of all, we show that there is a counterexample for this theorem with $\alpha = 1$. In other words, if we want ϕ to be bi-Lipschitz, condition (4.12) is not enough (unless of course $\epsilon = 0$).

[The notes after here are taken from [NVb]]

4.3 Explanatory example

In order to understand better the idea behind the improvement of the Reifenberg theorem, we use the famous snow-flake as a test case.



The construction of a snowflake of parameter $\eta > 0$ is well known (see for example [Mat95, section 4.13]). Take the unit segment $[0, 1] \times \{0\} \subseteq \mathbb{R}^2$, and replace the middle part $[1/3, 2/3] \times \{0\}$ with the top part of the isosceles triangle with base $[1/3, 2/3] \times \{0\}$ and of height $\eta \cdot$ lenght($[1/3, 2/3] \times \{0\}$). In other words, you are replacing the segment $[1/3, 2/3] \times \{0\}$ with the two segments joining $(1/3, 0)$ to $(1/2, \eta/3)$, and $(1/2, \eta/3)$ to $(2/3, 0)$. Then repeat this construction inductively on each of the 4 straight segments in the new set. Here on the left hand side you can see the very classical picture of the first three steps in the construction of the standard snowflake, with $\eta = \sqrt{3}/2$.

It is clear that the length of the curve at step i is equal to the length at step $i - 1$ times $2/3 + \sqrt{1 + 4\eta^2}/3$, so the lenght of the snowflake will be infinity for any $\eta > 0$. This is a simple application of the Pythagorean theorem, and the extra square power on η comes from the fact that at each step we are adding some lenght η to the curve, but in a direction perpendicular to it.

However, if we replace the fixed parameter η with a variable parameter η_i, we see immediately that the lenght of the limit curve will be finite if and only if $\sum \eta_i^2 < \infty$.

In particular, this implies that we cannot hope to prove that the snowflake is a $C^{0,1}$ image of a segment if η_i is just bounded, or even if $\eta_i \to 0$.

4.4 Technical Constructions toward New Reifenberg Results

In this section, we prove some technical lemmas needed for dealing with the relation between approximating spaces. These elementary results will be used in many of the estimates of subsequent sections.

4.5 Hausdorff distance and subspaces

We start by recalling some standard facts about affine subspaces in $\mathbb{R}^n$ and Hausdorff distance.

Definition 4.8. Given two linear subspaces $L, V \subseteq \mathbb{R}^n$, we define the Grassmannian distance between these two as

$$d_G(L,V) = d_H(L \cap B_1(0), V \cap B_1(0)) = d_H\left(L \cap \overline{B_1(0)}, V \cap \overline{B_1(0)}\right) . \tag{4.15}$$

Note that if $\dim(L) \neq \dim(V)$, then $d_G(L,V) = 1$.

For general subsets in $\mathbb{R}^n$, it is evident that $A \subseteq B_\delta(B)$ does not imply $B \subseteq B_{c\delta}(A)$. However, if A and B are affine spaces with the same dimension, then it is not difficult to see that this property holds. More precisely:

Lemma 4.9. *Let V, W be two k-dimensional affine subspaces in $\mathbb{R}^n$, and suppose that $V \cap B_{1/2}(0) \neq \emptyset$. There exists a constant $c(k,n)$ such that if $V \cap B_1(0) \subseteq B_\delta(W)$, then $W \cap B_1(0) \subseteq B_{c\delta}(V \cap B_1(0))$. Thus in particular $d_H(V \cap B_1(0), W \cap B_1(0)) \leq c\delta$.*

Proof. The proof relies on the fact that V and W have the same dimension. Let $x_0 \in V$ be the point of minimal distance from the origin. By assumption, we have that $\|x_0\| \leq 1/2$. Let $x_1, \cdots, x_k \in V \cap \overline{B_1(0)}$ be a sequence of points such that

$$\|x_i - x_0\| = 1/2 \quad \text{and for } i \neq j, \quad \left\langle x_i - x_0, x_j - x_0 \right\rangle = 0 . \tag{4.16}$$

In other words, $\{x_i - x_0\}_{i=1}^k$ is an affine base for V. Let $\{\tilde{y}_i\}_{i=0}^k \subseteq W$ be such that $d(x_i, \tilde{y}_i) \leq \delta$. This in particular implies that (for δ sufficiently small) $\|\tilde{y}_0\| \leq 2/3$ and $\|\tilde{y}_i\| \leq 1 + \delta$. Thus, there exists $\{y_i\}_{i=0}^k$ such that $y_i \in W \cap B_1(0)$ and $d(x_i, y_i) \leq 3\delta$. Then

$$\|y_i - y_0\| \geq 1/2 - 10\delta \quad \text{and for } i \neq j, \quad \left|\left\langle y_i - y_0, y_j - y_0 \right\rangle\right| \leq 10\delta + 10\delta^2 . \tag{4.17}$$

This implies that for $\delta \leq \delta_0(n)$, $\{y_i - y_0\}_{i=1}^k$ is an affine base for W and for all $y \in W$

$$y = y_0 + \sum_{i=1}^{k} \alpha_i (y_i - y_0), \quad |\alpha_i| \leq c(k) \|y - y_0\| . \tag{4.18}$$

Now let $y \in W \cap \overline{B_1(0)}$ be the point of maximum distance from V, and let π be the projection onto V and $\pi^\perp$ the projection onto $V^\perp$, which is the linear subspace orthogonal to V. Then

$$d(y, V) = d(y, \pi(y)) = \left\|\pi^\perp(y - x_0)\right\| \leq \left\|\pi^\perp(y_0 - x_0)\right\| + \sum_{i=1}^{k} |\alpha_i| \left\|\pi^\perp(y_i - y_0)\right\| \leq c'(n,k)\delta . \tag{4.19}$$

Since $y \in \overline{B_1(0)}$, then $\pi(y) \in V \cap B_{1+c'\delta}(0)$, and thus $d(y, V \cap B_1(0)) \leq 2c'\delta \equiv c\delta$, which proves the claim. □

Also:

Lemma 4.10. *Let V, W be two k-dimensional affine subspaces in* $\mathbb{R}^n$*, and suppose that* $V \cap B_{1/2}(0) \neq \emptyset$*. There exists a constant* $c(k,n)$ *(maybe even a fixed constant* c*) such that if* $d_H(V \cap B_1(0), W \cap B_1(0)) \leq \delta$*, then if* $\tilde{V}$ *and* $\tilde{W}$ *are the linear subspaces associated to V and W, we have*

$$d_G(\tilde{V}, \tilde{W}) \leq c\delta . \tag{4.20}$$

Proof. TO BE COMPLETED □

Next we will see that the Grassmannian distance between two subspaces is enough to control the projections with respect to these planes. In order to do so, we recall a standard estimate.

Lemma 4.11. *Let V, W be linear subspaces of a Hilbert space. Then* $d_G(V, W) = d_G(V^\perp, W^\perp)$*.*

Proof. We will prove that $d_G(V^\perp, W^\perp) \leq d_G(V, W)$. By symmetry, this is sufficient.

Take $x \in V^\perp$ such that $\|x\| = 1$, and consider that $d(x, W^\perp) = \|\pi_W(x)\|$. Let $z = \pi_W(x)$ and $y = \pi_V(z)$. We want to show that if $d_G(V, W) \leq \epsilon < 1$, then $\|z\| \leq \epsilon$. We can limit our study to the space spanned by x, y, z, and assume wlog that $x = (1, 0, 0)$, $y = (0, b, 0)$ and $z = (a, b, c)$. By orthogonality between z and $z - x$, we have

$$a^2 + b^2 + c^2 + (1 - a)^2 + b^2 + c^2 = 1 \quad \Longrightarrow \quad a = a^2 + b^2 + c^2 , \tag{4.21}$$

and since $z \in W$, we also have $\|z - y\| \leq \epsilon \|z\|$, which implies

$$a^2 + c^2 \leq \epsilon^2 \left(a^2 + b^2 + c^2\right) \quad \Longrightarrow \quad a^2 + c^2 \leq \frac{\epsilon^2}{1 - \epsilon^2} b^2 . \tag{4.22}$$

Since the function $f(x) = x^2/(1 - x^2)$ is monotone increasing for $0 \leq x < 1$, we can define $0 \leq \alpha < 1$ in such a way that

$$a^2 + c^2 = \frac{\alpha^2}{1 - \alpha^2} b^2 , \quad a = a^2 + b^2 + c^2 = \frac{1}{1 - \alpha^2} b^2 . \tag{4.23}$$

Note that necessarily we will have $\alpha \leq \epsilon$. Now we have

$$\frac{1}{(1-\alpha^2)^2} b^4 = a^2 \leq \frac{\alpha^2}{1-\alpha^2} b^2 \quad \implies \quad b^2 \leq \alpha^2 \left(1-\alpha^2\right) \quad \implies \quad \|z\|^2 = a^2 + b^2 + c^2 \leq \alpha^2 \leq \epsilon^2 . \quad (4.24)$$

This proves that $V^\perp \cap B_1(0) \subset B_\epsilon(W^\perp)$. In a similar way, one proves the opposite direction. □

As a corollary, we prove that the Grassmannian distance $d_G(V, W)$ is equivalent to the distance given by $\|\pi_V - \pi_W\|$.

Lemma 4.12. *Let V, W be linear subspaces of $\mathbb{R}^n$. Then for every $x \in \mathbb{R}^n$,*

$$\|\pi_V(x) - \pi_W(x)\| \leq 2 d_G(V, W) \|x\| . \quad (4.25)$$

In particular, if $x \in W^\perp$, then $\|\pi_V(x)\| \leq 2 d_G(V, W) \|x\|$.

Conversely, we have

$$d_G(V, W) \leq \sup_{x \in \mathbb{R}^n \setminus \{0\}} \left\{ \frac{\|\pi_V(x) - \pi_W(x)\|}{\|x\|} \right\} . \quad (4.26)$$

Proof. The proof is just a corollary of the previous lemma. Assume wlog that $\|x\| = 1$, and let $x = y + z$ where $y = \pi_V(x)$ and $z = \pi_{V^\perp}(x)$. Then

$$\|\pi_V(x) - \pi_W(x)\| = \|y - \pi_W(y) - \pi_W(z)\| \leq \|y - \pi_W(y)\| + \|z - \pi_{W^\perp}(z)\| = d(y, W) + d(z, W^\perp) . \quad (4.27)$$

Since $\|y\|^2 + \|z\|^2 = \|x\|^2 = 1$, by the previous lemma we get the first estimate.

The reverse estimate is an immediate consequence of the definition. □

4.6 Simple estimate with Reifenberg condition

Here we state and prove a simple estimate related to the uniform Reifenberg condition 4.6.

Lemma 4.13. *If 4.6 is valid, then for all $x, y \in S \cap B_{5/4}(0)$ and $0 < \sigma \leq \rho \leq 1$, if $B_\sigma(y) \subseteq B_\rho(x)$, then*

$$d_G(L(x,\rho), L(y,\sigma)) \leq c(n)\delta \frac{\rho}{\sigma} , \quad d(x, y + L(y,\sigma)) + d(y, x + L(x,\rho)) \leq c\delta\rho . \quad (4.28)$$

Sketch of proof, TO BE COMPLETED. The proof is a simple application of the triangle inequality and lemmas 4.9,4.10. By 4.6, we have

$$y + L(y,\sigma) \cap B_\sigma(y) \subset B_{\delta\sigma}(S \cap B_\sigma(y)) \subset B_{\delta\sigma}\left(S \cap B_\rho(x)\right) \subset B_{\delta(\sigma+\rho)}(x + L(x,\rho)) . \quad (4.29)$$

Translating the ball $B_\sigma(y)$ to the origin, and blowing it up to scale 1, we obtain

$$L(y,\sigma) \subseteq B_{\delta(1+\rho/\sigma)}(\tilde{x} + L(x,\rho)) , \quad (4.30)$$

where $\tilde{x}$ is the result of this transformation (translation+blowup). Thus, we obtain the conclusion from lemma 4.9 and lemma 4.10. □

4.7 bi-Lipschitz equivalences

In this subsection, we study a particular class of maps with nice local properties. These maps are a slightly modified version of the maps which are usually exploited to prove Reifenberg's theorem, see for example [Rei60, Tor95, DT12], [Mor66, section 10.5] or [Sim]. The estimates in this section are standard in literature.

We start by defining the functions σ. For some $0 < r \leq 1$, let $\{x_i\}$ be an $r/10$-separated subset of $\mathbb{R}^n$, i.e.,

i. $d(x_i, x_j) \geq r/10$.

Let also p_i be a points in $\mathbb{R}^n$ with

ii. $p_i \in B_{10r}(x_i)$

and let V_i be a sequence of k-dimensional linear subspaces.

By standard theory, it is easy to find a locally finite smooth partition of unity $\lambda_i : \mathbb{R}^n \to [0, 1]$ such that

iii. $\operatorname{supp}(\lambda_i) \subseteq B_{3r}(x_i)$ for all i,

iv. for all $x \in \bigcup_i B_{2r}(x_i)$, $\sum_i \lambda_i(x) = 1$ and $\sum_i \lambda_i(x') \leq 1$ for all $x' \in \mathbb{R}^n$,

v. $\sup_i \|\nabla\lambda_i\|_\infty \leq c(n)/r$,

vi. if we set $1 - \psi(x) = \sum_i \lambda_i(x)$, then ψ is a nonnegative smooth function with $\|\nabla\psi\|_\infty \leq c(n)/r$.

Note that by (iii), and since x_i is r-separated, there exists a constant $c(n)$ such that for all x, $\lambda_i(x) > 0$ for at most $c(n)$ different indexes.

For convenience of notation, set $\pi_V(v)$ to be the orthogonal projection onto the linear subspace V of the free vector v, and set

$$\pi_{p_i,V_i}(x) = p_i + \pi_{V_i}(x - p_i)\,. \tag{4.31}$$

In other words, π_{p_i,V_i} is the affine projection onto the affine subspace $p_i + V_i$. Recall that π_{V_i} is a linear map, and so the gradients of π_{V_i} and of π_{p_i,V_i} at every point are equal to π_{V_i}.

Definition 4.14. Given $\{x_i, p_i, \lambda_i\}$ satisfying (i) to (vi), and given a family of linear k-dimensional spaces V_i, we define a smooth function $\sigma : \mathbb{R}^n \to \mathbb{R}^n$ by

$$\sigma(x) = x + \sum_i \lambda_i(x)\pi_{V_i^\perp}(p_i - x) = \psi(x)x + \sum_i \lambda_i(x)\pi_{p_i,V_i}(x)\ . \tag{4.32}$$

By local finiteness, it is evident that σ is smooth. Moreover, if $\psi(x) = 1$, then $\sigma(x) = x$. It is clear that philosophically σ is a form of "smooth interpolation" between the identity and the projections onto the subspaces V_i. It stands to reason that if V_i are all close together, then this map σ is close to being an orthogonal projection in the region $\bigcup_i B_{2r}(x_i)$.

Lemma 4.15. *Suppose that there exists a k-dimensional linear subspace $V \subseteq \mathbb{R}^n$ and a point $p \in \mathbb{R}^n$ such that for all i*

$$d_G(V_i, V) \leq \delta\,, \quad d(p_i, p + V) \leq \delta\,. \tag{4.33}$$

Then the map σ restricted to the set $U = \psi^{-1}(0) = \left(\sum_i \lambda_i\right)^{-1}(1)$ can be written as

$$\sigma(x) = \pi_{p,V}(x) + e(x)\,, \tag{4.34}$$

and $e(x)$ is a smooth function with

$$\|e\|_\infty + \|\nabla e\|_\infty \leq c(n)\delta/r = c(n,r)\delta\,. \tag{4.35}$$

Remark 4.2. Thus, on U we have that σ is the affine projection onto V plus an error which is small in C^1.

Proof. On the set U, we can define

$$\begin{aligned} e(x) = \sigma(x) - \pi_{p,V}(x) &= -\pi_{p,V}(x) + \sum_i \lambda_i(x) \cdot \left(\pi_{p_i,V_i}(x)\right) \\ &= \sum_i \lambda_i(x) \cdot \left(p_i - p - \pi_V(p_i - p) + \pi_V(p_i) - \pi_{V_i}(p_i) + \pi_{V_i}(x) - \pi_V(x)\right)\,. \end{aligned} \tag{4.36}$$

By (4.33) and lemma 4.12, we have the estimates

$$\|p_i - p - \pi_V(p_i - p)\| \leq \delta\,, \quad \left\|\pi_V(x - p_i) - \pi_{V_i}(x - p_i)\right\| \leq 2\delta\,\|x - p_i\| \leq 20\delta r\,. \tag{4.37}$$

This implies

$$\|e\|_{L^\infty(U)} \leq c(n)(1 + 13r)\delta \leq c(n)\delta\,. \tag{4.38}$$

As for ∇e, we have

$$\nabla e = \sum_i \nabla\lambda_i(x) \cdot \left(p_i - p - \pi_V(p_i - p) + \pi_V(p_i) - \pi_{V_i}(p_i) + \pi_{V_i}(x) - \pi_V(x)\right) + \sum_i \lambda_i(x) \nabla\left(\pi_{V_i}(x) - \pi_V(x)\right)\,. \tag{4.39}$$

The first sum is easily estimated, and since $\langle \nabla(\pi_W)|_x, w\rangle = \pi_W(w)$, we can still apply lemma 4.12 and conclude:

$$\|\nabla e\|_{L^\infty(U)} \leq \frac{c(n)}{r}\delta\,. \tag{4.40}$$

□

As we have seen, σ is in some sense close to the affine projection to $p + V$. In the next lemma, which is similar in spirit to [Sim, squash lemma], we prove that the image through σ of a graph over V is again a graph over V with nice bounds.

Lemma 4.16 (squash lemma)**.** *Fix $\rho \leq 1$ and some $B_{r/\rho}(y) \subseteq \mathbb{R}^n$, let $I = \{x_i\} \cap B_{5r/\rho}(y)$ be an $r/10$-separated set and define σ as in Definition 4.14. Suppose that there exists a k-dimensional subspace V and some $p \in \mathbb{R}^n$ such that $d(y, p+V) \leq \delta r$ and for all i:*

$$d(p_i, p+V) \leq \delta r \quad \text{ and } \quad d_G(V_i, V) \leq \delta \,. \tag{4.41}$$

Suppose also that there exists a C^1 function $g : V \to V^\perp$ such that $G \subseteq \mathbb{R}^n$ is the graph

$$G = \{p + x + g(x) \ \text{ for } \ x \in V\} \cap B_{r/\rho}(y) \,,$$

and $r^{-1}\|g\|_\infty + \|\nabla g\|_\infty \leq \delta'$. There exists a $\delta_0(n) > 0$ sufficiently small such that if $\delta \leq \delta_0 \rho$ and $\delta' \leq 1$, then

- *i. $\forall z \in G$, $r^{-1}|\sigma(z) - z| \leq c(n)(\delta + \delta')\rho^{-1}$, and σ is a C^1 diffeomorphism from G to its image,*
- *ii. the set $\sigma(G)$ is contained in a C^1 graph $\{p + x + \tilde{g}(x)\,, \ x \in V\}$ with*

$$r^{-1}\|\tilde{g}\|_\infty + \|\nabla\tilde{g}\|_\infty \leq c(n)(\delta + \delta')\rho^{-1} \,. \tag{4.42}$$

- *iii. moreover, if U' is such that $B_{c(\delta+\delta')\rho^{-1}}(U') \subseteq \psi^{-1}(0)$, then the previous bound is independent of δ', in the sense that*

$$r^{-1}\|\tilde{g}\|_{L^\infty(U'\cap V)} + \|\nabla\tilde{g}\|_{L^\infty(U'\cap V)} \leq c(n)\delta\rho^{-1} \,. \tag{4.43}$$

  For example, if $\delta' \leq \delta_0(n)\rho^{-1}$, we can take $U' = \bigcup_i B_{1.5r}(x_i)$.

- *iv. the map σ is a bi-Lipschitz equivalence between G and $\sigma(G)$ with bi-Lipschitz constant $\leq 1 + c(n)(\delta + \delta')^2\rho^{-2}$.*

Proof. For convenience, we fix $r = 1$ and $p = 0$. By notation, given any map $f : \mathbb{R}^n \to \mathbb{R}^m$, $p \in \mathbb{R}^n$ and $w \in T_p(\mathbb{R}^n) = \mathbb{R}^n$, we will denote by $\nabla|_p f[w]$ the gradient of f evaluated at p and applied to the vector w.

Recall that

$$\sigma(x + g(x)) = \psi(z)(x + g(x)) + \sum_{x_i \in I} \lambda_i(z)\left(\pi_{p_i,V_i}(x + g(x))\right)\,, \quad 1 - \psi(x) = \sum_{x_i \in I}\lambda_i(x)\,, \tag{4.44}$$

where we have set for convenience $z = z(x) = x + g(x)$. Define $h(x)$ by

$$(1 - \psi(z))x + h(x) \equiv \sum_i \lambda_i(z)\left(\pi_{p_i,V_i}(x + g(x))\right)\,. \tag{4.45}$$

Set also $h^T(x) = \pi_V(h(x))$ and $h^\perp(x) = \pi_{V^\perp}(h(x))$. By projecting the function $\sigma(x + g(x))$ onto V and its orthogonal complement we obtain

$$\begin{aligned} \sigma(x + g(x)) &\equiv \sigma^T(x) + \sigma^\perp(x)\,, \\ \sigma^T(x) = x + h^T(x)\,, &\quad \sigma^\perp(x) = \psi(z)g(x) + h^\perp(x)\,. \end{aligned} \tag{4.46}$$

We claim that if $\delta' \leq 1$, then

$$\left\|h^T(x)\right\| + \left\|\nabla h^T(x)\right\| \leq \frac{c\delta}{\rho}\,, \tag{4.47}$$

where this bound is *independent* of δ' as long as $\delta' \leq 1$. Indeed, for all $x \in V$ we have

$$h^T(x) = \pi_V\left[\sum_i \lambda_i(z)\left(\pi_{p_i,V_i}(x+g(x)) - x\right)\right] = \sum_i \lambda_i(z)\pi_V\left[\left(\pi_{p_i,V_i}(x) - \pi_V(x)\right) + \pi_{V_i}(g(x))\right] \tag{4.48}$$

Given (4.41) and lemma 4.12, with computations similar to (4.37), we get $\left\|h^T(x)\right\| \leq c\delta(1+\rho^{-1}) \leq c\delta\rho^{-1}$. As for the gradient, we get for any vector $w \in V$

$$\nabla h^T|_x[w] = \pi_V\left[\sum_i \nabla\lambda_i|_z\left[w + \nabla g|_x[w]\right]\left(\pi_{p_i,V_i}(x+g(x)) - x\right) + \sum_i \lambda_i(z)\left(\pi_{V_i}\left(w + \nabla g[w]\right) - w\right)\right], \tag{4.49}$$

In particular, we obtain

$$\left\|\nabla h^T|_x[w]\right\| \leq \sum_i \left\|\nabla\lambda_i\right\|\left(1 + \left\|\nabla g\right\|\right)\left\|w\right\|\left\|\pi_{p_i,V_i}(x+g(x)) - x\right\| + \sum_i \lambda_i(z)\left(\left\|\pi_{V_i}(w) - w\right\| + \left\|\pi_{V_i}(\nabla g[w])\right\|\right). \tag{4.50}$$

For the first term, we can estimate

$$\left\|\nabla\lambda_i\right\| \leq c(n)\,, \quad \left\|\nabla g\right\| \leq \delta' \leq 1\,, \quad \left\|\pi_{p_i,V_i}(x+g(x)) - x\right\| \leq \left\|\pi_{p_i,V_i}(x) - x\right\| + \left\|\pi_{V_i}(g(x))\right\|. \tag{4.51}$$

Since $x \in V$ with $\|x\| \leq \rho^{-1}$, and $g(x) \in V^\perp$, by (4.41) and lemma 4.12 we obtain

$$\left\|\pi_{p_i,V_i}(x+g(x)) - x\right\| \leq c\delta\rho^{-1}\,. \tag{4.52}$$

As for the second term, we have

$$\left\|\pi_{V_i}(w) - w\right\| \leq c\delta\left\|w\right\|\,, \quad \left\|\pi_{V_i}(\nabla g[w])\right\| \leq c\delta\delta'\left\|w\right\| \leq c\delta\left\|w\right\|\,. \tag{4.53}$$

Summing all the contributions, we obtain (4.47) as wanted.

Thus we can apply the inverse function theorem on the function $\sigma^T(x) : V \to V$ and obtain a C^1 inverse Q such that for all $x \in V$, $\|Q(x) - x\| + \|\nabla Q - id\| \leq c(n)\delta\rho^{-1}$, and if $\psi(x+g(x)) = 1$, then $Q(x) = x$. So we can write that for all $x \in V$

$$\sigma(x+g(x)) = \sigma^T(x) + \tilde{g}(\sigma^T(x)) \quad \text{where} \quad \tilde{g}(x) = \sigma^\perp(Q(x)) = h^\perp(Q(x)) + \psi\left(z(Q(x))\right)g(Q(x))\,. \tag{4.54}$$

Arguing as above, we see that $h^\perp(x)$ is a C^1 function with

$$\left\|h^\perp(x)\right\| + \left\|\nabla h^\perp(x)\right\| \leq \frac{c\delta}{\rho}\,, \tag{4.55}$$

and this bound is independent of δ' (as long as $\delta' \leq 1$).

Thus the function $\tilde{g} : V \to V^\perp$ satisfies for all x in its domain

$$\|\tilde{g}(x)\| + \|\nabla\tilde{g}(x)\| \leq c(n)(\delta+\delta')\rho^{-1}\,, \tag{4.56}$$

Moreover, for those x such that $\psi(Q(x)+g(Q(x))) = 0$, the estimates on $\tilde{g}$ are independent of δ', in the sense that $\|\tilde{g}(x)\| + \|\nabla\tilde{g}(x)\| \leq c(n)\delta\rho^{-1}$. Note that by the previous bounds we have

$$\|Q(x)+g(Q(x))-x\| \leq c(\delta+\delta')\rho^{-1}\,, \tag{4.57}$$

and so if $B_{c(\delta+\delta')\rho^{-1}}(U') \subset \psi^{-1}(0)$, then for all $x \in U'\cap V$, $\psi(Q(x)+g(Q(x))) = 0$. This proves items (ii), (iii). As for item (i), it is an easy consequence of the estimates in (4.47), (4.55).

Now since both G and $\sigma(G)$ are Lipschitz graphs over V, it is clear that the bi-Lipschitz map induced by π_V would have the right bi-Lipschitz estimate. Since σ is close to π_V, it stands to reason that this property remains true. In order to check the estimates, we need to be a bit careful about the horizontal displacement of σ.

bi-Lipschitz estimates In order to prove the estimate in (iv), we show that for all $z = x+g(x) \in G$ and for all unit vectors $w \in T_z(G) \subset \mathbb{R}^n$, we have

$$\left|\|\nabla\sigma|_z[w]\|^2 - 1\right| \leq c(\delta+\delta')^2\,. \tag{4.58}$$

First of all, note that if $\psi(z) = 1$, then σ is the identity, and there's nothing to prove.

In general, we have that

$$\nabla\sigma|_z[w] = \underbrace{\left(\psi(z)w + \sum_i \lambda_i(z)\pi_{V_i}[w]\right)}_{:=A} + \underbrace{\left(z\nabla\psi[w] + \sum_i \pi_{p_i,V_i}(z)\nabla\lambda_i[w]\right)}_{:=B}\,. \tag{4.59}$$

Since $\psi(z) + \sum_i \lambda_i(z) = 1$ everywhere by definition, we have

$$\|B\| = \left\|\sum_i (\pi_{p_i,V_i}(z)-z)\nabla\lambda_i[w]\right\| \leq c\sup_i\left\{\left\|\pi_{p_i,V_i}(z)-z\right\|\right\} \leq c(\delta+\delta')\,. \tag{4.60}$$

This last estimate comes from the fact that G is the graph of g over V with $\|g\|_\infty \leq \delta'$. Moreover, we can easily improve the estimate for B in the horizontal direction using lemma 4.12. Indeed, since $\pi_{p_i,V}(z)-z = \pi_{V_i^\perp}(z-p_i)$, we have

$$\begin{aligned}\|\pi_V B\| &= \left\|\sum_i \pi_V\left(\pi_{p_i,V_i}(z)-z\right)\nabla\lambda_i[w]\right\| \leq c\sup_i\left\{\left\|\pi_V\left(\pi_{p_i,V_i}(z)-z\right)\right\|\right\} \\ &\leq c\sup_i\left\{\left\|\pi_V\left(\pi_{V_i^\perp}(x+g(x)) - \pi_{V_i^\perp}(p_i)\right)\right\|\right\} \leq c(\delta^2+\delta'\delta)\,.\end{aligned} \tag{4.61}$$

As for A, by adapting the proof of lemma 4.15, we get $\|A - \pi_V[w]\| \leq c\,(\delta + \delta')$. Moreover, also in this case we get better estimates for A in the horizontal direction. Indeed, we have

$$\|\pi_V(A) - \pi_V[w]\| = \left\|\psi(z)\pi_V[w] + \sum_i \left(\lambda_i(z)\pi_V[\pi_{V_i}[w]]\right) - \pi_V[w]\right\| = \left\|\sum_i \lambda_i(z)\left(\pi_V[\pi_{V_i}[w] - \pi_V[w]]\right)\right\| . \tag{4.62}$$

Now let $w = \pi_V[w] + \pi_{V^\perp}[w] = w_V + w_{V^\perp}$. Then we have

$$\begin{aligned} \|\pi_V(A) - \pi_V[w]\| &\leq \sum_i \lambda_i(z)\left(\left\|\pi_V[\pi_{V_i}[w_V] - w_V]\right\| + \left\|\pi_V[\pi_{V_i}[w_{V^\perp}]]\right\|\right) &\quad (4.63)\\ &= \sum_i \lambda_i(z)\left(\left\|\pi_V[\pi_{V_i^\perp}[w_V]]\right\| + \left\|\pi_V[\pi_{V_i}[w_{V^\perp}]]\right\|\right) . \end{aligned}$$

Since G is the Lipschitz graph of g over V with $\|\nabla g\| \leq c\delta'$, then $\|\pi_{V^\perp}[w]\| \leq c\delta'$. Then, by lemma 4.12, we have

$$\|\pi_V(A) - \pi_V[w]\| \leq c\sum_i \lambda_i(z)\left(\delta^2 + \delta\delta'\right) . \tag{4.64}$$

Summing up, since $\|\pi_V[w]\| \leq \|w\| = 1$, we obtain that

$$\begin{aligned} \left|\|\nabla\sigma|_z[w]\|^2 - 1\right| &= \left|\|\pi_{V^\perp}\nabla\sigma|_z[w]\|^2 + \|(\pi_V\nabla\sigma|_z[w] - \pi_V[w]) + \pi_V[w]\|^2 - 1\right| &\quad (4.65)\\ &\leq c(\delta+\delta')^2 + \left|\|\pi_V[w]\|^2 - 1\right| = c(\delta+\delta')^2 + \|\pi_{V^\perp}[w]\|^2 \leq c(\delta+\delta')^2 . &\quad (4.66) \end{aligned}$$

This proves that σ is *locally* a bi-Lipschitz equivalence. However, since both G and G' are Lipschitz graphs over a k-dimensional space, this also prove that σ is a global bi-Lipschitz equivalence on $B_r(0)$. EXPAND THIS LAST COMMENT.

□

We make also the following trivial observation.

Lemma 4.17. *There exists a* $\delta_0(n) > 0$ *such that if* $G \cap B_1(0)$ *is a graph over a subspace* $x + L$ *of a function* f *with*

$$|f| + |\nabla f| \leq \delta < \delta_0 , \tag{4.67}$$

and if $d(y, x + L) \leq \delta < \delta_0$ *and* $d_G(L, V) \leq \delta$*, then* $G \cap B_1(0)$ *is also a graph over the subspace* $y + V$ *of the function* $\tilde f$ *with*

$$\left|\tilde f\right| + \left|\nabla\tilde f\right| \leq c(n)\delta . \tag{4.68}$$

Proof. SAME TECHNIQUE AS IN THE SQUASH LEMMA, TO BE DONE □

4.8 Proof of the classical Reifenberg result

Here we prove the classical Reifenberg result. In order to do that, we will proceed by induction and prove that for all i there exists smooth maps σ_i and smooth manifolds T_i that approximate the set S at scale $\mathfrak{r}_i = 100^{-i}$. We will always assume that $0 < \delta(n) < 10^{-7} << 100^{-1}$. Actually, as $n \to \infty$, this construction will yield a $\delta(n) \to 0$.

Inductive construction In particular, we will show that for all $\mathfrak{r}_i = 100^{-i}$, we have a smooth manifold T_i and a map σ_i such that for all i:

1. $d_{\mathcal{H}}(S \cap B_1(0), T_i \cap B_1(0)) \leq c\delta\mathfrak{r}_i$ where $c = c(n)$
2. for all $x \in S$, $T_i \cap B_{10\mathfrak{r}_i}(x)$ is a graph over $x + L(x, \mathfrak{r}_i)$ of a smooth function f with
$$\mathfrak{r}_i^{-1}|f| + |\nabla f| \leq c\delta\,, \tag{4.69}$$
3. we define $T_i = \sigma_i(T_{i-1})$ for $i = 1, \cdots, \infty$
4. for all $x \in T_{i-1}$, $|\sigma_i(x) - x| \leq c\delta\mathfrak{r}_i$
5. for all $x \in T_{i-1}$ and all $v \in T_x T_{i-1}$, $||\nabla\sigma_i(x)[v]| - |v|| \leq c\delta\,|v|$
6. $\sigma_i : T_{i-1} \to T_i$ is locally a bi-Lipschitz equivalence with constant $\leq 1 + c\delta^2$.
7. $\sigma_i : T_{i-1} \to T_i$ is a global bi-Lipschitz equivalence with constant $\leq 1 + c\delta$, in the sense that for all $x, y \in T_{i-1}$, we have
$$(1 - c\delta)\,|x - y| \leq |\sigma_i(x) - \sigma_i(y)| \leq (1 + c\delta)\,|x - y|\ . \tag{4.70}$$

After all this is proved, we will also show that the maps

$$\phi_i = \sigma_{i-1} \circ \cdots \circ \sigma_1 = T_0 \to T_i \tag{4.71}$$

have uniform $C^{0,\alpha}$ bounds and they are injective with uniform $C^{0,\alpha}$ bounds for ϕ^{-1} as well.

First step For $i = 0$, we set $T_0 = 0 + L(0, 1)$ and $\sigma_0 = id$. These clearly satisfy all the properties we want. The only property worth considering is property (2). Obviously, the plane $L(0, 1) = 0 + L(0, 1)$ is a graph over $0 + L(0, 1)$. Now for all $x \in B_1(0) \cap S$, we have by Lemma 4.13 and 4.17 that $L(0, 1)$ is also a graph over $x + L(x, 1)$ with estimates as in (4.69). (ADD MORE DETAILS).

Inductive step THERE ARE SOME PARTS HERE JUST SKETCHED, BUT THE IDEAS ARE ALL HERE.

Assuming that everything holds for i, we build the map $\sigma_i : \mathbb{R}^n \to \mathbb{R}^n$ and use it to define $T_{i+1} = \sigma_i(T_i)$.

First of all, let $\left\{B_{\mathfrak{r}_{i+1}/3}\left(x_{i+1,j}\right)\right\}_{j\in I_{i+1}}$ be a maximal family of pairwise disjoint balls in $S\cap B_1(0)$. Since the family is maximal, we have that

$$S\cap B_1(0)\subset\bigcup_{j\in I_{i+1}}B_{2\mathfrak{r}_{i+1}/3}\left(x_{i+1,j}\right)\,. \tag{4.72}$$

As a consequence, by item (1) we have

$$T_i\cap B_1(0)\subset\bigcup_{j\in I_{i+1}}B_{\mathfrak{r}_{i+1}}\left(x_{i+1,j}\right)\,. \tag{4.73}$$

Now we build σ_i as in section 4.7. In particular, let $\left\{\lambda_j,\psi\right\}_{j\in I_{i+1}}$ be a partition of unity for $\mathbb{R}^n$ such that

1. $\lambda_j\geq 0$ and $\operatorname{supp}\left(\lambda_j\right)\subseteq B_{3\mathfrak{r}_{i+1}}\left(x_{i+1,j}\right)$ for all $j\in I_{i+1}$,
2. for all $x\in\bigcup_j B_{2\mathfrak{r}_{i+1}}\left(x_{i+1,j}\right)$, $\sum_j\lambda_j(x)=1$ and $\sum_i\lambda_i(x')\leq 1$ for all $x'\in\mathbb{R}^n$,
3. $\sup_j\left\|\nabla\lambda_j\right\|_\infty\leq c(n)\mathfrak{r}_{i+1}^{-1}$,
4. if we set $1-\psi(x)=\sum_j\lambda_j(x)$, then ψ is a nonnegative smooth function with $\|\nabla\psi\|_\infty\leq c(n)\mathfrak{r}_{i+1}^{-1}$.

Set for convenience $L_{i+1,j}=L(x_{i+1,j},\mathfrak{r}_{i+1})$, and define

$$\sigma_i(x)=x+\sum_{j\in I_{i+1}}\lambda_j(x)\pi_{L^\perp_{i+1,j}}\left(x_{i+1,j}-x\right)=\psi(x)x+\sum_{j\in I_{i+1}}\lambda_j(x)\pi_{x_{i+1,j},L_{i+1,j}}(x)\,. \tag{4.74}$$

First of all, note that the Reifenberg condition and Lemma 4.13 tells us that if $x,y\in S\cap B_1(0)$ and $y\in B_{50\mathfrak{r}_i}(x)$, then

$$d(y,L(x,70\mathfrak{r}_i))\leq c\delta\mathfrak{r}_i\,,\qquad d_G(L(x,70\mathfrak{r}_i),L(y,\mathfrak{r}_{i+1}))\leq c\delta\,. \tag{4.75}$$

By the triangle inequality, we also obtain

$$d(y,L(x,\mathfrak{r}_i))\leq c\delta\mathfrak{r}_i\,,\quad d_G(L(x,\mathfrak{r}_i),L(y,\mathfrak{r}_{i+1}))\leq c\delta\,,\quad d_G(L(x,\mathfrak{r}_i),L(y,\mathfrak{r}_i))\leq c\delta\,. \tag{4.76}$$

Now all of our properties will be corollaries of the squash lemma.

Inside $B_1(0)$ First we restrict ourselves to studying T_i and T_{i+1} inside $B_1(0)$.

Observe that by item (1), and by the triangle inequality, we have that $T_i\cap B_1(0)\subset\bigcup_j B_{r_{i+1}}\left(x_{i+1,j}\right)$. In other words, we are saying that $\psi=0$ over the whole set $T_i\cap B_1(0)$.

Now we focus on proving items (2) through (6). Take $x\in S$, by induction $T_i\cap B_{10\mathfrak{r}_i}(x)$ is a graph over $L(x,\mathfrak{r}_i)$ with bounds as in (4.69).

Now consider the ball $B_{20\mathfrak{r}_{i+1}}(x)\subseteq B_{10}(\mathfrak{r}_i)\,x$. Note that σ_{i+1} on this ball is determined only by the points $x_{i+1,j}\in B_{30\mathfrak{r}_{i+1}}(x)$, and by their relative planes $L_{i+1,j}$ (and the partition of unity λ_j).

To be more precise, the map σ_{i+1} restricted to this ball can be viewed as

$$\sigma_i(x)=x+\sum_{j\in J(y)}\lambda_j(x)\pi_{L^\perp_{i+1,j}}\left(x_{i+1,j}-x\right)=\psi(x)x+\sum_{j\in J(y)}\lambda_j(x)\pi_{x_{i+1,j},L_{i+1,j}}(x)\,, \tag{4.77}$$

where $J(y)$ is the set of $j \in I_{i+1}$ such that $\left|x_{i+1,j} - y\right| \leq 30\mathfrak{r}_{i+1}$. Indeed, if this is not the case, then $\lambda_j(y)$ is identically zero on the ball considered. Moreover, $\psi(x) = 0$ for all $x \in T_i \cap B_1(0)$, as already pointed out.

By the estimates in (4.76), we can apply the squash lemma to the ball $B_{20\mathfrak{r}_{i+1}}(x)$, and (since we are restricting ourselves to $T_i \cap B_1(0)$ where $\psi = 0$) we obtain that T_{i+1} is a graph over $L(x, \mathfrak{r}_{i+1})$ with the estimates in (4.69).

Note that the squash lemma guarantees that the estimates obtained here are *independent* of the step i of the induction. In particular, the constant $c(n)$ in (4.69) does NOT depend on i.

Items (4),(5),(6) are also a consequence of the squash lemma.

Item (1) In order to prove item (1), we just need to exploit (2) at the new scale. Indeed, since for all $x \in S \cap B_1(0)$ we know that $T_i \cap B_{\mathfrak{r}_{i+1}}(x) \cap B_1(0)$ is a graph over $x + L(x, \mathfrak{r}_{i+1})$ with (4.69), we also know that $d_H(T_i \cap B_{\mathfrak{r}_{i+1}}(x), x + L(x, \mathfrak{r}_{i+1}) \cap B_{\mathfrak{r}_{i+1}}(x)) \leq c\delta\mathfrak{r}_{i+1}$. Now we obtain item (1) just by exploiting this, the Reifenberg condition and the triangle inequality.

Estimates outside $B_1(0)$ As a first observation, note that technically we are not interested in what happens outside of the set $B_1(0)$ for the scope of our theorem.

THIS SECTION IS JUST A SKETCH.

In order to take care of the estimates outside $B_1(0)$, i.e., where ψ is not necessarily zero, we want to see that all the estimates (2) to (6) remain valid with an estimate independent of i.

Roughly speaking, the explanation is the following. Where $\psi \neq 0$, then it might be that the estimates in item (2) build up with the induction index i. However, for all $x \in T_i$, if $\psi(x) = 0$, then this point will never be changed again by a map σ_{i+j} for $j \geq 2$.

In order to be slightly more precise (but not much), suppose that $x \in T_i$ is such that $\psi_i(x) \in (0, 1)$. If $\psi_{i+1}(x) = 0$, we apply the previous construction. If $\psi_{i+1}(x) = 1$, then $\sigma_{i+1}(x) = x$ and there's nothing to prove. If $\psi_{i+1}(x) \in (0, 1]$, then this means that $x \in\notin \bigcup_j B_{2r_{i+1}}\left(x_{i+1,j}\right)$. Now if this is the case, since $S \subset \bigcup_j B_{2r_{i+1}/3}\left(x_{i+1,j}\right)$, then $x \notin B_{3r_{i+2}}(S)$, and thus $\psi_{i+2}(x) = 1$. This implies that the estimates that possibly "build up" are used at most once (so they don't actually "build up" with i).

Item (7) . As for item (7), if $x, y \in T_i$ satisfy $|x - y| \leq 10\mathfrak{r}_{i+1}$, this is the content of the squash Lemma. If this is not the case, then we can simply estimate

$$|\sigma_{i+1}(x) - \sigma_{i+1}(y)| = |\sigma_{i+1}(x) - x + x - y + y - \sigma_{i+1}(y)| \leq |x - y| + c\delta\mathfrak{r}_{i+1} \leq |x - y|(1 + c\delta), \tag{4.78}$$

and similarly for the other direction.

Global estimates Now we turn our attention to the maps ϕ_i. We want to prove global $C^{0,\alpha}$ estimates on these maps.

First of all, it is easy to see that these maps form a Cauchy sequence in C^0. Indeed, by item (4), we have

$$\left\|\phi_{i+j}(x) - \phi_i(x)\right\|_{L^\infty} \leq \sum_{k=i}^{i+j} c\delta\mathfrak{r}_k \leq c\delta\mathfrak{r}_i . \tag{4.79}$$

Now we claim that for all values of j, and all $x, y \in T_0$ we have

$$(1 - c\delta)^j |x - y| - c\delta \mathrm{r}_j \le |\phi(x) - \phi(y)| \le c\delta \mathrm{r}_j + (1 + c\delta)^j |x - y| \ . \tag{4.80}$$

This comes from the fact that for all j:

$$|\phi(x) - \phi(y)| = \left|\phi(x) - \phi_j(x) + \phi_j(x) - \phi_j(y) + \phi_j(y) - \phi(y)\right| \le c\delta \mathrm{r}_j + \left|\phi_j(x) - \phi_j(y)\right| \ . \tag{4.81}$$

Now by induction on $j = 0, 1, \cdots, \infty$, it is easy to see that

$$\left|\phi_j(x) - \phi_j(y)\right| \le (1 + c\delta)^j |x - y| \ . \tag{4.82}$$

Indeed, this is trivial for $j = 0$. If it is true for j, then

$$\left|\phi_j(x) - \phi_j(y)\right| = \left|\sigma_j(\phi_{j-1}(x)) - \sigma_j(\phi_{j-1}(y))\right| \le (1 + c\delta) \left|\phi_{j-1}(x) - \phi_{j-1}(y)\right| \le (1 + c\delta)^j |x - y| \ , \tag{4.83}$$

as desired.

The other inequality can be proved in a similar way.

With (4.80), we have the Hölder estimates. Indeed, let $j = -\log_{100}(|x - y|)$. For this j, we have

$$|\phi(x) - \phi(y)| \le c\delta |x - y| + (1 + c\delta)^{-\log_{100}(|x-y|)} |x - y| \le c\delta |x - y| + |x - y|^{1-\log_{100}(1+c\delta)} \le (1 + c\delta) |x - y|^{1-c\delta} \ . \tag{4.84}$$

Similarly for the other direction. This concludes the proof.

5 Generalized Reifenberg theorem

Here we want to have conditions for a better Reifenberg theorem. In order to do that, we first introduce a definition.

5.1 Definition of distortion / β-numbers

Definition 5.1. Let μ be a measure on B_2 with $r > 0$ and $k \in \mathbb{N}$. Then we define the k-dimensional displacement (sometimes referred to as Jones' β-2 number) by

$$D^k_\mu(x,r) \equiv \beta^k_\mu(x,r)^2 \equiv \inf_{L^k} r^{-(k+2)} \int_{B_r(x)} d^2(y, L^k)\, d\mu(y)\,, \tag{5.1}$$

where the inf's are taken over all k-dimensional affine subspaces $L^k \subseteq \mathbb{R}^n$. If $S \subseteq B_2$, then we can define its k-displacement $D^k_S(x,r)$ by associating to S the k-dimensional Hausdorff measure λ^k_S restricted to S.

Sometimes, we will omit the index k and the subscript μ or S when there can be no risk of confusion from the context. In particular, we will often write $D(x,r)$ for $D^k_\mu(x,r)$.

Remark 5.1. Notice that for all μ we have the very crude estimate

$$D^k_\mu(x,r) \leq \frac{\mu(B_r(x))}{r^k}\,. \tag{5.2}$$

Remark 5.2. Notice that the definitions are scale invariant. In particular, if we rescale $B_r \to B_1$ and let $\tilde{S}$ be the induced set, then $D^k_S(x,r) \to D^k_{\tilde{S}}(x,1)$.

Remark 5.3. Notice the monotonicity given by the following: If $\mu' \leq \mu$, then $D^k_{\mu'}(x,r) \leq D^k_\mu(x,r)$.

Remark 5.4. It is immediate to see from the definition that, up to dimensional constants, $D^k_\mu(x,r)$ is controlled on both sides by $D^k_\mu(x,r/2)$ and $D^k_\mu(x,2r)$. In particular, for all $y \in B_r(x)$, $D^k_\mu(x,r) \leq 2^{k+2} D^k_\mu(y,2r)$. As a corollary we have the estimate

$$D^k_\mu(x,r) \leq 2^{k+2} \fint_{B_r(x)} D^k_\mu(y,2r)d\mu(y)\,. \tag{5.3}$$

More generally, we also have that for all fixed $\rho \in (0,1]$, we can control

$$D^k_\mu(x,\rho\cdot r) \leq \rho^{-k-2} D^k_\mu(x,r)\,. \tag{5.4}$$

5.2 Examples

In this section we gather some examples regarding measures and β number estimates. We start with a basic example on the finiteness of the β number for graphs in order to help build an intuition.

Example 5.1. Let S be the graph of a C^2 function $f : \mathbb{R}^k \to \mathbb{R}^{n-k}$, so that $S \subset \mathbb{R}^n$. Evidently, one expects $\mathcal{H}^k \llcorner S$ to have nice β^2 bounds. In order to verify this, we compute the Taylor expansion of f around $(x_0, f(x_0)) \in S$:

$$|f(x) - f(x_0) - \nabla f(x_0) \cdot (x - x_0)| \leq \left\|\nabla^2 f\right\|_\infty |x - x_0|^2 \ . \tag{5.5}$$

In particular, using simple geometric considerations, this proves that

$$\beta^k_{\mathcal{H}^k \llcorner S}((x_0, f(x_0)), r)^2 \leq c(k, \|\nabla f\|_\infty) \left\|\nabla^2 f\right\|_\infty r \, . \tag{5.6}$$

This is because we can estimate

$$\beta^k_{\mathcal{H}^k \llcorner S}((x_0, f(x_0)), r)^2 \leq r^{-k-2} \int_{B_r((x_0, f(x_0)))} d^2(S, L_{x_0}) d\mathcal{H}^k \llcorner S \ , \tag{5.7}$$

where L_{x_0} is the tangent space of the graph $(x, f(x))$ at x_0, so it is the k-dimensional plane parametrized by

$$y \in \mathbb{R}^k \ \to \ (x_0, f(x_0)) + (y - x_0, df|_{x_0}(y - x_0)) = (y, f(x_0) + df|_{x_0}(y - x_0)) \, . \tag{5.8}$$

Thus we have the very crude estimate

$$\beta^k_{\mathcal{H}^k \llcorner S}((x_0, f(x_0)), r)^2 \leq C(k, \|\nabla f\|_\infty) r^{-k-2} \int_{\mathbb{R}^k \cap \{|y - x_0| \leq r\}} \left|f(y) - f(x_0) + df|_{x_0}(y - x_0)\right|^2 \leq \tag{5.9}$$

$$\leq C(k, \|\nabla f\|_\infty) \left\|\nabla^2 f\right\|_\infty r^4 \left[r^{-k-2} \int_{\mathbb{R}^k \cap \{|y - x_0| \leq r\}} 1 \right] \leq C(k, \|\nabla f\|_\infty) \left\|\nabla^2 f\right\|_\infty r^2 \, . \tag{5.10}$$

In turn, this implies the uniform pointwise bound

$$\int_0^1 \beta^k_{\mathcal{H}^k \llcorner S}((x_0, f(x_0)), r)^2 \frac{dr}{r} \leq c(k, \|\nabla f\|_\infty) \cdot \left\|\nabla^2 f\right\|_\infty \ . \tag{5.11}$$

It is easy to see that a similar computation holds for all graphs of $C^{1,\alpha}$ functions.

As opposed to the previous example, we show an easy case where the integral in (5.11) does not converge.

Example 5.2. Let S be the graph of the function $\alpha\,|x|$ in $\mathbb{R}^2$, where $\alpha \neq 0$. Then clearly $\beta^k_{\mathcal{H}^k \llcorner S}((0, 0), r)$ is constant in r, and thus $\int_0^1 \beta^k_{\mathcal{H}^k \llcorner S, 2}(r)^2 \frac{dr}{r} = \infty$. This proves that the estimate in (5.11) cannot hold in a pointwise sense for Lipschitz graphs.

It is worth noticing however that although the pointwise bound of (5.11) does not hold for Lipschitz graphs, an integral estimate follows from [Dor85, theorem 6] (see also [AS, theorem 1.2] and [DS93, Theorem 1.42]). For the sake of completeness, here we report the result (without proof):

Proposition 5.2. *Let $f : \mathbb{R}^k \to \mathbb{R}^{n-k}$ be a Lipschitz map with Lipschitz constant bounded by L, and let S be the graph of f. Then for all $x \in S$ and $r \geq 0$ we have*

$$\int_{B_r(x)} \left(\int_0^r \beta^k_{\mathcal{H}^k \llcorner S}(x, s)^2 \frac{ds}{s} \right) d\mathcal{H}^k(x) \leq C(k)(1 + L^2)^{k/2} L^2 r^k \, . \tag{5.12}$$

Example 5.3. Moebius strip example:

With respect to the conclusions of Theorem 5.5 there are two natural questions regarding how sharp they are. First, is it possible to obtain more structure from the set S than rectifiable? In particular, in Theorem 5.3 there are topological conclusions about the set, is it possible to make such conclusions in the context of Theorem 5.5? In the last example we saw this is not the case. Then a second question is to ask whether we can at least find a single rectifiable chart which covers the whole set S. This example taken from [DT12, counterexample 12.4] shows that the answer to this question is negative as well.

To build our examples let us first consider a unit circle $S^1 \subseteq \mathbb{R}^3$. Let $M^2 \supset S^1$ be a smooth Möbius strip around this circle, and let $S_\epsilon \subseteq M^2 \cap B_\epsilon(S^1) \equiv M^2_\epsilon$ be an arbitrary λ^2-measurable subset of the Möbius strip, contained in a small neighborhood of the S^1. In particular, $Area(S_\epsilon) \leq C\epsilon \to 0$ as $\epsilon \to 0$. It is not hard, though potentially a little tedious, to check that assumptions of Theorem 5.5 hold for $\delta \to 0$ as $\epsilon \to 0$.

However, we have learned two points from these example. First, since S_ϵ was an arbitrary measurable subset of a two dimensional manifold, we have that it is 2-rectifiable, however that is the most which may be said of S_ϵ. That is, structurally speaking we cannot hope to say better than 2-rectifiable about the set S_ϵ. More than that, since S_ϵ is a subset of the Möbius strip, we see that even though S_ϵ is rectifiable, we cannot even cover S_ϵ by a single chart from $B_1(0^2)$, as there are topological obstructions, see [DT12] for more on this.

5.3 New Reifenberg theorems

Before introducing the results which are really needed for the paper, it is worth mentioning the $W^{1,p}$-Reifenberg theorem obtained in [NVa]. This is a natural generalization of the Reifenberg and gives intuition and motivation for the rest of the statements, which are essentially more complicated versions of it.

Theorem 5.3 ($W^{1,p}$-Reifenberg)**.** *[NVa, theorem 3.2] For each $\epsilon > 0$ and $p \in [1, \infty)$ there exists $\delta(n, \epsilon, p) > 0$ such that the following holds. Let $S \subseteq B_4 \subseteq \mathbb{R}^n$ be a closed subset with $0^n \in S$, and assume for each $x \in S \cap B_1$ and $B_r(x) \subseteq B_4$ that*

$$\inf_{L^k} d_H(S \cap B_r(x), L^k \cap B_r(x)) < \delta r\,, \tag{5.13}$$

$$\int_{S\cap B_r(x)} \left(\int_0^r D^k_S(y, s)\, \frac{ds}{s} \right) d\lambda^k(y) < \delta^2 r^k\,. \tag{5.14}$$

Then the following hold:

1. *there exists a mapping $\phi : \mathbb{R}^k \to \mathbb{R}^n$ which is a $1 + \epsilon$ bi-$W^{1,p}$ map onto its image and such that $S \cap B_1(0^n) = \phi(B_{1+c\epsilon}(0^k)) \cap B_1(0^n)$.*
2. *$S \cap B_1(0^n)$ is countably k-rectifiable.*
3. *For each ball $B_r(x) \subseteq B_1$ with $x \in S$ we have*

$$(1 - \epsilon)\omega_k r^k \leq \lambda^k(S \cap B_r(x)) \leq (1 + \epsilon)\omega_k r^k\,. \tag{5.15}$$

Remark 5.5. Results (2) and (3) both follow from (1). We get (3) by applying the result of (1) to all smaller balls $B_r(x) \subseteq B_1$, since the assumptions of the theorem hold on these balls as well.

Remark 5.6. Note that, for $p > k$, a bi-$W^{1,p}$ map is a bi-C^α map, in particular we see that $\phi(B_1(0^k))$ is homeomorphic to the ball $B_1(0^k)$.

Remark 5.7. As it is easily seen, the requirement that S is closed is essential for this theorem, and in particular for the lower bound on the Hausdorff measure. As an example, consider any set $S \subseteq \mathbb{R}^k$ which is dense but has zero Hausdorff measure. In the following theorems, we will not be concerned with lower bounds on the measure, and we will be able to drop the closed assumption.

We are going to state another generalization of Reifenberg's theorem, more discrete in nature, which will be particularly important in the proof of the main theorems of this paper:

Theorem 5.4 (Discrete Reifenberg)**.** *[NVa, theorem 3.4] There exists $\delta(n) > 0$ and $C_R(n)$ such that the following holds. Let $\{B_{r_s}(x_s)\}_{s\in S} \subseteq B_2$ be a collection of disjoint balls with $x_s \in B_1(0)$, and let $\mu \equiv \sum_{s\in S} \omega_k r_s^k \delta_{x_s}$ be the associated measure. Assume that for each $B_r(x) \subseteq B_4$ with $\mu(B_r(x)) \geq \gamma_k r^k = \omega_k (r/40)^k$ we have*

$$\int_{B_r(x)} \left(\int_0^r D^k_\mu(y,t)\, \frac{dt}{t} \right) d\mu(y) < \delta^2 r^k \,. \tag{5.16}$$

Then we have the estimate

$$\sum_{s\in S} r_s^k < C_R(n)\,. \tag{5.17}$$

Remark 5.8. Instead of (5.16) we may assume the estimate

$$\sum_{\mathrm{r}_\alpha \leq r/2} \int_{B_r(x)} D^k_\mu(y, \mathrm{r}_\alpha)\, d\mu(y) < \delta^2 r^k \,. \tag{5.18}$$

In the applications, this will be the more convenient phrasing.

In order to prove rectifiability of the strata, we will also need the following version of Reifenberg's theorem. The proof of this theorem relies on the same ideas as the discrete-Reifenberg, for this reason we do not report it here and we refer the interested reader to [NVa, theorem 3.3].

Theorem 5.5 (Rectifiable-Reifenberg)**.** *[NVa, theorem 3.3] For every $\epsilon > 0$, there exists $\delta(n,\epsilon) > 0$ such that the following holds. Let $S \subseteq B_4 \subseteq \mathbb{R}^n$ be a λ^k-measurable subset, and assume for each $B_r(x) \subseteq B_4$ with $\lambda^k(S \cap B_r(x)) \geq \gamma_k r^k$ that*

$$\int_{S\cap B_r(x)} \left(\int_0^r D^k_S(y,s)\, \frac{ds}{s} \right) d\lambda^k(y) < \delta^2 r^k \,. \tag{5.19}$$

Then the following holds:

1. *For each ball $B_r(x) \subseteq B_1$ with $x \in S$ we have*
$$\lambda^k(S \cap B_r(x)) \leq (1+\epsilon)\omega_k r^k\,. \tag{5.20}$$
2. *$S \cap B_1(0^n)$ is countably k-rectifiable.*

Remark 5.9. Notice that for the statement of the theorem we do not need control over balls which already have small measure. This will be quite convenient for the applications.

5.4 Distance between L^2 best planes

Here we ue D to control the distance between bext approximating subspaces at different points and scales. In this section, ρ will be a fixed parameter. We will also introduce $\gamma_k = \omega_k 40^{-k}$, which is simply a threshold for the measure $\mu(B_1(0))$. Above this threshold, D will give us control over the set S, below this threshold it will not.

From now on, μ will be a non-negative Radon measure on $\mathbb{R}^m$. Moreover, we will denote by $V(x,r)$ the (or one of the) planes minimizing the quantity $\int_{B_r(x)} d(x,V)^2 d\mu(y)$.

If we want D to control the relation between the best planes V for μ at different points and scales, we need the measure μ to satisfy some properties. In particular, the measure should be "spread" over something roughly k-dimensional at the scale we're focussing. The next example illustrates this.

However, an upper bound on the measure is not enough to guarantee best L^2-planes are close, as the following example shows:

Example 5.4. Let V, V' be k-dimensional subspaces, $0 \in V \cap V'$, and set $S = (V \cap B_1(0) \setminus B_{1/10}(0)) \cup S'$, where $S' \subseteq V' \cap B_{1/10}(0)$ and $\mu = \lambda^k|_S$. Then evidently $D(0,1) \leq \lambda^k(S')$ and $D(0,1/10) = 0$, independently of V and V'. However, $V(0,1)$ will be close to V, while $V(0,1/10) = V'$. Thus, in general, we cannot expect $V(0,1)$ and $V(0,1/10)$ to be close if $\mu(B_{1/10}(0))$ is too small.

Example 5.5. It is easy to see that it is not just the amount of measure that matters, but also how it is spread. Consider the following example: let $k = 1$ and μ be the sum of 5 Dirac masses in $\mathbb{R}^2$
$$\mu = \delta_0 + \delta_{(1,0)} + \delta_{(-1,0)} + \delta_{(0,t)} + \delta_{(0,-t)}\,. \tag{5.21}$$
For $0 < t \leq 1/2$, it is easy to see that $V(0,1)$ is the x-axis, while $V(0,1/2)$ is the y-axis, and this is independent on the choice of t.

Moreover, we have
$$D(0,1/2) = 0\,, \quad D(0,1) = 2t^2\,. \tag{5.22}$$
As t approaches 0, the beta numbers clearly don't control the distance between $V(0,1)$ and $V(0,1/2)$ (which is constant in t and equal to 1).

So the geometry of the measure μ is essential to obtain the bound we want.

In particular, if we want to control the distance between best subspaces, we need the measure μ to be "spread out" on a set that looks k-dimensional. We can either ask for this condition directly and see where this leads us (see [ENV]), or we can ensure that this is condition is satisfied by controlling other properties of the measure μ. In the next lemma, we'll see that if $\mu(B_r(x)) \leq Mr^k$ for small r and all x, and if $\mu(B_1(0)) \geq \gamma_k$, then μ cannot be concentrating too close to a $k-1$-dimensional subspace. To be more precise:

Lemma 5.6. *Let $\gamma_k = \omega_k 40^{-k}$. There exists a $\rho_0(m, \gamma_k, M) = \rho_0(m, M)$ such that if for some $\rho \leq \rho_0$ and all $x \in B_1(0)$*

$$\mu(B_\rho(x)) \leq M\rho^k\,, \tag{5.23}$$

and if $\mu(B_1(0)) \geq \gamma_k$, then for every affine subspace $V \subseteq \mathbb{R}^m$ of dimension $\leq k-1$, there exists an $x \in S \cap B_1(0)$ such that $B_{10\rho}(x) \cap V = \emptyset$ and $\mu\left(B_\rho(x) \cap B_1(0)\right) \geq c(m,\rho) = c(m)\rho^m > 0$.

Proof. Let V be any $k-1$-dimensional subspace, and consider the set $B_{11\rho}(V)$. Let $B_i = B_\rho(x_i)$ be a sequence of balls that cover the set $B_{11\rho}(V) \cap B_1(0)$ and such that $B_i/2 \equiv B_{\rho/2}(x_i)$ are disjoint and $x_i \in B_{11\rho}(V) \cap B_1(0)$. If N is the number of these balls, then a standard covering argument gives

$$\begin{aligned} & N\omega_m \rho^m/2^m \leq \omega_{k-1}(1+\rho)^{k-1}\omega_{m-k+1}(12\rho)^{m-k+1} \leq 24^m \omega_{k-1}\omega_{m-k+1}\rho^{m-k+1} \\ \Longrightarrow \quad & N \leq 48^m \frac{\omega_{k-1}\omega_{m-k+1}}{\omega_m}\rho^{1-k}\,. \end{aligned} \tag{5.24}$$

By (5.23), the measure of the set $B_{11\rho}(V)$ is bounded by

$$\mu(B_{11\rho}(V)) \leq \sum_i \mu(B_i) \leq MN\rho^k \leq 48^m \frac{\omega_{k-1}\omega_{m-k+1}}{\omega_m} M\rho \leq 10^5 (50m)^m M\rho = c(m)M\rho\,. \tag{5.25}$$

where the next-to-last estimate is an extremely rough bound on the constants involved. Thus if

$$\rho \leq 10^{-5}(50m)^{-m}\gamma_k/(4M)\,, \tag{5.26}$$

then $\mu(B_{11\rho}(V)) \leq \gamma_k/4$. In particular, we get that there must be some point of S not in $B_{11\rho}(V)$. More effectively, let us consider the set $S \cap B_1(0) \setminus B_{11\rho}(V)$. This set can be covered by at most $c(m,\rho) = 4^m\rho^{-m}$ balls of radius ρ centered in x, and we also see that

$$\mu\left(B_1(0) \setminus B_{11\rho}(V)\right) \geq \frac{3\gamma_k}{4}\,. \tag{5.27}$$

Thus, there must exist at least one ball of radius ρ centered in x and disjoint from $B_{10\rho}(V)$ such that

$$\mu\left(B_\rho(x) \cap B_1(0)\right) \geq \frac{3\gamma_k}{4} 4^{-m}\rho^m \geq c(m)\rho^m\,. \tag{5.28}$$

□

Now if at two consecutive scales there are some balls on which the measure μ *effectively spans* k-dimensional subspaces, we show that these subspaces have to be close together.

Lemma 5.7. *Let μ be a positive Radon measure and assume $\mu(B_1(0)) \geq \gamma_k$ and that for each $y \in B_1(0)$ we have $\mu(B_{\rho^2}(y)) \leq M\rho^{2k}$, where $\rho \leq \rho_0$. Additionally, let $B_\rho(x) \subset B_1(0)$ be a ball such that $\mu(B_\rho(x)) \geq \gamma_k \rho^k$. Then if $A = V(0,1) \cap B_\rho(x)$ and $B = V(x,\rho) \cap B_\rho(x)$ are L^2-best subspace approximations of μ with $d(x,A) < \rho/2$, then*

$$d_H(A,B)^2 \leq c(n,\rho,M)\left(D^k_\mu(x,\rho) + D^k_\mu(0,1)\right). \tag{5.29}$$

Proof. Let us begin by observing that if $c(n,\rho,M) > 4\rho^2\delta^{-1}(n,\rho,M)$, which will be chosen later, then we may assume without loss of generality that

$$D^k_\mu(x,\rho) + D^k_\mu(0,1) \leq \delta = \delta(n,\rho,M)\,, \tag{5.30}$$

since otherwise (5.29) is trivially satisfied.

We will estimate the distance $d_H(A,B)$ by finding $k+1$ balls $B_{\rho^2}(y_i)$ which have enough mass and effectively span in the appropriate sense $V(x,\rho)$. Given the upper bounds on D^k_μ, we will then be in a position to prove our estimate.

Consider any $B_{\rho^2}(y) \subseteq B_1(0)$ with $\mu\left(B_{\rho^2}(y)\right) > 0$ and let $p(y) \in B_{\rho^2}(y)$ be the center of mass of μ restricted to this ball. Let also $\pi(p)$ be the orthogonal projection of p onto $V(x,\rho)$. By Jensen's inequality:

$$d(p(y),V(x,\rho))^2 = d(p(y),\pi(p(y)))^2 = d\left(\fint_{B_{\rho^2}(y)} z\, d\mu(z), V(x,\rho)\right)^2 \leq \frac{1}{\mu(B_{\rho^2}(y))}\int_{B_{\rho^2}(y)} d(z,V(x,\rho))^2 d\mu(z)\,. \tag{5.31}$$

Using this estimate and lemma 5.6 (or better its rescaled version applied to $B_\rho(x)$), we want to prove that there exists a sequence of $k+1$ balls $B_{\rho^2}(y_i)$ with $y_i \in B_\rho(x)$ such that

i. $\mu\left(B_\rho(x) \cap B_{\rho^2}(y_i)\right) \geq c(n,\rho,M) > 0$

ii. $\{\pi(p(y_i))\}_{i=0}^k \equiv \{\pi_i\}_{i=0}^k$ effectively spans $V(x,\rho)$. In other words for all $i = 1,\cdots,k$, $\pi_i \in V(x,\rho)$ and

$$\pi_i \notin B_{5\rho^2}\left(\pi_0 + \operatorname{span}(\pi_1 - \pi_0,\cdots,\pi_{i-1}-\pi_0)\right). \tag{5.32}$$

We prove this statement by induction on $i = 0,\cdots,k$. For $i = 0$, the statement is trivially true since $\mu(B_\rho(x)) \geq \gamma_k\rho^k$. In order to find y_{i+1}, consider the subspace $V^{(i)} = \pi_0 + \operatorname{span}(\pi_1-\pi_0,\cdots,\pi_i-\pi_0)$. By lemma 5.6 applied to the ball $B_\rho(x)$, there exists some $B_{\rho^2}(y_{i+1})$ such that $\mu\left(B_{\rho^2}(y_{i+1})\right) \geq c(n,\rho,M) > 0$, $y_{i+1} \in B_\rho(x)$ and

$$y_{i+1} \notin B_{10\rho^2}\left(\pi_0 + \operatorname{span}(\pi_1-\pi_0,\cdots,\pi_i-\pi_0)\right). \tag{5.33}$$

By definition of center of mass, it is clear that $d(y_{i+1},p(y_{i+1})) \leq \rho^2$. Moreover, by item (i) and equation (5.31), we get

$$d(p(y_{i+1}),V(x,\rho))^2 \leq c\int_{B_\rho(x)\cap B_{\rho^2}(y_{i+1})} d(z,V(x,\rho))^2 d\mu(z) \leq cD^k_\mu(x,\rho) \leq c\delta\,. \tag{5.34}$$

Thus by the triangle inequality we have $d(y_{i+1},\pi_{i+1})\leq 2\rho^2$ if $\delta\leq\delta_0(n,\rho,M)$ is small enough. This implies (5.32). Using similar estimates, we also prove $d(p(y_{i+1}),V(0,1))^2\leq c' D^k_\mu(0,1)$ for all $i=-1,0,\cdots,k-1$. Thus by the triangle inequality

$$d(\pi_{i+1},V(0,1))\leq d(\pi_{i+1},p(y_{i+1}))+d(p(y_{i+1}),V(0,1))\leq c(n,\rho,M)\left(D^k_\mu(x,\rho)+D^k_\mu(0,1)\right)^{1/2}\ . \tag{5.35}$$

Now consider any $y\in V(x,\rho)$. By item (ii), there exists a unique set $\{D_i\}_{i=1}^k$ such that

$$y=\pi_0+\sum_{i=1}^k D_i(\pi_i-\pi_0)\, ,\quad |D_i|\leq c(n,\rho)\left\|y-\pi_0\right\|\ . \tag{5.36}$$

Hence for all $y\in V(x,\rho)\cap B_\rho(x)$, we have

$$d(y,V(0,1))\leq d(\pi_0,V(0,1))+\sum_i |D_i|\left[d(\pi_i,V(0,1))+d(\pi_0,V(0,1))\right]\leq c(n,\rho,M)\left(D^k_\mu(x,\rho)+D^k_\mu(0,1)\right)^{1/2}\ . \tag{5.37}$$

By lemma 4.9, this completes the proof of (5.29). □

5.5 Comparison between L^2 and L^∞ planes

Given $B_r(x)$, we denote as before by $V(x,r)$ one of the k-dimensional subspace minimizing $\int_{B_r(x)} d(y,V)^2 d\mu$. Suppose that the support of μ satisfies a uniform one-sided Reifenberg condition, i.e. suppose that there exists a k-dimensional plane $L(x,r)$ such that $x\in L(x,r)$ and

$$\operatorname{supp}(\mu)\cap B_r(x)\subseteq B_{\delta r}(L(x,r))\ . \tag{5.38}$$

Then, by the same technique used in lemma 5.7, we can prove that

Lemma 5.8. *Let μ be a positive Radon measure with $\mu(B_1(0))\geq\gamma_k$ and such that for all $B_\rho(y)\subseteq B_1(0)$ we have $\mu(B_\rho(y))\leq M\rho^k$ and (5.38). Then*

$$d_H(L(0,1)\cap B_1(0)\, ,V(0,1)\cap B_1(0))^2\leq c(n,\rho,M)\left(\delta^2+D^k_\mu(0,1)\right)\ . \tag{5.39}$$

5.6 Uniform lower estimates with the classical Reifenberg [MAYBE THIS SHOULD BE MOVED TO A PREVIOUS SECTION]

With the classical Reifenberg theorem, we cannot give upper bounds on the measure of the set under consideration. However, we can give lower bounds.

Lemma 5.9. *Under the assumptions of Theorem 4.7, for all $x \in S$ such that $B_r(x) \subseteq B_1(0)$,*

$$\lambda^k(S \cap B_r(x)) \geq (1-\epsilon)\omega_k r^k \geq \omega_k r^k/7\,. \tag{5.40}$$

Proof. By scale invariance, we assume $x = 0$ and $r = 1$. The classic Reifenberg theorem proves that there exits a bi-Hölder continuous map $\phi : L \to \mathbb{R}^n$ where L is a k-dimensional plane and

1. $|\phi(x) - x| \leq c\delta$ for all $x \in L$
2. $\phi(x) = x$ for $|x| \geq 1 + c\delta$
3. $S \cap B_1(0) = \phi(L')$, where $B_{1-c\delta}(0) \cap L \subseteq L' \subseteq B_{1+c\delta}(0) \cap L$.

Now let $f = \pi_L \circ \phi : L \to L$. This map is continuous and it is the identity outside $B_{1+c\delta}(0)$, and thus by topological reasons (degree theory) it is also surjective from L to itself.

In particular, the set $A = B_{1-3c\delta}(0) \cap L$ is contained in the image of f. By the uniform Reifenberg condition, $\pi_L^{-1}(A) \cap S \subseteq B_{1-2c\delta}(0)$, and by the properties of ϕ, $f^{-1}(A) = \phi^{-1}(\pi_L^{-1}(A) \cap S) \subseteq B_{1-c\delta}(0)$. Thus $\phi(f^{-1}(A)) \subseteq S$.

Now, since π_L has Lipschitz constant 1, by a standard result (see [Fed69, 2.10.11])

$$\lambda^k(S) \geq \lambda^k(\phi(f^{-1}(A)) \geq \lambda^k(\pi_L \circ \phi(f^{-1}(A))) = \lambda^k(A) \geq (1-3c\delta)^k \omega_k r^k\,. \tag{5.41}$$

For δ small enough, we have the thesis. □

5.7 Proof of the discrete Reifenberg

Here we want to prove theorem 5.4. The proof is taken from [NVb]. The other theorems are proved using variants of this idea, and we will not give all the details here.

First of all, note that, by definition of μ, the statement of this theorem is equivalent to

$$\mu(B_1(0)) \leq D(n)/\omega_k\,. \tag{5.42}$$

In the proof, we will fix the constant $C_1(k) \leq 40^k \omega_k$ and therefore the positive scale $\rho(n, C_1(k)) = \rho(n) < 1$ according to lemma 5.6. For convenience, we will assume that $\rho = 2^q$, $q \in \mathbb{N}$. Moreover, as noted before, instead of (5.16) we may assume the estimate

$$\sum_{\mathrm{r}_\alpha \leq r/2} \int_{B_r(x)} D_\mu^k(y, \mathrm{r}_\alpha)\, d\mu(y) < \delta^2 r^k\,, \tag{5.43}$$

where $\mathrm{r}_\alpha = 2^{-\alpha}$.

The constant $C_1(k)$ will be defined by the end of the proof, however it is enough to know that it is itself bounded by $40^k \omega_k$. Moreover, also the constant ρ will be chosen later on.

Bottom scale In the proof, it will be convenient to assume that $r_s \geq \bar{r} > 0$. It is clear that, by means of a simple limiting argument, this assumption is not restrictive. In particular, fix any positive radius $\bar{r} = r_A = \rho^A$ for some $A \in \mathbb{N}$, and consider the measure $\mu_{\bar{r}} \leq \mu$ defined by

$$\mu_{\bar{r}} = \sum_{s \ \ s.t. \ \ r_s \geq \bar{r}} \omega_k r_s^k \delta_{x_s} \, . \tag{5.44}$$

Note that this is a finite sum if $\bar{r}$ is positive. By Remark 5.3, we see that $\mu_{\bar{r}}$ satisfies all the hypothesis of this theorem, and since $\mu_{\bar{r}} \nearrow \mu$, if we prove uniform bounds on $\mu_{\bar{r}}(B_1(0))$ which are *independent* of $\bar{r}$, we can conclude the theorem.

For this reason, in the rest of the proof we will assume for simplicity that $r_s \geq \bar{r} = \rho^A > 0$ for all s.

5.7.1 First induction: upwards

We are going to prove inductively on $j = A, \cdots, 0$ that for all $x \in B_1(0) \subset \mathbb{R}^n$ and $\mathfrak{r}_j = \rho^j \leq 1$, either $B_{\mathfrak{r}_j}(x)$ is contained in one of the balls $\{B_{r_s}(x_s)\}_{s \in S}$, or we have the bound

$$\mu\left(B_{\mathfrak{r}_j}(x)\right) \leq C_1(k)\mathfrak{r}_j^k \, . \tag{5.45}$$

Note that, for $j = A$, this bound follows from the definition of the measure μ and the assumption that $r_s \geq \bar{r}$. Note also that this implies $\mathfrak{r}_j \leq 2r_s$ for all $s \in S$ and $x_s \in \overline{B_{\mathfrak{r}_j}(x)}$.

Clearly, we can assume wlog that $\mu\left(B_{\mathfrak{r}_j}(x)\right) \geq \gamma_k \mathfrak{r}_j^k$, otherwise there is nothing to prove. This observation will be essential in order to apply lemma 5.7.

Moreover, as long as we are trying to prove (5.45), we can replace wlog μ with $\mu|_{B_{\mathfrak{r}_j}(x)}$. Indeed, by Remark 5.3, all the hypotheses of theorem 5.4 hold also for any restriction of μ, in particular equation (5.16). Thus, from now on, μ will indicate $\mu|_{B_{\mathfrak{r}_j}(x)}$ and $S = \operatorname{supp}(\mu) \subset B_{\mathfrak{r}_j}(x)$.

Remark 5.10. Note that if $\nu = \mu|_{B_{\mathfrak{r}_j}(x)}$, then all the D_ν^k on balls $B_s(y) \supset B_{\mathfrak{r}_j}(x)$ are controlled. Indeed, we have

$$D_\mu^k(x, \mathfrak{r}_j) = D_\nu^k(x, \mathfrak{r}_j) = \mathfrak{r}_j^{-k-2} \int_{B_{\mathfrak{r}_j}(x)} d(z, V(x, \mathfrak{r}_j))^2 d\mu(z) = \mathfrak{r}_j^{-k-2} \int_{B_s(y)} d(z, V(x, \mathfrak{r}_j))^2 d\nu(z) \geq \frac{s^{k+2}}{\mathfrak{r}_j^{k+2}} D_\nu^k(y, s) \, . \tag{5.46}$$

In particular, this implies that if $B_{\mathfrak{r}_j}(x) \subseteq B_s(y)$, then

$$D_\nu^k(y, s) \leq \left(\frac{\mathfrak{r}_j}{s}\right)^{k+2} D_\nu^k(x, \mathfrak{r}_j) = \left(\frac{\mathfrak{r}_j}{s}\right)^{k+2} D_\mu^k(x, \mathfrak{r}_j) \, . \tag{5.47}$$

In turn, as long as $\mu(B_{\mathfrak{r}_j}(x)) \leq c(n) r^k$, we also have the bound

$$\int_{\mathbb{R}^n} \left(\int_0^\infty D_\nu^k(z, t) \, \frac{dt}{t}\right) d\nu(z) < c\delta^2 s^k \, . \tag{5.48}$$

5.7.2 Rough estimate

Fix some j, and suppose that (5.45) holds on all scales below $\mathfrak{r}_j$, i.e., for all $y \in B_1(0)$ and $\bar{r} \le \mathfrak{r}_i \le \mathfrak{r}_j$, $\mu(B_{\mathfrak{r}_i}(y)) \le C_1(n)\mathfrak{r}_i^k$.

Let us first observe that we can easily obtain a bad upper bound on $\mu\left(B_{\chi\mathfrak{r}_j}(x)\right)$ for any fixed $\chi > 1$. Consider the points in $\{x_s\}_{s\in S} \cap B_{\chi\mathfrak{r}_j}(x)$, and divide them into two groups: the ones with $r_s \le \mathfrak{r}_j$ and the ones with $r_s > \mathfrak{r}_j$.

For the first group, cover them by balls $B_{\mathfrak{r}_j}(z_i)$ such that $B_{\mathfrak{r}_j/2}(z_i)$ are disjoint. Since there can be at most $c(n,\chi)$ balls of this form, and for all of these balls the upper bound (5.45) holds, we have an induced upper bound on the measure of this set.

As for the points with $r_s > \mathfrak{r}_j$, by construction there can be only $c(n,\chi)$ many of them, and we also have the bound $r_s \le 2\chi\mathfrak{r}_j$. Summing up the two contributions, we get the very rough estimate

$$\mu\left(B_{\chi\mathfrak{r}_j}(x)\right) \le C_2(n,\chi)\mathfrak{r}_j^k\,, \tag{5.49}$$

where $C_2 >> C_1$. Note that, as long as the inductive hypothesis holds, C_1 is independent of j. However, it is clear that successive repetitions of the above estimate will not lead to (5.45).

5.7.3 Second induction: downwards. Outline of the proof

Suppose that (5.45) is true for all $x \in B_1(0)$ and $i = j+1, \cdots, A$. Fix $x \in \mathbb{R}^n$, and consider the set $B = B_{\mathfrak{r}_j}(x)$. Recall that we always assume that B is not contained in one of the balls $B_{r_s}(x_s)$, otherwise the bound (5.45) might fail for trivial reasons. We are going to build by induction on $i \ge j$ a sequence of smooth maps $\sigma_i : \mathbb{R}^n \to \mathbb{R}^n$ and smooth k-dimensional manifolds $T_{j,i} = T_i$ which will serve as approximations for the support of μ at scale $\mathfrak{r}_i$. Let us outline the inductive procedure now, and introduce all the relevant terminology. Everything described in the remainder of this subsection will be discussed more precisely over the coming pages. To begin with, we will have at the first step that

$$\begin{aligned} \sigma_j &= id, \\ T_j &= V(x,\mathfrak{r}_j) \subset \mathbb{R}^n\,, \end{aligned} \tag{5.50}$$

where $V(x,\mathfrak{r}_j)$ is one of the k-dimensional affine subspaces which minimizes $\int_{B_{\mathfrak{r}_j}(x)} d^2(y,V)\,d\mu$. Thus, the first manifold T_j is a k-dimensional affine subspace which best approximates $B_{\mathfrak{r}_j}(x)$. At future steps we can recover T_{i+1} from T_i and σ_{i+1} from the simple relation

$$T_{i+1} = \sigma_{i+1}(T_i)\,. \tag{5.51}$$

We will see that σ_{i+1} is a diffeomorphism when restricted to T_i, and thus each additional submanifold T_{i+1} is also diffeomorphic to $\mathbb{R}^k$. As part of our inductive construction we will build at each stage a Vitali covering of T_i given by

$$B_{\mathfrak{r}_i}(T_i) \cap B_{\mathfrak{r}_j}(x) \sim \bigcup_{t=j}^{i}\left(\bigcup_{y\in I_b^t} B_{\mathfrak{r}_t}(y) \cup \bigcup_{x_s\in I_f^t} B_{r_s}(x_s)\right) \cup \bigcup_{y\in I_g^i} B_{\mathfrak{r}_i}(y)\;, \tag{5.52}$$

where I_g, I_b, and I_f represent the *good*, *bad*, and *final* balls in the covering. Final balls are balls belonging to the original covering $B_{r_s}(x_s)$ such that $r_s \in [\mathfrak{r}_i, \mathfrak{r}_{i-1})$, and the other balls in the covering are characterized as good or bad according to how much measure they carry. Good balls are those with large measure, bad balls the ones with small measure. More precisely, we have

$$\begin{aligned} &\mu(B_{\mathfrak{r}_i}(y)) \geq \gamma_k \mathfrak{r}_i^k\,, \quad \text{if} \quad y \in I_g^i\,,\\ &\mu(B_{\mathfrak{r}_i}(y)) < \gamma_k \mathfrak{r}_i^k\,, \quad \text{if} \quad y \in I_b^i\,. \end{aligned} \tag{5.53}$$

We will see that, over each good ball $B_{\mathfrak{r}_i}(y)$ in this covering, T_i can be written as a graph over the best approximating subspace $V(y, \mathfrak{r}_i)$ with good estimates.

Our goal in these constructions is the proof of (5.45) for the ball $B = B_{\mathfrak{r}_j}(x)$, and thus we will need to relate the submanifolds T_i, and more importantly the covering (5.52), to the set B. Indeed, this covering of T_i almost covers the set B, at least up to an excess set E_{i-1}. That is,

$$\operatorname{supp}(\mu) \cap B \subseteq E_{i-1} \cup \bigcup_{t=j}^{i} \left(\bigcup_{y \in I_b^t} B_{\mathfrak{r}_t}(y) \cup \bigcup_{x_s \in I_f^t} B_{r_s}(y) \right) \cup \bigcup_{y \in I_g^i} B_{\mathfrak{r}_i}(y)\ . \tag{5.54}$$

We will see that the set E_{i-1} consists of those points of B which do not satisfy a uniform Reifenberg condition. Thus in order to prove (5.45) we will need to estimate the covering (5.52), as well as the excess set E_{i-1}.

Let us now outline the main properties used in the inductive construction of the mapping $\sigma_{i+1} : \mathbb{R}^n \to \mathbb{R}^n$, and hence $T_{i+1} = \sigma_{i+1}(T_i)$. As is suggested in (5.52), it is the good balls and not the bad and final balls which are subdivided at further steps of the induction procedure. In order to better understand this construction let us begin by analyzing the good balls $B_{\mathfrak{r}_i}(y)$ more carefully. On each such ball we may consider the best approximating k-dimensional subspace $V(y, \mathfrak{r}_i)$. Since $B_{\mathfrak{r}_i}(y)$ is a good ball, one can check that most of $B \cap B_{\mathfrak{r}_i}(y)$ must satisfy a uniform Reifenberg and reside in a small neighborhood of $V(y, \mathfrak{r}_i)$. We denote those points which don't by $E(y, \mathfrak{r}_i)$, see (5.73) for the precise definition. Then we can define the next step of the excess set by

$$E_i = E_{i-1} \cup \bigcup_{y \in I_g^i} E(y, \mathfrak{r}_i)\,. \tag{5.55}$$

Thus our excess set represents all those points which do not lie in an appropriately small neighborhood of the submanifolds T_i. With this in hand we can then find a submanifold $T_i' \subseteq T_i$, which is roughly defined by

$$T_i' \approx T_i \setminus \left(\bigcup_{t=j}^{i+1} \bigcup_{y \in I_b^t} B_{\mathfrak{r}_t/6}(y) \cup \bigcup_{t=j}^{i+1} \bigcup_{x_s \in I_f^t} B_{r_s/6}(x_s) \right), \tag{5.56}$$

see (5.87) for the precise inductive definition, such that

$$\operatorname{supp}(\mu) \cap B \subseteq E_i \cup \bigcup_{t=j}^{i} \left(\bigcup_{y \in I_b^t} B_{\mathfrak{r}_s}(y) \cup \bigcup_{x_s \in I_f^t} B_{r_s}(y) \right) \cup B_{\mathfrak{r}_{i+1}/4}(T_i') \equiv R_i \cup \bigcup B_{\mathfrak{r}_{i+1}/4}(T_i')\,, \tag{5.57}$$

where R_i represents our remainder term, and consists of those balls and sets which will not be further subdivided at the next stage of the induction.

The basic idea is that if in our induction we find a bad ball or a final ball $B_{\mathfrak{r}}(x)$, we know that the measure carried by this ball is bounded by $C\mathfrak{r}^k$. Because of this upper bound, in order to get the final estimate on μ we do not need to further analyze the measure inside any of these balls. However, we do need to keep track of the measure carried by these balls in successive induction steps. This is why every time we find a bad or final ball, we create a corresponding "hole" in the manifold T_i, and obtain as a result T_i'. By construction, the k-dimensional measure of these holes is comparable to $\mu(B_{\mathfrak{r}}(x))$, and thus the k-dimensional measure of T_i (without holes) already "includes" the μ-measure of the final and bad balls at all bigger scales.

Now in order to finish the inductive step of the construction, we can cover $B_{\mathfrak{r}_{i+1}/4}(T_i')$ by some Vitali set

$$B_{\mathfrak{r}_{i+1}/4}(T_i') \subseteq \bigcup_{y\in I} B_{\mathfrak{r}_{i+1}}(y) \ , \tag{5.58}$$

where $y \in I \subseteq T_i'$. We may then decompose the ball centers

$$I^{i+1} = I_g^{i+1} \cup I_b^{i+1} \cup I_f^{i+1} \ , \tag{5.59}$$

based on (5.53). Now we will use Definition 4.14 and the best approximating subspaces $V(y, \mathfrak{r}_{i+1})$ to build $\sigma_{i+1} : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$\operatorname{supp}\{\sigma_{i+1} - Id\} \subseteq \bigcup_{y\in I_g^{i+1}} B_{3\mathfrak{r}_{i+1}}(y) \ . \tag{5.60}$$

In order to prove the final bounds, we need to track the measure of the approximating manifolds T_i as i goes to infinity. We can use the local bi-Lipschitz estimates for σ_i at scale $\mathfrak{r}_i$ and integrate them along each manifold T_i to obtain uniform bounds on $\lambda^k(T_i)$ as i goes to infinity. This completes the outline of the inductive construction.

5.7.4 First steps in the induction

In order to make the proof more understandable, we give in detail the proof of the first steps in the downwards induction, which contains most of the necessary ideas to carry out the whole construction.

Fix any $B_{\mathfrak{r}_j}(x)$. Without loss of generality, we assume that

$$\mu(B_{\mathfrak{r}_j}(x)) \geq 2\gamma_k \mathfrak{r}_j^k \,, \tag{5.61}$$

otherwise we clearly have the measure estimate we want to prove.

With this condition, it makes sense to talk about a best L^2 approximating subspace for the support of μ on $B_{\mathfrak{r}_j}(x)$. Denote this subspace by $V(x, \mathfrak{r}_j) \equiv T_j$.

Now, we want to cover the support of μ with balls of radius $9\mathfrak{r}_{j+1}/10$ (roughly one scale smaller) in such a way to have good k-dimensional packing estimates on these balls. The idea is that condition (5.16) will force the measure μ to be *almost* supported in a small tubular neighborhood of T_j, up to small measure. So we split our ball in

$$B_{\mathfrak{r}_j}(x) = \left(B_{\mathfrak{r}_j}(x) \cap B_{\mathfrak{r}_{j+1}/11}\left(V(x,\mathfrak{r}_j)\right)\right) \cup \left(B_{\mathfrak{r}_j}(x) \cap B_{\mathfrak{r}_{j+1}/11}\left(V(x,\mathfrak{r}_j)\right)^C\right) . \tag{5.62}$$

The second part in this splitting is in some sense preventing the measure μ to satisfy an L^∞ Reifenberg condition. We call this part *excess set*, in particular

$$E(x,\mathfrak{r}_j) = B_{\mathfrak{r}_j}(x) \cap B_{\mathfrak{r}_{j+1}/11}\left(V(x,\mathfrak{r}_j)\right)^C . \tag{5.63}$$

Although $\mu(E)$ can be positive, it cannot be too big. Indeed, we have the trivial estimate

$$\int_{B_{\mathfrak{r}_j}(x)\setminus E(x,\mathfrak{r}_j)} d(y,V(x,\mathfrak{r}_j))^2\, d\mu(y) + \mu\left(E(x,\mathfrak{r}_j)\right)(\mathfrak{r}_{j+1}/11)^2 \leq \tag{5.64}$$

$$\leq \int_{B_{\mathfrak{r}_j}(x)} d(y,V(x,\mathfrak{r}_j))^2\, d\mu(y) = \mathfrak{r}_j^{k+2} D^k_\mu(x,\mathfrak{r}_j) \leq c(n,\rho)\mathfrak{r}_j^{k+2}\delta\, . \tag{5.65}$$

Now, almost all of the measure μ must be concentrated in $B_{\mathfrak{r}_j}(x) \cap B_{\mathfrak{r}_{j+1}/11}\left(V(x,\mathfrak{r}_j)\right)$. In order to estimate this part, we build a covering with good overlapping properties of this set by balls of radius $\geq 9\mathfrak{r}_{j+1}/10$ centered on $V(x,\mathfrak{r}_j) \cap B_{\mathfrak{r}_j}(x)$.

This covering is built in the following way. First of all, we consider separately all the balls $\{B_{r_s}(x_s)\}_{s\in S}$ with $r_s \geq \mathfrak{r}_{j+1}$. Recall that by construction $r_s \leq 2\mathfrak{r}_j$ for all $s \in S$ with $x_s \in B_{\mathfrak{r}_j}(x)$, otherwise there's nothing to prove since $B_{\mathfrak{r}_j}(x)$ would be contained in $B_{r_s}(x_s)$ for some s. Note that, all of these balls are pairwise disjoint. We will call these balls *final balls*, and set I^f_j to be the set of centers of these balls.

Most of the times, the set of final balls will be empty or very small. We complete this partial covering of $B_{\mathfrak{r}_j}(x) \cap B_{\mathfrak{r}_{j+1}/11}\left(V(x,\mathfrak{r}_j)\right)$ with other balls centered on $V(x,\mathfrak{r}_j)$ of radius $9\mathfrak{r}_{j+1}/10$ in such a way that this covering have a Vitali property.

Thus we obtain

$$B_{\mathfrak{r}_j}(x) \cap B_{\mathfrak{r}_{j+1}/11}\left(V(x,\mathfrak{r}_j)\right) \subset \bigcup_{s\in I^f_j} B_{r_s}(x_s) \cup \bigcup_{q\in Q} B_{9\mathfrak{r}_{j+1}/10}\left(x_q\right) . \tag{5.66}$$

Now, we split the set Q according to how much measure is contained in $B_{\mathfrak{r}_{j+1}}\left(x_q\right)$. In particular, if $\mu\left(B_{\mathfrak{r}_{j+1}}\left(x_q\right)\right) \geq \gamma_k \mathfrak{r}^k_{j+1}$, we say that this is a *good ball*, otherwise we say that this is a *bad ball*.

Now we want to build a new best approximating manifold T_{j+1} at this scale. On good balls, we have a best approximating subspace $V(x_q,\mathfrak{r}_{j+1})$, and since these balls carry enough measure, we can apply lemma 5.7 and obtain a quantitative estimate on the distance between $V(x,\mathfrak{r}_j)$ and $V(x_q,\mathfrak{r}_{j+1})$. In turn, this will allow us to apply the construction in the squash lemma 4.16. In particular, we have a smooth map σ defined on $\mathbb{R}^n$, and moreover $\sigma(T_j) \equiv T_{j+1}$ is a diffeomorphism onto its image when restricted to T_j with good quantitative bi-Lipschitz estimates. The details of this construction are carried out in subsection 5.7.8.

As for bad balls, we don't have to worry too much about those, since they carry really small measure. In particular,

$$\mu\left(B_{\mathfrak{r}_{j+1}}\left(x_q\right)\right) < \lambda^k\left(T_{j+1} \cap B_{\mathfrak{r}_{j+1}/6}\left(x_q\right)\right) . \tag{5.67}$$

Thus, in order to keep track of the measure carried by bad balls, we can simply keep track of the k-dimensional measure of the approximating manifolds T_i. In some sense, every time we hit a bad ball, we can compare its measure μ to the k-dimensional Hausdorff measure of the "hole" $T_{j+1} \cap B_{\mathfrak{r}_{j+1}/6}\left(x_q\right)$. Since, as we will prove, for $i \geq j+1$, $T_i \cap B_{\mathfrak{r}_{j+1}/6}\left(x_q\right)$ and $T_j \cap B_{\mathfrak{r}_{j+1}/6}\left(x_q\right)$ are substantially equal, by estimating the measure of T_i we also estimate the total measure of all the bad balls.

Moreover, since this estimate covers the measure of the whole bad ball, from this step forward we do not have to worry about $\mu|_{B_{\mathfrak{r}_{j+1}}(x_q)}$ any longer in the induction. We can use a similar argument to track the measure of final balls.

This construction is carried out in Subsection 5.7.9

Evidently, we cannot hope to apply these considerations also to good balls in order to get the estimates we want, because the measure of good balls is not small (a priori it could be anything). Instead, on the new good balls, we start over the same construction we outlined here (excess set, construction of the new best approximating manifold, and so on) and keep going by induction. The inductive estimates on the k-dimensional measure are carried out in Subsection 5.7.10.

5.7.5 Second induction: details of the construction

Let us now describe precisely the proof of this inductive construction which will lead to (5.45). For $j \leq i \leq A$, we will define a sequence of approximating manifolds T_i for the support of μ and a sequence of smooth maps σ_i such that

i. $\sigma_j = id$, $T_j = V(x, \mathfrak{r}_j) \subset \mathbb{R}^n$,

ii. $T_i = \sigma_i(T_{i-1})$,

iii. for $i \geq j+1$ and $y \in T_{i-1}$,

$$d(\sigma_i(y), y) \leq c\delta\mathfrak{r}_i\,, \tag{5.68}$$

and $\sigma_i|_{T_{i-1}}$ is a diffeomorphism,

iv. for every $y \in T_i$, $T_i \cap B_{2\mathfrak{r}_i}(y)$ is the graph over some k-dimensional affine subspace of a smooth function f satisfying

$$\frac{\|f\|_\infty}{\mathfrak{r}_i} + \|\nabla f\|_\infty \leq c\delta\,. \tag{5.69}$$

As outlined before, the manifolds T_i will be good approximations of the set S up to some "excess" set of small measure. Moreover, we will also introduce the concept of *good, bad* and *final* balls (whose centers will be in the sets I^i_g, I^i_b and I^i_f), a *remainder set* R_i, and the manifolds $T'_i \subseteq T_i$. Before giving the precise definitions (which are in equations (5.84), (5.78), (5.75) and (5.87) respectively), let us group here all the properties that we will need (and prove) for these objects, so that the reader can always come back to this page to have a clear picture of what are the objectives of the proof.

v. for every $i \geq j+1$ and $y \in I_g^i$, $d(y, V(y,\mathfrak{r}_i)) \leq c\delta\mathfrak{r}_i$, the set $T_i \cap B_{1.5\mathfrak{r}_i}(y)$ is the graph over $V(y,\mathfrak{r}_i)$ of a smooth function f satisfying (5.69), where $V(y,r)$ is one of the k-dimensional affine subspaces minimizing $\int_{B_r(y)} d^2(y,V)\,d\lambda^k$,

vi. for all i, we have the inclusion

$$\operatorname{supp}(\mu) \cap B \subseteq B_{\mathfrak{r}_{i+1}/10}\left(T_i'\right) \cup R_i\,, \tag{5.70}$$

The last two properties needed are the key for the final volume estimates:

vii. we can estimate

$$\lambda^k(\sigma_i^{-1}(T_i' \cap B_{2\mathfrak{r}_j}(x))) + \#\left(I_b^i\right)\omega_k(\mathfrak{r}_i/10)^k + \omega_k \sum_{x_s \in I_f^i} (r_s/10)^k \leq \lambda^k(T_{i-1}' \cap B_{2\mathfrak{r}_j}(x))\,, \tag{5.71}$$

viii. we can estimate the excess set by

$$\mu(E(y,\mathfrak{r}_i))\mathfrak{r}_{i+1}^2 \leq C(n)\mathfrak{r}_i^{k+2} D_\mu^k(y, 2\mathfrak{r}_i)\,. \tag{5.72}$$

At the first step of our induction, we can assume wlog that $\mu(B_{\mathfrak{r}_j}(x)) \geq 2\gamma_k\mathfrak{r}_j^k$. We set $I_b^j = I_f^j = \emptyset$, $I_g^j = \{\bar{x}\}$, where $\bar{x}$ is the center of mass of $\mu|_{B_{\mathfrak{r}_j}(x)}$, $T_j = V(0,1) = T_j'$ and $\sigma_j = id$. We set $E(x,\mathfrak{r}_j)$ to be the excess set defined by (5.63), and $R_j = E(x,\mathfrak{r}_j)$. It is clear from these definitions that all the properties (i)-(viii) are satisfied.

Now we proceed by induction assuming that we have defined for all $t \in \{j,\cdots,i\}$ I_g^t, I_b^t, I_f^t, the maps σ_t and the manifolds T_t, T_t'. Moreover, we assume that we also have defined for all $t \in \{j,\cdots,i-1\}$ the excess sets $\{E(y,\mathfrak{r}_t)\}_{y\in I_g^t}$ and the remainder R_{i-1}.

In the induction step, we will first build $\{E(y,\mathfrak{r}_i)\}_{y\in I_g^i}$ and R_i, and then move on to the construction of $I_g^{i+1}, I_b^{i+1}, I_f^{i+1}$, σ_{i+1} and T_{i+1}, T_{i+1}'.

5.7.6 Excess set.

Let us begin by describing the construction of the excess set. We will only be interested here in a ball $B_{\mathfrak{r}_i}(x)$ which is a good ball, in the sense that $\mu(B_{\mathfrak{r}_i}(x)) \geq \gamma_k\mathfrak{r}_i^k$.

Thus define $V(x,r)$ to be (one of) the k-dimensional plane minimizing $\int_{B_r(x)} d(y,V)^2 d\mu$, and define also the excess set to be the set of points which are some definite amount away from the best plane V. Precisely,

$$E(x,\mathfrak{r}_i) = \left(B_{\mathfrak{r}_i}(x) \setminus B_{\mathfrak{r}_{i+1}/11}(V)\right) \cap S\,. \tag{5.73}$$

The points in $\operatorname{supp}(\mu) \cap E$ are in some sense what prevents the set S from satisfying a uniform one-sided Reifenberg condition at this scale. By construction, all points in E have a uniform lower bound on the distance from V, so that if we assume $\mu(B_{\mathfrak{r}_i}(x)) \geq \gamma_k \mathfrak{r}_i^k$, i.e. $B_{\mathfrak{r}_i}(x)$ is a good ball, then we can estimate

$$\int_{B_{\mathfrak{r}_i}(x)\setminus E(x,\mathfrak{r}_i)} d(y, V(x,\mathfrak{r}_i))^2 \, d\mu(y) + \mu(E(x,\mathfrak{r}_i))(\mathfrak{r}_{i+1}/11)^2 \leq \int_{B_{\mathfrak{r}_i}(x)} d(y, V(x,\mathfrak{r}_i))^2 \, d\mu(y) = \mathfrak{r}_i^{k+2} D_\mu^k(x,\mathfrak{r}_i)\,. \tag{5.74}$$

5.7.7 Good, bad and final balls

Inductively, let us define the remainder set to be the union of all the previous bad balls, final balls, and the excess sets:

$$R_i = \bigcup_{t=j}^{i} \left(\bigcup_{y \in I_b^t} B_{\mathfrak{r}_t}(y) \cup \bigcup_{x_s \in I_f^t} B_{r_s}(x_s) \cup \bigcup_{I_g^t} E(y,\mathfrak{r}_t) \right) . \tag{5.75}$$

The set R_i represents everything we want to throw out at the inductive stage of the proof. We will see later in the proof how to estimate this remainder set itself. Note that $R_j = E(x,\mathfrak{r}_j)$.

Now consider the points $x_s \in S$ outside the remainder set, and separate the balls $B_{r_s}(x_s)$ with radius $r_s \sim \mathfrak{r}_{i+1}$ from the others by defining for $y \in I_g^i$ the sets

$$I_f^{i+1}(y) = \{x_s \in (S \setminus R_i) \cap B_{\mathfrak{r}_i}(y) \quad s.t. \;\; r_s \in [\mathfrak{r}_{i+1}, \mathfrak{r}_i)\} \,, \tag{5.76}$$

and

$$J^{i+1}(y) = \{x_s \in (S \setminus R_i) \cap B_{\mathfrak{r}_i}(y) \quad s.t. \;\; r_s < \mathfrak{r}_{i+1}\} \,. \tag{5.77}$$

From this we can construct the sets

$$I_f^{i+1} = \cup_{y \in I_g^i} I_f^{i+1}(y) \quad \text{and} \quad J^{i+1} = \cup_{y \in I_g^i} J^{i+1}(y)\,. \tag{5.78}$$

Note that by construction and inductive item (vi), we have

$$S \setminus R_i = I_f^{i+1} \cup J^{i+1} \subset B_{\mathfrak{r}_{i+1}/10}\left(T_i'\right) . \tag{5.79}$$

Recall that, as long as we are trying to prove the estimate (5.45), we can assume wlog that $\mu = \mu|_{B_{\mathfrak{r}_j}(x)}$ and $S = \operatorname{supp}(\mu) \subset B_{\mathfrak{r}_j}(x)$.

Let us now consider a covering of (5.79) given by

$$S \setminus R_i \subseteq I_f^{i+1} \cup \bigcup_{z \in I} B_{9\mathfrak{r}_{i+1}/10}(z) \,, \tag{5.80}$$

where $I \subseteq T_i'$, and for any $p \neq q \in I_f^{i+1} \cup I$, $B_{r_p/5}(p) \cap B_{r_q/5}(q) = \emptyset$. Here we denote for convenience $r_p = \mathfrak{r}_{i+1}$ if $p \in I$, and $r_p = r_s$ if $p = x_s \in I_f^{i+1}$. Note that this second property is true by definition for

$p, q \in I_f^{i+1}$, we only need to complete this partial Vitali covering with other balls of the same size. To be precise, note that by (5.75) and (5.73)

$$(S \setminus R_i) \setminus \bigcup_{z \in I_f^{i+1}} B_{3r_z/5}(z) \subset \left(\bigcup_{y \in I_g^i} (B_{\mathfrak{r}_{i+1}/4}(V(y, \mathfrak{r}_i)) \cap B_{\mathfrak{r}_i}(y)) \right) \bigcap \left(\bigcup_{t=j}^{i+1} \bigcup_{z \in I_f^t} B_{3r_z/5}(z) \cup \bigcup_{t=j}^{i} \bigcup_{z \in I_b^t} B_{3\mathfrak{r}_t/5}(z) \right)^C . \tag{5.81}$$

Take a finite covering of this last set by balls $\{B_{3\mathfrak{r}_{i+1}/10}(y)\}_{y \in Y}$. Note that we can pick

$$Y \cap \left(\bigcup_{t=j}^{i+1} \bigcup_{z \in I_f^t} B_{3r_z/5}(z) \cup \bigcup_{t=j}^{i} \bigcup_{z \in I_b^t} B_{3\mathfrak{r}_t/5}(z) \right) = \emptyset . \tag{5.82}$$

By item (iv), T_i is locally a Lipschitz graph over some k-dimensional subspace with (5.69), and thus we can choose $Y \subset T_i$.

Consider a Vitali subcovering of this set, denote I the set of centers in this subcovering. Such a subcovering will have the property that the balls $\{B_{3\mathfrak{r}_{i+1}/10}(y)\}_{y \in I}$ will be pairwise disjoint. These balls will also be disjoint from $\bigcup_{z \in I_f^{i+1}} B_{r_z/5}(z)$ by (5.82). The (finite version of) Vitali covering theorem ensures that $\bigcup_{y \in I} B_{9\mathfrak{r}_{i+1}/10}(y)$ will cover the whole set in (5.81).

Now by construction of I_f and the remainder set, all the balls $\{B_{r_s}(x_s)\}_{s \in S}$ with $r_s \geq \mathfrak{r}_{i+1}$ have already been accounted for. This means that

$$(S \setminus R_i) \setminus \bigcup_{x_s \in I_f^{i+1}} B_{r_s}(x_s) \subset \bigcup_{y \in I} B_{9\mathfrak{r}_{i+1}/10}(y) , \tag{5.83}$$

as desired.

We split the balls with centers in I into two subsets, according to how much measure they carry. In particular, let

$$I_g^{i+1} = \left\{ y \in I \ \ s.t. \ \ \mu(B_{\mathfrak{r}_{i+1}}(y)) \geq \gamma_k \mathfrak{r}_{i+1}^k \right\} , \quad I_b^{i+1} = \left\{ y \in I \ \ s.t. \ \ \mu(B_{\mathfrak{r}_{i+1}}(y)) < \gamma_k \mathfrak{r}_{i+1}^k \right\} . \tag{5.84}$$

5.7.8 Map and manifold structure.

Let $\left\{ \lambda_s^{i+1} \right\} = \{\lambda_s\}$ be a partition of unity such that for each $y_s \in I_g^{i+1}$

- $\operatorname{supp}(\lambda_s) \subseteq B_{3\mathfrak{r}_{i+1}}(y_s)$
- for all $z \in \cup_{y_s \in I_g^{i+1}} B_{2\mathfrak{r}_{i+1}}(y_s)$, $\sum_s \lambda_s(z) = 1$
- $\max_s \|\nabla \lambda_s\|_\infty \leq C(n)/\mathfrak{r}_{i+1}$.

For every $y_s \in I_g^{i+1}$, let $V(y_s, \mathfrak{r}_{i+1})$ to be (one of) the k-dimensional subspace that minimizes $\int_{B_{\mathfrak{r}_{i+1}}(y_s)} d(z, V)^2 d\mu$. By Remark 5.4 we can estimate

$$\mathfrak{r}_{i+1}^{-k-2} \int_{B_{\mathfrak{r}_{i+1}}(y_s)} d(z, V(y_s, \mathfrak{r}_{i+1}))^2 d\mu(z) \leq D_\mu^k(y_s, \mathfrak{r}_{i+1}) . \tag{5.85}$$

Let $p_s \in B_{\mathfrak{r}_{i+1}}(y_s)$ be the center of mass of $\mu|_{B_{\mathfrak{r}_{i+1}}(y_s)}$. It is worth observing that $p_s \in V(y_s, \mathfrak{r}_{i+1})$.

Define the smooth function $\sigma_{i+1} : \mathbb{R}^n \to \mathbb{R}^n$ as in Definition 4.14, i.e.,

$$\sigma_{i+1}(x) = x + \sum_s \lambda_s^{i+1}(x) \pi_{V(y_s, \mathfrak{r}_{i+1})^\perp}(p_s - x) . \tag{5.86}$$

With this function, we can define inductively for $i \geq j$ the sets

$$T_j = V(x, \mathfrak{r}_j) , \quad T'_j = T_j \tag{5.87}$$

$$T_{i+1} = \sigma_{i+1}(T_i) , \quad T'_{i+1} = \sigma_{i+1}\left(T'_i \setminus \left(\bigcup_{y \in I_b^{i+1}} B_{\mathfrak{r}_{i+1}/6}(y) \cup \bigcup_{x_s \in I_f^{i+1}} B_{r_s/6}(x_s)\right)\right) . \tag{5.88}$$

Fix any $y \in I_g^{i+1}$, and let $z \in I_g^i$ be such that $y \in B_{9\mathfrak{r}_i/10}(z)$. By induction, $T_i \cap B_{10\mathfrak{r}_{i+1}}(y) \subseteq T_i \cap B_{2\mathfrak{r}_i}(z)$ is the graph of a C^1 function over $V(z, \mathfrak{r}_i)$. Consider the points $\{y_s\} = I_g^{i+1} \cap B_{6\mathfrak{r}_{i+1}}(y)$. By construction it is easy to see that $d(y_s, V(z, \mathfrak{r}_i)) \leq \mathfrak{r}_{i+1}/9$, and so we can apply the estimates in lemma 5.7 with $M = C_1$ by the first induction. Using condition (5.16), we obtain that for all y_s:

$$\mathfrak{r}_{i+1}^{-1} d_H\left(V(z, \mathfrak{r}_i) \cap B_{\mathfrak{r}_{i+1}}(y_s), V(y_s, \mathfrak{r}_{i+1}) \cap B_{\mathfrak{r}_{i+1}}(y_s)\right) \leq c\left(D_\mu^k(y_s, \mathfrak{r}_{i+1}) + D_\mu^k(z, \mathfrak{r}_i)\right)^{1/2} \leq c(n, \rho, C_1)\delta . \tag{5.89}$$

This implies that, if $\delta(n, \rho, C_1)$ is small enough, $T_i \cap B_{2\mathfrak{r}_{i+1}}(y)$ is a graph also over $V(y, \mathfrak{r}_{i+1})$ satisfying the same estimates as in (5.69), up to a worse constant c. That is, if δ is sufficiently small, we can apply lemma 4.16 and prove induction point (v).

Points (iii) and (iv) Points (iii) and (iv) are proved with similar methods. We briefly sketch the proofs of these two points.

Let $y \in T_{i-1}$, and recall the function $\psi_i \equiv 1 - \sum \lambda_s$. If $\psi_i|_{B_{2\mathfrak{r}_i}(y)}$ is identically 1, then $\sigma_i|_{B_{2\mathfrak{r}_i}(y)} = id$, and there is nothing to prove.

Otherwise, there must exist some $z' \in I_g^i \cap B_{5\mathfrak{r}_i}(y)$, and thus there exists a $z \in I_g^{i-1}$ such that $B_{8\mathfrak{r}_i}(y) \subseteq B_{1.5\mathfrak{r}_{i-1}}(z)$. By point (v) in the induction, $T_{i-1} \cap B_{1.5\mathfrak{r}_{i-1}}(z)$ is a Lipschitz graph over $V(z, \mathfrak{r}_{i-1})$. Proceeding as before, by the estimates in lemma 5.7 and lemmas 4.16, we obtain that $T_i \cap B_{2\mathfrak{r}_i}(y)$ is also a Lipschitz graph over $V(z, \mathfrak{r}_{i-1})$ with small Lipschitz constant, and that $\|\sigma_i(p) - p\| \leq c\delta\mathfrak{r}_i$ for all $p \in T_{i-1}$.

Moreover, $\sigma_i|_{T_{i-1}}$ is locally a diffeomorphism at scale $\mathfrak{r}_i$. From this we see that σ_i is a diffeomorphism on the whole T_{i-1}.

It is worth to remark a subtle point. In order to prove point (iv), we cannot use inductively (iv), we need to use point (v). Indeed, as we have seen, given any $z \in I_g^{i-1}$, then $T_{i-1} \cap B_{1.5\mathfrak{r}_{i-1}}(z)$ is a Lipschitz graph of a function f where $\|\nabla f\| \leq c\delta$, and this c is *independent* of the induction step we are considering by (iii) in lemma 4.16. If we tried to iterate directly the bound given by (iv), the constant c would depend on the induction step i, and thus we could not conclude the estimate we want.

5.7.9 Properties of the manifolds T_i'

Here we want to prove the measure estimate in (5.71). The basic idea is that bad and final balls correspond to holes in the manifold T_i, and each of these holes carries a k-dimensional measure which is proportionate to the measure inside the balls. In particular, let $B_r(y)$ be a bad or a final ball. In the first case, $r = \mathfrak{r}_{i+1}$, while in the second $r \in [\mathfrak{r}_{i+1}, \mathfrak{r}_i)$. In either case, we will see that y must be $\sim \mathfrak{r}_{i+1}$-close to T_i, which is a Lipschitz graph at scale $\mathfrak{r}_i$. This implies that $\mu(B_r(y) \cap T_i) \sim r^k$, and thus we can bound the measure of a bad or final ball with the measure of the hole we have created on T_i.

In detail, point (vi) is an immediate consequence of the definition of R_i.

In order to prove the volume measure estimate, consider that

$$T_i' \setminus \sigma_{i+1}^{-1}(T_{i+1}') \subseteq \left(\bigcup_{y\in I_b^{i+1}} B_{\mathfrak{r}_{i+1}/6}(y) \cup \bigcup_{x_s\in I_f^{i+1}} B_{r_s/6}(x_s) \right) . \tag{5.90}$$

Note that the balls in the collection $\{B_{\mathfrak{r}_{i+1}/5}(y)\}_{y\in I_f^{i+1}\cup I_b^{i+1}}$ are pairwise disjoint. Pick any $y \in I_b^{i+1}$, and let $z \in I_g^i$ be such that $y \in B_{\mathfrak{r}_i}(z)$. By definition, $y \in T_i'$ and $\mu(B_{\mathfrak{r}_{i+1}}(y)) < \gamma_k \mathfrak{r}_{i+1}^k < 10^{-k}\omega_k r_{i+1}^k$. Since $T_i \cap B_{2\mathfrak{r}_i}(z)$ is a graph over $V(z, \mathfrak{r}_i)$ with $y \in T_i$, and since $B_{\mathfrak{r}_{i+1}/6}(y) \cap T_i$ is disjoint from the "holes" of T_i', then

$$\lambda^k(T_i' \cap B_{\mathfrak{r}_{i+1}/6}(y)) \geq \omega_k 7^{-k} \mathfrak{r}_{i+1}^k . \tag{5.91}$$

A similar estimate holds for the final balls. The only difference is that if $x_s \in I_f^{i+1}$, then it is not true in general that $x_s \in T_i$. However, by (5.79), we still have that $d(x_s, V(z, \mathfrak{r}_i)) \leq \mathfrak{r}_{i+1}/10$. Given (5.69), we can conclude

$$\lambda^k(T_i' \cap B_{r_s/7}(x_s)) \geq \omega_k 10^{-k} r_s^k . \tag{5.92}$$

Now it is evident from the definition of T_{i+1}' that

$$\lambda^k(\sigma_{i+1}^{-1}(T_{i+1}') \cap B_{2\mathfrak{r}_j}(x)) + \#\left(I_b^{i+1}\right)\omega_k(\mathfrak{r}_{i+1}/10)^k + \omega_k \sum_{x_s\in I_f^{i+1}} (r_s/10)^k \leq \lambda^k(T_i' \cap B_{2\mathfrak{r}_j}(x)) . \tag{5.93}$$

5.7.10 Volume estimates on the manifold part

Here we want to prove that for every measurable $\Omega \subseteq T_i$

$$\lambda^k(\sigma_{i+1}(\Omega)) \leq \lambda^k(\Omega) + c(n, \rho, C_1) \int_{B_{\mathfrak{r}_j}(x)} D(p, 2\mathfrak{r}_{i+1}) d\mu(p) . \tag{5.94}$$

The main applications will be with $\Omega = T_i$ and $\Omega = T_i'$. In order to do that, we need to analyze in a quantitative way the bi-Lipschitz correspondence between T_i and T_{i+1} given by σ_{i+1}.

As we already know, $\sigma_{i+1} = id$ on the complement of the set $G = \cup_{y\in I_g^{i+1}} B_{5\mathfrak{r}_{i+1}}(y)$, so we can concentrate only on this set.

Using the same techniques as before, and in particular by lemmas 5.7 and 4.16, we can prove that for each $y \in I_g^{i+1}$, the set $T_i \cap B_{5\mathfrak{r}_{i+1}}(y)$ is a Lipschitz graph over $V(y, \mathfrak{r}_{i+1})$ with Lipschitz constant bounded by

$$c(n,\rho,C_1)\left(D(y,\mathfrak{r}_{i+1}) + \sum_{z\in I_g^i\cap B_{5\mathfrak{r}_i}(y)} D(z,\mathfrak{r}_i)\right)^{1/2} . \tag{5.95}$$

In a similar manner, we also have that $T_{i+1} \cap B_{5\mathfrak{r}_{i+1}}(y)$ is a Lipschitz graph over $V(y, \mathfrak{r}_{i+1})$ with Lipschitz constant bounded by

$$c(n,\rho,C_1)\left(\sum_{z\in I_g^{i+1}\cap B_{10\mathfrak{r}_{i+1}}(y)} D(z,\mathfrak{r}_{i+1}) + \sum_{z\in I_g^i\cap B_{5\mathfrak{r}_i}(y)} D(z,\mathfrak{r}_i)\right)^{1/2} . \tag{5.96}$$

Moreover, by the bi-Lipschitz estimates of lemma 4.16, we also know that σ_{i+1} restricted to $T_i \cap B_{5\mathfrak{r}_{i+1}}(y)$ is a bi-Lipschitz equivalence with bi-Lipschitz constant bounded by

$$L(y,5\mathfrak{r}_{i+1}) \leq 1 + c\left(\sum_{z\in I_g^{i+1}\cap B_{10\mathfrak{r}_{i+1}}(y)} D(z,\mathfrak{r}_{i+1}) + \sum_{z\in I_g^i\cap B_{5\mathfrak{r}_i}(y)} D(z,\mathfrak{r}_i)\right) . \tag{5.97}$$

In order to estimate this upper bound, we use (5.3) and the definition of good balls to write

$$D(z,r) \leq c \fint_{B_r(z)} D(p,2r)d\mu(p) \leq c(n,\rho,C_1) r^{-k} \int_{B_r(z)} D(p,2r)d\mu(p) . \tag{5.98}$$

Since by construction any point $x \in \mathbb{R}^n$ can be covered by at most $c(n)$ different good balls at different scales, we can bound

$$L(y,5\mathfrak{r}_{i+1}) \leq 1 + \frac{c(n,\rho,C_1)}{\mathfrak{r}_{i+1}^k} \int_{B_{5\mathfrak{r}_i}(y)} \left[D(p,2\mathfrak{r}_{i+1}) + D(p,2\mathfrak{r}_i)\right] d\mu(p) . \tag{5.99}$$

We can also badly estimate

$$D(p,2\mathfrak{r}_{i+1}) + D(p,2\mathfrak{r}_i) \leq c(n,\rho) D(p,2\mathfrak{r}_i) . \tag{5.100}$$

Now let P_s be a measurable partition of $\Omega \cap G$ such that for each s, $P_s \subseteq B_{5\mathfrak{r}_{i+1}}(y_s)$. By summing up the volume contributions of P_s, and since evidently $\lambda^k(P_s) \leq 7^k \omega_k \mathfrak{r}_{i+1}^k$, we get

$$\begin{aligned}
\lambda^k(\sigma_{i+1}(\Omega)) = \sum_s \lambda^k(\sigma_{i+1}(P_s)) &\leq \sum_s \lambda^k(P_s)\left(1 + \frac{c}{\mathfrak{r}_{i+1}^k}\int_{B_{5\mathfrak{r}_i}(y_s)} D(p,2\mathfrak{r}_i)d\mu(p)\right) \\
&\leq \lambda^k(\Omega) + c \int_{\bigcup_{y_s\in I_g^{i+1}} B_{5\mathfrak{r}_i}(y_s)} D(p,2\mathfrak{r}_i)d\mu(p) \\
&\leq \lambda^k(\Omega) + c(n,\rho,C_1)\int_{B_{\mathfrak{r}_j}(x)} D(p,2\mathfrak{r}_i)d\mu(p) .
\end{aligned} \tag{5.101}$$

5.7.11 Estimates on the excess set

In this paragraph, we estimate the total measure of the excess set, which is defined by

$$E_T = \bigcup_{i=j}^{A} \bigcup_{y\in I_g^i} E(y,\mathfrak{r}_i)\,. \tag{5.102}$$

At each y and at each scale, we have by (5.74) and (5.3)

$$\mu(E(y,\mathfrak{r}_i)) \le c(n,\rho)\mathfrak{r}_i^k D_\mu^k(y,\mathfrak{r}_i) \le c(n,\rho)\mathfrak{r}_i^k \fint_{B_{\mathfrak{r}_i}(x)} D_\mu^k(p,2\mathfrak{r}_i)d\mu(p)\,. \tag{5.103}$$

Since by definition of excess set $B_{\mathfrak{r}_i}(y)$ must be a good ball, then

$$\mu(E(y,\mathfrak{r}_i)) \le c(n,\rho)\int_{B_{2\mathfrak{r}_i}(y)} D_\mu^k(p,2\mathfrak{r}_i)d\mu(p)\,. \tag{5.104}$$

Now by construction of the good balls, there exists a constant $c(n)$ such that at each step i, each $x\in\mathbb{R}^n$ belongs to at most $c(n)$ good balls. Thus for each $i\ge j$, we have

$$\sum_{y\in I_g^i}\mu(E(y,\mathfrak{r}_i)) \le c(n,\rho)\int_{\cup_{y\in I_g^i}B_{\mathfrak{r}_i}(y)} D_\mu^k(p,2\mathfrak{r}_i)d\mu(p) \le c(n,\rho)\int_{B_{\mathfrak{r}_j}(x)} D_\mu^k(p,2\mathfrak{r}_i)d\mu(p)\,. \tag{5.105}$$

If we sum over all scales, we get

$$\mu\left(E_T\right) \le c(n,\rho)\sum_{i=j}^{A}\int_{B_{\mathfrak{r}_j}(x)} D_\mu^k(p,2\mathfrak{r}_i)d\mu(p)\,. \tag{5.106}$$

Since $\rho = 2^q$, it is clear that

$$\mu\left(E_T\right) \le c(n,\rho)\sum_{i=j}^{A}\int_{B_{\mathfrak{r}_j}(x)} D_\mu^k\left(p,2^{1-qi}\right)d\mu(p) \le c(n,\rho)\delta\mathfrak{r}_j^k\,, \tag{5.107}$$

since the sum in the middle is clearly bounded by (5.18).

This estimate is exactly what we want from the excess set.

5.7.12 Volume estimates.

By adding (5.94), with $\Omega\equiv\sigma_{i+1}^{-1}(T'_{i+1})$, and (5.71), we prove that for all $i=j,\cdots,A+1,\ldots$

$$\lambda^k(T'_{i+1}\cap B_{2\mathfrak{r}_j}(x)) + \#\left(I_b^{i+1}\right)\omega_k(\mathfrak{r}_i/10)^k + \omega_k\sum_{x_s\in I_f^{i+1}}(r_s/10)^k \le \tag{5.108}$$

$$\le \lambda^k(T'_i\cap B_{2\mathfrak{r}_j}(x)) + c(n,\rho,C_1)\int_{B_{\mathfrak{r}_j}(y)\cap B_1(0)} D(p,2\mathfrak{r}_{i+1})d\mu(p)\,. \tag{5.109}$$

Adding the contributions from all scales, by (5.16) we get

$$\begin{aligned}
&\lambda^k(T'_{i+1}\cap B_{2\mathrm{r}_j}(x)) + \sum_{t=j}^{i+1} \#\left(I^t_b\right)\omega_k(\mathrm{r}_t/10)^k + \sum_{t=j}^{i+1}\sum_{x_s\in I^t_f}\omega_k(r_s/10)^k \\
&\leq \lambda^k(T_j\cap B_{2\mathrm{r}_j}(x)) + c(n,\rho,C_1)\sum_{s=j}^{i+1}\int_{B_{\mathrm{r}_j}(y)\cap B_1(0)} D(p,2r_s)d\mu(p) \\
&\leq \lambda^k\left(T_j\cap B_{2\mathrm{r}_j}(x)\right)\left[1+c(n,\rho,C_1)\delta\right] \ , \qquad (5.110)
\end{aligned}$$

where in the last line we estimated $\lambda^k\left(T_j\cap B_{2\mathrm{r}_j}(x)\right)\sim \mathrm{r}_j^k$, since T_j is a k-dimensional subspace, and we bounded the sum using (5.18).

In the same way, we can also bound the measure of T_{i+1} by

$$\lambda^k(T_{i+1}\cap B_{2\mathrm{r}_j}(x)) \leq \lambda^k\left(T_j\cap B_{2\mathrm{r}_j}(x)\right)\left[1+c(n,\rho,C_1)\delta\right] \ . \qquad (5.111)$$

5.7.13 Upper estimates for μ.

Since we have assumed $r_s\geq \bar{r}=r_A$ for all $x_s\in S$, we know that for $i=A$ our construction ends, and the whole support S is contained in final and bad balls and excess sets. In other words, the sets $I^A_g=I^A_b=\emptyset$, and

$$\operatorname{supp}(\mu)\cap B = \operatorname{supp}(\mu)\cap B_{\mathrm{r}_j}(x)\subseteq R_A \ . \qquad (5.112)$$

This fact and the estimates in (5.107) and (5.110) imply

$$\begin{aligned}
\mu(B) &\leq \sum_{t=j}^{A}\#\left(I^t_b\right)\gamma_k\omega_k\mathrm{r}_t^k + \sum_{t=j}^{A}\sum_{x_s\in I^t_f}\omega_k r_s^k + \mu(E_T) \\
&\leq 10^k\left(\sum_{t=j}^{A}\#\left(I^t_b\right)\omega_k(\mathrm{r}_t/10)^k + \sum_{t=j}^{A}\sum_{x_s\in I^t_f}\omega_k(r_s/10)^k\right) + \mu(E_T) \leq C_3(k)(1+c(n,\rho,C_1)\delta)\mathrm{r}_j^k \ . \qquad (5.113)
\end{aligned}$$

In this last estimate, we can fix $C_3(k)=20^k\omega_k$, and $C_1(k)=2C_3(k)=2\cdot 20^k\omega_k\leq 40^k\omega_k$, and $\rho(n,C_1)$ according to lemma 5.6. Now, it is easy to see that if $\delta(n,\rho,C_1)$ is sufficiently small, then

$$\mu(B)\leq C_1(k)\mathrm{r}_j^k \, , \qquad (5.114)$$

which finishes the proof of the downward induction, and hence the actual ball estimate (5.45).

6 L^2-Best Approximation Theorems

Now we move back to our main focus: harmonic maps and their singular set. First of all, we are going to prove a theorem that will allow us to control (under suitable assumptions) that the beta-numbers defined in the previous section can be controlled using the normalized energy. This is the key to proving the final theorem on singular sets.

The idea is the following: as we have seen with compactness arguments, lemma 3.17 implies that if we have points $\{p_i\}_{1=0}^{m-2} \in B_1(0)$ such that

1. $\theta(p_i, 3/2) - \theta(p_i, 1/4) < \epsilon$
2. the points $\{p_i\}_{1=0}^{m-2}$ are in ρ-linearly independent,

then there are no singular points in $B_{1/2}(0)$. In particular, for all $\tau > 0$ there exists an $\epsilon(m, \rho, N, \Lambda) > 0$ such that the previous statement is true.

The objective of this section is to try to make the previous statement more quantitative. We don't simply want to see that for all τ there exists some ϵ, but we want to see how τ and ϵ are connected. We don't care about making the dependence on Λ, m, N explicit, just the dependence on τ.

Recall that if $\epsilon = 0$, then we are in the perfectly homogeneous case. In this case, if the points p_i span something $m-2$-dimensional (so in some sense $\tau = 0$ corresponds to $\epsilon = 0$).

Before we move to the actual statement, we make an example in dimension 3 that hopefully will clarify the estimates in general.

Half of this section is taken from [NVa].

6.1 Dimension 3

In order to fix some ideas, let's focus for a moment on the easy case, i.e., when $m = 3$. Note that in this case the standard dimension reduction argument proves that the singular set in discrete, and with a covering argument one can show directly that the number of singular points is bounded (see 3.12). Thus we don't actually need extra estimates for this case. However, it is instructive to see what happens here.

With the usual compactness argument (see the cone-splitting lemma 3.17), as seen above we know that given two points $x, y \in B_1(0)$ such that

$$W(x,1) \equiv \theta(x,3/2) - \theta(x,1/4) < \epsilon\,, \quad \text{and} \quad W(y,1) \equiv \theta(y,3/2) - \theta(y,1/4) < \epsilon\,, \tag{6.1}$$

then if ϵ is sufficiently small we have two options: either $B_{1/2}(0)$ contains no singular points, or $d(x,y) \leq \tau$.

In order to make the statement quantitative, observe the following. For convenience, let's take the points x, y of the form

$$x = (\tau/2, 0, 0)\,, \quad y = -x = (-\tau, 0, 0)\,, \tag{6.2}$$

so that their mutual distance is τ. Suppose that $\tau < 1/10$, and consider the annulus $A_{1/2,2}(0) = B_2(0) \setminus B_{1/2}(0)$. This annulus is contained in the intersection of $A_{3,1/4}(x) = B_3(x) \setminus B_{1/4}(x)$ and $A_{3,1/4}(y) = B_3(y) \setminus B_{1/4}(y)$.

Moreover, for all $z \in A_{1/2,2}0$, consider that obviously

$$x - y = x - z - (y - z)\,, \tag{6.3}$$

and so

$$\langle \nabla u(z), x - y\rangle = \langle \nabla u(z), x - z\rangle - \langle \nabla u(z), y - z\rangle\ . \tag{6.4}$$

MAYBE HERE A DRAWING IS A GOOD IDEA

Now if $W(x,1) = W(y,1) = 0$, this clearly implies that for all $z \in A$ we have $\langle \nabla u(z), x - y\rangle = 0$, and we also have the quantitative estimate

$$\int_{A_{1/2,2}(0)} |\langle \nabla u(z), x - y\rangle|^2 \leq 2\int_{A_{1/2,2}(0)} |\langle \nabla u(z), x - z\rangle|^2 + 2\int_{A_{1/2,2}(0)} |\langle \nabla u(z), y - z\rangle|^2 \leq \tag{6.5}$$

$$\leq c\left[W(x,1) + W(y,1)\right]\ . \tag{6.6}$$

Now by a simple compactness argument, we shall see in the next section that for all directions e (i.e., for all unit vectors e) in $\mathbb{R}^3$, if $B_{1/2}\,(0)$ contains some singular point than

$$\int_{A_{1/2,2}(0)} |\langle \nabla u(z), e\rangle|^2 \geq c_0(m, N, \Lambda)\,. \tag{6.7}$$

The precise statement will be given in the next section.

With this statement, since obviously $(x-y)/\,|x-y|$ is a norm 1 vector, we can conclude that:

$$c_0\,|x-y|^2 \leq \int_{A_{1/2,2}(0)} |\langle \nabla u(z), x - y\rangle|^2 \leq c\left[W(x,1) + W(y,1)\right]\,, \tag{6.8}$$

and so we have

$$|x-y|^2 \leq cc_0^{-1}\left[W(x,1) + W(y,1)\right]\ . \tag{6.9}$$

Thus, as desired, we have obtained an effective bound on the distance between x and y as a function of the energy density.

In general, we need to develop this idea also in higher dimension, which is not totally trivial. However, the technical difficulties of this part are far less challenging than in the previous sections.

6.2 Estimates in higher dimension

Now we move to higher dimensions and we prove with all details the estimates in full generality. For convenience, we set by definition

$$W_\alpha(x) \equiv W_{r_\alpha, r_{\alpha-3}}(x) \equiv \theta_{r_{\alpha-3}}(x) - \theta_{r_\alpha}(x) \geq 0\,, \tag{6.10}$$

where $r_\alpha = 2^{-\alpha}$.

First of all, we state the main theorem of this section.

Theorem 6.1 (L^2-Best Approximation Theorem)**.** *Let $f : B_{9r}(p) \subseteq \mathbb{R}^m \to N$ be a stationary harmonic map with $(9r)^{2-n} \int_{B_{9r}(p)} |\nabla f|^2 \leq \Lambda$, and let $8r = 2^{-\bar{\alpha}}$. There exist $\delta(n, N\Lambda)$ and $C(n, N, \Lambda) > 0$ such that the following is true. Suppose that for some $p' \in B_r(p)$,*

$$W_{\bar{\alpha}}(p') \leq \delta(n, N, \Lambda)\,, \tag{6.11}$$

and suppose that $B_r(p)$ contains some singular points. Then for any finite measure μ on $B_r(p)$ we have that

$$D^{m-3}_{\mu}(p, r) \leq \inf_{L^{m-3}} r^{-2-(m-3)} \int_{B_r(p)} d^2(x, L^k)\, d\mu(x) \leq C r^{-(m-3)} \int_{B_r(p)} W_{\bar{\alpha}}(x)\, d\mu(x)\,, \tag{6.12}$$

where the inf *is taken over all $(m-3)$-dimensional affine subspaces $L^{m-3} \subseteq \mathbb{R}^m$.*

Remark 6.1. It is easy to realize that the statement of this theorem is scale-invariant wrt r. Thus we will prove the statement for $r = 1$ without loss of generality.

6.3 Symmetry and Gradient Bounds

Here we prove the dichotomy mentioned before. In particuar, we want to show that if u is a minimizing harmonic map with at least one "pinched point" in the ball of radius 1, then either the energy relative to some $m-3$-dimensional subspace has some size, or there are no singular points in this ball. More precisely:

Lemma 6.2. *Let $u : B_9(0) \subseteq M \to N$ be a stationary harmonic map $⨍_{B_9(0)} |\nabla u|^2 \leq \Lambda$. Then there exists a $\delta(n, N, \Lambda) > 0$ such that if for some $p \in B_1(0)$ we have*

$$\theta(x, 8) - \theta(x, 1) \leq \delta\,, \tag{6.13}$$

then either $r_u(0) \geq 1$ (and so in particular $B_1(0)$ does not contain any singular point), or for all $m-2$ dimensional subspaces L we have

$$⨍_{A_{3,4}(0)} |\langle \nabla u, L\rangle|^2 \geq \delta\,. \tag{6.14}$$

Remark 6.2. Note that we use the notation $\langle \nabla u, L\rangle$ to replace

$$|\langle \nabla u, L\rangle|^2 = \sum_{i=1}^{m-3} |\langle \nabla u, v_i\rangle|^2\ , \tag{6.15}$$

where $\{v_i\}_{i=1}^{m-2}$ is an orthonormal basis for L. In other words, $|\langle \nabla u, L\rangle|^2$ is the energy in the direction of L.

Proof. For this proof, we don't need a quantitative estimate on δ, we just need to show its existence. Thus, we carry out this proof by contradiction.

Thus, for m, N, Λ fixed, let by contradiction u_i be a sequence of minimizing harmonic maps $u_i : B_9(0) \to N$ such that

1. for some $p_i \in B_1(0)$, we have $W_0(p_i) \leq i^{-1}$,
2. for some $m-2$-dimensional space L_i we have $\fint_{A_{3,4}(0)} |\langle \nabla u_i, L_i \rangle|^2 \leq i^{-1}$,
3. we have $r_u(0) < 1$.

Up to rotations, we can assume that $L_i = L$ is fixed, and up to passing to a subsequence we have $p_i \to p \in \overline{B_1(0)}$. Also, after passing to subsequences we have that

$$u_i \longrightarrow u : B_9(0) \to N\,, \tag{6.16}$$

where the convergence is strong in $H^1(B_8(()0))$ since u_i are minimizing maps. The energy in the L direction in the limit is going to be

$$\int_{A_{3,4}(0)} |\langle \nabla u, L \rangle|^2 = 0\,, \tag{6.17}$$

and the pinching condition implies that

$$W_0(p) = 0\,. \tag{6.18}$$

This and unique continuation show that u is 0-symmetric wrt p, and invariant wrt L. Thus u is an $m-2$-symmetric map, which in turn implies that $u = 0$. Since the convergence is strong, we arrive at a contradiction with $r_u(0) < 1$. □

6.4 Eigenvalue and eigenvectors of the matrix associated to a measure

Let us consider a probability measure μ with support in $B_1(0)$, and let x_{cm} be its center of mass, i.e.:

$$x_{cm} = x_{cm}(\mu) \equiv \int x\, d\mu(x)\,. \tag{6.19}$$

Consider the bilinear quadratic form $Q(v, w)$ defined by

$$Q(v, w) \equiv \int [(x - x_{cm}) \cdot v]\, [(x - x_{cm}) \cdot w]\; d\mu(x)\,. \tag{6.20}$$

In this section, we study the eigenvalue and eigenvectors of Q and their relations with the β_2 defined above.

Definition 6.3. Given a probability measure $\mu \in B_1(0)$, we set $\lambda_1(\mu), \cdots, \lambda_m(\mu)$ to be the eigenvalues of $Q(\mu)$ in decreasing order, and $v_1(\mu), \cdots, v_m(\mu)$ to be its eigenvectors. In case one eigenvalue has higher multiplicity, we take any choice of orthonormal eigenvectors inside the eigenspace.

Note that by definition of eigenvectors, we have

$$Q(v_k) = \lambda_k v_k = \int [(x - x_{cm}) \cdot v_k]\,(x - x_{cm})\,d\mu(x) \tag{6.21}$$

We also have a variational characterization of the eigenvalues given by

$$\lambda_1 = \lambda_1(\mu) \equiv \max_{|v|^2=1} \int |(x - x_{cm}) \cdot v|^2 \; d\mu(x)\,. \tag{6.22}$$

and $v_1 = v_1(\mu)$ is any of the norm 1 vectors obtaining this maximum. By induction, we also have

$$\lambda_{k+1} = \lambda_{k+1}(\mu) \equiv \max \left\{ \int |(x - x_{cm}) \cdot v|^2 \; d\mu(x) \;\; s.t. \;\; |v|^2 = 1 \text{ and } \forall i \le k\,,\; v \cdot v_i = 0 \right\}\,, \tag{6.23}$$

and v_{k+1} is a vector obtaining this maximum. Note that, by definition of v_k, the subspace $V_k = x_{cm} + \text{span}\{v_1, \ldots, v_k\}$ is the k-dimensional affine subspace (or one of the subspaces) achieving the maximum of

$$\max_{V \subseteq \mathbb{R}^m,\; \dim(V)=k} \int \|\pi_V(x - x_{cm})\|^2 \; d\mu(x) = \lambda_1(\mu) + \cdots + \lambda_k(\mu)\,. \tag{6.24}$$

Since we have

$$\|x - x_{cm}\|^2 = \|\pi_V(x - x_{cm})\|^2 + \|\pi_{V^\perp}(x - x_{cm})\|^2 = \|\pi_V(x - x_{cm})\|^2 + d(x, V)^2\,, \tag{6.25}$$

we can rephrase (6.24) and say that V_k is the subspace minimizing the distance

$$\min_{V \subseteq \mathbb{R}^m,\; \dim(V)=k} \int d^2(x, V)\,d\mu(x) = \min_{V \subseteq \mathbb{R}^m,\; \dim(V)=k} \left(\int \|x - x_{cm}\|^2 \; d\mu(x) - \int \|\pi_V(x - x_{cm})\|^2 \; d\mu(x) \right) = \tag{6.26}$$

$$\int d^2(x, V_k)\,d\mu(x) = \lambda_{k+1}(\mu) + \cdots + \lambda_m(\mu)\,. \tag{6.27}$$

Remark 6.3. Note that evidently V_k must pass through the center of mass of μ. This is an immediate corollary of Jensen's inequality (or Steiner's theorem).

By simple manipulations with λ_k and v_k, we obtain the following important estimate:

Proposition 6.4. *Let $u : B_9(0) \to N$ be a minimizing harmonic map, and let μ be a probability measure on $B_1(0)$ with $\lambda_k(\mu), v_k(\mu)$ defined as in Definition 6.3. Then there exists $C(m) > 0$ such that*

$$\lambda_k \int_{A_{3,4}(0)} |v_k \cdot \nabla u(z)|^2 \; dVol(z) \le C(m) \int W_0(x)\,d\mu(x)\,. \tag{6.28}$$

Proof. For simplicity, and without essential loss of generality, we assume that $x_{cm}(\mu) = 0$ (otherwise a simple translation will do the trick).

For any $z \in A_{3,4}$ and $k = 1, \cdots, m$, we take the scalar product of (6.21) with $\nabla u(z)$, and obtain

$$\lambda_k\,(v_k \cdot \nabla u(z)) = \int (x \cdot v_k)\,(\nabla u(z) \cdot x) \; d\mu(x)\,, \tag{6.29}$$

By definition of center of mass (see (6.19)), we have for all fixed z:

$$\int x \cdot z \, d\mu(x) = x_{cm} \cdot z = 0 \,. \tag{6.30}$$

Thus we can re-write (6.29) in the form:

$$\lambda_k \left(\nabla u(z) \cdot v_k\right) = \int (x \cdot v_k) \left[\nabla u(z) \cdot (x - z)\right] \, d\mu(x) \,. \tag{6.31}$$

A simple application of Hölder inequality tells us that for all fixed z:

$$\lambda_k^2 \left|\nabla u(z) \cdot v_k\right|^2 \leq \lambda_k \int \left|\nabla u(z) \cdot (x - z)\right|^2 \, d\mu(x) \,. \tag{6.32}$$

Note that we can evidently assume $\lambda_k > 0$, otherwise there's nothing to prove. By integrating both sides of the previous inequality on $A_{3,4}(0)$, we get

$$\lambda_k \int_{A_{3,4}(0)} \left|\nabla u(z) \cdot v_k\right|^2 \, d\mathrm{Vol}(z) \leq \int \int_{A_{3,4}(0)} \left|\nabla u(z) \cdot (x - z)\right|^2 \, d\mathrm{Vol}(z) \, d\mu(x) \leq \tag{6.33}$$

$$\leq \int \int_{A_{3,4}(0)} \frac{\left|\nabla u(z) \cdot (x - z)\right|^2}{|x - z|^m} \left|x - z\right|^m \, d\mathrm{Vol}(z) \, d\mu(x) \leq \tag{6.34}$$

$$\leq C(m) \int \int_{A_{1,8}(x)} \frac{\left|\nabla u(z) \cdot (x - z)\right|^2}{|x - z|^m} \, d\mathrm{Vol}(z) \, d\mu(x) \leq C(m) \int W_0(x) \, d\mu(x) \,, \tag{6.35}$$

as desired. □

6.5 Proof of Theorem 6.1

We are now in a position to prove Theorem 6.1. By rescaling, we can assume for convenience that $\mu(B_1(0)) = 1$. Since we have ordered λ_k to be decreasing in value, and by (6.26), we have

$$D^k_\mu(0, 1) = \lambda_{k+1}(\mu) + \cdots + \lambda_m(\mu) \leq (m - k)\lambda_{k+1} \,. \tag{6.36}$$

By applying Proposition 6.4 to each $j = 1, \cdots, k + 1$, we obtain

$$\sum_{j=1}^{k+1} \lambda_j \int_{A_{3,4}(0)} \left|\nabla u(z) \cdot v_j\right|^2 \, d\mathrm{Vol}(z) \leq (k + 1)C \int W_0(x) \, d\mu(x) \,. \tag{6.37}$$

Let $V^{k+1} = \mathrm{span}\,(v_1, \cdots, v_{k+1})$ be the linear part of the best $k + 1$-dimensional subspace of μ. Given that λ_j are decreasing in j (by definition), the last estimate leads to

$$\lambda_{k+1} \int_{A_{3,4}(0)} \left|V^{k+1} \cdot \nabla u(z)\right|^2 \, d\mathrm{Vol}(z) = \lambda_{k+1} \sum_{j=1}^{k+1} \int_{A_{3,4}(p)} \left|\nabla u(z) \cdot v_j\right|^2 \, d\mathrm{Vol}(z) \leq C \int W_0(x) \, d\mu(x) \,. \tag{6.38}$$

Since we assumed that there exists at least one pinched point in the ball and no singular points, by Lemma 6.2,

$$\int_{A_{3,4}(p)} \left|\nabla u(z) \cdot V^{m-2}\right|^2 \, d\mathrm{Vol}(z) \geq 8^{m-2}\delta(m, N, \Lambda) \,. \tag{6.39}$$

This allows us to estimate

$$c(m)\delta\lambda_{m-2} \leq \lambda_{m-2} \int_{A_{3,4}(0)} \left|\nabla u(z) \cdot V^{m-2}\right|^2 \, d\mathrm{Vol}(z) \leq C \int W_0(x)\, d\mu(x)\,. \tag{6.40}$$

Since δ is a positive constant depending only on (m, Λ, N), and by (6.36), we can conclude

$$D_\mu^{m-3}(0,1) \leq C(m,\Lambda,N) \int W_0(x)\, d\mu(x) \tag{6.41}$$

as desired.

□

7 Covering arguments

Now we are ready for the last part of our estimates. There are two issues that need to be addressed. First of all, there is the weird pinching assumption (6.11) that seems somewhat problematic to guarantee. In order to deal with this issue, the very rough intuition for the idea is the following: take $\mathcal{S} \subseteq B_1(0)$, and cover this set by balls $B_{r_i}(x_i)$ centered in $\mathcal{S}$ and such that

$$\theta(x_i, 1) - \theta(x_i, r_i) = \delta\,. \tag{7.1}$$

Now apply the discrete Reifenberg to the collection $\{B_{r_i}(x_i)\}_i$. Now every time you need to use theorem 6.1, the hypothesis of having a pinched point in the ball is guaranteed. Now start over the process on each ball $\{B_{r_i}(x_i)\}_i$ separately, and do the same kind of covering again. Every time you do one of this covering, you ensure that θ drops of a δ amount. Since θ is monotone, and $\theta(0,1) \leq \Lambda$, you will only need to do this some $\Lambda\delta^{-1}$ number of times, and this does not mess with the final estimates.

Needless to say, this is just an extremely rough and imprecise argument, but it might be of help in guiding through the next section.

The second issue is the following. If we want to apply the discrete Reifenberg, we need to know that

$$\int_{B_r(x)} \left(\int_0^r D_\mu^{m-3}(y,s)\,\frac{ds}{s} \right) d\mu(y) < \delta^2 r^k\,, \tag{7.2}$$

on all balls contained in $B_1(0)$. Theorem 6.1 tells us that

$$D_\mu^{m-3}(p,r) \leq C r^{-(m-3)} \int_{B_r(p)} W_s(x)\,d\mu(x)\,, \tag{7.3}$$

so we know that

$$\int_{B_r(x)} \left(\int_0^r D_\mu^{m-3}(y,s)\,\frac{ds}{s} \right) \leq C \int_0^r \left[s^{-(m-3)} \int_{B_r(x)} \left(\int_{B_s(y)} W_s(z)\,d\mu(z) \right) d\mu(y) \right] \frac{ds}{s}\,. \tag{7.4}$$

We would like to take the two integrals inside and turn them into a single integral. In particular, we would like to write something like

$$s^{-(m-3)} \int_{B_r(x)} \left(\int_{B_s(y)} W_s(z)\,d\mu(z) \right) d\mu(y) \leq C \int_{B_{r+s}(x)} W_s(y)\,d\mu(y)\,. \tag{7.5}$$

What we can do is the following. First of all, we exchange the order of integration. Since

$$\{(y,z) \text{ s.t. } \|y-x\| \leq r \text{ and } \|z-y\| \leq s\} \subset \{(y,z) \text{ s.t. } \|z-x\| \leq r+s \text{ and } \|y-z\| \leq s\}\,, \tag{7.6}$$

we can estimate

$$s^{-(m-3)} \int_{B_r(x)} \left(\int_{B_s(y)} W_s(z)\,d\mu(z) \right) d\mu(y) \leq s^{-(m-3)} \int_{B_{r+s}(x)} W_s(z) \left(\int_{B_s(y)} d\mu(y) \right) d\mu(z)\,. \tag{7.7}$$

Now IF we knew that

$$\left(\int_{B_s(y)} d\mu(y)\right) \leq C s^{m-3} \tag{7.8}$$

for some universal constant C, then we would have

$$\int_{B_r(x)} \left(\int_0^r D_\mu^{m-3}(y,s)\,\frac{ds}{s}\right) \leq C \int_{B_{r+s}(x)} \left[\int_0^r W_s(z)\frac{ds}{s}\right] d\mu(z)\,. \tag{7.9}$$

Now the integral in the middle is sort of a telescopic sum, since

$$\int_0^r W_s(z)\frac{ds}{s} \leq C\left(\theta(z,r)-\theta(0,0)\right)\,, \tag{7.10}$$

and thus maybe we can play some tricks with this estimate.

However, we still have the "IF" in (7.8) to take care of, and it looks like this claim is exactly what we want to prove in the end: uniform measure bounds on μ. And actually it is.

The point is the following: we are going to prove this claim inductively on scales going upwards, and at each step in the induction use the previous step to have ROUGH uniform bounds on the measure μ, and then exploit the Reifenberg theorem to refine these bounds to uniform bounds not depending on the induction step.

These are the main stumbling blocks, but as always there's also tons of other technicalities to be dealt with. We start with some extra lemmas needed in the proof of the covering arguments, then we move to the covering arguments, and we close with the proof of the main theorem.

7.1 Technical lemmas

Here we collect two technical lemmas needed in the following. Their proof can be carried out with an adaptation of the usual contradiction/compactness argument.

Proposition 7.1. *Let $\rho,\epsilon>0$ be fixed. There exists $\delta_3(m,\Lambda,\rho,\epsilon)>0$ such that the following holds. Let $F=\{y\in B_2(0)\ \text{s.t.}\ \theta(y,1)-\theta(y,\rho)<\delta_3\}$. If F ρ-effectively spans an $(m-3)$-dimensional subspace V, then*

$$\mathcal{S}\subseteq B_{2\rho}(V)\,. \tag{7.11}$$

Proof. EXPAND ☐

Lemma 7.2. *Let $u: B_3(0)\to N$ be a minimizing harmonic map. Let $\rho>0$ and $\eta>0$ be fixed, and assume that for all $y\in B_1(0)$, $\theta(y,1)\leq E$, then there exists $\delta_6=\delta_6(m,\Lambda,,\rho,\gamma,\eta)$ such that the following holds. Let $F(u,\delta_6)\subseteq\{y\in B_1(0)\ \text{s.t.}\ \theta(y,\rho)>E-\delta_6\}$. If F ρ-effectively spans a k-dimensional subspace L, then for all $x\in L\cap B_2(0)$, we have*

$$\theta(x,\rho)>E-\eta\,. \tag{7.12}$$

Moreover, if $k\geq m-2$, then $E\leq\eta$.

Proof. EXPAND ☐

7.2 Covering argument

In this subsection, we prove the inductive covering argument needed for the main theorem. We split this covering argument into two lemmas: in the first one, we keep refining inductively a covering by balls until all but a controlled amount of points in our balls have some definite drop in θ, and in the second one we show that this controlled amount of points without drop is small so that they can be "ignored".

Lemma 7.3 (Covering Lemma I)**.** *Let* $u : B_3(0) \subset \mathbb{R}^m \to N$ *be an minimizing harmonic map. Fix any* $0 < \rho < \rho(m) \leq 100^{-1}$*, and* $0 < r < R$*,* $0 < R \leq 1$ *arbitrary, set* $E = \sup_{x \in B_R(0) \cap \mathcal{S}} \theta(x, 1)$*, and assume the uniform bound* $E \leq \Lambda$*. There exists* $\delta = \delta(m, \Lambda, \rho) > 0$ *and* $C_V(m)$ *such that the following is true.*

For any subset $\mathcal{S} \subseteq \mathcal{S}(u)$ *there exists a finite covering of* $\mathcal{S} \cap B_R(0)$ *such that*

$$\mathcal{S} \cap B_R(0) \subseteq \bigcup_{x \in \mathcal{C}} B_{r_x}(x) \quad \text{with} \quad r_x \geq r \quad \text{and} \quad \sum_{x \in \mathcal{C}} r_x^k \leq C_V(m) R^k . \tag{7.13}$$

Moreover, for each $x \in \mathcal{C}$*, one of the following is verified*

1. $r_x = r$
2. *the set of points* $F_x \equiv \{y \in \mathcal{S} \cap B_{2r_x}(x) \ \ s.t.\ \ \theta(y, \rho r_x/10) > E - \delta\}$ *is contained in* $B_{\rho r_x/5}(L_x) \cap B_{2r_x}(x)$*, where* L_x *is some* $k - 1$ *dimensional affine subspace.*

Remark 7.1. By the scale-invariance properties of θ, this statement is equivalent to its version with $R = 1$ fixed.

Remark 7.2. Note that the set F_x may be empty.

Remark 7.3. For convenience, and without any loss of generality, we will assume in the proof that r is some (positive) power of ρ, and that ρ is some (negative) power of 2. In particular:

$$r = \rho^{\bar{j}} \quad \text{and} \quad \rho = 2^{-a} , \quad \text{with} \quad a, \bar{j} \in \mathbb{N} . \tag{7.14}$$

7.2.1 Proof of Lemma 7.3

The idea of the proof is the following. We are going to build inductively on i a covering of the set $\mathcal{S}$ by a family of balls of radius $r_i = \rho^i$. In the inductive step, we will look at each ball of radius r_i and determine if this is a "good" or a "bad" ball according to how many points inside this ball have $\theta(y, \rho r_i/10) \geq E - \delta$.

If this set of points "effectively span" some $(m - 3)$-dimensional affine subspace V, then we will apply Lemma 7.1 in order to see that the whole set $\mathcal{S} \cap B_{r_i}(x)$ is contained in small neighborhood of V. Moreover, using Lemma 7.2, we will see that we can cover the whole neighborhood of V by balls with uniform radius, and this covering will satisfy the assumptions of the discrete Reifenberg theorem. These balls are the good balls.

If this set of points is empty, or it does not "effectively span" something k-dimensional, then we will stop refining our covering, because by definition condition (2) is verified.

The uniform k-dimensional content estimates will follow from the discrete Reifenberg theorem 5.4 applied to the natural measure associated with this covering. The $D_\mu^{(m-3)}$ estimates needed to apply the Reifenberg theorem are a consequence of Section 6.

7.2.2 Inductive covering: first step

Consider the map $u : B_3(0) \to N$, let $\mathcal{S} \subseteq \mathcal{S}(u)$ be an arbitrary subset and define the set

$$F = \{y \in B_2(0) \cap \mathcal{S} \ \ s.t. \ \ \theta(y, \rho/10) > E - \delta\} \ . \tag{7.15}$$

If there exists a $k-1$-dimensional subspace L such that $F \subset B_{\rho/5}(L)$, then there's nothing to prove. In this case, we call $B_1(0)$ a *bad* ball.

Otherwise, we say that $B_1(0)$ is a *good* ball. In this second case, let V be a k-dimensional subspace which is $(\rho/10)$-effectively spanned by the set F. Thus by definition there exists $\left\{y_j\right\}_{j=0}^{k} \subset F$ that $(\rho/10)$-effectively span V. For δ sufficiently small, we can apply Lemma 7.1 to $B_1(0)$, we obtain that

$$\mathcal{S}(u) \cap B_1(0) \subset B_{\rho/5}(V) \ . \tag{7.16}$$

Consider a finite covering of $B_{\rho/5}(V) \cap B_1$ by balls $\left\{B_\rho(x)\right\}_{x\in\mathcal{C}}$ such that

1. $x \in V \cap B_1(0)$
2. if $x \neq y$, then $B_{\rho/5}(x) \cap B_{\rho/5}(y) = \emptyset$

Note that, by Lemma 7.2, we have for all $x \in \mathcal{C}$:

$$\theta(x, \rho/10) \geq E - \eta \,, \tag{7.17}$$

as long as δ is sufficiently small. Moreover, if $B_\rho(x)$ contains no singular point, we can simply discard it from our covering. Thus we can assume wlog that $\mathcal{S}(u) \cap B_\rho(x) \neq \emptyset$. This will be needed to use the theorem 6.1.

This completes the base step of the inductive covering we will be constructing in the next subsection. Now we will consider any of the balls $B_\rho(x)$ in this covering and start over the process.

7.2.3 Inductive step

We will build by induction a sequence of coverings

$$\mathcal{S} \subseteq \bigcup_{x\in\mathcal{C}^j} B_{r_x^j}(x) = \bigcup_{x\in\mathcal{C}_b^j} B_{r_x^j}(x) \cup \bigcup_{x\in\mathcal{C}_g^j} B_{r_x^j}(x) \equiv B_{r_x^j}\left(\mathcal{C}_b^j\right) \cup B_{r_x^j}\left(\mathcal{C}_g^j\right) \,, \tag{7.18}$$

where $\mathcal{C}_b^j$ will represent the centers of a collection of "bad balls" and $\mathcal{C}_g^j$ will represent the centers of a collection of "good balls" such that

1. If $x \in \mathcal{C}_b^j$ then $r_x^j \geq \rho^j$ and the set $F_x = \left\{y \in \mathcal{S} \cap B_{2r_x^j}(x) \ \ s.t. \ \ \theta(y, \rho r_x^j/10) \geq E - \delta\right\}$ is contained in some $B_{\rho r_x^j/5}(L_x)$, where L_x is a $k-1$-dimensional affine subspace.
2. If $x \in \mathcal{C}_g^j$ then $r_x^j \equiv \rho^j$ and the set $F_x = \left\{y \in \mathcal{S} \cap B_{2r_x^j}(x) \ \ s.t. \ \ \theta(y, \rho r_x^j/10) \geq E - \delta\right\}$ $(\rho r_x^j/10)$-effectively spans a k-dimensional affine subspace V_x.

3. For all $x \neq y \in \mathcal{C}^j$ we have $B_{r_x/5}(x) \cap B_{r_y/5}(y) = \emptyset$.
4. For all $x \in \mathcal{C}^j$ we have $\theta(x, r_x) \geq E - \eta$.
5. For all $x \in \mathcal{C}^j$, $\mathcal{S} \cap B_{r_x}(x) \neq \emptyset$.

Suppose that we have this covering for some j, and consider the set

$$R_j = \mathcal{S} \setminus \bigcup_{x \in \mathcal{C}_b^j} B_{r_x}(x) = \mathcal{S} \setminus B_{r_x}\left(\mathcal{C}_b^j\right) . \tag{7.19}$$

Note that by definition this set is contained in $B_{\rho^j}\left(\mathcal{C}_g^j\right)$. For each $x \in \mathcal{C}_g^j$, we know that F_x $[\rho^{j+1}/10]$-effectively spans a k-dimensional subspace V_x. As seen in the first inductive step, by Proposition 7.1 we have that

$$\mathcal{S}(u) \cap B_{2\rho^j}(x) \subset B_{\rho^{j+1}/5}(V_x) \tag{7.20}$$

for all $x \in \mathcal{C}_g^j$ as long as

$$\delta \leq \delta_3(m, \Lambda, , \gamma, \rho, \epsilon) . \tag{7.21}$$

In order to build an open covering of R_j, consider the set

$$A = \bigcup_{x \in \mathcal{C}_g^j} \left(B_{\rho^j}(x) \cap V_x\right) \setminus B_{r_x/2}\left(\mathcal{C}_b^j\right) . \tag{7.22}$$

By (7.19) and (7.20), and since $\rho < 100^{-1}$, we have

$$R_j \subseteq B_{\rho^{j+1}/5}(A) . \tag{7.23}$$

Now first note that by the definition of A and since $\rho \leq 100^{-1}$, all of these balls are disjoint from $B_{r_x/10}\left(\mathcal{C}_b^j\right)$. Moreover, by Lemma 7.2, if we choose δ sufficiently small, for all $y \in A$ we have

$$\theta\left(y, \rho^{j+1}/10\right) \geq E - \eta . \tag{7.24}$$

In particular, we need

$$0 < \delta \leq \delta_6(m, \Lambda, , \rho, \gamma, \eta) . \tag{7.25}$$

Now consider a (finite) Vitali subcovering of this set given by

$$R_j \subseteq \bigcup_{x \in \mathcal{C}^A} B_{\rho^{j+1}}(x) . \tag{7.26}$$

First of all, we discard from this covering all the balls that have empty intersection with the set $\mathcal{S} \subseteq \mathcal{S}(u)$. Since we want to cover this set, this is a reasonable thing to do.

We can classify all the balls in this covering into good and bad according to how spread their set F is. In particular, for all $x \in \mathcal{C}^A$ consider as above the set

$$F_x = \left\{ y \in \mathcal{S} \cap B_{2\rho^{j+1}}(x) \quad s.t. \;\; \theta\left(y, \rho^{j+2}/10\right) \geq E - \delta \right\} . \tag{7.27}$$

If F_x $[\rho^{j+2}/10]$-effectively spans a k dimensional subspace V_x, then we say that $B_{\rho^{j+1}}(x_t)$ is a good ball, and we put $x \in \mathcal{C}_g^A$. Otherwise, we say that $B_{\rho^{j+1}}(x)$ is a bad ball, and we put $x \in \mathcal{C}_b^A$.

We define

$$\mathcal{C}_b^{j+1} = \mathcal{C}_b^j \cup \mathcal{C}_b^A , \quad \mathcal{C}_g^{j+1} = \mathcal{C}_g^A . \tag{7.28}$$

Note that the set of bad balls contains *all* the bad balls encountered at any previous step. On the contrary, good balls get refined at each stage, and at each induction step the previous bad balls disappear from the set $\mathcal{C}_g$.

Now the induction is complete. Indeed, property 1 and 2 are a direct consequence of the definition of $\mathcal{C}_g$ and $\mathcal{C}_b$. Property 3 comes from the definition of A and the Vitali covering lemma. Finally, property 4 is a consequence of (7.24) and property 5 comes from the fact that we discard balls that have empty intersection with $\mathcal{S}$.

7.2.4 Volume estimates

Now we are in a position to prove the desired volume estimates, and in particular

$$\sum_{x \in \mathcal{C}} r_x^k \leq C_V(m) , \tag{7.29}$$

where $\mathcal{C} = \mathcal{C}^{\bar{j}}$ for $\bar{j}$ such that $\rho^{\bar{j}} = r$.

We will prove this estimate by an induction on the radius. For convenience, we define the measure

$$\mu = \omega_k \sum_{x \in \mathcal{C}} r_x^k \delta_x . \tag{7.30}$$

Upwards induction For all $t \in (0, 1]$, set $\mathcal{C}_t = \{x \in \mathcal{C} \;\; s.t. \;\; r_x \leq t\}$, and define the measure

$$\mu_t \equiv \omega_k \sum_{x \in \mathcal{C}_t} r_x^k \delta_x \leq \mu . \tag{7.31}$$

Now we want to prove inductively on $t = r, 2r, 2^2 r, 2^3 r, \cdots, 1/8$ that for some universal constant $C_R(m)$, for all $x \in B_3(0)$ and $s \geq r$ we have

$$\mu_t(B_t(x)) \equiv \left(\sum_{x \in \mathcal{C} \;\, s.t. \;\, r_x \leq t} \omega_k r_x^k \delta_x \right) (B_t(x)) \leq C_R(m) t^k . \tag{7.32}$$

Note that $C_R(m)$ is the constant in Theorem 5.4. Note also that $\mu_1 = \mu$, so at the last step of the induction we will have recovered an estimate for the whole μ, up to a covering of $B_1(0)$ by balls $B_{1/8}(p_i)$. In other words, we prove (7.29) with

$$C_V(m) = c(m)C_R(m)\,. \tag{7.33}$$

Note that the base step is easily seen to be true for $t = r$. Indeed, at this stage we have

$$\mu_r = \sum_{x\in\mathcal{C}_r} \omega_k r^k \delta_x\,, \tag{7.34}$$

where all $B_{r/5}(x_i)$ are disjoint. Thus we immediately have $\mu_r(B_r(x)) \le c(m)r^k$.

Now, suppose that we have proven (7.32) for $t \le 2^j r$, we will show that (7.32) holds also for $t = 2^{j+1}r$.

Rough estimate First of all, we note that by a very bad estimate we have for all $x \in B_1(0)$:

$$\mu_{2\bar r}(B_{2\bar r}(x)) \le c(m)C_R(m)(2\bar r)^k\,, \tag{7.35}$$

where for convenience we have set $\bar r = 2^j r$. Indeed, we can split $\mu_{2\bar r}$ into

$$\mu_{2\bar r} = \mu_{\bar r} + \tilde\mu_{2\bar r} \equiv \sum_{x\in\mathcal{C}_{\bar r}} \omega_k r_x^k \delta_x + \sum_{x\in\mathcal{C}\ s.t.\ r_x\in(\bar r, 2\bar r]} \omega_k r_x^k \delta_x\,. \tag{7.36}$$

Take a covering of $B_{2\bar r}(x)$ by balls $B_{\bar r}(y_i)$ such that $B_{\bar r/2}(y_i)$ are disjoint. The number of these balls has a universal bound $c(m)$, and by induction we have

$$\mu_{\bar r}(B_{2\bar r}(x)) \le \sum_i \mu_{\bar r}(B_{\bar r}(y_i)) \le c(m)C_R(m)\bar r^k\,. \tag{7.37}$$

As for the other part of μ, by definition of this measure all the balls $B_{r_x/5}(x)$ are pairwise disjoint, and so we get immediately

$$\tilde\mu_{2\bar r}(B_{2\bar r}(x)) \le c(m)(2\bar r)^k\,. \tag{7.38}$$

Reifenberg estimates We will show inductively that we can apply Theorem 5.4 to the measures $\mu_{2\bar r}$ on each fixed $B_{2\bar r}(x)$. For convenience, we set

$$\bar\mu = \mu_{2\bar r}|_{B_{2\bar r}(x)}\,. \tag{7.39}$$

Note that for all $x \in \operatorname{supp}(\mu)$, and all $s \in [r_x, 1]$, we have $\theta(x, s) - \theta(x, s/2) < \eta$ because $\theta(x_i, s) \le E$ by monotonicity of θ and by definition of E, and $\theta(x_i, s/2) \ge E - \eta$ by condition (4) of our constructed covering. Now we can choose η small enough so that for all $x \in \operatorname{supp}(\mu)$ and $0 < s \le 1$ we have the $D_\mu^{(m-3)}$ estimate

$$\beta_{2,\bar\mu}(x, s)^2 \overset{thm.6.1}{\le} C_1 s^{-k} \int_{B_s(x)} W_s(y)\, d\bar\mu(y)\,, \tag{7.40}$$

where we have set for all $x \in \operatorname{supp}(\mu)$:

$$W_s(x) = \begin{cases} W_s(x_i) & \text{if } s > r_x\,, \\ 0 & \text{if } s \le r_x\,. \end{cases} \tag{7.41}$$

Indeed, for $s \le r_x$, $\operatorname{supp}(\mu) \cap B_s(x) = \{x\}$, and there's nothing to prove. If $s \ge r_x$, then for all $y \in B_{r_x}(x)$, $r_y < s$ by construction of μ.

Recall that in all of the good balls, there exists at least a singular point. Thus we can apply Theorem 6.1, and we have the estimate (7.40) as desired.

Now we can prove that for all $y \in B_{2\bar r}(x)$, and $r \le 2\bar r$, we have

$$\int_{B_r(y)} \left(\int_0^r \beta^k_{2,\bar\mu}(z,s)^2 \, \frac{ds}{s} \right) d\bar\mu(z) < c(m) C_1 C_R^2 \eta r^k\,. \tag{7.42}$$

Indeed, by (7.40) we can estimate for all $s \le r$:

$$\int_{B_r(y)} \beta^k_{2,\bar\mu}(z,s)^2 \, d\bar\mu(z) \le C_1 s^{-k} \int_{B_r(y)} \left[\int_{B_s(z)} W_s(t) \, d\bar\mu(t) \right] d\bar\mu(z)\,. \tag{7.43}$$

Now, on $B_s(z)$, either $\bar\mu = \mu_s|_{B_{2\bar r}(x)}$, or there exists an $x \in \operatorname{supp}(\mu) \cap B_s(z)$ with $r_x > s$. Since $z \in \operatorname{supp}(\mu)$ as well, by construction we have $z = x = \operatorname{supp}(\mu) \cap B_s(z)$, and $W_s(z) = 0$. Thus in either case we have

$$\int_{B_r(y)} \beta^k_{2,\bar\mu}(z,s)^2 \, d\bar\mu(z) \le C_1 s^{-k} \int_{B_r(y) \cap B_{2\bar r}(x)} \left[\int_{B_s(z) \cap B_{2\bar r}(x)} W_s(t) \, d\mu_s(t) \right] d\mu_s(z)\,. \tag{7.44}$$

By induction, and by the rough estimates in (7.35), for all $s \in (0, 2\bar r]$ and $z \in B_1(0)$ we can estimate

$$\mu_s(B_s(z)) \le c(m) C_R s^k\,. \tag{7.45}$$

Thus we obtain

$$\int_{B_r(y)} \beta^k_{2,\bar\mu}(z,s)^2 \, d\bar\mu(z) \le c(m) C_1 C_R \int_{B_{r+s}(y) \cap B_{2\bar r}(x)} W_s(z) d\mu_s(z) = c(m) C_1 C_R \int_{B_{r+s}(y)} W_s(z) d\bar\mu(z)\,. \tag{7.46}$$

This yields

$$\int_{B_r(y)} \left(\int_0^r \beta^k_{2,\bar\mu}(z,s)^2 \, \frac{ds}{s} \right) d\bar\mu(z) \le c(m) C_1 C_R \int_{B_{2r}(y)} \left[\int_0^r W_s(z) \frac{ds}{s} \right] d\bar\mu(z)\,. \tag{7.47}$$

Note that for all $x \in \operatorname{supp}(\mu)$ and $r \le 2\bar r \le 1/8$, we have

$$\int_0^r W_s(x) \frac{ds}{s} = \int_{r_x}^r W_s(x) \frac{ds}{s} \le \int_{r_x}^{1/8} W_s(x) \frac{ds}{s} \le c \left[\theta(x,1) - \theta(x, r_x) \right] \le c\eta\,. \tag{7.48}$$

Thus, using again the induction hypothesis and the rough estimates (7.35), we prove (7.42).

If we choose η small enough, in particular

$$\eta \le \eta_1(m,,\Lambda) = c(m) \frac{\delta_R^2}{C_1 C_R^2}\,, \tag{7.49}$$

we can apply Theorem 5.4 to $\bar{\mu}$ and obtain (7.32) as wanted.

The only thing left to do is to choose $\delta = \delta(m, , \Lambda, \gamma, \rho, \epsilon) > 0$ in such a way that (7.24) is satisfied with

$$\eta \leq \min \{\eta_0, \eta_1, \delta_7\} \tag{7.50}$$

and also (7.21) is satisfied. Given (7.49), as noted above this is a simple application of Lemmas 7.2. This finishes the proof of Lemma 7.3.

□

7.2.5 Second covering lemma

By repeating this covering argument over bad balls, we obtain the following

Lemma 7.4 (Covering Lemma II). *Let $u : B_3(0) \to N$ be an minimizing harmonic map (as in the previous covering lemma). Fix any $\epsilon > 0$ and $0 < r \leq R$, $0 < R \leq 1$, set $E = \sup_{x \in B_{2R}(0) \cap \mathcal{S}} \theta(x, 1)$, and assume the uniform bound $E \leq \Lambda$. There exists $\delta = \delta(m, , \Lambda) > 0$ and $C_F(m)$ such that the following is true.*

For any subset $\mathcal{S} \subseteq \mathcal{S}(u)$, there exists a finite covering of $\mathcal{S} \cap B_R(0)$ such that

$$\mathcal{S} \cap B_R(0) \subseteq \bigcup_{x \in \mathcal{C}} B_{r_x}(x), \quad \text{with} \quad r_x \geq r \quad \text{and} \quad \sum_{x \in \mathcal{C}} r_x^k \leq C_F(m) R^k. \tag{7.51}$$

Moreover, for each $x \in \mathcal{C}$,

1. *either $r_x = r$*
2. *or we have the following uniform energy drop*

$$\forall y \in B_{r_x}(x) \cap \mathcal{S}, \quad \theta(y, r_x/10) \leq E - \delta. \tag{7.52}$$

Remark 7.4. As for the previous covering lemma, also in this case we can assume for simplicity and wlog that $R = 1$.

Proof. We need to refine the covering of the previous lemma. Recall that by lemma 7.3 we have a covering of $\mathcal{S} \cap B_1(0)$ given by

$$\mathcal{S} \cap B_1(0) \subseteq \bigcup_{x \in \mathcal{C}} B_r(x) \equiv \bigcup_{x \in \mathcal{C}_r} B_r(x) \cup \bigcup_{x \in \mathcal{C}_+} B_{r_x}(x) \quad \text{with} \quad r_x \geq r \quad \text{and} \quad \sum_{x \in \mathcal{C}_r \cup \mathcal{C}_+} r_x^k \leq C_V(m), \tag{7.53}$$

where we have set

$$\mathcal{C}_r = \{x \in \mathcal{C} \ s.t. \ r_x = r\} \quad \text{and} \quad \mathcal{C}_+ = \{x \in \mathcal{C} \ s.t. \ r_x > r\}, \quad \mathcal{C} = \mathcal{C}_r \cup \mathcal{C}_+. \tag{7.54}$$

We will of course keep $\mathcal{C}_r$ as part of our final covering, while we will refine the covering on each of the balls $\{B_{r_x}(x)\}_{x \in \mathcal{C}_+}$ in an inductive way. By item (2) of lemma 7.3, for each $x \in \mathcal{C}_+$ the set $F_x \equiv \{y \in \mathcal{S} \cap B_{2r_x}(x) \ \ s.t. \ \theta(y, \rho r_i/10) > E - \delta\}$ is close to a $k-1$-dimensional space. Assuming that $F_x = \emptyset$, all we need to do in order to achieve (7.52) is to re-cover $B_{r_x}(x)$ with balls $\left\{B_{\rho r_x}(y)\right\}_{y\in\mathcal{C}_x^{(1,f)}}$. These balls are the final covering we are looking for. Evidently, the number of these balls is bounded by a constant $C_f(m,\rho)$.

If $F_x \neq \emptyset$, we need to exploit the fact that we still know $F_x \subseteq B_{\rho r_x/5}(L_x) \cap B_{2r_x}(x)$, where L_x is at most $k-1$ dimensional. Thus we can cover $B_{r_x}(x) \setminus B_{\rho r_x}(F_x)$ as above, and cover $B_{\rho r_x}(F_x)$ separately by balls $\left\{B_{\rho r_x}(y)\right\}_{y\in\mathcal{C}_x^{(1,b)}}$. On these "bad balls", we will not be able to obtain any information on the energy drop over these new balls in the covering. However, their k-dimensional content is small since F_x behaves like a $k-1$ dimensional set. This will allow us to start over on each of these bad balls separately, and keep a uniform k-dimensional estimate on the content of the final covering. More precisely:

7.2.6 Re-covering of bad balls: Induction

In detail, we will build by induction on i a sequence of coverings of $\mathcal{S} \subseteq \mathcal{S}(u) \cap B_1(0)$ such that

1. For all $i = 1, 2, \cdots$
$$\mathcal{S} \subseteq \bigcup_{x\in\mathcal{C}^{(i,r)}} B_r(x) \cup \bigcup_{x\in\mathcal{C}^{(i,f)}} B_{r_x}(x) \cup \bigcup_{x\in\mathcal{C}^{(i,b)}} B_{r_x}(x) \,. \tag{7.55}$$
2. For all $x \in \mathcal{C}^{(i,r)}$, $r_x = r$. In other words, on these "r-balls" option (1) of our lemma is verified,
3. For all $x \in \mathcal{C}^{(i,f)}$ and all $z \in B_{2r_x}(x)$ we have $\theta(z, r_x/10) \leq E - \delta$. In other words, on these "final balls" option (2) of our lemma is verified,
4. for all $x \in \mathcal{C}^{(i,b)}$, $r < r_x \leq \rho^i$. On these "bad balls", none of the two stopping options is verified, thus we need to refine our covering here.
5. For some constant $C_F(m)$, we have the estimates
$$\sum_{x\in\mathcal{C}^{(i,r)}\cup\mathcal{C}^{(i,f)}} r_x^k \leq C_F(m)\left(\sum_{j=1}^{i} 2^{-j}\right), \qquad \sum_{x\in\mathcal{C}^{(i,b)}} r_x^k \leq 2^{-i}\,. \tag{7.56}$$

   Thus the estimates on r and final balls has uniform bounds, while our estimates on bad balls has exponentially decreasing bounds.

7.2.7 Re-covering of bad balls: First step in the induction

For $i = 1$, consider the covering (7.53) given by the previous lemma. We keep the balls $\{B_{r_x}(x)\}_{x\in\mathcal{C}_r}$ as they are, while for each $x \in \mathcal{C}_+$ consider two coverings of $B_{\rho r_x}(F_x)$ and its complement
$$B_{r_x}(x) \setminus B_{\rho r_x}(F_x) \subseteq \bigcup_{y\in\mathcal{C}_x^{(1,f)}} B_{\rho r_x}(y)\,, \quad B_{r_x}(x) \cap B_{\rho r_x}(F_x) \subseteq \bigcup_{y\in\mathcal{C}_x^{(1,b)}} B_{\rho r_x}(y)\,, \tag{7.57}$$
where $B_{\rho r_x/2}(y)$ are pairwise disjoint in both coverings.

By definition of F_x, for all $y \in \mathcal{C}_x^{(1,f)}$ the energy drop condition (7.52) is satisfied. Moreover we have the trivial estimates

$$\sum_{y \in \mathcal{C}_x^{(1,f)}} (\rho r_x)^k = (\rho r_x)^k \, \# \left\{ y \in \mathcal{C}_x^{(1,f)} \right\} \leq c(m) \rho^{k-m} r_x^k \equiv C_f(m, \rho) r_x^k \, . \tag{7.58}$$

Since the energy drop is verified on these balls, we define $\mathcal{C}^{(1,f)}$ to be the set of final balls at the step $i = 1$ by

$$\mathcal{C}^{(1,f)} = \bigcup_{x \in \mathcal{C}_+} \mathcal{C}_x^{(1,f)} \, . \tag{7.59}$$

For $y \in \mathcal{C}_x^{(1,b)}$, the energy drop condition is not verified. However, since there exists a $k-1$ dimensional space L_x such that

$$F_x = \{ y \in \mathcal{S} \cap B_{2r_x}(x) \quad s.t. \;\; \theta(y, \rho r_x/10) \geq E - \delta \} \subseteq B_{\rho r_x/5}(L_x) \, , \tag{7.60}$$

then we can estimate

$$\sum_{y \in \mathcal{C}_x^{(1,b)}} (\rho r_x)^k = \rho^k r_x^k \# \left\{ \mathcal{C}_x^{(1,b)} \right\} \leq c(m) \rho^{1-k} \rho^k r_x^k \equiv C_c(m) \rho r_x^k \, . \tag{7.61}$$

On these balls, we can either have the stopping condition $\rho r_x = r$, or we need to refine the covering further. Thus we define

$$\mathcal{C}^{(1,b)} = \bigcup_{x \in \mathcal{C}_+ , \;\; \rho r_x > r} \mathcal{C}_x^{(1,b)} \, , \qquad \mathcal{C}^{(1,r)} = \mathcal{C}_r \cup \bigcup_{x \in \mathcal{C}_+ , \;\; \rho r_x = r} \mathcal{C}_x^{(1,b)} \, . \tag{7.62}$$

$\mathcal{C}^{(1,b)}$ represents the set of “bad balls” where we need to refine our covering further.

By this and lemma 7.3, in particular by the estimates in (7.53), we obtain that

$$\sum_{y \in \mathcal{C}^{(1,b)}} r_y^k \leq C_c(m) \rho \sum_{x \in \mathcal{C}_+} r_x^k \leq C_V(m) C_c(m) \rho \, . \tag{7.63}$$

If we choose

$$0 < \rho(m) \leq \min \left\{ 100^{-1}, \frac{1}{2} C_V(m)^{-1} \cdot C_c(k)^{-1} \right\} \, , \tag{7.64}$$

we can rephrase the above estimates as

$$\sum_{y \in \mathcal{C}^{(1,b)}} r_y^k \leq \frac{1}{2} \, . \tag{7.65}$$

If we set

$$C_F(m) = 2 C_V(m) \left(C_f(m, \rho(m)) + C_c(m) \right) \, , \tag{7.66}$$

the estimates on the final and r-balls are

$$\sum_{y \in \mathcal{C}^{(1,r)} \cup \mathcal{C}^{(1,f)}} r_y^k = \# \{ \mathcal{C}_r \} \, r^k + \sum_{x \in \mathcal{C}_+} r_x^k \left(C_f(m, \rho) + C_c(m, \rho) \right) \leq C_V(m) \left(C_f(m, \rho) + C_c(m) \right) = \frac{1}{2} C_F(m) \, . \tag{7.67}$$

Note that clearly for all $y \in \mathcal{C}^{(1,b)}$, we have $r_y \leq \rho$.

7.2.8 Re-covering of bad balls: Induction step

Suppose that we have obtained our covering for i. It is clear that we need to improve our covering only on the balls $\{B_{r_x}(x)\}_{x\in\mathcal{C}^{(i,b)}}$. In order to do so, we consider each of these balls separately.

Since all the assumptions on lemma 7.3 are satisfied on each of the $B_{r_x}(x)$, we can apply again this lemma to each $B_{r_x}(x)$, and obtain that for all x there exists a covering

$$\mathcal{S}\cap B_{r_x}(x)\subseteq \bigcup_{y\in\mathcal{C}_{r,x}} B_r(y)\cup\bigcup_{y\in\mathcal{C}_{+,x}} B_{r_y}(y) \quad \text{with} \quad r_y\geq r \quad \text{and} \quad \sum_{y\in\mathcal{C}_{r,x}\cup\mathcal{C}_{+,x}} r_y^k\leq C_V(m)r_x^k\,. \tag{7.68}$$

Moreover, for each $y\in\mathcal{C}_{+,x}$, there exists a $k-1$ dimensional subspace L_y such that

$$F_y\equiv\left\{z\in\mathcal{S}\cap B_{2r_y}(y) \;\; s.t. \;\; \theta(z,\rho r_y/10)>E-\delta\right\}\subseteq B_{\rho r_y/5}\left(L_y\right)\cap B_{2r_y}(y)\,. \tag{7.69}$$

By applying exactly the same procedure described in the first step of the induction to each of the balls $\left\{B_{r_y}(y)\right\}_{y\in\mathcal{C}_{+,x}}$, we obtain the new desired covering. In particular, for each $y\in\mathcal{C}_{+,x}$ we can find a covering

$$B_{r_y}(y)\setminus B_{\rho r_y}\left(F_y\right)\subseteq\bigcup_{z\in\mathcal{C}_y^{(i+1,f)}} B_{\rho r_y}(z)\,, \quad B_{r_y}(y)\cap B_{\rho r_y}\left(F_y\right)\subseteq\bigcup_{z\in\mathcal{C}_y^{(i+1,b)}} B_{\rho r_y}(z)\,, \tag{7.70}$$

where for all $z\in\mathcal{C}_y^{(i+1,f)}$ and all $p\in\mathcal{S}\cap B_{2\rho r_y}(z)$, we have $\theta(p,\rho r_y)\leq E-\delta$, and we have the estimates

$$\sum_{z\in\mathcal{C}_y^{(i+1,f)}}(\rho r_y)^k\leq C_f(m,\rho)r_y^k\,, \qquad \sum_{z\in\mathcal{C}_y^{(i+1,b)}}(\rho r_y)^k\leq C_c(m)\rho r_y^k\,. \tag{7.71}$$

The new set $\mathcal{C}^{(i+1,f)}$ is now defined as the previous set of “final balls” $\mathcal{C}^{(i,f)}$ along with the new final balls $\mathcal{C}^{(i+1,f)}$ obtained with this covering, thus making

$$\mathcal{C}^{(i+1,f)}=\bigcup_{x\in\mathcal{C}^{(i,b)}}\bigcup_{y\in\mathcal{C}_{+,x}}\mathcal{C}_y^{(i+1,f)}\,, \qquad \mathcal{C}^{(i+1,f)}=\mathcal{C}^{(i,f)}\cup\mathcal{C}^{(i+1,f)}\,. \tag{7.72}$$

In a similar way for the r-balls, we obtain

$$\mathcal{C}^{(i+1,r)}=\bigcup_{x\in\mathcal{C}^{(i,b)}}\left(\mathcal{C}_{r,x}\cup\bigcup_{y\in\mathcal{C}_{+,x}\,,\;\rho r_y=r}\mathcal{C}_y^{(i+1,b)}\right)\,, \qquad \mathcal{C}^{(i+1,r)}=\mathcal{C}^{(i,r)}\cup\mathcal{C}^{(i+1,r)}\,. \tag{7.73}$$

However, evidently the new set of “bad balls” does not contain the bad balls at the previous scale, since those are the ones that were just re-covered. In particular

$$\mathcal{C}^{(i+1,b)}=\bigcup_{x\in\mathcal{C}^{(i,b)}}\bigcup_{y\in\mathcal{C}_{+,x},\;\rho r_y>r}\mathcal{C}_y^{(i+1,b)}\,. \tag{7.74}$$

The k-dimensional content estimate of our covering are obtained by iterating the estimates obtained in the first step. In detail, by arguing as in (7.63) and (7.65), and by choosing ρ according to (7.64), we obtain

$$\sum_{z\in\mathcal{C}^{(i+1,b)}} r_z^k\leq\sum_{x\in\mathcal{C}^{(i,b)}}\frac{1}{2}r_x^k=2^{-1-i}\,. \tag{7.75}$$

As for final and r-balls, arguing as in (7.67) we can estimate the contribution given by the new r and final balls by

$$\sum_{z\in\mathcal{C}^{(i+1,r)}\cup\mathcal{C}^{(i+1,f)}} r_z^k \leq \left(\frac{1}{2}C_F(m)\right)\sum_{z\in\mathcal{C}^{(i,b)}} r_z^k = 2^{-i-1}C_F(m)\,. \tag{7.76}$$

This yields the desired result (7.56), and in turn concludes the proof of the lemma.

It is worth noticing that at the i-th step of the induction, the radius of the biggest ball in the covering is smaller than ρ^i. Thus eventually $\rho^i \leq r$ and this induction will stop in a finite number of steps.

□

7.2.9 Keeping track of the constants

For the reader's convenience we record here how all the constants involved in the previous two lemmas were chosen.

First of all, note that $\epsilon > 0$ is arbitrary, as well as $r > 0$. However, it is of course important that all the constants here are *independent of* r.

$C_R(m)$ is the constant coming from the Reifenberg theorem 5.4, and it depends only on m. $C_V(m)$ is fixed in (7.33), and it is just a dimensional constant $c(m)$ (coming from a rough cover of $B_1(0)$ by balls of radius $1/8$) times $C_R(m)$. Thus $C_V(m)$ clearly depends only on m. $C_c(m)$ is fixed in (7.61), and is just another covering constant whose value depends only on m.

The parameter ρ, which was a free parameter in the first covering, is fixed once and for all in (7.64) as a constant depending only on m. For convenience, we can also pick a ρ satisfying (7.14). Once this choice has been fixed, also the constant $C_F(m)$ introduced in (7.66) depends only on m.

The parameter $\eta > 0$ is chosen according to (7.49), as explained in (7.50). At last, with this positive value of η fixed, we choose δ in such a way that (7.21) and (7.24) are all satisfied.

7.3 Proof of the main theorems

Now we are ready to prove our main thereoms on the singularities of harmonic maps. We recall the theorem here.

Theorem 7.5. *Let $u : B_2(0) \subseteq \mathbb{R}^m \to N$ be a minimizing harmonic map with $\theta(0,2) \leq \Lambda$. Its singular set $\mathcal{S}(u)$ satisfies*

$$\mathrm{Vol}\left(B_r\left(\mathcal{S}(u)\cap B_1(0)\right)\right) \leq C(m,\Lambda)r^3\,, \tag{7.77}$$

and $\mathcal{S}(u)$ is $m-3$-rectifiable.

This proof is basically a corollary of the covering Lemma 7.4. The proof is another induction, this time downwards and on the upper bound on the energy.

7.3.1 Volume estimate: Induction on energy upper bounds

Using the covering Lemma 7.4, we will prove by induction on $i = 0, 1, \cdots, \lfloor \delta^{-1} E \rfloor + 1$ that there exist coverings of $\mathcal{S}$ by balls $\{B_{r_x}(x)\}_{x \in \mathcal{C}^i}$ such that

$$\mathcal{S} \subseteq \bigcup_{c \in \mathcal{C}^i} B_{r_x}(x) , \quad \sum_{x \in \mathcal{C}^i} r_x^k \leq (c(m) C_F(m))^i . \tag{7.78}$$

Moreover, for all i we have

$$r_x \leq r \quad \text{or} \quad \forall y \in \mathcal{S} \cap B_{2r_x}(x) , \ \theta(y, r_x) \leq E - i\delta . \tag{7.79}$$

It is clear that if we pick $i = \lfloor \delta^{-1} E \rfloor + 1$, then the second condition cannot be true anywhere, which means that the first condition must be true, which will complete the construction of the covering.

In particular, this implies that at the step $i = \lfloor \delta^{-1} E \rfloor + 1$ we have a covering of $\mathcal{S}(u)$ by balls of radius r and such that their number is bounded by

$$\exp\left(\log(c(m) C_F(m)) \Lambda \delta(m, \Lambda)^{-1}\right) r^{3-m} . \tag{7.80}$$

As a corollary, we immediately obtain that (7.77) is satisfied.

Now we turn to the proof of the covering. Note that this covering is trivial for $i = 0$, since $\mathcal{S} \subseteq B_1(0)$ does the trick at this stage.

By induction, suppose that (7.78) and (7.79) are true for i. Pick any $x \in \mathcal{C}^i$, and consider $B_{r_x}(x)$. By the covering lemma 7.4 (or better, by an r_x-rescaled version of this lemma), there exists a covering $\mathcal{C}_x$ of $\mathcal{S} \cap B_{r_x}(x)$ such that

$$\mathcal{S} \cap B_{r_x}(x) \subseteq \bigcup_{y \in \mathcal{C}_x} B_{r_y}(y) , \quad r_y \leq \rho r_x \leq \rho^i , \quad \sum_{y \in \mathcal{C}_x} r_y^k \leq C_F(m) r_x^k . \tag{7.81}$$

Moreover, for all $y \in \mathcal{C}_x$, we have

$$\text{either } r_y = r \quad \text{or} \quad \forall z \in \mathcal{S} \cap B_{2r_y}(y) , \ \theta(z, \rho r_y/10) \leq E - i\delta - \delta = E - (i+1)\delta . \tag{7.82}$$

By covering each $B_{r_y}(y)$ again by a minimal set of balls of radius $\rho(m) r_y \leq r_y/10$, we obtain a covering $\mathcal{C}_x$ such that

$$\mathcal{S} \cap B_{r_x}(x) \subseteq \bigcup_{y \in \mathcal{C}_x} B_{r_y}(y) , \quad r_y \leq \rho r_x \leq \rho^i , \quad \sum_{y \in \mathcal{C}_x} r_y^k \leq c(m) C_F(m) r_x^k . \tag{7.83}$$

Moreover, for all $y \in \mathcal{C}_x$, we have

$$\text{either } r_y \leq r \quad \text{or} \quad \forall z \in \mathcal{S} \cap B_{2r_y}(y) , \ \theta(z, \rho r_y) \leq E - i\delta - \delta = E - (i+1)\delta . \tag{7.84}$$

By summing all the contributions coming from balls $\{B_{r_x}(x)\}_{x \in \mathcal{C}^i}$, we obtain

$$\mathcal{C}^{i+1} = \bigcup_{x \in \mathcal{C}^i} \mathcal{C}_x , \quad \sum_{y \in \mathcal{C}^{i+1}} r_y^k = \sum_{x \in \mathcal{C}^i} \left(\sum_{y \in \mathcal{C}_x} r_y^k \right) \leq (c(m) C_F(m))^{i+1} , \tag{7.85}$$

as desired.

7.3.2 Rectifiability

By the volume estimates in (7.77), we have $\lambda^k\left(\mathcal{S}(u) \cap B_1(0)\right) \leq C$. By applying the same estimates on any ball $B_r(x)$ with $x \in B_1(0)$ and $r \leq 1$, we obtain that

$$\lambda^k\left(\mathcal{S}(u) \cap B_r(x)\right) \leq C r^k\,, \tag{7.86}$$

in other words, $\mathcal{S}(u)$ is upper-Ahlfors regular.

We will prove that for all measurable subsets $\mathcal{S} \subseteq \mathcal{S}(u) \cap B_1(0)$, there exists a k-measurable subset $E \subset \mathcal{S}$ with $\lambda^k(E) \leq 7^{-1}\lambda^k(\mathcal{S})$ such that $\mathcal{S} \setminus E$ is k-rectifiable. Since $\mathcal{S}$ is an arbitrary measurable subset, this is enough to prove rectifiability by a standard density argument.

Consider any $\mathcal{S} \subseteq \mathcal{S}(u) \cap B_1(0)$. We can assume wlog that $\lambda^k(\mathcal{S}) > 0$, otherwise there is nothing to prove. Consider the function $f(x,r) = \theta(x,r) - \theta(x,0)$ on $B_1(0)$. This function is monotone nondecreasing in r, uniformly bounded for all $x \in B_1(0)$ and $r \leq 1$, and pointwise converging to 0 as $r \to 0$.

Thus, by dominated convergence, for all $\delta > 0$, there exists a radius $\bar{r} > 0$ such that

$$\fint_{\mathcal{S}} f(x, 10\bar{r})\, d\lambda^k(x) \leq \delta^2\,. \tag{7.87}$$

Let $E \subset \mathcal{S}$ be a measurable subset with $\lambda^k(E) \leq \delta\lambda^k(\mathcal{S})$ and such that $f(x, 10\bar{r}) \leq \delta$ for all $x \in F \equiv \mathcal{S} \setminus E$.

Now cover F by a finite number of balls $B_{\bar{r}}(x_i)$ centered on F. We want to show that, if δ is chosen small enough, then on each of these balls we can apply Theorem 5.5 to $F \cap B_{\bar{r}}(x_i)$ for all i, and thus proving that F is k-rectifiable as desired.

Reifenberg estimates The estimates here are basically equivalent to the estimates carried out in Section 7.2.4. Actually, since we already know that (7.86) holds, we do not even need the upper induction part of that argument. For this reason, we will only sketch the main passages in the estimates.

Fix any i, and consider the set $F \cap B_{\bar{r}}(x_i)$. For convenience, we rescale the ball $B_{\bar{r}}(x_i)$ to $B_1(0)$. With an abuse of notation, we will keep denoting by u, θ, $\mathcal{S}(u)$ and F also the rescaled objects.

By definition of $F \subset \mathcal{S}(u)$, we have that $\theta(x,10) - \theta(x,0) \leq \delta$ for all $x \in F$. By an estimate analogous to (7.40), we have for all $x \in F$ and $s \leq 1$

$$\beta_{2,\lambda^k|_F}(x,s)^2 \leq C_1 s^{-k} \int_{B_s(x)} W_s(y)\, d\lambda^k|_F(y) \tag{7.88}$$

By integrating, and by (7.86), we obtain for all $x \in B_1(0)$ and $s \leq r \leq 1$:

$$\int_{B_r(x)} \beta_{2,\lambda^k|_F}(z,s)^2\, d\lambda^k|_F(z) \leq C_1 s^{-k} \int_{B_r(y)} \left[\int_{B_s(z)} W_s(t)\, d\lambda^k|_F(t)\right] d\lambda^k|_F(z) \leq C_1 C \int_{B_{r+s}(y)} W_s(z) d\lambda^k|_F(z)\,. \tag{7.89}$$

Integrating again in s, we finally get for all $x \in B_1(0)$ and $r \leq 1$:

$$\int_{B_r(x)} \left[\int_0^s \beta_{2,\lambda^k|_F}(z,s)^2 \frac{ds}{s}\right] d\lambda^k|_F(z) \leq C_1 C \int_{B_{2r}(x)} [\theta(x,8r) - \theta(x,0)] d\lambda^k|_F(z) \leq c(m) C_1 C^2 \delta r^k\,. \tag{7.90}$$

By choosing

$$\delta \leq \frac{\delta_R^2}{c(m)C_1C^2}\,, \tag{7.91}$$

we can apply Theorem 5.5 to the set $F \cap B_1(0)$, thus proving that it is k-rectifiable. This concludes the proof.

References

[AS] J. Azzam and R. Schul, An analyst's traveling salesman theorem for sets of dimension larger than one, *preprint, arXiv 1609.02892*. Available at `https://arxiv.org/abs/1609.02892`.

[CN13] J. Cheeger and A. Naber, Quantitative stratification and the regularity of harmonic maps and minimal currents, *Comm. Pure Appl. Math.* **66** (2013), 965–990. MR 3043387. Zbl 1269.53063. `http://dx.doi.org/10.1002/cpa.21446`. Available at `http://arxiv.org/abs/1107.3097`.

[CG89] J.-M. Coron and R. Gulliver, Minimizing p-harmonic maps into spheres, *J. Reine Angew. Math.* **401** (1989), 82–100. MR 1018054. Zbl 0677.58021. `http://dx.doi.org/10.1515/crll.1989.401.82`.

[DS93] G. David and S. Semmes, *Analysis of and on uniformly rectifiable sets*, *Mathematical Surveys and Monographs* **38**, American Mathematical Society, Providence, RI, 1993. MR 1251061. Zbl 0832.42008. `http://dx.doi.org/10.1090/surv/038`.

[DT12] G. David and T. Toro, Reifenberg parameterizations for sets with holes, *Mem. Amer. Math. Soc.* **215** (2012), vi+102. MR 2907827. Zbl 1236.28002. `http://dx.doi.org/10.1090/S0065-9266-2011-00629-5`.

[DL08] C. De Lellis, *Rectifiable sets, densities and tangent measures*, *Zurich Lectures in Advanced Mathematics*, European Mathematical Society (EMS), Zürich, 2008. MR 2388959. `http://dx.doi.org/10.4171/044`. Available at `http://dx.doi.org/10.4171/044`.

[Dor85] J. R. Dorronsoro, A characterization of potential spaces, *Proc. Amer. Math. Soc.* **95** (1985), 21–31. MR 796440. Zbl 0577.46035. `http://dx.doi.org/10.2307/2045567`.

[ENV] N. S. Edelen, A. Naber, and D. Valtorta, Quantitative Reifenberg Theorem for Measures, arXiv:1612.08052. Available at `https://arxiv.org/abs/1612.08052`.

[Fed69] H. Federer, *Geometric measure theory*, *Die Grundlehren der mathematischen Wissenschaften, Band 153*, Springer-Verlag New York Inc., New York, 1969. MR 0257325. Zbl 0176.00801.

[Lin87] F.-H. Lin, A remark on the map $x/|x|$, *C. R. Acad. Sci. Paris Sér. I Math.* **305** (1987), 529–531. MR 916327. Zbl 0652.58022.

[Mat95] P. Mattila, *Geometry of sets and measures in Euclidean spaces*, *Cambridge Studies in Advanced Mathematics* **44**, Cambridge University Press, Cambridge, 1995, Fractals and rectifiability. MR 1333890. Zbl 0819.28004. `http://dx.doi.org/10.1017/CBO9780511623813`.

[Mor66] C. B. Morrey, Jr., *Multiple integrals in the calculus of variations*, *Die Grundlehren der mathematischen Wissenschaften, Band 130*, Springer-Verlag New York, Inc., New York, 1966. MR 0202511. Zbl 0142.38701.

[Mos05] R. Moser, *Partial regularity for harmonic maps and related problems*, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2005. MR 2155901. Zbl 1246.58012. `http://dx.doi.org/10.1142/9789812701312`.

[NVa] A. Naber and D. Valtorta, Rectifiable-reifenberg and the regularity of stationary and minimizing harmonic maps, *accepted on Annals of Mathematics*. Available at `https://arxiv.org/abs/1505.03428`.

[NVb] ______, The singular structure and regularity of stationary and minimizing varifolds. Available at `https://arxiv.org/abs/1505.03428`.

[NVc] ______, Stratification for the singular set of approximate harmonic maps, *preprint*. Available at `https://arxiv.org/abs/1611.03008`.

[NVV] A. Naber, D. Valtorta, and G. Veronelli, Quantitative regularity for p-harmonic maps, *Preprint*. Available at `http://arxiv.org/abs/1409.8537`.

[Rei60] E. R. Reifenberg, Solution of the Plateau Problem for m-dimensional surfaces of varying topological type, *Acta Math.* **104** (1960), 1–92. MR 0114145. Zbl 0099.08503. Available at `http://link.springer.com/article/10.1007%2FBF02547186`.

[Sim] L. Simon, Reifenberg's topological disc theorem. Available at `http://www.math.uni-tuebingen.de/ab/analysis/pub/leon/reifenberg/reifenberg.dvi`.

[Tor95] T. Toro, Geometric conditions and existence of bi-Lipschitz parameterizations, *Duke Math. J.* **77** (1995), 193–227. MR 1317632. Zbl 0847.42011. `http://dx.doi.org/10.1215/S0012-7094-95-07708-4`.