

The regularity of minimal surfaces in higher codimension

Camillo De Lellis

ABSTRACT. In this paper we review the regularity theory for area minimizing m -dimensional currents in codimension higher than 1, which bounds the dimension of the singular set with $m - 2$. In recent joint works with Emanuele Spadaro we have revisited the pioneering program of Almgren, bringing some new techniques from metric analysis and some new ideas to deal with the most intricate aspects of the proof.

CONTENTS

1. Introduction	154
2. The Federer-Fleming theory of integral currents	157
3. First considerations in the regularity theory	165
4. The regularity theory in codimension 1	172
5. Federer's theorem and the failure of ε -regularity in codimension $\bar{n} \geq 2$	179
6. Almgren's stratification	182
7. Multiple valued functions minimizing the Dirichlet energy	184
8. The frequency function	193
9. Approximation with multiple valued graphs	198
10. A first attempt to prove Theorem 3.4	205
11. The center manifold	209
12. The approximation on the normal bundle of the center manifold	215
13. The frequency function again	220
14. The persistence of singularities	224
References	226

2010 *Mathematics Subject Classification.* 49Q15, 49N60, 49Q05.

Key words and phrases. integer rectifiable currents, regularity, area minimizing, multiple valued functions.

1. Introduction

The Plateau's problem investigates those surfaces of least area spanning a given contour. It is one of the most classical problems in the calculus of variations, it lies at the crossroad of several branches of mathematics and it has generated a large amount of mathematical theory in the last one hundred years. The problem itself and its various generalizations have found fundamental applications in several mathematical and scientific branches. Since it is a prototype of a vast family of questions with geometric and physical significance, the techniques developed to analyze it have proved to be very useful in a variety of other situations.

The original formulation is attributed to the Belgian physicist Plateau, although it was considered earlier by Lagrange, and it regards 2-dimensional surfaces spanning a given one-dimensional contour γ in the 3-dimensional space: among these surfaces one is interested in those which minimize the area (and, more in general, on the critical points of the area, although in this survey we will restrict our attention to "absolute" minimizers). Plateau considered such "minimal surfaces" to model soap films. However it is very natural to generalize the question and look for surfaces of dimension m which minimize the m -dimensional volume among those spanning a given contour of dimension $m - 1$ in \mathbb{R}^{m+n} , or in more general ambient spaces. Such generalizations have not only an intrinsic mathematical beauty, but they have proved to be very fruitful. In this note we will restrict ourselves to ambient spaces which are complete oriented Riemannian manifolds Σ and since all the considerations will be of a local nature we will often assume that Σ itself is isometrically embedded in some euclidean space (of dimension $m+n$). In this way the competitor surfaces (classical or generalized) spanning the contour γ will always be (suitable generalizations of) subsets of the standard euclidean space, constrained to be subsets of Σ . Although this is not very elegant from a geometric point of view, it allows us to avoid a lot of technicalities.

The very formulation of the Plateau's problem has proved to be a quite challenging mathematical question. In particular, how general are the surfaces that one should consider? What is the correct concept of "spanning" and the correct concept of " m -dimensional volume" that one should use? The author believes that there are no final answers to these two questions: many different significant ones have been given in the history of our subject and, depending upon the context, the features of one formulation might be considered more important than those of the others.

The several different formulations of the Plateau's problem could be subdivided in three large classes.

- The parametric formulations: the competitor surfaces are supposed to be images of maps defined on a given domain (or on a class of domains) and the volume can then be computed through the parametrization using the standard "area formula". The notion of

“spanning” reduces to imposing that the trace of the maps on the chosen domain is a parametrization of the contour γ . This was the first successful attempt at giving a suitable existence theory for 2-dimensional area minimizing surfaces, pioneered by Douglas and Rado in the thirties ([42], [58]; see the monograph [41] for a modern introduction to the subject).

- The set-theoretical formulations: the competitor surfaces are supposed to be just (closed) sets and the notion of m -dimensional volume is then given by the Hausdorff measure (or some other measure-theoretic generalization of the m -dimensional volume of a submanifold). The first to pioneer and implement successfully such formulations was Reifenberg in [59]. The notion of spanning is usually much more involved in this context; among the most recent proposals let us mention the elegant one of Harrison (cf. [54]) and the ones of David (cf. [18]). For a rather general compactness principle which could be useful in several set-theoretic frameworks we refer to the recent papers [26] and [38].
- The functional-analytic formulations: the surfaces are mostly viewed as objects acting on a given (linear) space of smooth test functions, usually via integration. In this note we address the most popular functional analytic formulations, De Giorgi’s theory of sets of finite perimeter ([19, 20]; we refer to [23] for a recent English translation) and Federer and Fleming’s theory of integral currents ([47]), and we will discuss extensively the corresponding existence and regularity theories.

In all these approaches there are two first fundamental issues that a satisfactory variational theory needs to address: existence and regularity. Of course these are by no means the *only* important aspects of Plateau’s problem: however almost all the other necessarily build on these two important pieces of information, namely that

- (a) there is a minimizer for a large class of boundaries;
- (b) the minimizer is sufficiently regular, so that one can compute interesting geometric quantities and infer additional conclusions.

The success of the Federer and Fleming’s theory is due to the vast applicability of its existence part in all dimensions and codimensions. Thanks to the efforts of several outstanding mathematicians a rather far-reaching (and satisfactory) regularity theory was achieved in the seventies in codimension 1 (see for instance [53]). This theory has been digested by the subsequent generations of scholars working in differential geometry and PDEs, leading ultimately to many breakthroughs in different problems in geometry, PDEs and mathematical physics. On the contrary the most important conclusion in the higher codimension case can be attributed to the monumental work of a single person, F. J. Almgren Jr. ([5]). Unlike the codimension one case, only a relatively small portion of the monstrous proof of Almgren has been

truly understood. In a recent series of papers Emanuele Spadaro and the author have given a new, much shorter, account of Almgren's regularity program, relying on the several advances in geometric measure theory of the last two decades and on some new ideas. The aim of this note is to give a rather detailed picture of the several issues that this program must face and of how they are resolved.

We will start by reviewing the basic definitions and the most important results of the theory of Federer and Fleming, showing how it gives a natural and very elegant existence theory for the Plateau's problem in any dimension and codimension (see Section 2). As part of our exposition we will isolate the features of the codimension 1 case, where in fact the Federer-Fleming theory is equivalent to De Giorgi's theory of sets of finite perimeter. In Section 3 we will review the first considerations in the regularity theory and summarize the state of the art in the subject. We will then review the regularity theory for minimizers in codimension 1, focusing on its most important steps, cf. Section 4. In Section 5 we will discuss what are the obstructions to a direct generalization of the codimension 1 regularity theory to higher codimension. In the (several!) remaining sections we will discuss the details of the proofs of Almgren's theorem as presented in the papers [29, 30, 33, 31, 32].

This survey has been conceived to be accessible also to those scholars who do not have any knowledge of the theory of currents: I will only assume that the reader is familiar with some basic concepts of functional analysis, measure theory and differential geometry. The reader who is already familiar with geometric measure theory (and more precisely with the theory of currents) is instead encouraged to read the survey [25] or the recent lecture notes [72]. In particular the next three sections can be thought as a gentle introduction to some of the classical works of the fifties, sixties and seventies regarding what we have called the "functional analytic" approach to the Plateau's problem. Obviously the understanding of many issues will become increasingly difficult for the novice as we go deeper in the description of Almgren's program: nonetheless I tried to make all the exposition self contained providing, when necessary, a brief account of the main technical tools used in the statements.

It is the belief of the author that there is a general misconception of geometric measure theory, namely that it is a highly technical and obscure subject. Although I certainly agree that the proofs are long and difficult, the main concepts are in fact rather elegant and in most cases they can be introduced without invoking too much theory. The main reason behind the misconception is the lack of introductory textbooks in the area. The books available are conceived for experts and they aim at developing a far-reaching and self-contained theory from the very beginning: motivated by this necessity they introduce a lot of terminology and notation to deal with all possible technicalities in the very first chapters, sacrificing the readability and the intuitive picture at a very early stage. On the other hand many "gentle" introductions to geometric measure theory are "invitations" for very young

students and they have the opposite tendency of neglecting any discussion of the most complicated parts. In this survey I have tried to avoid as much as possible the introduction of a lot of notation and of many concepts, keeping such things to the absolute minimum needed to understand the main points of the arguments. On the other hand I will deal with the most complicated parts of the proofs, trying to single out the key ideas that overcome the main difficulties.

2. The Federer-Fleming theory of integral currents

2.1. Currents, boundaries and mass. The idea of treating (oriented) surfaces as linear functionals on a suitable space of smooth objects dates back at least to De Rham, cf. [40]. More precisely, consider a smooth m -dimensional oriented surface Γ (with or without boundary) in \mathbb{R}^{m+n} and denote by \mathcal{D}^m the space of smooth, compactly supported m -forms. Γ defines naturally a linear map

$$(1) \quad \mathcal{D}^m \ni \omega \quad \mapsto \quad \int_{\Gamma} \omega.$$

This motivates the following

DEFINITION 2.1 (Current, De Rham, cf. [68, Definition 26.1]). An m -dimensional current T is a continuous linear map $T : \mathcal{D}^m \rightarrow \mathbb{R}$. Here the continuity condition must be understood in the following sense: $T(\omega^k) \rightarrow T(\omega)$ whenever $\{\omega^k\} \subset \mathcal{D}^m$ is a sequence such that

- (a) there is an open set $\Omega \subset \subset \mathbb{R}^{m+n}$ with $\text{spt}(\omega^k) \subset \Omega$ for every k ;
- (b) $\omega^k \rightarrow \omega$ in $C^j(\Omega)$ for every j .

If we regard smooth functions as 0-forms, then 0-dimensional currents are simply the usual distributions. The reader familiar with them (or in general with certain type of functional analysis) will recognize that the definition above follows a rather standard path and will not be surprised if we introduce a corresponding “dual topology” on the space of currents: more precisely we will say that a sequence of currents T^k (of the same dimension m) converges to T if $T^k(\omega) \rightarrow T(\omega)$ for every $\omega \in \mathcal{D}^m$ (cf. [68, eq. 26.12]).

Very naturally the concept of boundary is defined “enforcing” Stokes’ Theorem, pretty much as the derivative of a distribution is defined “enforcing” integration by parts formulae.

DEFINITION 2.2 (Boundary, De Rham, cf. [68, eq. 26.3]). We say that an $(m-1)$ -dimensional current S is the boundary of an m -dimensional current T if

$$(2) \quad T(d\omega) = S(\omega) \quad \text{for every } \omega \in \mathcal{D}^{m-1}.$$

S will then be denoted by ∂T .

The class of smooth oriented submanifolds $\Gamma \subset \mathbb{R}^{m+n}$ with smooth boundaries can then naturally be viewed as a subset of the space of m -dimensional currents. In order to distinguish between any such Γ and its

“action” as linear functional via integration of forms, we will use the notation $\llbracket \Gamma \rrbracket$ for the current, namely $\llbracket \Gamma \rrbracket(\omega) := \int_{\Gamma} \omega$ (cf. [68, eq. 26.2]). If we consider points $P \in \mathbb{R}^{m+n}$ as 0-dimensional submanifolds, consistently with our convention we will denote by $\llbracket P \rrbracket$ the usual Dirac delta at P , although a more common symbol would be δ_P .

Observe also that the definitions above generalize to any Riemannian ambient manifold Σ in place of \mathbb{R}^{m+n} . However, if Σ is embedded in \mathbb{R}^{m+n} another way to introduce currents T in Σ is to consider them as currents T in the euclidean space \mathbb{R}^{m+n} whose support $\text{spt}(T)$ is contained in Σ . As usual the support $\text{spt}(T)$ is the complement of the maximal open set U for which $T(\omega) = 0$ whenever $\text{spt}(\omega) \subset U$ (cf. [45, Section 4.1]).

We next need to introduce a suitable concept of m -dimensional volume. When Γ is a smooth surface, its volume $\text{Vol}^m(\Gamma)$, as defined in the usual textbooks on integration of forms, can be computed with the area formula and coincides with its Hausdorff m -dimensional measure (cf. [68, Section 2.8]). There is however a definition of volume which exploits the natural duality between forms and submanifolds. Recall that a simple m -vector is an element of $\Lambda_m(\mathbb{R}^{m+n})$ of the form $v_1 \wedge \dots \wedge v_m$. Moreover there is a natural definition of *length* of a simple m -vector: $|v_1 \wedge \dots \wedge v_m|$ is the m -dimensional Hausdorff measure of the parallelogram spanned by the vectors v_1, \dots, v_m .

DEFINITION 2.3 (Comass, cf. [45, Section 1.8]). Let $\omega \in \mathcal{D}^m$. Then the comass of ω is the norm

$$\|\omega\|_c := \max \left\{ \langle \omega(p), v_1 \wedge \dots \wedge v_m \rangle : |v_1 \wedge \dots \wedge v_m| = 1, p \in \mathbb{R}^{m+n} \right\}.$$

Now, it can be easily checked that, when Σ is a smooth submanifold of \mathbb{R}^{m+n} , then

$$(3) \quad \text{Vol}^m(\Sigma) = \sup \left\{ \int_{\Sigma} \omega : \omega \in \mathcal{D}^m \text{ and } \|\omega\|_c \leq 1 \right\}.$$

This motivates the

DEFINITION 2.4 (Mass, cf. [45, Section 4.1.7]). Given an m -dimensional current T we denote by $\mathbf{M}(T)$ its mass, namely the quantity

$$(4) \quad \mathbf{M}(T) = \sup \{ T(\omega) : \|\omega\|_c \leq 1 \}.$$

An important feature of this definition is that the mass can be localized. More precisely for any open set Ω we can define

$$(5) \quad \|T\|(\Omega) = \sup \{ T(\omega) : \text{spt}(\omega) \subset \Omega \text{ and } \|\omega\|_c \leq 1 \}.$$

When T is a current of finite mass, namely $\mathbf{M}(T) < \infty$, $\|T\|$ turns out to be a Radon measure and obviously $\mathbf{M}(T) = \|T\|(\mathbb{R}^{m+n})$ (cf. [45, Sections 4.1.5 and 4.1.7]).

It is now a simple exercise in functional analysis to prove the following

THEOREM 2.5. *Let Z be an $(m-1)$ -dimensional current for which there is \bar{T} with $\partial\bar{T} = Z$ and $\mathbf{M}(\bar{T}) < \infty$. Then there is a T_0 such that $\partial T_0 = Z$ and*

$$\mathbf{M}(T_0) = \min\{\mathbf{M}(T) : \partial T = Z\}.$$

If $\text{spt}(\bar{T}) \subset K$ for some closed set K we also have the existence of a T_0 such that $\partial T_0 = Z$, $\text{spt}(T_0) \subset K$ and

$$\mathbf{M}(T_0) = \min\{\mathbf{M}(T) : \partial T = Z \text{ and } \text{spt}(T) \subset K\}.$$

Theorem 2.5 is however not very satisfactory from the “classical” point of view, as it rather often gives many minimizers which are not classical surfaces.

EXAMPLE 2.6. Consider for instance the south and north poles S and N in the standard sphere $\mathbb{S}^2 \subset \mathbb{R}^3$ and let Z be the 0-dimensional current $\llbracket N \rrbracket - \llbracket S \rrbracket$. For any meridian γ joining S to N the corresponding current $\llbracket \gamma \rrbracket$ is a minimizer of the mass among all currents T with $\partial T = Z$ and $\text{spt}(T) \subset \mathbb{S}^2$. However the same holds for any convex combination $\lambda \llbracket \gamma \rrbracket + (1 - \lambda) \llbracket \eta \rrbracket$ where η is any other meridian and $\lambda \in [0, 1]$. In fact one can push this idea even further. Let us parametrize the meridians as $\{\gamma_t\}_{t \in \mathbb{S}^1}$, where t is the intersection of γ_t with the equator $\{x_3 = 0\} \cap \mathbb{S}^2$. If μ is a probability measure on \mathbb{S}^1 , then the current

$$T_0(\omega) := \int_{\mathbb{S}^1} \llbracket \gamma_t \rrbracket(\omega) d\mu(t)$$

is also a minimizer of the mass (among those currents T with $\text{spt}(T) \subset \mathbb{S}^2$ and $\partial T = Z$).

This does not seem a serious issue as there are anyway “classical minimizers” in the example above. However, we have the following remarkable theorem (for a very short and elegant proof we refer to [78]).

THEOREM 2.7 (Lavrentiev gap, Young [82]). *For every smooth closed embedded curve γ in \mathbb{R}^4 define*

$$(6) \quad M(\gamma) := \inf \{ \text{Vol}^2(\Sigma) : \Sigma \text{ is immersed, oriented and } \partial \Sigma = \gamma \}$$

$$(7) \quad m(\gamma) := \min \{ \mathbf{M}(T) : \partial T = \llbracket \gamma \rrbracket \}.$$

Then there are γ 's for which $M(\gamma) > m(\gamma)$.

Before going on with our discussions we stop a moment to introduce some simple concepts and notations which will be very useful later on. First of all we observe that, if F is a smooth *proper* map between two Euclidean spaces \mathbb{R}^N and \mathbb{R}^k , the pullback $F^\sharp \omega$ of an element $\omega \in \mathcal{D}^m(\mathbb{R}^k)$ is an element of $\mathcal{D}^m(\mathbb{R}^N)$. By duality this gives naturally a notion of pushforward of currents, namely $F_\# T(\omega) = T(F^\sharp \omega)$ (cf. [68, eq. 26.20]). The assumption that F is proper is needed to guarantee that $F^\sharp \omega$ has compact support if ω has compact support. This is however not needed when $\text{spt}(T)$ is compact:

since this will be mostly the case, in several occasions we will push currents forward via maps which are just smooth.

Next, if $\Omega \subset \mathbb{R}^N$ is an open set and we denote by $\mathcal{D}^m(\Omega)$ the space of smooth m -forms with compact support in Ω , we can naturally define a concept of *current* in Ω (with corresponding “local” notions of boundary and mass), cf. [45, Section 4.1.7]). Moreover, when $\Omega \subset \Omega'$ are open and T is a current on Ω' we can easily restrict its action to forms supported in Ω : for such restriction we use the notation $T \llcorner \Omega$. Observe that, whenever T has finite mass, $T \llcorner \Omega$ has also finite mass and we have the relation $\|T\|(\Omega) = \mathbf{M}(T \llcorner \Omega)$ (cf. [68, eq. 26.9]). It is also the case that, when T has finite mass, $T \llcorner \Omega$ can be given a meaning as a current in the original domain Ω' : of course this might “add” some extra boundary “located” in $\partial\Omega$. Since we will always consider currents of (at least locally) finite mass, we will not insist too much on their domain of definition and always assume they are defined in duality with forms $\omega \in \mathcal{D}^m(\mathbb{R}^{m+n})$.

In the future, to make our discussion simple, we will say that two currents T and S agree on an open subset Ω if $T \llcorner \Omega = S \llcorner \Omega$. We use a similar notation on Borel measures μ (which we always assume to be defined on some subset K of the Euclidean space): given a μ -measurable $E \subset K$, we use the symbol $\mu \llcorner E$ for the measure $\mu \llcorner E(A) := \mu(A \cap E)$.

Another operation which is well defined on currents is the product (cf. [68, Definition 26.16]). In turn this allows, for any given current S in \mathbb{R}^{m+n} and any given point $P \in \mathbb{R}^{m+n}$, to define the current T which is the cone with base S and vertex P (cf. [68, eq. 26.26]). The common notation for such T will be $P \rtimes S$. The construction is not complicated but requires a certain amount of notation and terminology: we will instead give an intuitive definition in the special case of integer rectifiable currents later.

2.2. The Federer-Fleming theory. If we want to rule out minimizers as in Example 2.6 it seems desirable to introduce a restricted class of currents which, roughly speaking, “allows only integer multiplicities”.

DEFINITION 2.8 (Integral currents, Federer-Fleming, cf. [45, Definition 27.1]). A current T is *integer rectifiable* if there are a sequence of oriented C^1 surfaces $\Sigma_i \subset \mathbb{R}^{m+n}$, a sequence of pairwise disjoint closed subsets $K_i \subset \Sigma_i$ and a sequence of positive integers k_i such that

$$(8) \quad \sum_i k_i \text{Vol}^m(K_i) < \infty$$

$$(9) \quad T(\omega) = \sum_i k_i \int_{K_i} \omega \quad \forall \omega \in \mathcal{D}^m.$$

T is *integral* if both T and ∂T are integer rectifiable.

REMARK 2.9. It is not very hard to see that under the assumptions above $\mathbf{M}(T) = \sum_i k_i \text{Vol}^m(K_i)$ and thus it is implicit in the definition that integer rectifiable currents have finite mass. This is however not always a

desirable feature: for instance an entire smooth graph would not be an integer rectifiable current according to the definition above. For this reason, it is customary to consider integer rectifiable those currents T for which a representation as above can be found for the restriction $T \llcorner \Omega$ to any bounded open set $\Omega \subset \mathbb{R}^{m+n}$. In the rest of the note we will then use the term integer rectifiable (and integral) for this larger class.

Observe that the space of integer rectifiable currents is not any more a *linear space* and there is no simple functional-analytic principle which provides a good compactness property. A fundamental result in the theory of Federer and Fleming is that, nonetheless, the space of integral currents is compact in a suitable sense.

THEOREM 2.10 (Compactness of integral currents, Federer-Fleming, cf. [68, Theorem 32.2]). *If $\{T^k\}$ is a sequence of integral m -dimensional currents such that*

$$\sup_k (\mathbf{M}(T^k) + \mathbf{M}(\partial T^k)) < \infty,$$

then there is a subsequence, not relabeled, and an integral m -dimensional current T such that $T_k \rightarrow T$.

As a corollary we achieve

COROLLARY 2.11. *Let Z be an $(m-1)$ -dimensional integer rectifiable current and \bar{T} an m -dimensional integral current with $\partial \bar{T} = Z$ and $\mathbf{M}(\bar{T}) < \infty$. Then there is an integer rectifiable current T_0 such that $\partial T_0 = Z$ and*

$$\mathbf{M}(T_0) = \min\{\mathbf{M}(T) : T \text{ is int. rect. and } \partial T = Z\}.$$

If $\text{spt}(\bar{T}) \subset K$ for some closed set K we also have the existence of an integer rectifiable current T_0 such that $\partial T_0 = Z$, $\text{spt}(T_0) \subset K$ and

$$\mathbf{M}(T_0) = \min\{\mathbf{M}(T) : T \text{ is int. rect., } \partial T = Z \text{ and } \text{spt}(T) \subset K\}.$$

Indeed both Theorem 2.10 and Corollary 2.11 can be extended to a fairly large class of metric spaces, cf. [9].

Theorem 2.10 does not exhaust the major results of the foundational paper of Federer and Fleming. Indeed we wish to mention three other important cornerstones. First of all, the rectifiability of the boundary can be recovered from that of the current under the only assumption that the boundary has finite mass.

THEOREM 2.12 (Boundary rectifiability, Federer-Fleming, cf. [68, Theorem 30.3]). *If T is integer rectifiable and $\mathbf{M}(\partial T) < \infty$, then T is integral.*

Secondly, the “Lavrentiev gap” phenomenon of Lemma 2.7 is not present: any integral current can be suitably approximated by a sequence of “polyhedral chains” with integer coefficients. This is the content of the so-called Deformation lemma (see [68, Theorem 29.1 and Corollary 29.3]). Its precise statement would require the introduction of some terminology and goes beyond the scopes of this note. We record, however, a rather useful corollary of the Deformation Lemma.

THEOREM 2.13 (Isoperimetric inequality, Federer-Fleming, cf. [68, Theorem 30.1]). *There are constants $C(m, n)$ with the following property. Assume S is an integer rectifiable m -dimensional current in \mathbb{R}^{m+n} with $\partial S = 0$. Then there is an integral current T with $\partial T = S$ and $\mathbf{M}(T) \leq C(\mathbf{M}(S))^{(m+1)/m}$.*

Observe the following interesting corollary of the latter inequality: if S is an integer rectifiable cycle (namely $\partial S = 0$), then it is a boundary. In fact a major achievement of the Federer-Fleming theory is that the integral homology theory which derives from the chain complex of integral currents is equivalent to the standard homology theories with integer coefficients in all Riemannian manifolds (cf. [45, Section 4.4.1]). As a consequence, each integral homology class in a compact smooth Riemannian manifold can be represented by an integral cycle with least mass. Moreover, the theory can be generalized to other coefficient groups (cf. [45, Section 4.46]).

We finish this section by introducing a few other objects which will be very convenient in the rest of the note. First of all looking back at the Definition 2.8 it is tempting to introduce a “density” for an integer rectifiable current at every point x belonging to any of the sets K_i appearing in Definition 2.8. The natural choice would be the number k_i . This can indeed be done, but it raises the question whether the corresponding density depends only on the current T and not instead on the chosen decomposition. In fact it is not difficult to show that this definition of density is unique in a suitable measure-theoretic sense and we record an important characterization of it in the next lemma. Here we denote by ω_m the m -dimensional volume of the unit m -dimensional ball and by \mathcal{H}^m the Hausdorff m -dimensional measure.

LEMMA 2.14. *If T is an integer rectifiable current, then the number*

$$(10) \quad \Theta(T, p) := \lim_{r \downarrow 0} \frac{\|T\|(\mathbf{B}_r(p))}{\omega_m r^m}$$

exists and it is a positive integer for $\|T\|$ -a.e. p . Moreover, if the sets K_i and the integers k_i are as in Definition 2.8, then $\|T\| = \sum_i k_i \mathcal{H}^m \llcorner K_i$ and $\Theta(T, p) = k_i$ for \mathcal{H}^m -a.e. $p \in K_i$.

Next it is tempting to attach a tangent plane to T at the points $p \in K_i$: a natural candidate would be the tangent plane to the oriented C^1 submanifold Σ_i . Again this raises the question whether such definition is truly intrinsic. As for the density we can answer this question through a characterization which follows a “blow-up procedure”. In order to give the corresponding statement we introduce two conventions:

- First of all, we will consider all m -dimensional planes π as oriented. Thus, for each π we have a unique integral current $\llbracket \pi \rrbracket$.
- Given a current T we will denote by $T_{p,r}$ the result of translating it so that p becomes the origin and enlarging it of a factor r^{-1} . Formally, if $\iota_{p,r}$ denotes the map $x \mapsto (x - p)/r$, then $T_{p,r} :=$

$(\iota_{p,r})_{\#}T$. Note that when $T = \llbracket \Gamma \rrbracket$ for some smooth surface Γ then $T_{p,r} = \llbracket \iota_{p,r}(\Gamma) \rrbracket$.

The procedure of “zooming in a particular point” and taking limits of (subsequences of) the corresponding rescalings is called “blowing-up” the current T and it is a device which is employed recurrently in the regularity theory.

LEMMA 2.15. *If T is an integer rectifiable current, then for $\|T\|$ -a.e. p there is a unique plane $\pi(p)$ such that*

$$(11) \quad T_{p,r} \rightarrow \Theta(T, p) \llbracket \pi(p) \rrbracket \quad \text{as } r \downarrow 0.$$

$\pi(p)$ will then be called the tangent plane to T at p . Moreover, if the sets K_i and the submanifolds Σ_i are as in Definition 2.8, $\pi(p) = T_p \Sigma_i$ (the classical oriented tangent to Σ_i at p) for \mathcal{H}^m -a.e. $p \in K_i$.

If $\pi(p)$ is as above and e_1, \dots, e_m is a positively oriented orthonormal base for it, it is customary to introduce the simple m -vector

$$\vec{T}(p) := e_1 \wedge \dots \wedge e_m :$$

this defines a Borel map from \mathbb{R}^{m+n} into $\Lambda_m(\mathbb{R}^{m+n})$, the m -th exterior product of the standard euclidean space. We will use \vec{T} extensively for such map and observe that we have the simple identity

$$T(\omega) = \int \langle \omega(p), \vec{T}(p) \rangle d\|T\|(p) = \int \langle \omega(p), \vec{T}(p) \rangle \Theta(T, p) d\mathcal{H}^m(p),$$

where \langle, \rangle denotes the usual duality pairing between m -vectors and m -covectors. This measure theoretic representation allows us to restrict integer rectifiable currents to subsets $E \subset \mathbb{R}^{m+n}$ which are just Borel measurable (rather than open). Namely we define

$$T \llcorner E(\omega) := \int_E \langle \omega(p), \vec{T}(p) \rangle d\|T\|(p).$$

REMARK 2.16. Since $\Lambda_m(\mathbb{R}^{m+n})$ can be endowed with a natural scalar product (namely through $\langle e_1 \wedge \dots \wedge e_m, f_1 \wedge \dots \wedge f_m \rangle = \det \langle e_i, f_j \rangle$: here we use \langle, \rangle for the Euclidean scalar product!), T can be seen as a measure taking values in $\Lambda_m(\mathbb{R}^{m+n})$: $\|T\|$ is then the total variation measure of T and $\vec{T}\|T\|$ its “polar (or Radon-Nikodým) decomposition” (cf. [68, eq. 26.7 and Definition 27.1]). However, this will not play any crucial role in our discussion and the reader who is not familiar with Radon-Nikodým decompositions can safely ignore this remark.

The extra structure given by Definition 2.8 to integer rectifiable currents allows to generalize readily several computations which are valid for C^1 submanifolds. The simple recipe is just to use the decomposition of Definition 2.8 to *chop* an integer rectifiable current in pieces K_i of C^1 submanifolds: the corresponding computations can then be carried on each K_i and patched suitably. The reader who is not familiar with such a procedure does not have

to worry whether the corresponding result depends or not on the chosen decomposition: in essentially all important cases it does not, although the corresponding proof might be delicate.

An instructive example is the cone with basis T and vertex p , which from now on will be denoted by $p \times T$. As already mentioned, there is an intrinsic definition for this object whatever the current T is (see [45, Section 4.1.1]). In the case of integer rectifiable currents, however, we can use the idea above to reduce the definition of $p \times T$ to the standard one for immersed C^1 surfaces. First of all, if $F : \Omega \rightarrow \mathbb{R}^{m+n}$ is a C^1 map with $\Omega \subset \mathbb{R}^m$ compact, we can define the m -dimensional current

$$T(\Omega) := F_{\#} [\![\Omega]\!] (\omega) := \int_{\Omega} F_{\#} \omega.$$

If $p \in \mathbb{R}^{m+n}$ is a point, then we can define the map $G : \Omega \times [0, 1]$ by

$$(x, t) \mapsto (1 - t)p + tF(x)$$

and the $(m + 1)$ -dimensional current $G_{\#} [\![0, 1] \times \Omega]\!]$: this current coincides obviously with our intuitive picture of the cone $p \times (F_{\#} [\![\Omega]\!])$. Now, given an integer rectifiable current T , we can decompose it into compact subsets of C^1 embedded surfaces as in Definition 2.8, define the corresponding cone with vertex p for each such piece and then sum them all: the result is $p \times T$.

2.3. The codimension 1 case: Sets of finite perimeter. Integral currents of codimension 1 have a special feature: they can be seen, locally, as boundaries of integral currents of “top dimension”. By definition, integer rectifiable currents T of dimension $m + n$ are represented by $\sum k_i [\![E_i]\!]$, where the E_i ’s are pairwise disjoint closed subsets of \mathbb{R}^{m+n} : the action of $[\![E_i]\!]$ on a “top form” $fdx_1 \wedge \dots \wedge dx_{m+n}$ is then given by the standard Lebesgue integral

$$[\![E_i]\!] (fdx_1 \wedge \dots \wedge dx_{m+n}) = \int_{E_i} f(x) dx.$$

The current ∂T can then be thought as $\sum_i k_i \partial [\![E_i]\!]$, i.e. as an integral combination of “boundaries of sets”. In this form the Federer-Fleming theory for codimension 1 currents existed already since few years before the appearing of the foundational paper [47]: the corresponding objects, introduced by De Giorgi in [19] and [23] following some pioneering ideas of Caccioppoli (see [12]), are called *sets of finite perimeter* or *Caccioppoli sets*. The relevant definition is

DEFINITION 2.17 (Caccioppoli sets, De Giorgi, cf. [8, Definition 3.35]). A measurable set $E \subset \mathbb{R}^{m+1}$ with finite Lebesgue measure is a set of finite perimeter if its indicator function $\mathbf{1}_E$ is a function of bounded variation, namely if

$$\mathbf{P}(E) := \sup \left\{ \int_E \operatorname{div} X : X \in C_c^\infty(\mathbb{R}^{m+1}, \mathbb{R}^{m+1}) \text{ and } \|X\|_{C^0} \leq 1 \right\} < \infty.$$

$\mathbf{P}(E)$ is called the *perimeter* of E .

Like the mass, the perimeter can be localized to define a Radon measure. In fact such measure coincides with the total variation of the distributional derivative $D\mathbf{1}_E$, which is usually denoted by $\|D\mathbf{1}_E\|$ (cf. [8, Theorem 3.3.6]). If E is a smooth set and Ω is an open set, then $\|D\mathbf{1}_E\|(\Omega)$ is the m -dimensional volume of that portion of ∂E which lies in Ω . For this reason it is customary to use the notation $\mathbf{P}(E, \Omega)$ for the *relative perimeter* $\|D\mathbf{1}_E\|(\Omega)$ when Ω is an open set and E is a Caccioppoli set (again cf. [8, Definition 3.3.5]).

The fundamental link between the two theories is then given by the following

PROPOSITION 2.18 (cf. [68, Theorem 27.6 and Corollary 27.8]). *Let E be a measurable subset of \mathbb{R}^{m+1} with finite Lebesgue measure. $\llbracket E \rrbracket$ is then an integral current if and only if E is a Caccioppoli set. Moreover, in this case $\mathbf{M}(\partial \llbracket E \rrbracket) = \mathbf{P}(E)$.*

Consider next an m -dimensional integer rectifiable current T with finite mass and let $\mathbf{B}_\rho(x) \subset \mathbb{R}^{m+1}$ be such that $\text{spt}(\partial T) \cap \mathbf{B}_\rho(x) = \emptyset$. Then there are countably many sets of finite perimeter E_i and positive integer k_i such that

- (i) $\|T\|(\mathbf{B}_\rho(x)) = \sum_i k_i \mathbf{P}(E_i, \mathbf{B}_\rho(x))$;
- (ii) $T = \sum k_i \partial \llbracket E_i \rrbracket$ on $\mathbf{B}_\rho(x)$.

Theorem 2.10, Corollary 2.11, Theorem 2.12 and Theorem 2.13 are all generalizations of theorems proved by De Giorgi for sets of finite perimeter (see [8, Sections 3.3 and 3.5]).

3. First considerations in the regularity theory

Going back to the Plateau's problem, a current T as in the second statement of Corollary 2.11 must have the following local minimality property:

- (AM) If S is an integral current of dimension $m+1$ supported in K , then
- $$\mathbf{M}(T + \partial S) \geq \mathbf{M}(T).$$

Such currents will be called *area minimizing in K* (cf. [68, Definition 33.1]) and in the rest of the paper K will always be some (sufficiently smooth) embedded Riemannian manifold of dimension $m + \bar{n}$, denoted by Σ . The number \bar{n} will be considered the *codimension* of the area minimizing current T .

The rest of this note will be dedicated to the question: what kind of regularity is implied by (AM)? Since we will always argue at the local level and the problem is scaling invariant, we can assume all sorts of nice properties upon Σ (for instance that it is a global graph of a smooth function with good bounds on its C^k norms): it is therefore natural to expect that the answer will not depend upon the nature of Σ , but rather upon the codimension \bar{n} . We next state the best theorems proved so far concerning the regularity of area minimizing currents, but before coming to them we summarize the relevant definitions in the following

DEFINITION 3.1. Let $\Omega \subset \mathbb{R}^{m+n}$ be open and $\Sigma \subset \mathbb{R}^{m+n}$ be a smooth complete submanifold without boundary of dimension $m + \bar{n}$. We say that an m -dimensional integer rectifiable current T is area minimizing in $\Sigma \cap \Omega$ if

- $\text{spt}(T) \subset \Sigma$;
- $\mathbf{M}(T + \partial S) \geq \mathbf{M}(T)$ for every $(m+1)$ -dim. integral S with $\text{spt}(S) \subset \Sigma \cap \Omega$.

DEFINITION 3.2. We say that $p \in \text{spt}(T) \setminus \text{spt}(\partial T)$ is an interior regular point if there is a positive radius $r > 0$, a smooth embedded submanifold Γ of Σ and a positive integer Q such that $T \llcorner \mathbf{B}_r(p) = Q \llbracket \Gamma \rrbracket$. The set of interior regular points, which of course is relatively open in $\text{spt}(T) \setminus \text{spt}(\partial T)$, is denoted by $\text{Reg}(T)$. Its complement $\text{spt}(T) \setminus (\text{spt}(\partial T) \cup \text{Reg}(T))$, the interior singular set of T , is denoted by $\text{Sing}(T)$.

The first theorem summarizes the achievements of several outstanding mathematicians from the end of the sixties till the nineties: De Giorgi, Almgren, Fleming, Simons, Federer, Bombieri, Giusti and Simon. It is fair to say that, with the notable exception of Simon's rectifiability result, the various aspects of the following theorem have been well digested in the mathematical communities of elliptic PDEs and geometric analysis.

THEOREM 3.3 (Regularity in codimension 1). *Assume that Ω , Σ and T are as in Definition 3.1 and that $\bar{n} = 1$. Then*

- (i) *For $m \leq 6$ $\text{Sing}(T) \cap \Omega$ is empty (Fleming & De Giorgi ($m=2$), Almgren ($m=3$), Simons ($4 \leq m \leq 6$), see [21, 48, 22, 3, 71] and also [60, 75]);*
- (ii) *For $m = 7$ $\text{Sing}(T) \cap \Omega$ consists of isolated points (Federer, see [46]);*
- (iii) *For $m \geq 8$ $\text{Sing}(T) \cap \Omega$ has Hausdorff dimension at most $m - 7$ (Federer, [46]) and it is countably $(m-7)$ -rectifiable, namely, up to a set of \mathcal{H}^{m-7} -measure zero, it can be covered by countably many C^1 surfaces of dimension $m - 7$ (Simon, [69]);*
- (iv) *The above results are optimal, namely for every $m \geq 7$ there are area minimizing integral currents T in the euclidean space \mathbb{R}^{m+1} for which $\text{Sing}(T)$ has positive \mathcal{H}^{m-7} measure (Bombieri-De Giorgi-Giusti, [11]).*

As already mentioned, after discussing the features of the codimension 1 case, the rest of the note will be devoted to the understanding of the higher codimension, i.e. $\bar{n} \geq 2$. For this case the best results are the following.

THEOREM 3.4 (Regularity in codimension $\bar{n} \geq 2$). *Assume that Ω , Σ and T are as in Definition 3.1 and that $\bar{n} \geq 2$. Then*

- (i) *For $m = 1$ $\text{Sing}(T) \cap \Omega$ is empty;*
- (ii) *For $m \geq 2$ $\text{Sing}(T) \cap \Omega$ has Hausdorff dimension at most $m - 2$ (Almgren, [5]);*

- (iii) *The above result is optimal, namely for every $m \geq 2$ there are area minimizing integral currents T in \mathbb{R}^{m+2} for which $\text{Sing}(T)$ has positive \mathcal{H}^{m-2} measure (Federer, [44]).*

Almgren's result was subsequently sharpened by Chang (cf. [13]) for 2-dimensional area minimizing currents.

THEOREM 3.5 ($m = 2$, $\bar{n} \geq 2$). *Assume that Ω , Σ and T are as in Definition 3.1, that $\bar{n} \geq 2$ and $m = 2$. Then $\text{Sing}(T) \cap \Omega$ consists of isolated points.*

In Section 5 we will discuss extensively the difficulties that any argument for (ii) must face. Almgren's original typewritten proof was more than 1700 pages long and was published posthumously thanks to the efforts of his students Scheffer and Taylor in a book of almost 1000 pages. In this note we will describe the main steps of Almgren's program following the papers [29, 30, 33, 31, 32] by Emanuele Spadaro and the author.

Chang's result builds heavily on Almgren's book. Moreover Chang's paper [13] does not provide the proof of one major step of the argument, the existence of a "branched center manifold": the construction of such object requires the understanding of 4/5 of Almgren's monograph and a suitable modification of its most obscure and involved part, which gives the construction of the "non-branched center manifold" (cf. Sections 11 and 12 below). Building upon [29, 30, 33, 31, 32], in joint papers with Emanuele Spadaro and Luca Spolaor we will give the first proof of the existence of a "branched center manifold" and extend Chang's theorem to a large class of objects which are almost minimizing in a suitable sense, cf. [37, 34, 35, 36]. That proof (and Chang's theorem) will however not be discussed in this survey.

In the rest of this section we will delve into the preliminaries of the regularity theory, namely the monotonicity formula and its consequences, which are common to both the codimension 1 and the higher codimension cases.

Often, we will discuss generalizations of the regularity theorems to *stationary* and sometimes to *stable* objects. We give thus their formal definition here. Recall that we can push-forward currents through maps. In what follows, given a smooth compactly supported vector field X on \mathbb{R}^{m+n} we will consider the one parameter family Φ_t of diffeomorphisms generated by X , in other words the flux of X :

$$(12) \quad \begin{cases} \frac{d}{dt} \Phi_t(x) = X(\Phi_t(x)) \\ \Phi_0(x) = x. \end{cases}$$

If the vector field X is tangent to a given submanifold Σ , then obviously Φ_t maps Σ into itself. For a current T supported in Σ , an *admissible* one parameter family of deformations is then given by $T_t = (\Phi_t)_\# T$ where the generator X is assumed to be tangent to Σ and to vanish on $\text{spt}(\partial T)$. Notice

in particular that, for an area minimizing current T in Σ , we would have

$$(13) \quad \frac{d}{dt} \mathbf{M}(T_t) = 0$$

and

$$(14) \quad \frac{d^2}{dt^2} \mathbf{M}(T_t) \geq 0.$$

DEFINITION 3.6. An integer rectifiable current T with $\text{spt}(T) \subset \Sigma$ is called

- *stationary* if (13) holds for any admissible deformation described above;
- *stable* if both (13) and (14) hold for any admissible deformation described above.

3.1. Compactness. A first basic fact about area minimizing currents is that, under the same assumptions of the compactness theorem of Federer and Fleming, they are also a compact class. It is not difficult to form an intuition about this: assume that a sequence of integral currents T_k has a uniform bound on the mass in a certain open set Ω , namely $\sup_k \|T_k\|(\Omega) < \infty$ and that ∂T_k vanishes in Ω . By the compactness theorem of Federer and Fleming (in fact a suitable localized version of Theorem 2.10), we can assume that T_k converges to an integer rectifiable current T in Ω , which has no boundary in Ω . The mass is lower semicontinuous and thus

$$(15) \quad \mathbf{M}(T) \leq \liminf_k \|T_k\|(\Omega).$$

Next assume that each T_k is area minimizing in \mathbb{R}^{m+n} . If we had a strict inequality, we could imagine to use T as a competitor for T_k , after “gluing T_k and T ” close to $\partial\Omega$: the gain in mass from T to T_k is a certain positive number, whereas the “gluing” costs less and less as $k \rightarrow \infty$ because the currents T_k come “closer” to T . The result would contradict the minimizing property of T_k for k large enough and we conclude therefore that the \liminf in the right hand side of (15) is a limit and that equality holds.

The discussion above is correct, but a rigorous proof requires all the power of the Federer-Fleming theory, in particular of the Deformation lemma. It is also clear that for the same reason it should not be possible to lower the mass of T by perturbing it in a compact subset of Ω , i.e. T is area minimizing in Ω . A similar conclusion holds also in case the T_k ’s are area minimizing in a given smooth submanifold Σ . Actually we could allow the ambient manifold to vary with k : if we denote it with Σ_k and we assume that it converges smoothly (C^2 suffices) to a smooth Σ , then the limiting current will be area minimizing in Σ . We summarize our discussion in the following

THEOREM 3.7 (Compactness of area minimizing currents, cf. [68, Theorem 34.5]). *Let Σ_k be a sequence of C^2 submanifolds of \mathbb{R}^{m+n} of dimension*

$m + \bar{n}$ which converge in C^2 to Σ and let T_k be a sequence of integer rectifiable area minimizing currents in Σ_k of dimension m with $\sup_k \mathbf{M}(T_k) < \infty$. Assume that $\partial T_k = 0$ on some open set Ω and that $T_k \llcorner \Omega \rightarrow T$. Then

- T is area minimizing in $\Omega \cap \Sigma$;
- $\|T_k\| \llcorner \Omega \xrightarrow{*} \|T\|$ in the sense of Radon measures.

3.2. The monotonicity formula and its consequences. A pivotal role in the regularity theory for area minimizing currents and, more in general, for all known generalizations of the concept of critical point for the area functional, is the so-called monotonicity formula. We start with the following simple observation: assume a certain area minimizing current T in \mathbb{R}^{m+n} is in fact a smooth surface Γ (namely $T = \llbracket \Gamma \rrbracket$) and fix a point $p \in \Gamma \setminus \partial\Gamma$ and a radius $r < \text{dist}(p, \partial\Gamma)$. Assume moreover that $\partial\mathbf{B}_r(p)$ intersects Γ transversally. If we replace Γ in the ball $\mathbf{B}_r(p)$ with the cone having vertex p and boundary $\Gamma \cap \partial\mathbf{B}_r(p)$ we must increase the volume of Γ . Namely

$$(16) \quad \text{Vol}^m(\Gamma \cap \mathbf{B}_r(p)) \leq \frac{r}{m} \text{Vol}^{m-1}(\Gamma \cap \partial\mathbf{B}_r(p)).$$

On the other hand the coarea formula implies that

$$(17) \quad \text{Vol}^{m-1}(\partial\mathbf{B}_r(p) \cap \Gamma) \leq \left. \frac{d}{dt} \right|_{t=r} \text{Vol}^m(\mathbf{B}_t(p) \cap \Gamma)$$

and we reach easily the conclusion that

$$(18) \quad \frac{d}{dr} \frac{\text{Vol}^m(\Gamma \cap \mathbf{B}_r(p))}{r^m} \geq 0.$$

In fact this is a very crude argument: a more careful computation using the stationarity of Γ , i.e. the vanishing of the first variation of the area functional, gives the much more precise formula

$$(19) \quad \frac{\text{Vol}^m(\Gamma \cap \mathbf{B}_r(p))}{r^m} - \frac{\text{Vol}^m(\Gamma \cap \mathbf{B}_s(p))}{s^m} = \int_{\Gamma \cap (\mathbf{B}_r(p) \setminus \mathbf{B}_s(p))} \frac{|(x-p)^\perp|^2}{|x-p|^{m+2}} d\text{Vol}^m(x),$$

where $(x-p)^\perp$ denotes the component of the vector $x-p$ which is orthogonal to the tangent space $T_x\Gamma$. A similar formula, which includes a further correction due to the second fundamental form of Σ , is valid for minimal surfaces in smooth submanifolds $\Sigma \subset \mathbb{R}^{m+n}$.

The formula (19), in the framework above an *a-priori estimate*, is indeed valid for area minimizing currents as well (and in general for a very powerful generalization of the concept of “stationary surface”, called stationary varifold, see [1]; cf. also [68, Section 4.3]). Observe indeed that for an integer-rectifiable current T it is rather easy to make sense of the right hand side of (19): since at $\|T\|$ -a.e. x we have a well defined tangent plane, we can define $(x-p)^\perp$ for $\|T\|$ -a.e. x .

A first obvious consequence of the monotonicity formula is that the density of an area minimizing current is in fact defined at *every* point: thus

from now on we will use $\Theta(T, p)$ (cf. (10)) as a well defined quantity for every $p \notin \text{spt}(\partial T)$. Indeed it is a simple exercise to show that Θ is an upper semicontinuous function (and this because, modulo some small technicalities, the map $p \rightarrow \|T\|(\mathbf{B}_r(p))$ is continuous at each fixed r).

We next describe another crucial consequence of the monotonicity formula. Let T be an area minimizing current (in the euclidean space, to simplify our discussion) and $p \in \text{spt}(T) \setminus \text{spt}(\partial T)$. Recall the homothetic rescalings $T_{p,r}$. It is obvious that each $T_{p,r}$ is an area minimizing current. Observe also that:

- For each bounded open set Ω , we clearly have $\text{spt}(\partial T_{p,r}) \cap \Omega = \emptyset$, provided r is small enough;
- For each $R > 0$ we have a uniform bound for $\|T_{p,r}\|(\mathbf{B}_R(0))$: the latter is indeed the number $r^{-m}\|T\|(\mathbf{B}_{Rr}(p))$, which is bounded independently of r thanks to the monotonicity formula.

Thus, by Theorem 3.7, for every fixed bounded open Ω we can extract a subsequence $\{T_{p,r_k}\}_{r_k \downarrow 0}$ which converges in Ω to an area minimizing current T_0 . Actually, by a standard diagonal argument we can find a “global” limit current T_0 which is an integral current on each bounded open subset of \mathbb{R}^{m+n} , which has no boundary and whose restriction to any bounded open set Ω is area minimizing. Although strictly speaking T_0 violates our definition of area minimizing current (because it does not have finite mass on the *entire* euclidean space), we will still use the term area minimizing by a slight abuse of notation.

The monotonicity formula and Theorem 3.7 imply that $\|T_0\|(\mathbf{B}_r(p)) = \Theta(T, p)r^m$ for every $r > 0$. If T_0 were regular, namely a classical surface Γ (with multiplicity), we would conclude that the right hand side of (19) vanishes when $p = 0$: note that this is only possible if for every $x \in \Gamma$ the segment joining x with the origin is contained in Γ , namely if Γ is a *cone* with vertex at the origin. The same conclusion can be drawn without knowing any a-priori regularity for T_0 . It is customary to use the term *area minimizing cone* for such objects and to call them *tangent cones to T at p* if there is a sequence $r_k \downarrow 0$ such that $T_{p,r_k} \rightarrow T_0$. This motivates the following

DEFINITION 3.8. An area minimizing cone of dimension m is an integer rectifiable current S of dimension m with the following properties

- $\partial S = 0$ and $S_{0,r} = S$ for every positive r ;
- $S \llcorner \Omega$ is area minimizing for any bounded open set Ω .

Next, if T and S are two currents such that, for some $p \in \text{spt}(T)$ and some $r_k \downarrow 0$, T_{p,r_k} converges to S , we then say that S is *tangent to T at p* .

Rather than giving the precise formulation of the monotonicity formula for area minimizing currents in a submanifold Σ (i.e. with the exact remainder), we record in the following proposition all the most important consequences.

THEOREM 3.9 (Tangent cones, cf. [68, Section 7.3]). *Let T be an area minimizing integral current of dimension m in a C^2 submanifold Σ . Then*

- (i) *$r \mapsto e^{Cr-m}\|T\|(\mathbf{B}_r(p))$ is a monotone function for each $p \notin \text{spt}(\partial T)$, provided $r \in]0, \text{dist}(p, \text{spt}(\partial T)[$ and C is a suitable constant, which only depends on the size of the second fundamental form of Σ ;*
- (ii) *The density $\Theta(T, p)$ is well defined at every $p \notin \text{spt}(\partial T)$, it is at least 1 at each point $p \in \text{spt}(T) \setminus \text{spt}(\partial T)$ and it is upper semicontinuous;*
- (iii) *For every $p \notin \text{spt}(\partial T)$ and every sequence $r_k \downarrow 0$ there is a subsequence, not relabeled, and an area minimizing cone T_0 such that $T_{p, r_k} \rightarrow T_0$; $T_0 \neq 0$ if and only if $p \in \text{spt}(T)$.*

Note that Lemma 2.15 already guarantees that at $\|T\|$ -a.e. $p \in \text{spt}(T) \setminus \text{spt}(\partial T)$ there is a *unique* tangent cone, which is an integer multiple (such multiple being $\Theta(T, p)$) of an m -dimensional plane $\pi(p)$. In order to make our discussion shorter, from now on a tangent cone will be called *flat* if it is a multiple of an m -dimensional plane (note that the multiple will necessarily turn out to be a nonzero integer). It is obvious that at every $p \in \text{Reg}(T)$ there is a unique tangent cone and it is flat. On the other hand if there is even a single tangent cone at p which is not flat, then necessarily $p \in \text{Sing}(T)$! At first glance a pretty plausible conjecture is that regular points coincide indeed with those points where at least one tangent cone is flat. We will see that this is true in codimension 1 (the first deep regularity theorem, due to De Giorgi [21]), but not necessarily in higher codimension. In codimension 1 the theorem of De Giorgi gives right away that $\mathcal{H}^m(\text{Sing}(T)) = 0$, thanks to the following elementary, but powerful, corollary of Theorem 3.9.

COROLLARY 3.10 (cf. [68, Sections 4.3 and 7.3]). *Let T be an area minimizing current of dimension m . Then at \mathcal{H}^m -a.e. $p \in \text{spt}(T) \setminus \text{spt}(\partial T)$ $\Theta(T, p) \in \mathbb{N} \setminus \{0\}$ and there is a unique flat tangent cone to T at p .*

Moreover, the convergence of area minimizing currents can be improved in the following sense: If T_k , Σ_k and Ω are as in Theorem 3.7, then $\text{spt}(T_k)$ converges to $\text{spt}(T)$ locally (i.e. on every compact subset of Ω) in the sense of Hausdorff.

Before coming to the next section, let us look at what is perhaps the most intriguing open problem in the regularity theory of minimal surfaces. Observe that Theorem 3.9 does not imply the *uniqueness* of the tangent cone at a given point, namely it leaves the possibility that, for two different sequences $r_k \downarrow 0$ and $s_k \downarrow 0$, the limits of T_{p, r_k} and T_{p, s_k} do not coincide. This seems wasteful from the point of view of area, but proving it turns out to be the most challenging open problem in the field. More precisely the following conjecture is widely open

CONJECTURE 3.11. *The tangent cone to an area minimizing current T is unique at every point $p \in \text{spt}(T) \setminus \text{spt}(\partial T)$.*

The conjecture has been proved for 2-dimensional currents T in *any* codimension by White in his remarkable paper [77] and it has been shown by Simon in codimension 1 at any isolated singularity in the fundamental work [67]. The latter result is indeed a consequence of a remarkably general approach, which applies to other variational problems (such as the uniqueness of tangent maps to energy minimizing maps) but also to the study of the asymptotic behaviour of solutions to parabolic equations, see [67].

4. The regularity theory in codimension 1

The first breakthrough in the regularity theory is due to De Giorgi: he realized in his fundamental work [21] that the existence of one flat tangent plane at p is enough to conclude that p is a regular point in codimension 1. His theorem was then extended to any codimension by Almgren in [2] (see also [65]) under an important assumption on the density which we will discuss extensively in a moment (indeed, it is possible to extend the validity even to general Hilbert spaces, cf. [7]). In fact Almgren's statement covers many more geometric functionals, which satisfy an appropriate ellipticity assumption. In the framework of minimal surfaces the most important generalization of De Giorgi's ε -regularity theorem is due to Allard in [1] (cf. also [68, Chapter 4] and [24]): his theorem, valid for a far reaching generalization of classical stationary surfaces (namely integer rectifiable varifolds with sufficiently summable generalized mean curvature) is the starting point of a variety of applications of the minimal surface theory to geometric and topological problems.

We will state here the De Giorgi-Almgren ε -regularity theorem in all dimensions and codimensions and we will later emphasize why its consequences are much stronger in codimension 1. As already mentioned, we are dealing with an ε -regularity theorem: under the assumption that a certain particular quantity is sufficiently small at a given scale, we will conclude the regularity of the current at a smaller scale. The quantity which plays such a pivotal role is the *excess* of the current T , which we now define:

DEFINITION 4.1. Let T be an integer rectifiable m -dimensional current and π be an m -dimensional plane, oriented by the unit simple m -vector $\vec{\pi}$. The excess of T in the ball $\mathbf{B}_\rho(p)$ with respect to π is the quantity

$$(20) \quad \mathbf{E}(T, \mathbf{B}_\rho(p), \pi) := \frac{1}{\omega_m \rho^m} \int_{\mathbf{B}_\rho(p)} |\vec{T}(x) - \vec{\pi}|^2 d\|T\|(x).$$

The *excess* in $\mathbf{B}_\rho(p)$ is

$$(21) \quad \mathbf{E}(T, \mathbf{B}_\rho(p)) := \min\{\mathbf{E}(T, \mathbf{B}_\rho(p), \pi) : \pi \text{ is an oriented } m\text{-plane}\}.$$

The excess is then an integral measure of the oscillation of the tangent plane to the current. We use the notation A_Σ for the second fundamental form of Σ and the standard $[\cdot]_{0,\alpha}$ for the Hölder seminorms (cf. [52]). Finally, we will often deal with m -dimensional balls in m -dimensional planes π and we introduce therefore the notation $B_r(p, \pi)$ for the set $\mathbf{B}_r(p) \cap (p + \pi)$.

THEOREM 4.2. *Let T be an m -dimensional integer-rectifiable area minimizing current in a C^2 submanifold Σ of dimension $m + \bar{n}$. There are constants $\alpha > 0$, $\varepsilon > 0$ and C , depending only upon m and \bar{n} , such that the following holds. Assume that for some $\rho > 0$ and some m -dimensional plane π we have*

- (a) $\partial T \llcorner \mathbf{B}_{2\rho}(p) = 0$;
- (b) $\Theta(T, p) = Q$ and $\Theta = Q \|T\|$ -a.e. on $\mathbf{B}_{2\rho}(p)$, for some positive integer Q ;
- (c) $\|T\|(\mathbf{B}_{2\rho}(p)) \leq (Q\omega_m + \varepsilon)(2\rho)^m$;
- (d) $E := \mathbf{E}(T, \mathbf{B}_{2\rho}(p), \pi) < \varepsilon$ and $\rho \mathbf{A} := \rho \max_{\Sigma \cap \mathbf{B}_{2\rho}(p)} |A_\Sigma| < \varepsilon$.

Then $T \llcorner \mathbf{B}_\rho(p) = Q \llbracket \Gamma \rrbracket$ for a surface Γ which is the graph of a suitable $C^{1,\alpha}$ function $u : B_r(p, \pi) \rightarrow \pi^\perp$. Moreover $[Du]_{0,\alpha} \leq C(E^{1/2} + \rho \mathbf{A})\rho^{-\alpha}$.

Since in the future we will deal very often with oriented graphs of functions and the corresponding currents, we will use the following notation: $\text{Gr}(u)$ will denote the set-theoretic graph of the function u and \mathbf{G}_u will denote the induced current (for the latter to be well defined we need some regularity for u , which will be discussed in detail later).

4.1. De Giorgi's idea. The crucial point of the proof of Theorem 4.2 is that, under the above assumptions, the current T is close to the graph of an harmonic function. The implementation of this idea is not at all trivial, since it is not at all obvious how we should approximate T with a graph in the first place. Secondly, the various assumptions play a key role and we will see that, in codimension higher than 1, none of them can be dropped: in particular, in higher codimension assumption (b) is crucial to be able to find a (single valued) graph which is sufficiently close to T , cf. Example 5.3.

However, leaving these points aside, assume for the moment that T is the graph \mathbf{G}_u of a Lipschitz function $u : p + \pi \rightarrow \pi^\perp$. For every $\Omega \subset \pi$ we can compute the mass of T in the cylinder $\mathbf{C} := \Omega \times \pi^\perp$ using the area formula:

$$\mathbf{M}(T \llcorner \mathbf{C}) = \int_\Omega \sqrt{1 + |Du|^2 + \sum_{k \geq 2} \sum [\det(M^k(Du))]^2}$$

where with $M^k(Du)$ we denote an arbitrary $k \times k$ minor of Du .

The assumption that the excess $\mathbf{E}(T, \mathbf{B}_\rho(p), \pi)$ is small is similar to the requirement that the derivative Du is small (and certainly it does imply that $|Du|$ is small on most of $B_r(p, \pi)$). The Taylor expansion of the integrand then gives

$$\mathbf{M}(T \llcorner \mathbf{C}) = |\Omega| + \frac{1}{2} \int_\Omega |Du|^2 + O(|Du|^4).$$

In addition it is not difficult to see that $\mathbf{M}(T \llcorner \mathbf{C}) - |\Omega| = \frac{1}{2} \int_{\mathbf{C}} |\vec{T} - \vec{\pi}|^2 d\|T\|$. Thus we can assume that u is rather close to a minimizer of the Dirichlet energy, i.e. that it is close to an harmonic function.

Following similar computations we can compare $\mathbf{E}(T, \mathbf{B}_{2\rho}, \pi)$ to the average integral

$$\oint_{B_{2\rho}(p, \pi)} |Du|^2.$$

and $\mathbf{E}(T, \mathbf{B}_\rho(p))$ to a similar “optimized” quantity

$$\min_A \oint_{B_\rho(p, \pi)} |Du - A|^2 = \oint_{B_\rho(p, \pi)} \left| Du - \oint_{B_\rho(p, \pi)} Du \right|^2.$$

For harmonic functions v we have the following decay estimate, which could be proved using the expansion of the trace $v|_{\partial B_\rho(p, \pi)}$ in spherical harmonics (see [74, Chapter 5, Section 2])

$$(22) \quad \oint_{B_\rho(p, \pi)} \left| Dv - \oint_{B_\rho(p, \pi)} Dv \right|^2 \leq \frac{1}{4} \oint_{B_{2\rho}(p, \pi)} |Dv|^2.$$

We could then hope to transfer such decay to the current in the form

$$(23) \quad \mathbf{E}(T, \mathbf{B}_\rho(p)) \leq 2^{-2+2\delta} \mathbf{E}(T, \mathbf{B}_{2\rho}(p), \pi),$$

where the constant $\delta > 0$ takes into account (quite a few) error terms.

Note however that we could optimize on the plane in the right hand of (23) to achieve

$$(24) \quad \mathbf{E}(T, \mathbf{B}_\rho(p)) \leq 2^{-2+2\delta} \mathbf{E}(T, \mathbf{B}_{2\rho}(p)).$$

In turn this latter estimate would imply that the assumption (d) of Theorem 4.2 holds also in the ball $\mathbf{B}_\rho(p)$. Since all other assumptions are automatically satisfied at any scale smaller than 2ρ (the monotonicity formula plays a crucial role here), we could then iterate the argument to obtain the decay

$$\mathbf{E}(T, \mathbf{B}_r(p)) \leq Cr^{2-2\delta}.$$

Given our intuition that $\mathbf{E}(T, \mathbf{B}_r(p))$ is essentially a mean square oscillation of the tangent plane, the latter decay is a Morrey-type estimate which suggests $C^{1, 1-\delta}$ regularity of the current.

REMARK 4.3. The above analysis leads to guess that the exponent α in Theorem 4.2 can be taken arbitrarily close to 1, at the price of making the threshold ε suitably small and the constant C fairly large. This is indeed the case and the interested reader can check [28, Corollary 2.4 and Appendix A] for a proof which follows closely the argument outlined above.

The subsequent generalizations of Almgren [2, 4], Allard [1] and other authors (cf. for instance [10] and [65]) of De Giorgi’s ε -regularity statement have lost the feature of implying directly (24) and seem to need a more careful argument to reach the conclusion that α is arbitrarily close to 1: a sacrifice to flexibility, since the latter results can be applied to much more general objects and situations. Moreover, any $C^{1, \alpha}$ graph which is stationary for the area functional enjoys higher regularity as a simple consequence of the Schauder estimates. Thus a loss of sharpness in the exponent α would anyway play no important role in the classical results.

We will see however that in codimension higher than 1 De Giorgi's original (variational) idea and its important byproduct that the decay of the excess can be assumed to be "almost quadratic" play both a crucial role in the proof of Theorem 3.4.

4.2. First consequences of the ε -regularity theorem. It is rather simple to see that the conditions (a), (c) and (d) will be met at a sufficiently small radius ρ as soon as $p \in \text{spt}(T) \setminus \text{spt}(\partial T)$ and there is at least *one* flat tangent cone at p . However condition (b) discriminates severely between the codimension 1 case ($\bar{n} = 1$) and the higher codimensions. Indeed, Proposition 2.18 shows that in codimension 1 a current without boundary can be described as a "superposition" of boundaries of finitely many Caccioppoli sets E_i (in the case of currents in Riemannian manifolds, a similar statement holds as well).

In fact from Proposition 2.18(i) it is not difficult to conclude that, when T is area minimizing, each current $\partial \llbracket E_i \rrbracket$ in the decomposition is as well area minimizing. Intuitively one does not expect integer multiplicities higher than 2 for boundaries of sets, at least not at most points: for instance it is not difficult to prove that the density $\Theta(\partial \llbracket E \rrbracket, p)$ equals 1 at $\|\partial \llbracket E \rrbracket\|$ -a.e. p when E is a set of finite perimeter. Ultimately it is then possible to prove the following corollary (cf. [68, Section 7.37]).

COROLLARY 4.4. *If T is an area minimizing current of dimension m in a C^2 submanifold Σ of dimension $m + 1$, then any point p at which there is a flat tangent cone is a regular point.*

In particular we conclude immediately that $\|T\|(\text{Sing}(T)) = 0$ and, by Corollary 3.10, that $\mathcal{H}^m(\text{Sing}(T)) = 0$. In higher codimension the arguments above do not apply and we will see that indeed Corollary 4.4 fails. It is nonetheless possible to conclude a weaker statement because, by the upper semicontinuity of the density and an elementary topological argument, Assumption (b) of Theorem 4.2 can still be verified in a rather large set.

COROLLARY 4.5 (cf. [68, Theorem 36.2]). *If T is an area minimizing current of dimension m in a C^2 submanifold Σ of dimension larger than $m + 1$, then $\text{Reg}(T)$ is dense in $\text{spt}(T) \setminus \text{spt}(\partial T)$.*

Indeed this statement has been recently extended to any Hilbert space, cf. [7].

Remarkably, Corollary 4.5 was the best regularity result available before the appearance of Almgren's manuscript [5] with its proof of Theorem 3.4. In fact this is the current situation for *stationary* integer rectifiable m -dimensional varifolds: Allard's theorem gives the regularity up to a meager closed set, even in codimension 1, and this is up to now the best regularity result available in the literature for stationary objects. In particular it is not known that the singular set is \mathcal{H}^m -negligible, not even in the simplest setting of stationary 2-dimensional varifolds in 3 dimensions. For stable hypersurfaces a rather satisfactory theory is instead available thanks to the

pioneering works of Schoen - Simon - Yau [66] and Schoen - Simon [64] and to the recent ones of Wickramasekera [80].

4.3. Full regularity for $m \leq 6$ and $\bar{n} = 1$: Simons' theorem. Let us now focus on the case of codimension $\bar{n} = 1$. Corollary 4.4 naturally leads to discuss the existence of area minimizing (hyper-)cones which are not flat. On the one hand, their nonexistence would imply via Corollary 4.4 that $\text{Sing}(T)$ is empty. On the other hand, it is rather easy to see that non-flat area minimizing cones must be necessarily singular at the origin.

The investigations upon these questions were started by De Giorgi and Fleming who could show full regularity for $m = 2$, cf. [48] and [22]. Moreover, De Giorgi showed that the problem of deciding whether every codimension 1 area minimizing cone in \mathbb{R}^{m+1} is flat is equivalent to decide whether any entire minimal (hyper-) graph in \mathbb{R}^{m+2} is affine, the so-called Bernstein problem. The result of De Giorgi and Fleming was subsequently improved by Almgren ([3], $m = 3$) and finally by Simons in [71] to show full regularity for $m \leq 6$.

It must be noticed the following: if one shows that in dimension $m \leq m_0$ there is no singular area minimizing hypercone, then any area minimizing hypercone of dimension $m + 1$ is necessarily regular except possibly at the origin: thus the cross-section is a minimal (i.e. stationary) embedded hypersurface of the standard $m + 1$ -dimensional sphere (this fact will be discussed in a couple of sections as the starting point of the so-called *Federer's reduction argument*). On the other hand, any cone whose cross section is a minimal hypersurface of the standard $m + 1$ -dimensional sphere is stationary for the area functional in \mathbb{R}^{m+2} . However if such a cone is area minimizing, then it must also be *stable*, in the sense of Definition 3.6. The famous theorem of Simons is the following statement.

THEOREM 4.6 (Simons). *Let $2 \leq m \leq 6$. Any stable minimal hypersurface of \mathbb{R}^{m+1} which is a cone over a minimal submanifold of $\partial\mathbf{B}_1(0)$ is necessarily an m -dimensional plane.*

4.4. Simons' cone and the Theorem of Bombieri-De Giorgi-Giusti. In his celebrated paper [71] Simons provided also an example which showed the optimality of his theorem. More precisely he showed that the cone over $\mathbb{S}^3 \times \mathbb{S}^3 \subset \mathbb{S}^7$, namely

$$(25) \quad S := \{x \in \mathbb{R}^8 : x_1^2 + x_2^2 + x_3^2 + x_4^2 = x_5^2 + x_6^2 + x_7^2 + x_8^2\}$$

is stationary and *stable*. The surface is usually called Simons' cone in the literature which followed [71]. Later Bombieri, De Giorgi and Giusti in [11] showed that S is indeed an area minimizing cone and were thus able to settle the Bernstein problem in all dimensions.

THEOREM 4.7 (Bombieri-De Giorgi-Giusti). *S in (25) is an area minimizing current in \mathbb{R}^8 and therefore for any $n \geq 8$ there are functions $u : \mathbb{R}^n \rightarrow \mathbb{R}$ which satisfy the minimal surface equation and are not affine.*

We refer the reader to the recent paper [39] for an elegant and simple proof of Theorem 4.7

4.5. $m \geq 7$ and Federer's reduction argument. We have already mentioned that any area minimizing cone of dimension 7 in \mathbb{R}^8 is necessarily smooth outside the origin because the only area minimizing cones in \mathbb{R}^7 are (multiple of) hyperplanes. We wish to illustrate this statement as an introduction to a powerful idea of Federer, which has been applied to several problems in geometric analysis.

Consider an m -dimensional current T without boundary in $\mathbb{R}^{m+\bar{n}}$ and assume it is invariant under translation in a given direction v , which for simplicity we set to be the first vector e_1 of the standard orthonormal basis: we then say that S “splits off a line”. It is not difficult to see that, in this case, the current T is the product of a current S of dimension $m-1$ in $\mathbb{R}^{m-1+\bar{n}}$ and a line, namely the 1-dimensional current on \mathbb{R} which is given by integration of the top 1-form and we denote by $[\mathbb{R}]$: $[\mathbb{R}](f dx_1) = \int f(x_1) dx_1$. It is rather intuitive that $[\mathbb{R}] \times S$ is locally area minimizing (i.e. its restriction to bounded open sets is area minimizing) if and only S is.

Consider now an area minimizing cone T in \mathbb{R}^8 and let $p \in \text{spt}(T) \setminus \{0\}$. Then it is not difficult to see that any tangent cone to T at p splits off a line, because it will be invariant under translations in the direction p . We summarize our discussion in the following key

LEMMA 4.8 (cf. [68, Lemma 35.5 and proof of Theorem 35.3]). *If S_0 is a tangent cone to an area minimizing m -dimensional cone S in $\mathbb{R}^{m+\bar{n}}$ at a point $p \neq 0$, then S_0 splits off a line, namely $S_0 = [\mathbb{R}] \times Z$ for some $(m-1)$ -dimensional area minimizing cone Z in $\mathbb{R}^{m-1+\bar{n}}$.*

Fix now an area minimizing hypercone S of dimension 7 (i.e. $m = 7$ and $\bar{n} = 1$). Fix S_0 as in the lemma above and observe that Z is an area minimizing hypercone of dimension 6: as such we conclude from Simons' theorem that S_0 must then be a plane. Thus we can apply Corollary 4.4 and conclude that either S is regular or it has an isolated singularity at the origin.

This in turn shows that the singular set $\text{Sing}(T)$ of an area minimizing 7-dimensional current T in \mathbb{R}^8 is discrete (cf. [45, Section 5.4.17]). Indeed, let $p \in \text{Sing}(T)$ and consider any tangent cone S at p . Let $r > 0$ and $T_{p,r}$ be a rescaling of S which is sufficiently close to S . By the regularity of S , there is a radius $\rho > 0$ such that the excess of S in $\mathbf{B}_{2\rho}(q)$ is smaller than $\varepsilon/2$ at any point q with $|q| = 1$. If $T_{p,r}$ is sufficiently close to S , the excess of $T_{p,r}$ in $\mathbf{B}_{2\rho}(q)$ will be smaller than ε and we can apply the ε -regularity theorem to conclude that $T_{p,r}$ is regular in $\mathbf{B}_\rho(q)$. In fact the rigorous argument must treat also the conditions (b) and (c) of Theorem 4.2 on the multiplicity: these can be settled thanks to the codimension 1 assumption. Our discussion leads naturally to the following statement, which requires just an appropriate compactness argument on the set of tangent cones at p : there are positive constants $r, \rho > 0$ such that, if $0 < |p-q| < r$, then T is regular in $\mathbf{B}_{\rho|p-q|}(q)$.

We in fact highlight a general important principle behind the above discussion, a very well-known and widely used effect of ε -regularity statements:

PROPOSITION 4.9. *Let T , Σ , T_k and Σ_k be as in Theorem 3.7 and assume that the codimension \bar{n} is 1. If T is regular in Ω , then for any open set $\Gamma \subset \subset \Omega$, T_k is regular in Γ for k large enough.*

For the sake of our future discussions we will rephrase the proposition above in the following equivalent way, underlying that “singularities persist in the limit”: we will stress later on that this persistence can be seen as the major difference between the codimension 1 and the higher codimension.

PROPOSITION 4.10 (Persistence of singularities in codimension 1). *Let Ω , T , Σ , T_k and Σ_k , be as in Theorem 3.7 and assume that the codimension \bar{n} is 1. If $p_k \in \text{Sing}(T_k)$ and $p_k \rightarrow p \in \Omega$, then $p \in \text{Sing}(T)$.*

The basic ideas that singularities must persist in the limit and that repeated “blow-ups” reduce the dimension have been used by Federer to give a first rough description of $\text{Sing}(T)$ when $\bar{n} = 1$ and $m > 7$. He used the resulting “scheme”, called *Federer’s reduction argument* (cf. [46] and [68, Appendix A]), to prove the following

THEOREM 4.11 (Federer). *Let m , \bar{n} , T and Σ be as in Theorem 3.3. If $m = 7$ then $\text{Sing}(T)$ is discrete. If $m \geq 8$, then $\mathcal{H}^{m-7+\alpha}(\text{Sing}(T)) = 0$ for every $\alpha > 0$, namely $\text{Sing}(T)$ has Hausdorff dimension at most $m - 7$.*

The following is a rough sketch of Federer’s argument. Assume the existence of an area minimizing current T of dimension $m \geq 8$ in \mathbb{R}^{m+1} such that, for some positive α , $\mathcal{H}^{m-7+\alpha}(\text{Sing}(T)) > 0$. An elementary measure theoretic argument shows the existence of many points p for which

$$\limsup_{r \downarrow 0} \frac{\mathcal{H}^{m-7+\alpha}(\text{Sing}(T) \cap \mathbf{B}_r(p))}{r^{m-7+\alpha}} > 0$$

(in fact the above property holds for $\mathcal{H}^{m-7+\alpha}$ -a.e. $p \in \text{Sing}(T)$, cf. [68, Theorem 3.2]).

We can thus assume the existence of an area minimizing cone T_0 and of a subsequence of rescalings T_{p,r_k} converging to it for which

$$\mathcal{H}^{m-7+\alpha}(\text{Sing}(T_{p,r_k}) \cap \bar{\mathbf{B}}_1) \geq \eta$$

for some positive η . After taking a further subsequence, not relabeled, we can assume that $\text{Sing}(T_{p,r_k}) \cap \bar{\mathbf{B}}_1$ converges to some compact set F in the Hausdorff distance: Proposition 4.9 implies then that $F \subset \text{Sing}(T_0)$. We would like to infer that $\mathcal{H}^{m-7+\alpha}(\text{Sing}(T_0)) \geq \eta > 0$. However the Hausdorff measures are *not* upper semicontinuous under convergence in the Hausdorff distance. This is resolved by using a suitable variant, the $\mathcal{H}_\infty^{m-7+\alpha}$ measure: the latter turns out to be upper semicontinuous while it has the same null-sets as the $\mathcal{H}^{m-7+\alpha}$ measure (and the same “density property” used above; cf. [68, Appendix A]).

Summarizing, from the existence of an area minimizing current T of dimension m with a singular set of positive $\mathcal{H}^{m-7+\alpha}$ measure we have concluded the existence of an m -dimensional area minimizing cone T_0 with the same property. We can now repeat this argument again with T_0 in place of T , blowing up at some point q distinct from the origin. We conclude that, for some appropriate tangent cone S to T_0 at q , $\mathcal{H}^{m-7+\alpha}(\text{Sing}(S)) > 0$. On the other hand S splits off a line and it is easy to see that this implies the existence of an area minimizing current Z of dimension $m - 1$ in \mathbb{R}^m such that $\mathcal{H}^{m-8+\alpha}(\text{Sing}(Z)) > 0$.

The process can be iterated until we end up with a 7-dimensional area minimizing current \bar{T} in \mathbb{R}^8 which has a singular set of positive \mathcal{H}^α measure. Since $\alpha > 0$, this contradicts what we have already proved, namely that in this case \bar{T} has (at most) isolated singularities.

4.6. Simon's rectifiability result. We complete our survey of the regularity results in codimension 1 by mentioning Simon's spectacular achievement: combining his fundamental theorem about the uniqueness of tangent cones at isolated singularities with several additional innovative ideas, he was able to show that, when $\bar{n} = 1$, $\text{Sing}(T)$ can be covered, up to a set of \mathcal{H}^{m-7} -measure zero, by a countable collection of C^1 $(m - 7)$ -dimensional submanifolds, cf. [69]. A new proof of Simon's theorem, which avoids the discussion of the uniqueness of tangent cones at isolated singularities, has been very recently found by Naber and Valtorta, see [57]. This is till now the best description available for the behavior of the singular set in codimension 1.

5. Federer's theorem and the failure of ε -regularity in codimension $\bar{n} \geq 2$

5.1. Holomorphic subvarieties as area minimizing currents. We start by recalling that holomorphic subvarieties of \mathbb{C}^{k+j} , namely zeros of holomorphic maps $u : \mathbb{C}^{k+j} \rightarrow \mathbb{C}^j$ (k and j being, respectively, the complex dimension and codimension of the variety) can be given a natural orientation. In what follows we identify \mathbb{C}^{k+j} with \mathbb{R}^{2k+2j} in the usual way: if z_1, \dots, z_{k+j} are complex coordinates and $x_j = \text{Re } z_j$, $y_j = \text{Im } z_j$, we let $x_1, y_1, \dots, x_{k+j}, y_{k+j}$ be the standard coordinates of \mathbb{R}^{2k+2j} . Recall then that an holomorphic subvariety Γ of \mathbb{C}^{k+j} of complex dimension k is a (real analytic) submanifold of $\mathbb{R}^{2k+2j} \setminus \text{Sing}(\Gamma)$ of (real) dimension $m = 2k$, where $\text{Sing}(\Gamma)$ is an holomorphic subvariety of complex dimension $k - 1$.

Furthermore, at each point $p \in \Gamma \setminus \text{Sing}(\Gamma)$, the (real) tangent $2k$ -dim. plane $T_p\Gamma$ can be identified with a complex k -dimensional plane of \mathbb{C}^n . If v_1, \dots, v_k is a complex basis of $T_p\Gamma$, we can then define a canonical orientation for $T_p\Gamma$ using the simple $2k$ -vector

$$\text{Re } v_1 \wedge \text{Im } v_1 \wedge \dots \wedge \text{Re } v_k \wedge \text{Im } v_k.$$

This allows us to define the current $[\Gamma]$ by integrating forms over the oriented submanifold $\Gamma \setminus \text{Sing}(\Gamma)$. It is also easy to check that $\partial [\Gamma] = 0$, the

reason being that the “singular set” $\text{Sing}(\Gamma)$ is a set of (locally) finite \mathcal{H}^{2k-2} measure.

The discussion can be “localized” to holomorphic subvarieties in open subsets Ω of \mathbb{C}^{k+j} (and more generally in complex hermitian manifolds). Note also that, if Ω' is a bounded open subset of the domain Ω where Γ is defined, then $[\Gamma]$ has finite mass in Ω' and it is thus an integer rectifiable current. The following fundamental observation is due to Federer and is based on a classical computation of Wirtinger ([81]).

THEOREM 5.1 (Federer, cf. [45, Section 5.4.19]). *Let $\Gamma_1, \dots, \Gamma_N$ be holomorphic subvarieties of complex dimension k in $\Omega \subset \mathbb{C}^{k+j}$ and let k_1, \dots, k_N be positive integers. Then the current $T := k_1 [\Gamma_1] + \dots + k_N [\Gamma_N]$ is area minimizing in the sense that $\mathbf{M}(T \llcorner \Omega') \leq \mathbf{M}(T \llcorner \Omega' + \partial S)$ for any open $\Omega' \subset \subset \Omega$ and any $2k+1$ -dimensional integral current S with $\text{spt}(S) \subset \Omega'$.*

Indeed the above theorem holds in general Kähler manifolds, cf. [45, 5.4.19].

5.2. Branching phenomena. Before giving an idea of why Theorem 5.1 holds we want to illustrate the deep consequences that it has in the regularity theory for area minimizing currents in codimension higher than 1. Holomorphic subvarieties give easy counterexamples to Corollary 4.4 when $\bar{n} > 1$: assumption (b) in Theorem 4.2 is absolutely crucial in this case. As a byproduct even Proposition 4.10 fails and singularities might disappear in the limit when we deal with sequences of area minimizing currents: in the rest of this note we will see that the core difficulty in the proof of Theorem 3.4 is precisely this phenomenon of “disappearance of singularities”. We illustrate these points with three explicit examples.

EXAMPLE 5.2. Let $\delta > 0$ be a small number and consider the holomorphic curve

$$\Gamma_\delta := \{(z, w) \in \mathbb{C}^2 : z^2 = \delta w\}$$

and the plane

$$(26) \quad \pi := \{(z, w) \in \mathbb{C}^2 : z = 0\}.$$

There is no neighborhood of 0 where Γ_δ is the graph of a function $z = f(w)$, in spite of the fact that $\mathbf{E}([\Gamma_\delta], \mathbf{B}_1(0), \pi)$ converges to 0 as $\delta \downarrow 0$. In fact the conclusion of Theorem 4.2 does not apply: although each Γ_δ is smooth and it is graphical in $\mathbf{B}_\rho(0)$ for any ρ , there is no uniform control of the $C^{1,\alpha}$ norm of the graph in terms of the excess. Observe that the currents Γ_δ do not satisfy the condition (c) in Theorem 4.2, although they satisfy (a), (b) and (d).

EXAMPLE 5.3. Consider the holomorphic curve

$$\Gamma := \{(z, w) \in \mathbb{C}^2 : z^2 = w^3\}.$$

The origin belongs to $\text{Sing}([\Gamma])$. On the other hand:

- The unique tangent cone at 0 is given by $2[\pi]$ for π as in (26).

- The density of $\llbracket \Gamma \rrbracket$ equals 2 at 0;
- $\lim_{r \downarrow 0} \mathbf{E}(\llbracket \Gamma \rrbracket, \mathbf{B}_r(0), \pi) = 0$.

Therefore:

- Corollary 4.5 is false for 2-dimensional area minimizing currents in \mathbb{R}^4 : Γ is singular at the origin in spite of the existence of a flat tangent cone there.
- Again Theorem 4.2 does not apply in any ball $\mathbf{B}_{2\rho}(0)$. Note however that the only missing assumption is (b): the density $\Theta(\llbracket \Gamma \rrbracket, p)$ equals 1 at every point $p \in \Gamma \setminus \{0\}$ and equals 2 at $p = 0$.
- Proposition 4.10 fails for 2-dimensional area minimizing currents in \mathbb{R}^4 . Indeed 0 is a singular point for $\llbracket \Gamma \rrbracket_{0,r}$ for every positive $r > 0$. On the other hand $\llbracket \Gamma \rrbracket_{0,r} \rightarrow 2 \llbracket \pi \rrbracket$ and thus 0 is *not* a singular point of the limit: the singularity “has disappeared”.

EXAMPLE 5.4. Consider finally the holomorphic curve

$$\Xi := \{(z, w) \in \mathbb{C}^2 : (z - w^2)^2 = w^{2015}\}.$$

All the considerations valid for the holomorphic curve Γ of Example 5.3 are also valid for Ξ . Ξ does not add much for the moment to our discussion, but it will play a crucial role later: observe that 0 is a singular point in spite of the fact that Ξ is an almost imperceptible perturbation of the smooth current $2 \llbracket \{z = w^2\} \rrbracket$.

We close this section by remarking that Theorem 5.1 gives also a great abundance of singular area minimizing cones in higher codimension: the zero set of any homogeneous polynomial $P(z_1, \dots, z_{k+1})$ in $k+1$ complex variables is an area minimizing cone of dimension $2k$ in \mathbb{R}^{2k+2} . More generally, for any projective subvariety of $\mathbb{P}^k \mathbb{C}$ with complex dimension j we can construct a corresponding area-minimizing cone in \mathbb{R}^{2k+2} of dimension $2j + 2$. These cones are singular except when the corresponding algebraic subvarieties are affine. The easiest example of a singular area minimizing cone is thus the union of an arbitrary number of complex lines in \mathbb{C}^2 . Such cones might however be considered “mildly” singular: in \mathbb{C}^3 the generic cone associated to a projective curve of $\mathbb{P}^2 \mathbb{C}$ has a singular set which behaves in rather complicated way.

5.3. Calibrations and the proof of Theorem 5.1. We illustrate here the simple, yet deep, principle lying behind Theorem 5.1. Recall first the notion of comass of a form, given in Definition 2.3.

DEFINITION 5.5 (Calibrations, cf. [55]). A *calibration* ω is a closed m -form such that $\|\omega\|_c \leq 1$. An integer rectifiable current T is said to be calibrated by a calibration ω if $\langle \omega_p, \vec{T}(p) \rangle = 1$ for $\|T\|$ -a.e. p .

Observe in particular that $\mathbf{M}(T) \geq T(\omega)$ whenever ω is a calibration and that the equality sign holds if and only if T is calibrated by ω . The following is then a trivial fact.

LEMMA 5.6. *If T is calibrated by a calibration ω , then T is an area minimizing current.*

PROOF. Let S be an $(m+1)$ -dimensional integral current. Then

$$\mathbf{M}(T) = T(\omega) = T(\omega) + S(d\omega) = (T + \partial S)(\omega) \leq \mathbf{M}(T + \partial S). \quad \square$$

Holomorphic subvarieties are the primary example of calibrated currents and this observation dates back essentially to Wirtinger. More precisely, if $z_\ell = x_\ell + iy_\ell$ are the standard coordinates in \mathbb{C}^{k+j} , consider the Kähler form

$$\omega := dx_1 \wedge dy_1 + \dots + dx_{k+j} \wedge dy_{k+j}.$$

Wirtinger's theorem can then be stated in the following form

THEOREM 5.7 (Wirtinger, cf. [81]). *If ω is the Kähler form and*

$$\omega^k = \frac{1}{k!} \underbrace{\omega \wedge \dots \wedge \omega}_{k \text{ times}}$$

then ω^k is a calibration. Moreover, $\langle \omega^k, v_1 \wedge \dots \wedge v_{2k} \rangle = |v_1 \wedge \dots \wedge v_{2k}|$ if and only if v_1, \dots, v_{2k} is a positively oriented $(\mathbb{R}-)$ base of a complex plane.

Calibrations and calibrated submanifolds are a rich source of interesting geometries: we refer the reader to [55] for several important examples.

6. Almgren's stratification

From now on we will mostly have in mind the case of codimension \bar{n} strictly larger than 1 and we proceed with the investigations leading to Theorem 3.4. One first simple step in the analysis of the singular set of the area minimizing currents is an elegant generalization of Federer's reduction argument.

We start by taking a second look at Federer's argument, roughly sketched in Section 4.5. Given an area minimizing m -dimensional cone S we define its *spine* as the vector space V of maximal dimension for which S can be written as $S' \times \llbracket V \rrbracket$, where S' is an area minimizing cone of dimension $m - \dim(V)$. Equivalently, V is the subset of those vectors v such that S is invariant under translations in direction v and it is a simple exercise (using the monotonicity formula) to show that V can be characterized as the subset of those points $p \in \text{spt}(S)$ such that $\Theta(S, p) = \Theta(S, 0)$ or also as the subset of those points q such that $S_{q,1}$ is a cone with vertex 0 (cf. [68, Proof of Lemma 35.5]).

At the intuitive level it is clear that S must have a certain “asymmetry” in the directions which are transversal to V . The dimension of the spine of S is called the *building dimension of the cone S* (cf. [79]). Note that such building dimension equals m if and only if S is an integer multiple of an m -dimensional plane, namely if and only if S is flat.

Consider now the situation where p and q are two points in the support of an area minimizing current T such that at a scale r comparable to $|p - q|$, $T_{p,r}$ is close to some cone S and $T_{q,r}$ is close to a certain other cone S' . Then

$p - q$ is “almost in the spine of T ”, because $(T_{q,r})_{p-q,1} = T_{p,r}$ is “almost a cone”. If the building dimension of S is a certain number \bar{m} and V is the corresponding spine of S , we therefore conclude that $(p - q)/|p - q|$ must be very close to V . Summarizing:

- If at a given point p the current T is sufficiently close to a cone at scale r , then all the points q surrounding p (i.e. at distance at most r) and at which the current is close to some cone at scale r must be contained in a neighborhood of size εr of the spine V of T .

This suggests to introduce the following stratification of points in the support of T :

DEFINITION 6.1. A point $p \in \text{spt}(T)$ belongs to the stratum $\mathcal{S}_k(T)$ if every tangent cone to T at p has building dimension at most k and if there is *at least one* tangent cone to T at p with building dimension k .

Moreover, after “discretizing” all possible scales, we can subdivide further the stratum $\mathcal{S}_k(T)$ in a countable number of subsets according to the scale at which the current T starts looking sufficiently close to a cone. The consideration above implies that each such subset is contained, at all scales smaller than a given one, in a small neighborhood of some k -dimensional plane. It is therefore not difficult to imagine that we can bound the Hausdorff dimension of $\mathcal{S}_k(T)$ with k .

The discussion above is essentially the content of Almgren’s generalization of Federer’s argument, which we state in the following theorem.

THEOREM 6.2 (Almgren’s stratification, cf. [79]). *For any given area minimizing current T the stratum $\mathcal{S}_k(T)$ has Hausdorff dimension at most k and $\mathcal{S}_0(T)$ is a discrete set.*

Observe that the discussion of Section 4.5 proves that:

- (F) Given any area minimizing cone S of codimension $\bar{n} = 1$, either such cone is a multiple of an m -dimensional plane, or its building dimension is at most $m - 7$.

As a corollary we conclude that for $\bar{n} = 1$ the strata

$$\mathcal{S}_{m-1}(T), \mathcal{S}_{m-2}(T), \dots, \mathcal{S}_{m-6}(T)$$

are all empty. Next, at any point p in the top stratum $\mathcal{S}_m(T)$ there is a flat tangent cone and thus, by Corollary 4.4, we actually know that $\mathcal{S}_m(T) = \text{Reg}(T)$ (we stress again that this holds only under the assumption that $\bar{n} = 1$: Example 5.3 gives a counterexample as soon as $\bar{n} = 2$ and $m = 2$). We therefore conclude that $\text{Sing}(T) = \mathcal{S}_0(T) \cup \dots \cup \mathcal{S}_{m-7}(T)$ and thus Theorem 3.3(ii)&(iii) is a corollary of Theorem 6.2.

Unfortunately from Section 5 we know that the identity $\text{Reg}(T) = \mathcal{S}_m(T)$ does not hold anymore when the codimension $\bar{n} > 1$. On the other hand we surely have $\text{Reg}(T) \subset \mathcal{S}_m(T)$. We could call “branch points” for T those points $p \in \mathcal{S}_m(T) \setminus \text{Reg}(T)$. The major concern in the rest of the note will be to estimate the Hausdorff dimension of $\mathcal{S}_m(T) \setminus \text{Reg}(T)$. A simple

consequence of Theorem 6.2 is that, in order to prove Theorem 3.4(ii), the “only” concern is truly to bound the Hausdorff dimension of the set of branch points by $m - 2$, because of the following lemma.

LEMMA 6.3 (cf. [68, Theorem 35.3]). *The stratum $\mathcal{S}_{m-1}(T)$ is empty in any codimension \bar{n} .*

PROOF. It suffices to show that the building dimension of an area minimizing cone of dimension m cannot be $m - 1$. On the other hand if there were an m -dimensional area minimizing cone with building dimension $m - 1$, then there would be a 1-dimensional cone S' which is not flat. Now, it is rather easy to show that any 1-dimensional cone S' can be written as

$$S' = \sum_{i=1}^Q [\ell_i] - \sum_{j=Q+1}^{2Q} [\ell_j] ,$$

where each ℓ_k is a half-line starting at the origin oriented so that $\partial [\ell_k] = [0]$. Observe moreover that $Q = \Theta(S, 0)$. If we choose one i between 1 and Q and one j between $Q+1$ and $2Q$, we can write $S' = Z + S''$ where $S'' = [\ell_i] - [\ell_j]$: since $\partial S'' = \partial Z = 0$ and $\|S''\|(\mathbf{B}_1) + \|Z\|(\mathbf{B}_1) = \|S'\|(\mathbf{B}_1)$, S'' must be itself area minimizing. On the other hand it is very simple to show that S'' is area minimizing if and only if ℓ_i and ℓ_j form, together, a single straight line ℓ passing through the origin. But then $S' = Q [\ell]$, contradicting the assumption that S' is singular at 0. \square

The proof of Theorem 6.2 is rather elementary and “soft”. In spite of this the idea is powerful and can be applied to several different problems in geometric analysis; for instance, we refer the reader to Simon’s work on the singularities of harmonic maps, [70], to White’s far-reaching generalization of Theorem 6.2 and its applications to the mean-curvature flow, [79], and to recent results about Riemannian manifolds with one-sided curvature bounds, see for instance [15]. Recently, in a series of works (cf. [16, 14, 17]), the method of Almgren has been extended to deal with the Minkowski content, see also [49] for an abstract general version of this.

We finally mention that the cones with building dimension $m - 2$ can be actually further characterized: it is not difficult to see that such cones are necessarily unions of multiples of m -dimensional planes. The spines of such cones are $(m - 2)$ -dimensional subspaces. Due to the remarkable work of White, [77], when $m = 2$ there is one such unique tangent cone at every point $p \in \mathcal{S}_0(T)$. However, for $m \geq 3$ the same uniqueness result is not yet proved and in fact it is not even known whether at points $p \in \mathcal{S}_m(T) \setminus \text{Reg}(T)$ the flat tangent cone is the unique one!

7. Multiple valued functions minimizing the Dirichlet energy

As already noticed, in codimension 1 the regularity in a neighborhood of a point with integer multiplicity Q where at least one tangent cone is flat can be reduced to the case of multiplicity $Q = 1$, whereas the discussions of

Section 5 show that this reduction is impossible in codimension larger than 1. Indeed, in the Examples 5.3 and 5.4 even the starting point of De Giorgi's strategy as described in Section 4.1 fails dramatically: no matter how small is the neighborhood U of the origin that we choose, it is simply not possible to approximate efficiently the corresponding current T in U with the graph of a (single valued) function. However, in each of these examples the current turns out to be a "multivalued" graph, where the number of values is in fact determined by the multiplicity $Q = \Theta(T, 0)$. This discussion motivates the starting idea of Almgren's monograph: in order to go beyond an Allard's type statement (namely regularity in a dense relatively open subset of $\text{spt}(T) \setminus \text{spt}(\partial T)$) we need to develop an efficient theory for "multiple valued functions" minimizing a suitable generalization of the Dirichlet energy, where we can (and we will) consider the multiplicity to be a constant preassigned positive integer Q .

7.1. The metric space of unordered Q -tuples. The obvious model case to keep in mind is the following. Given two integers k, Q with $\text{MCD}(k, Q) = 1$, look at the set valued map which assigns to each point $z \in \mathbb{C}$ the set $M(z) := \{w^k : w^Q = z\} \subset \mathbb{C}$. Obviously for each z we can choose some arbitrary ordering $\{u_1(z), \dots, u_Q(z)\}$ of the elements of the set $M(z)$. However, it is not possible to do it in such a way that the resulting "selection maps" $z \mapsto u^i(z)$ are continuous: even at the local level, this is impossible in every neighborhood of the origin.

Our example motivates the following definition. Given an integer Q we define a Q -valued map from a set $E \subset \mathbb{R}^m$ into \mathbb{R}^n as a function which to each point $x \in E$ associates an unordered Q -tuple of vectors in \mathbb{R}^n . Following Almgren, we consider the group \mathcal{P}_Q of permutations of Q elements and we let $\mathcal{A}_Q(\mathbb{R}^n)$ be the set $(\mathbb{R}^n)^Q$ modulo the equivalence relation

$$(v_1, \dots, v_Q) \equiv (v_{\pi(1)}, \dots, v_{\pi(Q)}) \quad \forall \pi \in \mathcal{P}.$$

Hence a multiple valued map is simply a map taking values in $\mathcal{A}_Q(\mathbb{R}^n)$. There is a fairly efficient formulation of this definition which will play a pivotal role in our discussion, because the set $\mathcal{A}_Q(\mathbb{R}^n)$ can be naturally identified with a subset of the set of measures (cf. [5] and [29, Definition 0.1]).

DEFINITION 7.1 (Unordered Q -tuples). Denote by $\llbracket P_i \rrbracket$ the Dirac mass in $P_i \in \mathbb{R}^n$. Then,

$$\mathcal{A}_Q(\mathbb{R}^n) := \left\{ \sum_{i=1}^Q \llbracket P_i \rrbracket : P_i \in \mathbb{R}^n \text{ for every } i = 1, \dots, Q \right\}.$$

Observe that with this definition each element of $\mathcal{A}_Q(\mathbb{R}^n)$ is in fact a 0-dimensional integral current. This set has also a natural metric structure; cf. [5] and [29, Definition 0.2] (the experts will recognize the well-known Wasserstein 2-distance, cf. [76]).

DEFINITION 7.2. For every $T_1, T_2 \in \mathcal{A}_Q(\mathbb{R}^n)$, with $T_1 = \sum_i \llbracket P_i \rrbracket$ and $T_2 = \sum_i \llbracket S_i \rrbracket$, we set

$$(27) \quad \mathcal{G}(T_1, T_2) := \min_{\sigma \in \mathcal{P}_Q} \sqrt{\sum_i |P_i - S_{\sigma(i)}|^2}.$$

REMARK 7.3. Since we will often need to compute $\mathcal{G}(T, Q \llbracket 0 \rrbracket)$ we introduce the special notation $|T|$ for the latter quantity. Observe, however, that $\mathcal{A}_Q(\mathbb{R}^n)$ is *not* a linear space except for the special case $Q = 1$: the map $T \rightarrow |T|$ is not a norm.

7.2. The generalized Dirichlet energy. Using the metric structure on $\mathcal{A}_Q(\mathbb{R}^n)$ one defines obviously measurable, Lipschitz and Hölder maps from subsets of \mathbb{R}^m into $\mathcal{A}_Q(\mathbb{R}^n)$. However, if we want to approximate area minimizing currents with multiple valued functions and “linearize” the area functional in the spirit of De Giorgi, we need to define a suitable concept of Dirichlet energy. We will now show how this can be done naturally. However, the approach outlined below is not the one of Almgren.

Consider again the model case of $Q = 2$ and assume $u : \Omega \rightarrow \mathcal{A}_2(\mathbb{R}^n)$ is a Lipschitz map. If, at some point x , $u(x) = \llbracket P_1 \rrbracket + \llbracket P_2 \rrbracket$ is “genuinely 2-valued”, i.e. $P_1 \neq P_2$, then there exist obviously a ball $B_r(x) \subset \Omega$ and two Lipschitz functions $u_1, u_2 : B_r(x) \rightarrow \mathbb{R}^n$ such that $u(y) = \llbracket u_1(y) \rrbracket + \llbracket u_2(y) \rrbracket$ for every $y \in B_r(x)$ (in this and similar situations, we will then say that there is a *regular selection* for u in $B_r(x)$, cf. [29, Definition 1.1]). For each separate function u_i , the classical Theorem of Rademacher ensures the differentiability almost everywhere.

Recall that our ultimate goal is to define the Dirichlet energy so that it is a suitable approximation of the area of the graph of u . The “graph of u over $B_r(x)$ ” is simply the union of the graphs of the two functions u_i . When the derivatives Du_i are close to 0, the area of each graph is close to

$$\int_{B_r(x)} \left(1 + \frac{1}{2} |Du_i|^2 \right).$$

Thus, the only suitable definition of Dirichlet energy of u on the domain $B_r(x)$ is given by

$$\int_{B_r(x)} |Du|^2 := \int_{B_r(x)} (|Du_1|^2 + |Du_2|^2).$$

By an obvious localization procedure, this definition can be extended to the (open!) set $\Omega_2 \subset \Omega$ where u is “genuinely” 2-valued.

For each element z in the complement set $\Omega_1 := \Omega \setminus \Omega_2$, $u(z)$ is a single point counted with multiplicity 2. Then there is a Lipschitz map $v : \Omega_1 \rightarrow \mathbb{R}^n$ such that $u(z) = 2 \llbracket v(z) \rrbracket$ for every $z \in \Omega_1$. Again in view of our goal, the only suitable definition of the Dirichlet energy of u over Ω_1 is twice the Dirichlet energy of v . We thus are left with only one possibility for the

Dirichlet energy on the global set Ω :

$$\text{Dir}(u, \Omega) := \int_{\Omega_2} (|Du_1|^2 + |Du_2|^2) + 2 \int_{\Omega_1} |Dv|^2.$$

This analysis can be obviously generalized to any positive integer Q , leading to a general definition of Dirichlet energy for Lipschitz multiple valued functions. The graphs of Lipschitz multiple valued functions carry naturally a structure of integer rectifiable currents (see [30] and cf. [72, Section 3.2] for a brief explanation). It is not difficult to see that, when the Lipschitz constant is small, the Dirichlet energy defined in this section is the second order approximation of the area of the corresponding graph, cf. [30, Corollary 3.3].

Having established the correct notion of Dirichlet energy for Lipschitz functions, one could define the Sobolev space $W^{1,2}(\Omega, \mathcal{A}_Q(\mathbb{R}^n))$ through a “completion strategy”: a measurable map $v : \Omega \rightarrow \mathcal{A}_Q(\mathbb{R}^n)$ is in $W^{1,2}$ if and only if there is a sequence of Lipschitz maps u_k converging to v a.e. and enjoying a uniform bound $\text{Dir}(\Omega, u_k) \leq C$. The Dirichlet energy of v is then defined via a “relaxation procedure”: $\text{Dir}(\Omega, v)$ is the infimum of all constants C for which there is a sequence with the properties above.

7.3. The intrinsic approach to $W^{1,2}$ Q -valued maps. Although the definition above is certainly very natural and gives a good geometric intuition for the Dirichlet energy, it turns out that it is rather complicated to work with it, in particular if one wants to recover the usual statements of the Sobolev space theory for classical functions.

Instead, a rather efficient way to achieve such statements is to rely on a more abstract definition of Dirichlet energy and Sobolev functions, as proposed in [29]. A very general theory has been developed in the literature for Sobolev maps taking values in abstract metric spaces, following the pioneering works of Ambrosio [6] and Reshetnyak [62, 61]. The careful reader will notice, however, that there is a crucial difference between the definition of Dirichlet energy in [62] and the one given below.

DEFINITION 7.4 (Sobolev Q -valued functions, cf. [29, Definition 0.5]). A measurable $f : \Omega \rightarrow \mathcal{A}_Q$ is in the Sobolev class $W^{1,p}$ ($1 \leq p \leq \infty$) if there exist m functions $\varphi_j \in L^p(\Omega; \mathbb{R}^+)$ such that

- (i) $x \mapsto \mathcal{G}(f(x), T) \in W^{1,p}(\Omega)$ for all $T \in \mathcal{A}_Q$;
- (ii) $|\partial_j \mathcal{G}(f, T)| \leq \varphi_j$ a.e. in Ω for all $T \in \mathcal{A}_Q$ and for all $j \in \{1, \dots, m\}$.

It is not difficult to show the existence of minimal functions $\tilde{\varphi}_j$ fulfilling (ii), i.e. such that, for any other φ_j satisfying (ii), $\tilde{\varphi}_j \leq \varphi_j$ a.e. (cf. [29, Proposition 4.2]). Such “minimal bounds” will be denoted by $|\partial_j f|$ and we note that they are characterized by the following property (see again [29, Proposition 4.2]): for every countable dense subset $\{T_i\}_{i \in \mathbb{N}}$ of \mathcal{A}_Q and for every $j = 1, \dots, m$,

$$(28) \quad |\partial_j f| = \sup_{i \in \mathbb{N}} |\partial_j \mathcal{G}(f, T_i)| \quad \text{almost everywhere in } \Omega.$$

We are now ready to define the Dirichlet energy.

DEFINITION 7.5 (cf. [29, Definition 0.6]). The function $|Df|^2$ is defined to be the sum of $|\partial_j f|^2$. The Dirichlet energy of $f \in W^{1,2}(U; \mathcal{A}_Q)$ is then defined by $\text{Dir}(f, U) := \int_U |Df|^2$.

As already mentioned, this definition is equivalent to the one proposed in the previous section (cf. [29, Proposition 4.4]). The main feature is that, however, essentially all the conclusions of the usual Sobolev space theory for single valued functions can be proved to be valid by routine modifications of the arguments: among them we mention Sobolev and Morrey embeddings, compact embeddings, Poincaré inequalities, semicontinuity results, trace properties (cf. [29, Chapter 4]).

One tool which instead is not available in the multivalued setting is the usual regularization by convolution. However in several instances this can be replaced by “gradient truncations” to produce regularizations that are Lipschitz (cf. [29, Section 4.2]). This will be discussed in details in later sections, because it will play an important role. From now on, we will use often standard tools available in the Sobolev space theory: unless we explicitly mention that there is some extra work to do, the reader can safely assume that the corresponding statements can be proved to be valid via “abstract nonsense”.

One important point to be made is about the existence of “selections”. A selection for a Q -valued function u is given by Q classical single valued functions u_1, \dots, u_Q such that $u(x) = \sum_{i=1}^Q \llbracket u_i(x) \rrbracket$, cf. [29, Definition 1.1]. If the u_i are measurable, continuous, Lipschitz, etc. the selection will be called measurable, continuous, Lipschitz, etc. It is rather easy to show that a measurable selection exists for any measurable u , cf. [29, Proposition 0.4]. Incidentally, this will be used repeatedly as we write

$$\sum_i \llbracket u_i \rrbracket$$

for any given measurable Q -valued map u , tacitly assuming to have chosen some measurable selection.

However continuous maps (resp. Sobolev, Lipschitz) do not possess in general selections which are continuous (resp. Sobolev, Lipschitz). The primary examples are the maps stemming from holomorphic subvarieties already discussed at length. Only maps defined on 1-dimensional intervals are a notable exception: in this case continuous, Hölder, Lipschitz and Sobolev multivalued maps have always correspondingly regular selections: indeed there is a linear bound relating the regularity of the selection to that of the initial map in all these cases. For the case of Sobolev and Lipschitz maps the proof is very elementary, cf. [29, Proposition 1.2]. For continuous and Hölder maps the proof turns out to be much harder, cf. [5, Proposition 1.10] and the simpler (and more general) approach of [27]. In the proof of Theorem 3.4 only the existence of Sobolev and Lipschitz selections play a role.

In many instances, although we are not able to find a regular selection of the Q -valued map u , we might be able to split it into two regular Q_i -valued maps, namely

$$u(x) = u_1(x) + u_2(x)$$

where u_1 and u_2 are as regular as u and Q_i -valued for some positive integers with $Q_1 + Q_2 = Q$. In this case we say that u “splits” or “decomposes” into simpler maps, cf. [29, Definition 1.1].

7.4. The cornerstones of the theory of Dir-minimizers. We are now ready to state the main results in the theory of Dir-minimizing maps. In what follows, Ω is always assumed to be a bounded open set with a sufficiently regular boundary (in fact, in order to give a complete account, we should have defined the trace at $\partial\Omega$ of $W^{1,2}$ multiple valued functions; we have avoided to enter in the details to keep our presentation short: the interested reader can consult, for instance, [29, Definition 0.7]).

THEOREM 7.6 (Existence for the Dirichlet Problem, cf. [29, Theorem 0.8]). *Let $g \in W^{1,2}(\Omega; \mathcal{A}_Q)$. Then there exists a Dir-minimizing $f \in W^{1,2}(\Omega; \mathcal{A}_Q)$ such that $f|_{\partial\Omega} = g|_{\partial\Omega}$.*

THEOREM 7.7 (Hölder regularity, cf. [29, Theorem 0.9]). *There is a positive constant $\alpha = \alpha(m, Q)$ with the following property. If $f \in W^{1,2}(\Omega; \mathcal{A}_Q)$ is Dir-minimizing, then $f \in C^{0,\alpha}(\Omega')$ for every $\Omega' \subset\subset \Omega \subset \mathbb{R}^m$. For two-dimensional domains, we have the explicit constant $\alpha(2, Q) = 1/Q$.*

For the second regularity theorem we need the definition of the singular set of f .

DEFINITION 7.8 (Regular and singular points, cf. [29, Definition 0.10]). A Dir-minimizing f is regular at a point $x \in \Omega$ if there exists a neighborhood B of x and Q analytic functions $f_i : B \rightarrow \mathbb{R}^n$ such that

$$(29) \quad f(y) = \sum_i \llbracket f_i(y) \rrbracket \quad \text{for every } y \in B$$

and either $f_i(y) \neq f_j(y)$ for every $y \in B$, or $f_i \equiv f_j$. The singular set $\text{Sing}(f)$ is the complement of the set of regular points.

THEOREM 7.9 (Estimate of the singular set, cf. [29, Theorem 0.11]). *Let f be Dir-minimizing. Then, the singular set $\text{Sing}(f)$ is relatively closed in Ω . Moreover, if $m = 2$, then $\text{Sing}(f)$ is at most countable, and if $m \geq 3$, then the Hausdorff dimension of $\text{Sing}(f)$ is at most $m - 2$.*

Note in particular the striking similarity between the estimate of the size of the singular set in the case of multiple valued Dir-minimizers and in that of area minimizing currents. It will be discussed later that, even in the case of Dir-minimizers, there are singular solutions (which are no better than Hölder continuous).

Complete and self-contained proofs of these theorems can be found in [29]. The key tool for the estimate of the singular set is the celebrated

frequency function (cf. with [29, Section 3.4]), which has been indeed used in a variety of different contexts in the theory of unique continuation of elliptic partial differential equations (see for instance the papers [50], [51]). This is the central tool of our proofs as well. However, our arguments manage much more efficiently the technical intricacies of the problem and some aspects of the theory are developed in further details. For instance, we present in [29, Section 3.1] the Euler-Lagrange conditions derived from first variations in a rather general form. This is to our knowledge the first time that these conditions appear somewhere in this generality.

Largely following ideas of [13] and of White, we improved the second regularity theorem to the following optimal statement for planar maps.

THEOREM 7.10 (Improved 2-dimensional estimate, cf. [29, Theorem 0.12]). *Let f be Dir-minimizing and $m = 2$. Then $\text{Sing}(f)$ is discrete.*

This result was announced in [13]. However, to our knowledge the proof has never appeared before [29].

A new addition to the regularity theory, which will have a lot of importance in the subsequent discussions, is the following higher integrability result.

THEOREM 7.11 (Higher integrability of Dir-minimizers, cf. [33, Theorem 5.1]). *Let $\Omega' \subset \subset \Omega \subset \subset \mathbb{R}^m$ be open domains. Then, there exist $p > 2$ and $C > 0$ such that*

$$(30) \quad \|Du\|_{L^p(\Omega')} \leq C \|Du\|_{L^2(\Omega)} \quad \forall \text{Dir-minim. } u \in W^{1,2}(\Omega, \mathcal{A}_Q(\mathbb{R}^n)).$$

We believe that several intricate arguments and complicated constructions in the third chapter of Almgren's monograph can be reinterpreted as rather particular cases of this key observation (see for instance [5, Section 3.20]). Surprisingly, this higher integrability can be proved in a very simple way by deriving a suitable reverse Hölder inequality and using a (nowadays) very standard version of the classical Gehring's Lemma.

Theorem 7.11 has been stated and proved for the first time in [33]. The relevant reverse Hölder inequality has been derived using a comparison argument and hence relying heavily on the minimality of the Dir-minimizers. A second proof, exploiting the Euler-Lagrange conditions to give a Caccioppoli-type inequality, has been given in [73]. This last proof still uses the regularity theory for Dir-minimizers. However, this occurs only at one step: one could hope to remove this restriction and generalize the higher integrability to "critical" points of the Dirichlet energy.

In [73] a yet different proof for the planar case is proposed, yielding the optimal range of exponents p for which (30) holds. The optimality of this result, as well as the optimality of Theorems 7.7 and 7.10, is shown by another remarkable observation of Almgren. Besides giving area minimizing currents, holomorphic varieties are locally graphs of Dir minimizing Q -valued functions. In [5, Section 2.20] Almgren proves this statement appealing to

his powerful approximation results for area minimizing currents. However this is unnecessary and a rather elementary proof can be found in [73].

7.5. Hölder continuity of Dir-minimizers. The entire Section 8 will be dedicated to the proof of Theorem 7.9, since it contains, in a simplified setting, several of the themes of the proof of Theorem 3.4.

In this paragraph we will instead discuss briefly the ideas behind the proof of Theorem 7.7. We first assume that u is a classical (single valued) function and, for simplicity, that $m \geq 3$ (the case $m = 2$ is somewhat special and can be handled in a simpler way). Assume that $u : B_1(x) \rightarrow \mathbb{R}^n$ is harmonic and compare its energy to the energy of the 0-homogeneous extension v of its trace on $\partial B_1(x)$: we achieve the following crude inequality

$$(31) \quad \int_{B_1(x)} |Du|^2 \leq \int_{B_1(x)} |Dv|^2 \leq \frac{1}{m-2} \int_{\partial B_1(x)} |Du|^2.$$

A scaling-invariant version of the above inequality can be combined with Fubini's theorem to give the following differential inequality:

$$(32) \quad \int_{B_r(x)} |Du|^2 \leq \frac{r}{m-2} \frac{d}{dr} \int_{B_r(x)} |Du|^2,$$

which in turn gives the bound $\text{Dir}(u, B_r(x)) \leq Cr^{m-2}$. If we could improve the constant in (32) to $\frac{1}{m-2+2\varepsilon}$, the same reasoning would give the estimate $\text{Dir}(u, B_r(x)) \leq Cr^{m-2+2\varepsilon}$, which by a standard Morrey-Campanto argument implies the ε -Hölder continuity of u . Now, for a single valued function u the first inequality in (31) is certainly strict, since v does not satisfy the Euler-Lagrange conditions of a minimizer. It is not difficult to see that the very same conclusion can be drawn in the multivalued setting, where the 0-homogeneous extension is also well defined. The problem is to gain, in the factor of the right hand side of (31), a constant $\varepsilon > 0$ which is independent of the function (and, more importantly, of the central point x).

We can therefore focus on improving the constant in the right hand side of (31) and without loss of generality we can assume $x = 0$. It is easy to see that we can assume, again without loss of generality, that the Dirichlet energy in $B_1(0)$ is normalized to 1. When u is single valued we can also assume that u has average 0 after subtracting a second suitable constant: the “uniform gain” from $\frac{1}{m-2}$ to $\frac{1}{m-2+2\varepsilon}$ in (32) is then a simple consequence of the standard compactness of Sobolev maps (via Poincaré inequality). However, although there is a Poincaré inequality for multivalued maps, we cannot “subtract” constant values in general. The only well defined operation is the subtraction of the *same* value p from *all* Q sheets, namely given $u = \sum_i \llbracket u_i \rrbracket$ we can set

$$v(x) = \sum_{i=1}^Q \llbracket u_i(x) - p \rrbracket.$$

In particular we cannot expect compactness when we only control the Dirichlet energy: for a general Sobolev map some sheets might be very far apart on a large subset and be very close on another, very small, subset. However, it can be shown that if the average separation between some sheets of a Dir-minimizer v is too large compared to its Dirichlet energy on a given domain, then v must split into simpler functions in a smaller domain. This allows to prove that there is a uniform gain in the constant of the right hand side of the inequality of (31). The gain will depend upon Q , but this is not an artifact of the proof: it can be shown that the Hölder exponent in Theorem 7.7 does deteriorate to 0 as $Q \rightarrow \infty$.

7.6. Almgren's extrinsic maps. The metric \mathcal{G} on $\mathcal{A}_Q(\mathbb{R}^n)$ is “locally euclidean” at most of the points. Consider for instance the model case $Q = 2$ and a point $P = \llbracket P_1 \rrbracket + \llbracket P_2 \rrbracket$ with $P_1 \neq P_2$. Then, obviously, in a sufficiently small neighborhood of P , the metric space $\mathcal{A}_2(\mathbb{R}^n)$ is isometric to (an open subset of) the Euclidean space \mathbb{R}^{2n} . This fails instead in any neighborhood of a point of type $P = 2 \llbracket P_1 \rrbracket$. On the other hand, if we restrict our attention to the closed subset $\{2 \llbracket X \rrbracket : X \in \mathbb{R}^n\}$, we obtain a close subset isometric to \mathbb{R}^n .

A remarkable observation of Almgren is that $\mathcal{A}_Q(\mathbb{R}^n)$ is biLipschitz equivalent to a deformation retract of the Euclidean space (cf. [5, Section 1.3]). For a simple presentation of this fact we refer the reader to [29, Section 2.1].

THEOREM 7.12. *There exists $N = N(Q, n)$ and an injective $\xi : \mathcal{A}_Q(\mathbb{R}^n) \rightarrow \mathbb{R}^N$ such that:*

- (i) $\text{Lip}(\xi) \leq 1$;
- (ii) if $\mathcal{Q} = \xi(\mathcal{A}_Q)$, then $\text{Lip}(\xi^{-1}|_{\mathcal{Q}}) \leq C(n, Q)$.

Moreover there exists a Lipschitz map $\rho : \mathbb{R}^N \rightarrow \mathcal{Q}$ which is the identity on \mathcal{Q} .

In fact much more can be said: the set \mathcal{Q} is a cone and a polytope. On each separate face of the polytope the metric structure induced by \mathcal{G} is euclidean, essentially for the reasons outlined a few paragraphs above (cf. again [5, Section 1.3] or [33, Section 6.1]). A simple, yet important, observation of White is that the map ξ can be easily constructed so that the Dirichlet energy of $\xi \circ u$ (as classical Euclidean map) coincides with that of u (as multivalued map) for any $u \in W^{1,2}$.

Later on a more complicated version of the map ρ will play a rather important role. As already mentioned, for $Q > 1$ the space $\mathcal{A}_Q(\mathbb{R}^n)$ is not linear and we cannot regularize Q -valued maps by convolution. Nonetheless we will need a way to smooth $W^{1,2}$ maps suitably with a procedure which retains some of the basic estimates available for convolutions with a standard mollifier (in particular when computing the energy of the regularizations). A possible approach is to smooth the euclidean map $\xi \circ u$ and then “project” it back onto \mathcal{Q} using ρ . However, projecting back might be rather costly in terms of the energy since the Lipschitz constant of ρ is indeed rather far from 1.

To bypass this problem, we follow Almgren and prove the existence of “almost” projections, denoted by ρ_δ^* , which are $(1 + \mu)$ -Lipschitz in the δ -neighborhood of $\xi(\mathcal{A}_Q(\mathbb{R}^n))$. These maps cannot be the identity on \mathcal{Q} , but they are at a uniform distance η from it. Almgren’s original proof is rather complicated. In [33, Proposition 6.2] we have proposed a different proof which uses heavily Kirszbraun’s extension theorem and seems to yield a better estimate of μ and η in terms of δ (in particular in the version of [33] these are suitable positive powers of δ).

8. The frequency function

In this section we review the ideas behind the proof of Theorem 7.9. As already mentioned the argument will serve as a prototype for the argument of Theorem 3.4 and for this reason we will be quite detailed.

8.1. First variations. There are two natural types of variations that can be used to perturb Dir-minimizing Q -valued functions. The first ones, which we call inner variations, are generated by right compositions with diffeomorphisms of the domain. The second, which we call outer variations, correspond to “left compositions”. More precisely, let f be a Dir-minimizing Q -valued map.

(IV) Given $\varphi \in C_c^\infty(\Omega, \mathbb{R}^m)$, for ε sufficiently small, $x \mapsto \Phi_\varepsilon(x) = x + \varepsilon\varphi(x)$ is a diffeomorphism of Ω which leaves $\partial\Omega$ fixed. Therefore,

$$(33) \quad 0 = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \int_{\Omega} |D(f \circ \Phi_\varepsilon)|^2.$$

(OV) Given $\psi \in C^\infty(\Omega \times \mathbb{R}^n, \mathbb{R}^n)$ such that $\text{spt}(\psi) \subset \Omega' \times \mathbb{R}^n$ for some $\Omega' \subset\subset \Omega$, we set $\Psi_\varepsilon(x) = \sum_i [f_i(x) + \varepsilon\psi(x, f_i(x))]$ and derive

$$(34) \quad 0 = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \int_{\Omega} |D\Psi_\varepsilon|^2.$$

The identities (33) and (34) lead to interesting first variation conditions in integral form. In order to state them we need anyway a suitable notation to handle the “differential” of a multivalued map $f = \sum_i [f_i]$. Following the discussion in Section 7.2 it is possible, for a Lipschitz multivalued map, to introduce a suitable notion of multivalued differential, which will be denoted by Df . This will be a multiple valued map taking values in $\mathcal{A}_Q(\mathbb{R}^{m \times n})$, which roughly speaking gives at each point the unordered Q -tuple of the differentials of the different branches. There is a coherent way of finding a measurable selection for both f and Df : the coherence has, as a consequence, that when the map splits locally into single valued maps, the i -selection g_i of $Df = \sum_i [g_i]$ corresponds to the differential Df_i of the i -selection f_i . However, this is just one consequence: the coherence can be stated generally even when differential selections do not exist and we refer the reader to [29, Remark 1.11] for the precise definition. This representation allows to

derive chain rules for multivalued maps which are just the analog of the corresponding chain rules for classical maps, cf. [29, Section 1.3.1].

PROPOSITION 8.1 (First variations, cf. [29, Proposition 3.1]). *Let $f : \Omega \rightarrow \mathcal{A}_Q(\mathbb{R}^n)$ be a Dir-minimizer. For every $\varphi \in C_c^\infty(\Omega, \mathbb{R}^m)$, we have*

$$(35) \quad 2 \int \sum_i \langle Df_i : Df_i \cdot D\varphi \rangle - \int |Df|^2 \operatorname{div} \varphi = 0.$$

For every $\psi \in C^\infty(\Omega_x \times \mathbb{R}_u^n, \mathbb{R}^n)$ such that

$$\operatorname{spt}(\psi) \subset \Omega' \times \mathbb{R}^n \quad \text{for some } \Omega' \subset \subset \Omega,$$

and

$$(36) \quad |D_u \psi| \leq C < \infty \quad \text{and} \quad |\psi| + |D_x \psi| \leq C(1 + |u|),$$

we have

$$(37) \quad \int \sum_i \langle Df_i(x) : D_x \psi(x, f_i(x)) \rangle dx \\ + \int \sum_i \langle Df_i(x) : D_u \psi(x, f_i(x)) \cdot Df_i(x) \rangle dx = 0.$$

8.2. The monotonicity of the frequency function. (35) and (36) give particularly interesting identities when tested with functions which depend on $|x|$. In what follows, ν will always denote the outer unit normal on the boundary ∂B of a given ball. The following proposition gives the relevant identities when we test with the singular functions $\varphi(y) = \mathbf{1}_{B_r(x)}(y)y$ and $\psi(x, u) = u\mathbf{1}_{B_r(x)}(y)$ (the proof follows from a standard regularization of these φ and ψ).

PROPOSITION 8.2 (cf. [29, Proposition 3.1]). *Let $x \in \Omega$ and $f : \Omega \rightarrow \mathcal{A}_Q(\mathbb{R}^n)$ be Dir-minimizing. Then, for a.e. $0 < r < \operatorname{dist}(x, \partial\Omega)$, we have*

$$(38) \quad (m-2) \int_{B_r(x)} |Df|^2 = r \int_{\partial B_r(x)} |Df|^2 - 2r \int_{\partial B_r(x)} \sum_i |\partial_\nu f_i|^2,$$

$$(39) \quad \int_{B_r(x)} |Df|^2 = \int_{\partial B_r(x)} \sum_i \langle \partial_\nu f_i, f_i \rangle.$$

We next introduce Almgren's frequency function and state his celebrated monotonicity estimate, which is a straightforward consequence of the identities (38) and (39). Recall the notation $|f|$ for the function $\mathcal{G}(f, Q[0])$.

DEFINITION 8.3 (The frequency function, cf. [29, Definition 3.13]). *Let f be a Dir-minimizing function, $x \in \Omega$ and $0 < r < \operatorname{dist}(x, \partial\Omega)$. We define the functions*

$$(40) \quad D_{x,f}(r) = \int_{B_r(x)} |Df|^2, \quad H_{x,f}(r) = \int_{\partial B_r} |f|^2 \quad \text{and} \quad I_{x,f}(r) = \frac{r D_{x,f}(r)}{H_{x,f}(r)}.$$

$I_{x,f}$ is called the *frequency function*.

When x and f are clear from the context, we will often use the shorthand notation $D(r)$, $H(r)$ and $I(r)$.

THEOREM 8.4 (Monotonicity of the frequency function, cf. [29, Theorem 3.15]). *Let f be Dir-minimizing and $x \in \Omega$. Either there exists $\varrho > 0$ such that $f|_{B_\varrho(x)} \equiv 0$ or $I_{x,f}(r)$ is an absolutely continuous nondecreasing positive function on $]0, \text{dist}(x, \partial\Omega)[$. This function takes a constant value α if and only if $f(y)$ is α -homogeneous in $y - x$.*

This monotonicity is the main ingredient in the proof of both Theorems 3.4 and 7.9. An important observation, which was first made in [32], is that the frequency function can be thought as a “singular limit” of smoother objects, i.e. of regularized frequency functions, which are also monotone. This simple remark (which is not present in Almgren’s monograph) gives an important advantage: the regularized frequency functions enjoy better continuity properties in terms of f .

DEFINITION 8.5 (Regularized frequency functions). Assume ϕ is a Lipschitz nonnegative nonincreasing compactly supported function on $[0, 1[$ which is constant and positive in a neighborhood of 0 and define

$$\begin{aligned} \mathbf{D}_{0,f}(r) &:= \int \phi\left(\frac{|x|}{r}\right) |Df|^2(x) dx \\ \mathbf{H}_{0,f}(r) &:= - \int \phi'\left(\frac{|x|}{r}\right) \frac{|f|^2(x)}{|x|} dx \\ \mathbf{I}_{0,f}(r) &:= \frac{r \mathbf{D}_{0,f}(r)}{\mathbf{H}_{0,f}(r)}. \end{aligned}$$

THEOREM 8.6. *Let f be Dir-minimizing and $0 \in \Omega$. Either there exists $\varrho > 0$ such that $f|_{B_\varrho(0)} \equiv 0$ or $\mathbf{I}_{0,f}(r)$ is an absolutely continuous nondecreasing positive function on $]0, \text{dist}(x, \partial\Omega)[$. This function takes a constant value α if and only if $f(x)$ is α -homogeneous in x .*

We do not have a reference for the latter theorem, which follows from a straightforward adaption of the arguments used in the proof of Theorem 8.4. A special case of Theorem 8.6, namely for a special choice of the cut-off ϕ , is hidden in the computations of [32, Theorem 3.2] (cf. in particular [32, Eq. (3.13)]).

8.3. The two fundamental consequences of the monotonicity formula. Theorem 8.4 has two crucial consequences, when “blowing-up” a given Dir-minimizing function. More precisely, consider a Dir-minimizing f taking $Q > 1$ values and a point p in its domain. Without loss of generality we can assume that $p = 0$. If the support of $f(0)$ contains two different points, then, by continuity, in a neighborhood U of 0 f splits into two separate functions u_1 and u_2 which are both $W^{1,2}$ and continuous. It is simple to see that both must be minimizers of the Dirichlet energy in U . 0 is then a good point, where we have reduced the complexity of the problem. For

instance, if Q were 2 we would know that u_1 and u_2 are two classical (single valued) harmonic functions. The “problematic points” are then those p where $f(p) = Q \llbracket q \rrbracket$.

We can therefore assume that $f(0) = Q \llbracket q \rrbracket$ for some $q \in \mathbb{R}^n$. Now, according to our definition of the singular set $\text{Sing}(f)$, we have two possibilities:

- (a) f equals Q copies of a classical harmonic function in a neighborhood of 0;
- (b) 0 is a singular point for f .

In general, an interesting object to look at is the average of the sheets of $f = \sum_i \llbracket f_i \rrbracket$, namely $\frac{1}{Q} \sum_i f_i$. For this average we fix the notation $\boldsymbol{\eta} \circ f$. It is not difficult to see that $\boldsymbol{\eta} \circ f$ is a classical harmonic function. Indeed, if we define

$$\bar{f} := \sum_i \llbracket f_i - \boldsymbol{\eta} \circ f \rrbracket ,$$

it is immediate to see that $\text{Dir}(f) = \text{Dir}(\bar{f}) + Q \text{Dir}(\boldsymbol{\eta} \circ f)$. In particular it is not difficult to conclude that \bar{f} is also a Dir-minimizer, cf. [29, Lemma 3.23]. Looking at the latter function we can thus restate the alternative as: either $\bar{f} \equiv Q \llbracket 0 \rrbracket$ in a neighborhood of the origin, or 0 is a singular point for \bar{f} (and thus a singular point of f !).

The discussion above leads to the consideration that, without loss of generality, we can assume $\boldsymbol{\eta} \circ f \equiv 0$. Assume further that the (more interesting!) alternative (b) above holds. Then f does not vanish identically and therefore both $D_{0,f}(r)$ and $H_{0,f}(r)$ are positive for some r . Using Theorem 7.7 it is not difficult to see that, under the assumption $f(0) = Q \llbracket 0 \rrbracket$, we have a uniform bound of the form

$$(41) \quad H_{0,f}(r) \leq Cr D_{0,f}(r) \quad \forall r \in \left] 0, \frac{\text{dist}(0, \partial\Omega)}{2} \right[,$$

where the constant C is independent of f . The obvious consequence of Theorem 8.4 is that there is also a reverse control

$$(42) \quad r D_{0,f}(r) \leq \bar{C} H_{0,f}(r)$$

although the latter constant \bar{C} depends upon the point (0 in this case) and the function f . Indeed such constant approaches, for $r \downarrow 0$, the limit $I_0(f) := \lim_{\rho \downarrow 0} I_{0,\rho}(f)$, which by (41) is bounded away from 0 and by Theorem 8.4 is finite: on the other hand we have no explicit (neither universal!) upper bound, we insist that $I_0(f)$ depends upon f and the particular point (0 in this case) where we are “blowing-up”.

Consider now the rescaled functions $f_{0,r}(x) := f(rx)$ and their renormalized versions

$$u_{0,r}(x) := \frac{f_{0,r}}{\text{Dir}(f_{0,r}, B_1)^{1/2}} .$$

In particular the energy of $u_{0,r}$ is 1 in $B_1(0)$. However the L^2 norm of $|u_{0,r}|$ is also under control because of (41). We then have compactness for the family

$\{u_{0,r}\}_r$. Fix a map \bar{u} which is the limit of any subsequence u_{0,r_k} with $r_k \downarrow 0$. It is not difficult to see that a sequence of minimizers with such uniform controls converge *strongly* in $W^{1,2}$ in any compact subset: namely the Dirichlet energy of the limiting function is the limit of the Dirichlet energy of the corresponding functions on any subdomain Ω which is compactly contained in $B_1(0)$, cf. [29, Proposition 3.20]. However the minimizing property alone does not guarantee strong convergence on the *whole* domain $B_1(0)$.

To understand the latter statement, consider for instance the planar (single valued!) harmonic functions

$$f_k(x_1, x_2) = \operatorname{Re}(x_1 + ix_2)^k$$

and their normalizations

$$u_k := f_k / \operatorname{Dir}(f_k, B_1(0)).$$

It is very elementary to see that u_k converges to 0 in $B_1(0)$: in fact most of the Dirichlet energy of u_k lies in a thin layer around the boundary $\partial B_1(0)$. For k large the layer becomes thinner and thinner and all the energy is “pushed” towards the boundary $\partial B_1(0)$. On the other hand it is easy to see that the ratio

$$\frac{D_{0,u_k}(1)}{H_{0,u_k}(1)} = \frac{1}{H_{0,u_k}(1)}$$

explodes, namely that the L^2 norm of u_k on $\partial B_1(0)$ converges to 0.

This highlights the first important consequence of the frequency function: the “reverse Poincaré” inequality (42) excludes that the energy of $u_{0,r}$ concentrates towards the boundary. Any limit \bar{u} of a sequence u_{0,r_k} must therefore have energy equal to 1. Since Theorem 7.7 guarantees uniform convergence, we also conclude that $\bar{u}(0) = Q \llbracket 0 \rrbracket$. Moreover, $\boldsymbol{\eta} \circ \bar{u} \equiv 0$ because $\boldsymbol{\eta} \circ u_{0,r} \equiv 0$.

Thus 0 must be a singular point of \bar{u} as well: the only way \bar{u} could be regular around 0 would be to take the value $Q \llbracket 0 \rrbracket$ identically in a neighborhood of 0. However notice that $I_{0,\bar{u}}(r) = I_{0,f}(0) =: \alpha$ for every r . But then Theorem 8.4 implies that \bar{u} is α -homogeneous, and if \bar{u} would vanish in a neighborhood of 0, then it would vanish on the entire ball $B_1(0)$, contradicting the fact that the Dirichlet energy of \bar{u} is indeed 1.

The conclusion is that the singularity has *persisted* in the limit. Recalling that our main concern in proving Theorem 3.4 was the disappearance of singular points along sequences of converging currents, the reader will understand why the monotonicity of the frequency function is such an exciting discovery. It must also be noticed that the monotonicity of the frequency function was unknown even for classical single valued harmonic functions before [5]: the shear observation that Almgren was able to discover a new fundamental fact for classical harmonic functions around 1970 gives in my opinion the true measure of his genius.

The second fundamental consequence of the monotonicity of the frequency function is that $I_{0,\bar{u}}(r)$ is indeed constant in r and equals $\alpha := I_{0,f}(0)$,

which, as already noticed, gives that \bar{u} is α -homogeneous. In particular when the domain is 2-dimensional, it is not difficult to classify all α -homogeneous Dir-minimizers and to show that their only singularity is at the origin, cf. [29, Proposition 5.1].

The careful reader will recognize the formal analogy with the two ingredients of Federer's reduction argument illustrated in Section 4.5: pretty much the same reasoning gives the proof of Theorem 7.9. There is however one important difference: for a Q -valued minimizer f on a 2-dimensional connected domain we do not conclude the discreteness of $\text{Sing}(f)$, but rather the weaker statement that

- either “multiplicity Q points” of f are isolated;
- or f collapses to $Q \llbracket \eta \circ f \rrbracket$.

Only in the case $Q = 2$ the statement above is equivalent to discreteness of the singular set of f . When, for instance, $Q = 3$, we have not ruled out that singular points with “2-sheeted branching” could converge towards a singular point with a “3-sheeted” branching.

Thus, the argument sketched above gives, in the 2-dimensional case, that $\text{Sing}(f)$ is countable, but it does not imply its discreteness. The proof of Theorem 7.10 needs much more work and in particular it passes through the important additional conclusion that the tangent functions \bar{u} analyzed above are unique, namely the renormalized blown-up functions $u_{0,r}$ have a unique limit as $r \downarrow 0$, cf. [29, Theorem 5.3]. At present this uniqueness is an open problem when the dimension of the domain is higher than 2.

9. Approximation with multiple valued graphs

Following the intuition that a “sufficiently flat” area minimizing current is close to the graph of a Dir-minimizing multivalued function, we wish therefore to find a first approximation of the current with a Lipschitz multivalued graph.

9.1. Multivalued graphs as currents, projections and slices.

One first technical detail that we have to tackle is the integer rectifiable current induced by multivalued maps. Assume therefore to have fixed a measurable map $u : \mathbb{R}^m \supset \Omega \rightarrow \mathcal{A}_Q(\mathbb{R}^n)$, $u = \sum_i \llbracket u_i \rrbracket$. The “set-theoretic” graph of u is clearly

$$\text{Gr}(u) := \{(x, y) \in \mathbb{R}^m \times \mathbb{R}^n : y = u_i(x) \text{ for some } i \in \{1, \dots, Q\}\},$$

or equivalently $\text{Gr}(u) = \{(x, y) : y \in \text{spt}(u(x))\}$ (recall that $u(x)$ is a 0-dimensional current).

When u is sufficiently regular, we want to give to $\text{Gr}(u)$ a structure as integer rectifiable current. Following the discussion of Section 7.2 it is not difficult to see that, when u is Lipschitz, $\text{Gr}(u)$ can be decomposed in a countable union of graphs of single valued Lipschitz functions, defined over domains which might be very irregular (Borel sets, in general). In turn, a

classical theorem of Whitney (cf. [45, Theorem 3.1.14]) allows to decompose any Lipschitz graph in a countable union of C^1 graphs modulo sets of (Hausdorff) measure zero. We thus can give very naturally a structure of integer rectifiable current (which we will denote by \mathbf{G}_u) to $\text{Gr}(u)$, adjusting the multiplicity in a coherent fashion: in particular, if $u = Q \llbracket v \rrbracket$ for some classical Lipschitz function v , then $\mathbf{G}_u = Q\mathbf{G}_v$ (cf. [25] and [72, Section 3.2]).

In a similar fashion we can define the graph of a Lipschitz multivalued function which is defined over a submanifold Σ of \mathbb{R}^m . Now, if Ω is a smooth open set, u is a Lipschitz multiple valued function and v denotes the restriction to the boundary $\partial\Omega$, we expect $\partial\mathbf{G}_u = \mathbf{G}_v$. In the single valued case this is a rather simple fact, since we can use Stokes' theorem when u is smooth and then conclude for general Lipschitz u via regularization. In the multivalued setting this road cannot be followed because there is no regularization of u , but an elementary proof can be found in [30] (cf. the more general Theorem 2.1 therein).

Next we want to find under which conditions an integral current without boundary in a given cylinder can be efficiently approximated by a Lipschitz graph. To be more precise, we will denote by $\mathbf{C}_r(p)$ the cylinder $B_r(x) \times \mathbb{R}^n$ when $p = (x, y) \in \mathbb{R}^m \times \mathbb{R}^n$. In fact in the future we wish to consider cylinders with bases parallel to different m -dimensional planes: having fixed an m -dimensional plane π , we set $B_r(p, \pi) := \mathbf{B}_r(p) \cap (p + \pi)$ and $\mathbf{C}_r(p, \pi) = B_r(p, \pi) + \pi^\perp$. The notation π_0 will be reserved for the “horizontal plane” $\mathbb{R}^m \times \{0\}$ and we will use \mathbf{p}_π and \mathbf{p}_π^\perp for the orthogonal projections onto π and π^\perp .

If T is an integral current without boundary in $\mathbf{C}_r(0)$, a Lipschitz $u : B_r(0) \rightarrow \mathcal{A}_Q(\mathbb{R}^n)$ is an efficient approximation if $\mathbf{M}(T - \mathbf{G}_u)$ is small (compared to r^m). Since \mathbf{G}_u is, in a “loose” sense, a Q -fold cover of $B_r(0)$, we obviously expect that a well-approximated current T is also a Q -fold cover. There is a very efficient way to express this concept in the theory of currents: assuming $\text{spt}(T)$ is bounded, we can define the current $(\mathbf{p}_{\pi_0})_\# T$, which is the push-forward of T on the horizontal plane π_0 . It is rather obvious that this should be an m -dimensional integral current, with no boundary in $\mathbf{C}_r(0)$ and thus should be an integer multiple of $\llbracket B_r(0, \pi_0) \rrbracket$ (incidentally, an integral m -dimensional current with no boundary and supported in a smooth, connected m -dimensional submanifold Γ must be an integer multiple of $\llbracket \Gamma \rrbracket$: this is called the Constancy theorem in the literature, cf. [45, Sections 4.1.4 and 4.1.7]). The condition that T covers Q times the base of the cylinder $\mathbf{C}_r(0)$ can then be expressed by

$$(43) \quad (\mathbf{p}_{\pi_0})_\# T \llcorner \mathbf{C}_r(0) = Q \llbracket B_r(0, \pi_0) \rrbracket .$$

When T is given by a smooth submanifold Γ , the number Q can be computed using the classical degree theory in the following way: given a generic point $y \in B_r(0)$ we consider the finitely many points p_1, \dots, p_N in which Γ intersects the fiber $\{y\} \times \mathbb{R}^n$ (transversally) and assign $\epsilon(p_i) = 1$ if

$T_{p_i}\Gamma \times \pi_0^\perp$ has the same orientation as \mathbb{R}^{m+n} and $\epsilon(p_i) = -1$ otherwise. Then $Q = \sum_{p_i} \epsilon(p_i)$.

There is a way to formalize this concept for a generic current: $\sum_i \epsilon(p_i) \llbracket p_i \rrbracket$ is the “slice” of the current $\llbracket \Gamma \rrbracket$. The latter object can be defined for general (integer rectifiable) currents T and it is usually denoted by $\langle T, \mathbf{p}_{\pi_0}, y \rangle$: roughly speaking it is the intersection of the current T with $\mathbf{p}_{\pi_0}^{-1}(\{y\})$, cf. [68, Section 6.28]. \mathbf{p}_{π_0} might be replaced by a generic (Lipschitz) map whose target is k -dimensional for some $k \leq m$: the resulting slices will then be $(m-k)$ -dimensional currents. The graph of a Lipschitz function u always intersects $\mathbf{p}_{\pi_0}^{-1}(\{y\}) = \{y\} \times \mathbb{R}^n$ “positively”, since the tangents to \mathbf{G}_u do not tilt much compared to the horizontal plane (this is obvious for a single valued function but rather elementary even for a multivalued function). Our discussion motivates then the introduction of the cylindrical excess. For reasons which will be clear later, we also introduce the height of a current in any given set.

DEFINITION 9.1 (Cylindrical excess). Given an integer rectifiable m -dimensional current T in \mathbb{R}^{m+n} with finite mass and compact support and m -planes π, π' , we define the *excess* of T in the cylinder $\mathbf{C}_r(x, \pi)$ compared to π' as

$$(44) \quad \mathbf{E}(T, \mathbf{C}_r(x, \pi), \pi') := (2\omega_m r^m)^{-1} \int_{\mathbf{C}_r(x, \pi)} |\vec{T} - \vec{\pi}'|^2 d\|T\|.$$

If $\pi = \pi'$, then we write $\mathbf{E}(T, \mathbf{C}_r(x, \pi))$.

The *height function* in a set $A \subset \mathbb{R}^{m+m}$ with respect to an m -dimensional plane π is

$$\mathbf{h}(T, A, \pi) := \sup_{x, y \in \text{spt}(T) \cap A} |\mathbf{p}_{\pi^\perp}(x) - \mathbf{p}_{\pi^\perp}(y)|.$$

9.2. The main approximation theorem. We are now ready to state the main approximation theorem needed to carry on our program. To simplify our notation \mathbf{p}_{π_0} and $\mathbf{p}_{\pi_0}^\perp$ will be denoted by \mathbf{p} and \mathbf{p}^\perp .

ASSUMPTION 9.2. $\Sigma \subset \mathbb{R}^{m+n}$ is a C^2 submanifold of dimension $m + \bar{n} = m + n - l$, which is the graph of an entire function $\Psi : \mathbb{R}^{m+\bar{n}} \rightarrow \mathbb{R}^l$ and satisfies the bounds

$$(45) \quad \|D\Psi\|_0 \leq c_0 \quad \text{and} \quad \mathbf{A} := \|A_\Sigma\|_0 \leq c_0,$$

where c_0 is a positive (small) dimensional constant. T is an integral current of dimension m with bounded support contained in Σ and which, for some open cylinder $\mathbf{C}_{4r}(x)$ (with $r \leq 1$) and some positive integer Q , satisfies

$$(46) \quad \mathbf{p}_\# T \llcorner \mathbf{C}_{4r}(x) = Q \llbracket B_{4r}(x) \rrbracket \quad \text{and} \quad \partial T \llcorner \mathbf{C}_{4r}(x) = 0.$$

THEOREM 9.3 (Strong approximation, cf. [33, Theorem 1.4]). *There exist constants $C, \gamma_1, \varepsilon_1 > 0$ (depending on m, n, \bar{n}, Q) with the following property. Assume that T is area minimizing, satisfies Assumption 9.2 in*

$\mathbf{C}_{4r}(x)$ and $E = \mathbf{E}(T, \mathbf{C}_{4r}(x)) < \varepsilon_1$. Then, there is a map $f : B_r(x) \rightarrow \mathcal{A}_Q(\mathbb{R}^n)$, with $\text{Gr}(f) \subset \Sigma$, and a closed set $K \subset B_r(x)$ such that

$$(47) \quad \text{Lip}(f) \leq C E^{\gamma_1} + C \mathbf{A} r,$$

$$(48) \quad \begin{aligned} \mathbf{G}_f \llcorner (K \times \mathbb{R}^n) &= T \llcorner (K \times \mathbb{R}^n) \quad \text{and} \\ |B_r(x) \setminus K| &\leq C E^{\gamma_1} (E + r^2 \mathbf{A}^2) r^m, \end{aligned}$$

$$(49) \quad \begin{aligned} &\left| \|T\|(\mathbf{C}_{\sigma r}(x)) - Q \omega_m (\sigma r)^m - \frac{1}{2} \int_{B_{\sigma r}(x)} |Df|^2 \right| \\ &\leq C E^{\gamma_1} (E + r^2 \mathbf{A}^2) r^m \quad \forall 0 < \sigma \leq 1. \end{aligned}$$

If in addition $\mathbf{h}(T, \mathbf{C}_{4r}(x), \pi_0) \leq r$, then

$$(50) \quad \text{osc}(f) \leq C \mathbf{h}(T, \mathbf{C}_{4r}(x), \pi_0) + C(E^{1/2} + r \mathbf{A}) r,$$

where $\text{osc}(f) := \sup\{|p - q| : p \in \text{spt}(f(x)), q \in \text{spt}(f(y)), x, y \in B_r(x)\}$.

We note that the theorem is scaling invariant and thus it suffices to prove it in the case $r = 1$. Moreover, for simplicity we will mostly ignore Σ and often assume that T is area minimizing in the whole euclidean space: this will be of great help in illustrating the main ideas behind the proof, avoiding some technicalities.

An elementary computation shows that, under Assumption 9.2,

$$\mathbf{E}(T, \mathbf{C}_r(x)) = \frac{\|T\|(\mathbf{C}_r(x))}{\omega_m r^m} - Q.$$

It is then natural to introduce the following “excess measure”:

DEFINITION 9.4 (Excess measure, cf. [33, Definition 1.2]). For a current T as in Assumption 9.2 we define the *excess measure* \mathbf{e}_T and its *density* \mathbf{d}_T :

$$\mathbf{e}_T(A) := \|T\|(A \times \mathbb{R}^n) - Q |A| \quad \text{for every Borel } A \subset B_r(x),$$

$$\mathbf{d}_T(y) := \limsup_{s \rightarrow 0} \frac{\mathbf{e}_T(B_s(y))}{\omega_m s^m} = \limsup_{s \rightarrow 0} \mathbf{E}(T, \mathbf{C}_s(y)).$$

9.3. BV estimate for slices and first approximation. It is rather clear that the smallness of the cylindrical excess prevents the tangent plane to T at p to have negative intersection with $\{\mathbf{p}(p)\} \times \mathbb{R}^n$ at most points p in $\text{spt}(T)$. In fact this is a simple measure-theoretic fact: even without assuming that T is area minimizing, it remains true that, under Assumption 9.2, most slices $\langle T, \mathbf{p}, y \rangle$ will be elements of $\mathcal{A}_Q(\mathbb{R}^n)$. The exceptions y to this property will form a set of small measure.

It is instructive to see what happens if $Q = 1$ and T is assumed to be a-priori the graph of a classical map v , assuming a Lipschitz bound like $\text{Lip}(v) \leq 1$. The cylindrical excess E is then comparable, up to constants, to the L^2 norm of Dv . It is a classical statement for a (single valued) Sobolev map that a Lipschitz control holds on the restriction of the map on a fairly large closed set, cf. for instance [43, Section 6.6.3]. Indeed a way to identify a

good set on which such Lipschitz bound holds is to look at those points where the Hardy-Littlewood maximal function of $|Dv|$ is suitably small. Under our idealized situation, $|Dv|^2$ is indeed comparable to the excess density \mathbf{d}_T introduced above. This motivates the introduction of a maximal function in our setting

DEFINITION 9.5 (Maximal function of the excess measure, cf. [33, Definition 2.1]). Given a current T as in Assumption 9.2 we introduce the “non-centered” maximal function of \mathbf{e}_T :

$$\mathbf{me}_T(y) := \sup_{y \in B_s(w) \subset B_{4r}(x)} \frac{\mathbf{e}_T(B_s(w))}{\omega_m s^m} = \sup_{y \in B_s(w) \subset B_{4r}(x)} \mathbf{E}(T, \mathbf{C}_s(w)).$$

Going on with our Sobolev space analogy, if we denote by E the square of the L^2 norm of $|Dv|$ (normalized by r^m) and we let K be the set where the maximal function of $|Dv|^2$ lies below the threshold $E^{2\gamma_1}$, then the restriction of v to K will have Lipschitz constant E^{γ_1} and the size of the complement of K can be estimated with $r^m E^{1-2\gamma_1}$. Of course we can then extend $v|_K$ outside K to a Lipschitz function with essentially the same Lipschitz bound. Neglecting the effect of Ψ , it is then clear that, relying solely on Assumption 9.2, we can hope for estimate (47) if we replace the superlinear $E^{1+\gamma_1}$ in (48) and (49) with, respectively, $E^{1-2\gamma_1}$ and E .

This heuristic discussion can be in fact made rigorous in a very direct way relying on some recent developments in geometric measure theory. Regarding the slicing map $\langle T, \mathbf{p}, \cdot \rangle$ as a map taking values into the space of 0-dimensional currents (endowed with a suitable metric) and using the formalism introduced by Ambrosio in [6] for BV maps with metric targets, Jerrard and Sonner have given in [56] a rather elementary way to prove that such map is a function of bounded variation, with norm which can be controlled with the mass of T and the mass of its boundary. Ambrosio and Kirchheim used then this idea in [9] to develop part of their general theory of metric currents and give a rather efficient and general approach to the Federer-Fleming compactness theorem. The resulting computations must be suitably adjusted to our setting. However the theory allows a quite direct proof of the following

PROPOSITION 9.6 (Lipschitz approximation, cf. [33, Proposition 2.2]). *There exists a constant $C > 0$ with the following property. Let T and Ψ be as in Assumption 9.2 in the cylinder $\mathbf{C}_{4s}(x)$. Set $E = \mathbf{E}(T, \mathbf{C}_{4s}(x))$, let $0 < \delta_{11} < 1$ be such that $16^m E < \delta_{11}$, and define*

$$K := \{\mathbf{me}_T < \delta_{11}\} \cap B_{3s}(x).$$

Then, there is $u \in \text{Lip}(B_{3s}(x), \mathcal{A}_Q(\mathbb{R}^n))$ such that $\text{Gr}(u) \subset \Sigma$ for every $y \in B_{3s}(x)$ and

$$\text{Lip}(u) \leq C (\delta_{11}^{1/2} + \|D\Psi\|_0), \quad \text{osc}(u) \leq C \mathbf{h}(T, \mathbf{C}_{4s}(x), \pi_0) + Cs \|D\Psi\|_0,$$

$$\mathbf{G}_u \llcorner (K \times \mathbb{R}^n) = T \llcorner (K \times \mathbb{R}^n),$$

$$(51) \quad |B_r(x) \setminus K| \leq \frac{10^m}{\delta_{11}} \mathbf{e}_T \left(\{ \mathbf{m} \mathbf{e}_T > 2^{-m} \delta_{11} \} \cap B_{r+r_0 s}(x) \right) \quad \forall r \leq 3s,$$

where $r_0 = 16 \sqrt[m]{E/\delta_{11}} < 1$.

From Proposition 9.6 one derives immediately a version of Theorem 9.3 where the bound (47) is correct, whereas in the bound (48) the factor $E^{1+\gamma_1}$ must be replaced by $E^{1-2\gamma_1}$ and in the bound (49) $E^{1+\gamma_1}$ must be replaced by E . In the rest of this section we will discuss why the area minimizing assumption, which so far we have not yet used, allows to improve the bounds to achieve Theorem 9.3.

9.4. Superlinear gain. Going back to our heuristic idea, in which T is replaced by the graph of a single valued function v and the excess by the square of the L^2 norm, the “maximal function truncation” described in the previous paragraph would deliver the desired superlinear estimates if we knew that the $L^{2+\beta}$ norm of Dv were controlled by $E^{1/2}$, namely by the L^2 norm of Dv , for some $\beta > 0$. This amounts to a reverse Hölder inequality of the form

$$(52) \quad \|Dv\|_{L^p} \leq C \|Dv\|_{L^2} \quad \text{for some } p > 2.$$

In our setting one possible translation would be: the excess measure \mathbf{e}_T is absolutely continuous and its density \mathbf{d}_T enjoys the estimate

$$(53) \quad \|\mathbf{d}_T\|_{L^{1+\varepsilon}(B_{2r}(x))} \leq C \mathbf{e}_T(B_{4r}(x)) \leq C r^m E.$$

This is certainly not correct under the only Assumption 9.2: it is clear that in order to hope for such a bound we need to use the hypothesis that T is area minimizing. We do not know whether (53) is correct under the additional assumption that T be area minimizing: even if it is, we expect that its proof is rather difficult, see the discussion below. However, the cornerstone of our approach to Theorem 9.3 is that the following slightly weaker form of (53) is correct and can be achieved with a moderate effort.

THEOREM 9.7 (Gradient L^p estimate, cf. [33, Theorem 1.3]). *There exist constants $p_1 > 1$ and $C, \varepsilon_{10} > 0$ (depending on m, n, \bar{n}, Q) with the following property. Let T be as in Assumption 9.2 in the cylinder \mathbf{C}_4 . If T is area minimizing and $E = \mathbf{E}(T, \mathbf{C}_4) < \varepsilon_{10}$, then*

$$(54) \quad \int_{\{\mathbf{d} \leq 1\} \cap B_2} \mathbf{d}_T^{p_1} \leq C E^{p_1-1} (E + \mathbf{A}^2).$$

From Theorem 9.7 and Proposition 9.6 we cannot conclude directly Theorem 9.3 because we lack control on the set where \mathbf{d}_T is rather high (and on the singular part of the measure \mathbf{e}_T !). We would rather need an estimate which controls the regions where the tangent to T has high slope (compared to π_0). Theorem 9.7 can be indeed used to prove something of that kind:

THEOREM 9.8 (Almgren's strong excess estimate, cf. [33, Theorem 6.1]). *There are constants $\varepsilon_{11}, \gamma_{11}, C > 0$ (depending on m, n, \bar{n}, Q) with the following property. Assume T satisfies Assumption 9.2 in \mathbf{C}_4 and is area minimizing. If $E = \mathbf{E}(T, \mathbf{C}_4) < \varepsilon_{11}$, then*

$$(55) \quad \mathbf{e}_T(A) \leq C \left(E^{\gamma_{11}} + |A|^{\gamma_{11}} \right) (E + \mathbf{A}^2) \quad \text{for every Borel } A \subset B_{\frac{9}{8}}.$$

Actually, in the case of a classical single valued map Theorem 9.8 could be concluded directly by comparing the mass of the current T with that of a suitable convolution of the approximating Lipschitz map. The effect of the convolution is to smear high gradients and show that they are energetically not favorable. As already discussed in Section 7.6 there is a surrogate of convolution for multivalued maps u , but it is not as energetically favorable as the classical convolution. In particular, to keep under control how much the convolution of $\xi \circ u$ falls off the set \mathcal{Q} (cf. Section 7.6) a crucial role is played by Theorem 9.7.

9.5. Higher integrability and harmonicity. Going back to our analogy, we know that if T were the graph of a function, the minimality assumption and the smallness of the excess should imply that v is close to an harmonic function. Of course for single valued harmonic functions the reverse Hölder inequality (52) is true for any exponent $p > 2$. On the other hand we already discussed that, for a suitable choice of p , the same reverse Hölder inequality does hold in the multivalued setting as well, cf. Theorem 7.11. This suggests that in order to prove Theorem 9.7 we could first show that the Lipschitz map of Proposition 9.6 is almost Dir-minimizing. Looking at Theorem 9.3 it is rather intuitive that the “almost Dir-minimality” of f should correspond to have a $o(E)$ in place of $E^{1+\gamma_1}$ in (49), where $o(E)$ is any function of E which vanishes faster than E at 0. Now, using an energetic comparison, such a gain would correspond to show that

$$\int_{B_r(x) \setminus K} |Df|^2 = r^m o(E).$$

If this were not true we could run a contradiction argument over a sequence of currents T_k with vanishing excess E_k and look at the normalized approximations $u_k := f_k / E_k^{1/2}$. We could also rescale the corresponding balls to have radius 1 and center 0. The m -dimensional volume of the corresponding bad sets $B_1 \setminus K_k$ is converging to 0 and in spite of that

$$\liminf_k \int_{B_1 \setminus K_k} |Du_k|^2 \geq \eta$$

for some positive η . If we assume that u_k is converging in L^2 to some u , the Dirichlet energy of u would then satisfy

$$\liminf_k \int_{B_1} |Du_k|^2 \geq \int_{B_1} |Du|^2 + \eta.$$

But then the graph of $E_k^{1/2}u$ must have less mass than T_k and we could hope to modify it and gain a comparison current which would contradict the minimality of T_k , at least for k sufficiently large.

Recalling Section 7.5 there is a delicate point to address, namely that for multiple valued functions a uniform control on the Dirichlet energy of a sequence does not imply compactness, since the separation between sheets could explode along the sequence. Nonetheless a careful analysis shows that this program can be carried on. Incidentally it also shows that the approximation of Theorem 9.3 is close to a Dir-minimizer, which we record in the following theorem (for the notation $(u, \Psi(x, u))$, whose meaning should be intuitively clear to the reader, we refer to [29, 33]).

THEOREM 9.9 (Harmonic approximation). *Let γ_1 be the constant of Theorem 9.3. Then, for every $\bar{\eta}, \bar{\delta} > 0$, there is a positive constant $\bar{\varepsilon}_1$ with the following property. Assume that T is as in Theorem 9.3, $E := \mathbf{E}(T, \mathbf{C}_{4r}(x)) < \bar{\varepsilon}_1$ and $r\mathbf{A} \leq E^{1/4+\bar{\delta}}$. If f is the map in Theorem 9.3 and we fix suitable coordinates, then there exists a Dir-minimizing function $u : B_r(x) \rightarrow \mathcal{A}_Q(\mathbb{R}^{\bar{n}})$ such that $w := (u, \Psi(y, u))$ satisfies*

$$(56) \quad \begin{aligned} r^{-2} \int_{B_r(x)} \mathcal{G}(f, w)^2 + \int_{B_r(x)} (|Df| - |Dw|)^2 \\ + \int_{B_r(x)} |D(\boldsymbol{\eta} \circ f) - D(\boldsymbol{\eta} \circ w)|^2 \leq \bar{\eta} E r^m. \end{aligned}$$

10. A first attempt to prove Theorem 3.4

In this section we summarize what we have achieved so far and propose a first strategy to show Theorem 3.4. After resolving the first important issues, we will have to face a major obstacle: more than half of Almgren's monograph is in fact dedicated to overcome this point and even in the proof given by [29, 30, 33, 31, 32] the same phenomenon is responsible for roughly one quarter of the combined length, namely paper [31].

The strategy to prove Theorem 3.4 starts similarly to Federer's reduction argument. Assume that there is an area minimizing current T of dimension $m \geq 2$, in a sufficiently smooth Riemannian manifold Σ , which has a large singular set $\text{Sing}(T)$: more precisely we assume that, for some $\alpha > 0$, $\mathcal{H}^{m-2+\alpha}(\text{Sing}(T)) > 0$.

From Theorem 6.2 and Lemma 6.3 we conclude immediately that at $\mathcal{H}^{m-2+\alpha}$ -a.e. $p \in \text{Sing}(T)$ there is one flat tangent plane and the multiplicity is integral. Let us introduce the notation $D_Q(T)$ for those points in $\text{spt}(T) \setminus \text{spt}(\partial T)$ where the density of T is the positive integer Q . Similarly, we set $\text{Sing}_Q(T) := D_Q(T) \cap \text{Sing}(T)$. We then know that $\text{Sing}_1(T)$ is empty. Indeed the assumptions (a), (b) and (c) in Theorem 4.2 follow from the monotonicity formula when ρ is sufficiently small. The second assumption in (d) is also fulfilled: since we can assume that the second fundamental form of Σ is bounded, for ρ sufficiently small we obviously have $\rho\mathbf{A} < \varepsilon$. It would remain

to prove that the excess with respect to some plane is suitably small at a sufficiently small scale. This is however not difficult since all tangent cones at a point p with $\Theta(T, p) = 1$ must be necessarily flat: it can be shown that the only area minimizing m -dimensional cones S with $\|S\|(\mathbf{B}_1(0)) = \omega_m$ are m -dimensional planes counted with multiplicity 1.

We stop for a moment to observe the following interesting consequence of the above discussion. Let p be a point in $\text{spt}(T)$ where the multiplicity is Q and assume that the surrounding points in $\text{spt}(T)$ have the same multiplicity at a sufficiently small scale, say in $\mathbf{B}_\rho(p)$. Then $S := T/Q$ is a well defined integer-rectifiable area minimizing current in $\mathbf{B}_\rho(p)$ and moreover $p \in D_1(S)$. Thus S is regular in a neighborhood of p . We summarize the outcome of the latter discussion in the following

COROLLARY 10.1. *If S is an area minimizing cone with $\Theta(S, 0) = 1$, then S is a flat plane with multiplicity 1.*

Let T be an area minimizing current in a C^2 Riemannian manifold Σ . If $p \in D_Q(T)$ and there is a neighborhood U of p where the density is Q at $\|T\|$ -a.e. point, then p is a regular point.

We next recover our discussion and look at the current T which should contradict Theorem 3.4. We infer from Corollary 10.1 that there must be an integer $Q > 1$ such that $\text{Sing}_Q(T)$ has positive $\mathcal{H}^{m-2+\alpha}$ -measure. Now, recalling the approach of Federer's reduction argument, we know that for $\mathcal{H}_\infty^{m-2+\alpha}$ -a.e. $p \in \text{Sing}_Q(T)$ we have

$$(57) \quad \liminf_{r \downarrow 0} \frac{\mathcal{H}_\infty^{m-2+\alpha}(\text{Sing}(T) \cap \mathbf{B}_r(p))}{r^{m-2+\alpha}} > 0.$$

Moreover, by Theorem 6.2 we can assume that at least one tangent cone at p is flat. We thus have a sequence of rescalings T_{p,s_k} which are converging to a flat plane and a sequence of rescalings T_{p,r_k} for which (by (57))

$$(58) \quad \lim_{k \uparrow \infty} \mathcal{H}_\infty^{m-2+\alpha}(\text{Sing}(T_{p,r_k}) \cap \mathbf{B}_1(0)) = \eta > 0.$$

Of course the sequence $\{s_k\}$ does not necessarily coincide (or is comparable to) $\{r_k\}$. However, it can be shown that, w.l.o.g., the two sequences can be assumed to coincide (cf. [32] and [72, Section 4]). More precisely

PROPOSITION 10.2 (cf. [32, Proposition 1.3]). *If Theorem 3.4 were false then there would be an area minimizing current T in a smooth Riemannian manifold Σ , a point $p \in \text{Sing}_Q(T)$ and a sequence of rescalings T_{p,r_k} converging to a flat plane of multiplicity Q and such that (58) holds.*

We will see in a moment the simple idea behind Proposition 10.2. Taking it for granted, one could at this point hope to carry the following program:

- (A) We apply Theorem 9.3 to construct a sequence of Lipschitz maps f_k whose graphs approximate efficiently T_{p,r_k} ;

- (B) After normalizing suitably f_k , we apply Theorem 9.9 and, up to extraction of a further subsequence, assume that it converges to a Dir-minimizing multivalued map f_∞ ;
- (C) We finally use (58) to show that f_∞ has a singular set of positive $\mathcal{H}^{m-2+\alpha}$ measure: this would contradict Theorem 7.9.

(C) is again a “persistence of singularity” statement. Unfortunately it is not difficult to see that it is false in this form and thus the rough strategy outlined above must be suitably adjusted. After dealing with Proposition 10.2 we will discuss in detail why (C) fails and propose therefore a new strategy to prove Theorem 3.4.

10.1. The existence of a good sequence. The proof of Proposition 10.2 is still a suitable modification of Federer’s reduction argument. By the discussion above, we first choose a point $p \in \text{Sing}_Q(T)$ and a sequence $r_k \downarrow 0$ where (58) holds. Assume without loss of generality that $p = 0$. If T_{0,r_k} converges to a flat plane of multiplicity Q we are done. Otherwise we can assume that it converges to some tangent cone S , which is singular and such that $\Theta(S, 0) = Q$. We now wish to show that $\mathcal{H}_\infty^{m-2+\alpha}(\text{Sing}_Q(S)) > 0$.

First of all, by the monotonicity formula, $\mathcal{H}_\infty^{m-2+\alpha}(D_Q(S)) > 0$. Now, if

$$(59) \quad \mathcal{H}_\infty^{m-2+\alpha}(\text{Sing}_Q(S)) = 0$$

then many of the points in $D_Q(S)$ should be regular: let us denote by $\text{Reg}_Q(S)$ the set of such points. Note that $\text{Reg}_Q(S)$ is relatively open. If $S \llcorner (\text{Reg}_Q(S))$ has nonempty boundary, then such boundary consists of elements in $\text{Sing}_Q(S)$ and we can expect that it has positive \mathcal{H}^{m-1} measure. The latter statement can in fact be made rigorous and (provided $\alpha < 1$), (59) would imply that $S' = S \llcorner (\text{Reg}_Q(S))$ has no boundary. Hence S' is an area minimizing cone with multiplicity Q at every $p \in \text{spt}(S') \setminus \{0\}$ and with multiplicity no larger than Q at the tip 0 (because $\|S'\| \leq \|S\|$). But the upper semicontinuity of the density implies that $\Theta(S', 0) \geq Q$: thus $\|S\|(\mathbf{B}_1(0)) = Q\omega_m = \|S'\|(\mathbf{B}_1(0))$. This is possible if and only if S and S' coincide. We thus conclude that Corollary 10.1 is applicable to S , which must be flat, contrary to our assumption.

Having found that S is another area minimizing current with large $\text{Sing}_Q(S)$, we can apply the discussion above to some point $p \in \text{spt}(S) \setminus \{0\}$. We thus find a sequence S_{p,r_k} such that $\lim_k \mathcal{H}_\infty^{m-2+\alpha}(\text{Sing}_Q(S_{p,r_k})) > 0$. As above, S_{p,r_k} can be assumed to be converging to some tangent cone Z : if it is flat, we then have achieved the conclusion of Proposition 10.2. Otherwise $\mathcal{H}^{m-2+\alpha}(\text{Sing}_Q(Z)) > 0$ and we can restart with Z in place of S' : this time, however, Z splits off a line. Iterating this procedure we keep “splitting off” lines, until eventually we must reach a sequence as in the statement of Proposition 10.2.

10.2. Persistence of multiplicity Q points. Having proved Proposition 10.2, we are now in the position to attempt the strategy outlined

few paragraphs above. Point (A) is obvious and we have to face point (B). Let us fix a sequence as in Proposition 10.2 that it is converging to $Q \llbracket \pi_0 \rrbracket$ where $\pi_0 = \mathbb{R}^m \times \{0\}$. Thus, for a sufficiently large k , Theorem 9.3 applies to T_{0,r_k} in the cylinder $\mathbf{C}_4(0)$. Let f_k be the corresponding approximating maps $f_k : B_1(0, \pi_0) \rightarrow \mathcal{A}_Q(\mathbb{R}^n)$. It is not obvious, apriori, that we can apply Theorem 9.9, since the excess $E_k := \mathbf{E}(T_{0,r_k}, \mathbf{C}_4(0))$ might converge to zero too fast compared to $r_k \mathbf{A}$. Let us leave this technical problem aside: we then could assume that $u_k := f_k / E_k^{1/2}$ is converging to a Dir-minimizing map u_∞ .

Next, we can intuitively expect that u_∞ has many points of multiplicity Q , in particular all the ones which are limits of sequences lying in $\mathbf{p}_{\pi_0}(\mathrm{D}Q(T_{0,r_k}))$. Namely, we expect that points in $\mathbf{p}_{\pi_0}(\mathrm{D}Q(T_{0,r_k}))$ cluster towards points where $u_\infty = Q \llbracket \boldsymbol{\eta} \circ u_\infty \rrbracket$. This intuition is correct. In fact we can first prove the following

THEOREM 10.3 (Persistence of Q -points, cf. [33, Theorem 1.7]). *For every $\hat{\delta}, C^* > 0$, there is $\bar{s} \in]0, \frac{1}{2}[$ such that, for every $s < \bar{s}$, there exists $\hat{\varepsilon}(s, C^*, \hat{\delta}) > 0$ with the following property. If T is as in Theorem 9.3, $E := \mathbf{E}(T, \mathbf{C}_{4r}(x)) < \hat{\varepsilon}$, $r^2 \mathbf{A}^2 \leq C^* E$ and $\Theta(T, (p, q)) = Q$ at some $(p, q) \in \mathbf{C}_{r/2}(x)$, then the approximation f of Theorem 9.3 satisfies*

$$(60) \quad \int_{B_{sr}(p)} \mathcal{G}(f, Q \llbracket \boldsymbol{\eta} \circ f \rrbracket)^2 \leq \hat{\delta} s^m r^{2+m} E.$$

For the proof of the latter theorem we refer to [33]: for a short explanation, the reader might consult [72, Section 7.2.1] or [25].

Looking back at our u_∞ , which is the strong L^2 limit of $u_k = f_k / E_k^{1/2}$, when p is a point in the domain of u_∞ which is the limit of (the projections onto π_0 of) a sequence of $(p_k, q_k) \in \mathrm{spt}(T_k)$ with $\Theta(T_k, (p_k, q_k)) = Q$, we then have

$$\lim_{r \rightarrow 0} \int_{B_r(p)} \mathcal{G}(u_\infty, Q \llbracket \boldsymbol{\eta} \circ u_\infty \rrbracket)^2 = 0.$$

Since u_∞ is Dir-minimizing and, hence, continuous, we have $u_\infty(p) = Q \llbracket \boldsymbol{\eta} \circ u_\infty(p) \rrbracket$. Now, we must have a set of points p with positive $\mathcal{H}^{m-2+\alpha}$ measures where this occurs. Since Theorem 7.9 tells us that the singular set of u_∞ has dimension at most $m - 2$, the only alternative left is that u_∞ is a classical harmonic single valued function counted Q times. That is, once again the singularities have failed to survive in the limit. If we could exclude this disappearance of the singularities, we would have reached a contradiction and hence proved Theorem 3.4.

Let us look of what happens if we apply the analysis above when the current T is the holomorphic curve of Example 5.4 in a neighborhood of 0. It is obvious that (in complex coordinates) the procedure above will deliver the map $u_\infty(z) = 2 \llbracket z^2 \rrbracket$: although the currents $T_{0,r}$ are singular at the origin, u_∞ is regular. If we compare our situation with the proof of Theorem 7.9 outlined in Section 8, it is quite obvious why we failed to capture the singularity in the limit: we have not subtracted the “average of the sheets”,

namely the regular part of our multiple valued function. The latter has much higher energy than the branching singularity, which is a very small perturbation: if we do not normalize our approximations in some way, we fail to capture the singular behavior in the limit.

11. The center manifold

Summarizing the discussion of the previous section, we are confronted with the following problem. Given a Q -valued Dir-minimizing function $u = \sum_i \llbracket u_i \rrbracket$, the average of its sheets, namely $\boldsymbol{\eta} \circ u := \frac{1}{Q} \sum_i u_i$, is a classical harmonic function and after subtracting it from u we find a new Dir-minimizing Q -valued function

$$\bar{u} := \sum_i \llbracket u_i - \boldsymbol{\eta} \circ u \rrbracket.$$

When \bar{u} is nontrivial, a point p where $\bar{u}(p) = Q \llbracket 0 \rrbracket$ is necessarily singular. Loosely speaking we could say that \bar{u} is the “well-balanced part” of u . If an area minimizing current T satisfies Assumption 9.2 we would like to have a replacement for the average of the sheets $\boldsymbol{\eta} \circ u$ and a replacement for the well-balanced Q -valued map \bar{u} .

One possibility would be to apply Theorem 9.3: we then gain a corresponding approximating Lipschitz map f : the average of its sheets, namely $\boldsymbol{\eta} \circ f$, and its well-balanced part \bar{f} , are both well defined. However, we wish to use these objects in a blow-up procedure: obviously $\boldsymbol{\eta} \circ f$ and \bar{f} do not serve our purposes, since f is a good approximation of the current only at the scale of a certain given cylinder

We would rather like to localize the idea above. Clearly, this is only possible in those regions (and those scales) where the current is sufficiently flat. On the other hand we might not worry about those portions, or those scales, at which the current is not sufficiently flat: in the blow-up procedure we wish to capture the limiting behavior of the current around those points belonging to $\text{Sing}_Q(T)$ and we already saw in the previous section, namely in the proof of Proposition 10.2, that when a lot of points $D_Q(T)$ cluster at a certain scale, the current is necessarily rather flat.

11.1. The construction algorithm. Localizing the basic idea above is a very delicate issue, which involves several parameters. First of all, to fix ideas we will assume that our center manifold will be constructed in a cylinder \mathbf{C} of size comparable to 1 (namely the radius will be a fixed geometric constant, certainly larger than 1), centered at the origin and with basis parallel to $\pi_0 := \mathbb{R}^m \times \{0\}$. We will assume that in the cylinder \mathbf{C} the ambient manifold Σ is very close to be flat: this “almost flatness” is measured in a suitable norm (see [31, Assumption 1.3] for the precise definition): recalling that the excess $\mathbf{E}(T, \mathbf{C})$ is a “quadratic” quantity, the square of the latter norm will be compared to the excess and we will denote by \mathbf{m}_0 the maximum of these two quantities. \mathbf{m}_0 is the first parameter we encounter: it will be assumed to be “suitably small”, namely smaller than some ε_2 . The

latter is however the last constant to be chosen: it can always be (safely) assumed to be sufficiently small, no matter what the other parameters will be required to satisfy.

We further assume that the origin is in the support of the current and it is a point of multiplicity Q . The classical “height bound” for area minimizing currents implies then that the height of T in (a slightly smaller cylinder than) \mathbf{C} is comparable to $\mathbf{m}_0^{1/2m}$. This is due to Almgren himself in his first generalization of De Giorgi’s ε -regularity theorem, see [2], and it is proved nowadays by a very elementary argument which uses the isoperimetric inequality and the monotonicity formula. An important technical generalization of this can be found in [31, Appendix A]: with a minor modification of the usual proof, if one drops the assumption that there is a point of multiplicity Q , we can prove the following “layering theorem”. To avoid lengthy technicalities we give here a rough statement, referring to [31, Theorem A.1] for the precise one.

THEOREM 11.1 (Layered height bound). *Let T satisfy Assumption 9.2 in the cylinder $\mathbf{C}_{4r}(0, \pi_0)$ and let $E := \mathbf{E}(T, \mathbf{C}_{4r}(0, \pi_0))$. In a slightly smaller cylinder $\mathbf{C}_\rho(0, \pi_0)$ the current is then supported in $k \leq Q$ disjoint layers of the form $B_\rho(0, \pi_0) \times B_{C_\rho E^{1/2m}}(y_i, \pi_0^\perp)$: in each layer the density $\Theta(\cdot, T)$ does not exceed a certain integer Q_i by much and $\sum_i Q_i = Q$.*

The “scales” of the construction will be discretized using a suitable Whitney decomposition. We start by subdividing the square $[-1, 1]^m \subset \pi_0$ into $2^{-N_0 m}$ cubes L : for each cube we let $\ell(L)$ be half of the sidelength and $x_L \in \mathbb{R}^m$ denote its center, so that $L = [x_L^1 - \ell(L), x_L^1 + \ell(L)] \times \dots \times [x_L^m - \ell(L), x_L^m + \ell(L)]$. We have just encountered the second of the parameters of the construction, N_0 , which measures the fineness of our starting grid. To each cube we associate a ball \mathbf{B}_L , whose choice is specified in the following way. First the center of the ball \mathbf{B}_L is an arbitrarily chosen $p_L \in \text{spt}(T) \cap (\{x_L\} \times \pi_0^\perp)$. This might seem rather arbitrary but, when \mathbf{m}_0 is very small compared to N_0 , Theorem 11.1 guarantees that the relative distance between points of $\text{spt}(T) \cap (\{x_L\} \times \pi_0^\perp)$ is extremely small compared to $\ell(L)$. Since the radius is going to be $64r_L := 64M_0\sqrt{m}\ell(L)$, where M_0 is a very large constant, the fact that the center of \mathbf{B}_L might “wiggle” slightly depending on the chosen p_L has no effect on the rest of the discussion. Note that we have encountered a third parameter M_0 , whose choice has priority upon N_0 : if M_0 is very large, N_0 should be chosen accordingly large, so to guarantee that at least our ball \mathbf{B}_L is contained in the original cylinder \mathbf{C} .

We next set up a refinement procedure: starting with the initial grid of dyadic cubes, denoted by \mathcal{C}^{N_0} , we subdivide each of them into 2^m cubes H of sidelength $2\ell(H) = \ell(L)$. Now, given our starting hypothesis, we know that both the excess and the height in \mathbf{B}_L are small compared to π_0 . We wish to keep refining the cubes H as long as the height and the excess in the corresponding balls \mathbf{B}_H keep sufficiently small. However, the reference plane might tilt as we refine the scales and we wish to keep track of this. For this

reason, we denote by π_L a given plane for which $\mathbf{E}(T, \mathbf{B}_L) = \mathbf{E}(T, \mathbf{B}_L, \pi_L)$ and we define the height of T in the ball \mathbf{B}_L as $\mathbf{h}(T, \mathbf{B}_L) := \mathbf{h}(T, \mathbf{B}_L, \pi_L)$.

We then stop the refining procedure at some dyadic cube L if

- (EX) The excess has become too large in \mathbf{B}_L ;
- (HT) The height has become too large in \mathbf{B}_L , although the excess has remained small;
- (NN) The refinement stopped at some cube J which touches L and has double sidelength, although in L itself both the height and the excess would be small enough to keep refining.

The reader familiar with Whitney's (or Calderon-Zygmund) decompositions will recognize that the latter condition is enforced to guarantee that, after all the steps of the refinement procedure have been carried on, all nearby cubes have comparable sides. Unfortunately the conditions (EX) and (HT) taken alone do not guarantee this outcome (the troubles are indeed caused by condition (EX)) and the extra (NN) is a source of a few technical complications.

What do we expect from the decomposition above? Fix a point $x \in [-1, 1]^m$. Only two situations might occur:

- (NS) x does not belong to any cube where the refinement procedure stopped;
- (S) $x \in L$ for some cube L where the refinement procedure stopped.

Let $\mathbf{\Gamma}$ be the set of points as in (NS). If the stopping conditions (EX) and (HT) are sufficiently severe, we can expect that $T \llcorner (\mathbf{\Gamma} \times \pi_0^\perp)$ is a Q multiple of a portion of a (rather smooth) single valued graph.

Consider next a point $x \in L$ as in (S). Although we stopped the refinement at L , at a slightly larger scale we still have a very small excess and a very small height: both $\mathbf{E}(T, \mathbf{B}_L)$ and $\mathbf{h}(T, \mathbf{B}_L)$ are still rather small. We thus can hope to apply the approximation Theorem 9.3 in a suitable cylinder $\mathbf{C}_{32r_L}(p_L, \pi_L)$: we can then construct a good Lipschitz Q -valued approximation $f_L : B_{8r_L}(p_L, \pi_L) \rightarrow \mathcal{A}_Q(\pi_L^\perp)$, which will be called the π_L -approximation in \mathbf{B}_L . Finally we take its average $\eta \circ h_L$ and smooth it by convolution with a kernel (which we take to be radial, although the importance of this assumption will become clear only in the next section). The scale of the regularization cannot be larger than r_L , otherwise all the information would be lost: we choose indeed $\ell(L)$ as scale for the convolution. The regularized map is denoted by h_L and will be called the *tilted interpolating function*. Note that in the procedure above one point should be clarified: the cylinder $\mathbf{C}_{32r_L}(p_L, \pi_L)$ is not contained in \mathbf{B}_L . However, through a careful inductive argument, $\text{spt}(T) \cap \mathbf{C}_{32r_L}(p_L, \pi_L)$ can be shown to be contained in \mathbf{B}_L .

We wish to patch the graphs of the various tilted interpolating functions h_L in a single submanifold (and we also hope that this submanifold will glue smoothly with the portion of the current lying over $\mathbf{\Gamma}$!). However, since the graphs of h_L are relative to different systems of coordinates, we need to

parametrize them as graphs in a common coordinate system. Given condition (EX), we can hope that along the refinement procedure the planes π_L did not tilt much. If this is the case the graph of each h_L can be described by the graph of some g_L in the “original coordinates” $\pi_0 \times \pi_0^\perp$. We could then glue the various g_L together using a partition of unity.

11.2. Identification of the scales. Of course we need to give a precise quantification for the conditions (EX) and (HT). An intuition which has guided our understanding is the following: if we are under the assumption $Q = 1$, when the ε -regularity theorem is applicable and therefore T is a Lipschitz graph, so the refinement procedure should never stop at any cube and $\mathbf{\Gamma}$ should in fact coincide with $[-1, 1]^m$. We know from the discussion in Section 4.1 that the decay can be expected to be almost quadratic. We thus set the condition in (EX) as

$$(61) \quad \mathbf{E}(T, \mathbf{B}_L) > C_e \mathbf{m}_0 \ell(L)^{2-2\delta_2},$$

where δ_2 is a small number and C_e a large one.

We then expect the height to be comparable to $\ell(L)^{1+1/m}$ because of Theorem 11.1. We set therefore the height condition as

$$(62) \quad \mathbf{h}(T, \mathbf{B}_L) > C_h \mathbf{m}_0^{1/2m} \ell(L)^{1+\beta_2},$$

where β_2 is a positive parameter smaller than $\frac{1}{m}$. It must be noticed that, if on the (EX) condition it is vital to choose δ_2 rather small and use the full power of De Giorgi’s idea, cf. Remark 4.3, in the height condition we do not need β_2 to be close to $\frac{1}{m}$: it suffices that β_2 can be chosen sufficiently large, in particular compared to δ_2 .

We have now introduced all the parameters of the construction and we want to specify their choice. The exponent β_2 is the first to be chosen and it must be positive but smaller than $\frac{1}{m}$. In fact in [31] the exponent is chosen even smaller, compared to the exponent γ_1 of Theorem 9.3: this is however needed only much later, specifically in the proof of Theorem 13.2. The exponent δ_2 is chosen next: it must be positive but small, compared to β_2 and (more importantly, see below) compared to γ_1 in Theorem 9.3. After δ_2 we can choose M_0 : M_0 must be chosen very large, depending on δ_2 . The reason is the following: if we consider two balls \mathbf{B}_L and \mathbf{B}_H , where H is a “son” of L (namely it is contained in L and has half sidelength), we would like to treat them as concentric, although the reason will come only in the next section. Since the distance between the centers is comparable to $\ell(L)$, the balls will be “close to concentric” provided M_0 is very large.

Next, N_0 should be chosen larger depending on M_0 in such a way that the balls \mathbf{B}_L are anyway not large compared to the initial cylinder \mathbf{C} . Finally, since the stopping conditions are written in terms of $\ell(L)$ rather than the radii of the balls \mathbf{B}_L , C_e and C_h must be taken large so that the refining procedure goes on for at least a few steps before stopping. However C_h will be chosen even larger compared to C_e . The reason is the following. Thanks to

Theorem 11.1, if C_h is relatively large compared to C_e , at any cube L where the refining procedure has stopped by the (HT) condition the current will form at least *two* separate “layers” of thickness smaller than $\ell(L)^{1+(1-\delta_2)/m}$ and parallel to π_L . In turn, this “layered” structure will be inherited by the graph of the π_L approximation, which plays a crucial role in some estimates.

We could summarize the discussion above in the following theorem (again, the statement is very far from being precise; the reader might consult [72, Section 5.2] or [25] for a more thorough explanation and [31, Theorem 1.17] for all the details).

THEOREM 11.2 (Center manifold). *The construction outlined above yields a function $\varphi : \pi_0 \rightarrow \pi_0^\perp$ which is $C^{3,\kappa}$ for some positive $\kappa > 0$ and has small $\|\cdot\|_{C^{3,\kappa}}$ norm (indeed $\|D\varphi\|_{C^{2,\kappa}} \leq C\mathbf{m}_0^{1/2}$). The graph of φ is the center manifold relative to π_0 in the cylinder \mathbf{C} .*

To keep our discussion simple we have avoided several subtle points. We just mention one that plays a crucial role: for currents T which are area minimizing in some submanifold Σ we will need that the graph of φ is contained in Σ .

11.3. The C^3 estimate for the center manifold. Theorem 11.2 is probably the most complicated part of the proof of Theorem 3.4. However it is important to notice that Theorem 11.2 alone does not encode the full strength of the construction described above: we will see in the next section that the graph of φ is indeed a very good substitute for the “average of the sheets of a Q -valued graph”. For instance, the algorithm can be applied even under the assumption that the density of T equals Q $\|T\|$ -almost everywhere in \mathbf{C} : in this case the refinement procedure never stops, Γ equals $[-1, 1]^m$ and finally $\text{spt}(T) \cap [-1, 1]^m \times \pi_0^\perp \subset \text{Gr}(\varphi)$. Namely, by the constancy theorem $T \llcorner ([-1, 1]^m \times \pi_0^\perp) = Q\mathbf{G}_\varphi$ and thus we have gained two more derivatives in the conclusion of Theorem 4.2.

The latter surprising conclusion, namely that with a “purely geometric construction” it is possible to improve the classical ε -regularity theorem by 2 derivatives, is already observed in the introduction of Almgren’s monograph. It is however possible to find a rather fast shortcut to this conclusion since multiple valued functions are not needed. A self-contained “elementary” proof has in fact been given in [28]. The latter reference contains, in a highly simplified setting, the most important estimates which hide behind Theorem 11.2. In this paragraph we will give a rough idea of the C^3 regularity of φ , but we will touch the aspect only superficially. The interested reader can consult the much deeper discussion given in the survey [25].

We start by introducing the notation \mathscr{W} for the dyadic cubes where the refinement procedure has stopped. Coming to Theorem 11.2, it is obvious that we need to prove a uniform $C^{3,\kappa}$ estimate on g_L for any $L \in \mathscr{W}$. This alone will not suffice: an elementary inspection of the partition of unity used

to “glue” the g_L ’s together in φ show that, when L and H are two neighboring cubes in \mathcal{W} , we need the estimate $\|D^l(g_L - g_H)\|_{C^0} \leq C\ell(H)^{3-l+\kappa}$ for every $l \in \{0, 1, 2, 3\}$. However, leaving aside the “interaction” between nearby cubes, let us focus on $\|g_L\|_{C^3}$ and, to simplify the matter even further, let us in fact look at $\|h_L\|_{C^3}$. It is rather plausible that if we could prove a uniform bound on $\|h_L\|_{C^3}$ for the tilted interpolating functions h_L , this will not be destroyed by the change to the coordinates $\pi_0 \times \pi_0^\perp$.

Let us therefore fix $L \in \mathcal{W}$ and let $L = L_i \subset L_{i+1} \subset \dots \subset L_{N_0}$ be a chain of dyadic cubes where L_{j-1} is the father of L_j for every j . Now $h_{L_{N_0}}$ is the convolution of a Lipschitz function at a scale which is obviously comparable with 1 (since $\ell(L_{N_0}) = 2^{1-N_0}$ and N_0 is a fixed “constant”, although rather large). Thus $\|D^k h_{L_{N_0}}\|_{C^0}$ is in fact bounded a-priori with a constant depending only on k (and on all the other parameters of the construction). We next want to study how the norm $\|D^k h_{L_j}\|_{C^0}$ might increase compared to $\|D^k h_{L_{j-1}}\|_{C^0}$: the hope is that this can be bounded by some power of $\ell(L_j)$, leading in turn to a convergent geometric series when $k \in \{0, 1, 2, 3\}$. This would then give a uniform bound on $\|h_L\|_{C^3}$.

Ideally we would like to compute $\|D^k(h_{L_j} - h_{L_{j-1}})\|_{C^0}$. This is however not really possible, since the two functions are defined according to two different coordinate systems (namely $\pi_{L_j} \times \pi_{L_j}^\perp$ and $\pi_{L_{j-1}} \times \pi_{L_{j-1}}^\perp$). Let us however assume, for the sake of argument, that $\pi_{L_j} = \pi_{L_{j-1}} =: \pi$. Moreover, to simplify the notation let us denote L_j by J and L_{j-1} by H .

Under this assumption, h_J is the convolution of $\eta \circ f_J$ and h_H the convolution of $\eta \circ f_H$. We know that both f_H and f_J approximate very accurately the area minimizing current T , at two scales which are comparable by a factor 2. Thus, for both graphs \mathbf{G}_{f_J} and \mathbf{G}_{f_H} the first variation is close to 0, which in turn, given the smallness of the excess of the current at that scale, should imply that both $\eta \circ f_J$ and $\eta \circ f_H$ are almost harmonic.

The latter discussion is correct but must be quantified. It is not difficult to see that it can be translated into an estimate for $\Delta(\eta \circ f_J)$ and $\Delta(\eta \circ f_H)$ in some negative Sobolev space (more precisely we use the $W^{-1,1}$ norm, cf. [31, Proposition 5.1]). To simplify the matter even further, let us assume that both $\eta \circ f_H$ and $\eta \circ f_J$ are in fact harmonic. The regularization by convolution will then not change the two functions, because the convolution kernel is radial. Thus we would have $\|D^k(h_H - h_J)\|_{C^0} = \|D^k(\eta \circ f_H - \eta \circ f_J)\|_{C^0}$. On the other hand, again by the mean-value formula for harmonic function, we could estimate

$$(63) \quad \|D^k(h_H - h_J)\|_{C^0} \leq \frac{C}{\ell(H)^{m+k}} \|\eta \circ f_H - \eta \circ f_J\|_{L^1}.$$

Let us now recall that the graph of f_H coincides with the current T , except for a set of measure $C\ell(H)^m \mathbf{E}(T, \mathbf{B}_H)^{1+\gamma_1}$. In turn the excess $\mathbf{E}(T, \mathbf{B}_H)$ is of size $\ell(H)^{2-2\delta_1}$. Since an analogous consideration holds for J in place of H , we conclude that the two maps f_H and f_J coincides except for a set of measure at most $\ell(H)^{m+(2-2\delta_1)(1+\gamma_1)}$. On the other hand the “heights of

both maps" is of order $\ell(H)^{1+\beta_2}$, i.e. the available estimate for the heights $\mathbf{h}(T, \mathbf{B}_H)$ and $\mathbf{h}(T, \mathbf{B}_L)$. Combining these observations, we conclude that

$$\|\boldsymbol{\eta} \circ f_H - \boldsymbol{\eta} \circ f_J\|_{L^1} \leq C\ell(H)^{m+3+\kappa}.$$

Inserting the latter inequality in (63) we would then conclude

$$\|D^k(h_H - h_J)\|_{C^0} \leq C\ell(H)^{3+\kappa-k}.$$

12. The approximation on the normal bundle of the center manifold

To carry on our program for proving Theorem 3.4 by "blow-up", we now need to approximate again our area minimizing current in a cylinder where the excess is small, taking advantage of the center manifold. Let $\mathcal{M} = \text{Gr}(\varphi)$ be the center manifold constructed in the previous section and let us make some first considerations.

First of all, by the $C^{3,\kappa}$ estimates, we know that in a sufficiently small neighborhood \mathbf{U} of \mathcal{M} there is a $C^{2,\kappa}$ orthogonal projection $\mathbf{p} : \mathbf{U} \rightarrow \mathcal{M}$ which to each $p \in \mathbf{U}$ assigns the unique point $q = \mathbf{p}(p)$ such that $p - q$ is normal to $T_q\mathcal{M}$. In fact, since $\|D\varphi\|_{C^{2,\kappa}} \leq C\mathbf{m}_0^{1/2}$ and \mathbf{m}_0 can be chosen arbitrarily small, the "thickness" of \mathbf{U} can be assumed to be of any given size, say 1. In turn, since the height of T in \mathbf{C} is of order $\mathbf{m}_0^{1/2m}$, we can certainly assume that $\text{spt}(T \llcorner \mathbf{C}_{1/2}(0)) \subset \mathbf{U}$. It is also not difficult to see that T is a Q -fold covering of \mathcal{M} , namely $(\mathbf{p}_\#(T \llcorner \mathbf{C}_{1/2})) \llcorner \mathbf{C}_{1/4} = Q \llbracket \mathcal{M} \cap \mathbf{C}_{1/4} \rrbracket$ (we need to restrict slightly the radius of the cylinder to avoid "boundary effects"). We could define a "curvilinear excess" compared to \mathcal{M} with the following procedure: at each point $p \in \text{spt}(T)$ we measure the distance between $\vec{T}(p)$ and the oriented tangent plane to \mathcal{M} at the projection $\mathbf{p}(p)$. We then integrate the square of this quantity over $\text{spt}(T) \cap \mathbf{C}_{1/2}$. The corresponding formula is

$$\int_{\mathbf{C}_{1/2}} |\vec{T}(p) - T_{\mathbf{p}(p)}\mathcal{M}|^2 d\|T\|(p).$$

It is no surprise that the latter is controlled by the "straight excess" in the cylinder \mathbf{C} , simply because the tilt between $T_q\mathcal{M}$ and π_0 is controlled by $\|D\varphi\|_{C^0}$, for which in turn we have the bound $C\mathbf{m}_0^{1/2}$. Thus, as it happens with the "straight excess" we can expect to be able to approximate T efficiently with a multivalued map N defined on \mathcal{M} and taking values in the normal bundle of \mathcal{M} .

To be more precise, we are looking for an approximating map $F : \mathcal{M} \rightarrow \mathcal{A}_Q(\mathbb{R}^{m+n})$ with the properties that

- $F(q) = \sum_i \llbracket F_i(q) \rrbracket$ where $F_i(q) \in \mathbf{U}$ and $\mathbf{p}(F_i(q)) = q$ for every $q \in \mathcal{M}$;
- for most q 's, namely for q 's belonging to a certain closed subset $\mathcal{K} \subset \mathcal{M}$, we have " $F(q) = \text{spt}(T) \cap (q + T_q\mathcal{M})^\perp$ "; the latter identity is under quotation marks because the real requirement is that the

normal slice of the current T coincides with $F(q)$, namely $F(q) = \langle T, \mathbf{p}, q \rangle$ (cf. Section 9.3).

The normal part of the map F is then the multivalued map $N(q) := \sum_i \llbracket F_i(q) - q \rrbracket$: the latter is our *normal approximation*. We will require that the map F be Lipschitz and to its image (or, equivalently, to the graph of N) we can associate naturally an integer rectifiable current, which will be denoted by \mathbf{T}_F . We wish not only to construct the map N , but also to keep various related quantities under control: certainly its Lipschitz constant, its height, the size of \mathcal{K} and the difference in mass between \mathbf{T}_F and T . Moreover, we certainly expect that these estimates will depend, locally, upon the scale at which the refining of the Whitney decomposition described in the previous section stopped. We wish therefore to introduce “Whitney” regions on \mathcal{M} : they will be denoted by \mathcal{L} and, loosely speaking, each such \mathcal{L} is a suitable enlargement of the portion of the graph of \mathcal{M} lying over a cube $L \in \mathcal{W}$.

However, before coming to the precise statements concerning these estimates, we should make one further important consideration. Our plan is to show that N is close to a Dir-minimizing Q -valued map and that a large singular set for T induces a large singular set on the latter map. The first point was, in “straight coordinates”, an effect of the Taylor expansion of the area functional of a graph. Thus, it makes sense to compare the mass of \mathbf{T}_F with that of $Q \llbracket \mathcal{M} \rrbracket$.

THEOREM 12.1 (Expansion of $\mathbf{M}(\mathbf{T}_F)$, cf. [30, Theorem 3.2]). *If \mathcal{M} , F and N are as above and the Lipschitz constant of N is sufficiently small, then*

$$(64) \quad \mathbf{M}(\mathbf{T}_F) = Q \mathcal{H}^m(\mathcal{M}) - Q \int_{\mathcal{M}} \langle H, \boldsymbol{\eta} \circ N \rangle + \frac{1}{2} \int_{\mathcal{M}} |DN|^2 + \text{H.O.T.},$$

where H is the mean curvature vector of \mathcal{M} and H.O.T. contains higher order terms, namely expressions that are at least trilinear in N or bilinear in N but multiplied by a small factor.

Notice in particular that there is a nonvanishing linear term in the expansion. In order to show that N is quasi-harmonic, we therefore need $\langle H, \boldsymbol{\eta} \circ N \rangle$ to be much smaller than $|DN|^2$. We can certainly expect so, given that the center manifold should be “well-centered”. However, this adds an additional quantity which should be kept under control in a special way, namely the average $\boldsymbol{\eta} \circ N$.

12.1. The construction of N . In order to simplify the next discussions, we will denote by Φ the map $x \mapsto (x, \boldsymbol{\varphi}(x))$, i.e. the graphical parametrization of the center manifold \mathcal{M} . A simple consequence of the refining algorithm used to construct the center manifold is that on $\Phi(\Gamma)$ the current T coincides completely with \mathcal{M} . More precisely, $\langle T, \mathbf{p}, p \rangle = Q \llbracket p \rrbracket$ whenever $p \in \Phi(\Gamma)$. Hence we will obviously set $N(p) = Q \llbracket 0 \rrbracket$ for any such

p : it remains therefore to define N on each region of the form $\Phi(L)$ when $L \in \mathscr{W}$.

Let us fix $L \in \mathscr{W}$. Going back to the construction of the center manifold, we discover that we already have a rather accurate graphical approximation of T in the region of our interest, since we already defined the π_L -approximation f_L . If f_L were a classical single valued function, we could simply parametrize its graph on the normal bundle of \mathcal{M} . Indeed the tangent planes to the graph of f_L are certainly close to π_L and on the other hand \mathcal{M} is constructed patching a suitable smoothing of the average $\eta \circ f_L$. Thus, in the cylinder $\mathbf{C}_{6r_L}(p_L, \pi_L)$, the angle between a generic tangent plane to the graph of f_L and a generic tangent plane to \mathcal{M} is rather small. In the Q -valued setting “reparametrizing” graphs is a much more subtle issue than in the classical single valued setting. However it is not very hard to prove a suitable theorem (see [30, Theorem 5.1]) which allows to describe the graph of f_L through a Lipschitz map N_L on the normal bundle of \mathcal{M} . Note moreover that N_L can be defined on a domain much larger than $\Phi(L)$.

We next have to face a new difficulty: if L and H are two nearby cubes, the maps N_L and N_H do not necessarily agree on the intersection of their domains. However, recall that the graphs of f_L and f_H coincide with the current T except for two sets of small measure. Thus the values of N_L and N_H coincide on a very large portion of the intersection of their domains. In turn, the construction algorithm ensures that each H intersects only a finite number of other cubes in \mathscr{W} : such number is bounded a-priori by a geometric constant (for instance, when $m = 2$ each square of \mathscr{W} can intersect at most 12 other squares of \mathscr{W}). So, after removing from each $\Phi(L)$ all those points where N_L does not coincide with all the N_H related to neighboring cubes, we reach a uniquely defined map N on a rather large subset of $\mathcal{M} \setminus \Phi(\Gamma)$.

We now wish to extend this map to a Lipschitz one defined on all \mathcal{M} . It is not difficult to see that N is already *globally* Lipschitz on its domain of definition and that it approaches the value $Q \llbracket 0 \rrbracket$ on sequences of points converging to $\Phi(\Gamma)$: this is because the height of f_L (and thus that of N_L) is controlled by $C\ell(L)^{1+\beta_2}$ and cubes in \mathscr{W} close to Γ necessarily have small sidelength. However, it does not serve our purposes to give a *global* Lipschitz extension of N which does not respect the local properties of the map. In particular we desire an extension that on each Whitney region \mathcal{L} has small Lipschitz constant: the smaller the scale, the smaller should be the Lipschitz constant. To achieve this property we follow an elementary idea, which we next describe in the special case of dimension $m = 2$. As a matter of fact, since Φ is Lipschitz we can, for the sake of our discussion, assume that $\Phi(L)$ is flat and coincides with the square L . We fix the four points A, B, C and D which are the four vertices of the square. We first wish to extend N to four small neighborhoods of these points; such neighborhoods will be called $U(A), U(B), U(C)$ and $U(D)$ and we will fix them to be disks centered at the respective points with radius $\ell(L)/4$.

Take for instance A : the latter point is a common vertex for four squares H_1, H_2, H_3, L of \mathscr{W} . We then consider the intersection of the domain of N with $H_1 \cup H_2 \cup H_3 \cup L$: if we restrict the map N to this set and consider its Lipschitz constant, we can then use the Lipschitz extension theorem for multivalued functions to extend it to a neighborhood $U(A)$ of A , without increasing such constant by much. We proceed and extend the map separately to neighborhoods of A, B, C and D . However when we extend the map to the neighborhood $U(B)$ we disregard what we did in the neighborhood of A and we only take into consideration the “original” N : having chosen such neighborhoods sufficiently small the distance between two points p and q lying in different ones (say, $U(A)$ and $U(B)$) is larger than $\ell(L)/2$ and the height of the extension is no larger than $C\ell(L)^{1+\beta_2}$, thus providing automatically a good Lipschitz bound on $U(A) \cup U(B)$. This procedure can then be repeated for all squares and we have a new map N' which is evidently defined in a neighborhood of any vertex of any $L \in \mathscr{W}$.

With the same principle we extend N' to neighborhoods $U(AB), U(BC), U(CD), U(DA)$ of the corresponding sides of the square L . This seems more problematic because, for instance, $U(AB)$ and $U(BC)$ intersect in a neighborhood of B . However N' is already defined on $U(A) \cup U(B) \cup U(C) \cup U(D)$. Thus we need to extend it to the sets $V(AB) = U(AB) \setminus (U(A) \cup U(B))$ and $V(BC) = U(BC) \setminus (U(B) \cup U(C))$. The latter can be now assumed to be disjoint and separated by a distance of the order $c_0\ell(L)$ for some $c_0 > 0$: it just suffices to choose the thickness of the neighborhoods $U(AB)$ and $U(BC)$ much smaller than the thickness of the neighborhoods $U(A), U(B)$ and $U(C)$. We can then literally argue as above and gain a second extension of the map, namely N'' , which is defined on the boundary of any square L of \mathscr{W} . At this point the third (and final) extension is achieved by considering each square separately.

12.2. The estimates on each Whitney region \mathcal{L} . We are now ready to state the main estimates on the approximating map N .

THEOREM 12.2 (Local estimates for the \mathcal{M} -normal approximation, cf. [31, Theorem 2.4]). *Let $\gamma_2 := \frac{\gamma_1}{4}$, with γ_1 the constant of [33, Theorem 1.4]. The \mathcal{M} -normal approximation can be constructed so to satisfy the following estimates on every Whitney region \mathcal{L} associated to a cube $L \in \mathscr{W}$, with constants $C = C(\beta_2, \delta_2, M_0, N_0, C_e, C_h)$:*

$$(65) \quad \text{Lip}(N|_{\mathcal{L}}) \leq C\mathbf{m}_0^{\gamma_2}\ell(L)^{\gamma_2} \quad \text{and} \quad \|N|_{\mathcal{L}}\|_{C^0} \leq C\mathbf{m}_0^{1/2m}\ell(L)^{1+\beta_2},$$

$$(66) \quad |\mathcal{L} \setminus \mathcal{K}| + \|\mathbf{T}_F - T\|(\mathbf{p}^{-1}(\mathcal{L})) \leq C\mathbf{m}_0^{1+\gamma_2}\ell(L)^{m+2+\gamma_2},$$

$$(67) \quad \int_{\mathcal{L}} |DN|^2 \leq C\mathbf{m}_0\ell(L)^{m+2-2\delta_2}.$$

Moreover, for any $a > 0$ and any Borel $\mathcal{V} \subset \mathcal{L}$, we have (for $C = C(\beta_2, \delta_2, M_0, N_0, C_e, C_h)$)

$$(68) \quad \int_{\mathcal{V}} |\eta \circ N| \leq C \mathbf{m}_0 \left(\ell(L)^{3+\beta_2/3} + a \ell(L)^{2+\gamma_2/2} \right) |\mathcal{V}| + \frac{C}{a} \int_{\mathcal{V}} \mathcal{G}(N, Q[\eta \circ N])^{2+\gamma_2}.$$

The three estimates (65), (66) and (67) are indeed all simple consequences of the analogous estimate for the π_L -approximation f_L , taking into account that in \mathbf{B}_L we have a specified control on the size of the excess and of the height. The estimate (68) is instead more subtle, but its reason is also rather obvious: the center manifold \mathcal{M} is indeed very close to the graph of $\eta \circ f_L$ by construction and thus $\eta \circ N_L$ must be very small.

12.3. The splitting before tilting phenomenon. The final important conclusions on the center manifold are given by two estimates which are somewhat of a dual nature to those of Theorem 12.2. Ignoring for the moment the special cubes L which stopped because of the condition (NN), we must remember that all other cubes must have stopped for one of the following two very precise reasons:

- (h) either the eight in \mathbf{B}_L exceeds $C_h \mathbf{m}_0^{1/2m} \ell(L)^{1+\beta_2}$;
- (e) or the excess in \mathbf{B}_L exceeds $C_e \mathbf{m}_0 \ell(L)^{2-2\delta_2}$.

In the first occasion recall the layered height bound of Theorem 11.1. This theorem shows that T in (a suitable subset of) \mathbf{B}_L splits in at least two currents which are contained in two layers parallel to π_L and separated by a distance comparable to the height in \mathbf{B}_L , say $\frac{1}{2} C_h \mathbf{m}_0^{1/2m} \ell(L)^{1+\beta_2}$. Thus, *anywhere* in a region of diameter $C\ell(L)$ which includes \mathcal{L} (where C will depend on the constant M_0) we can expect $|N|$ to be at least as large as $\frac{1}{4} C_h \mathbf{m}_0^{1/2m} \ell(L)^{1+\beta_2}$. Thus, on every region Ω of measure $c\ell(L)^m$ close to \mathcal{L} , the size of $\int_{\Omega} |N|^2$ is at least $C\ell(L)^{m+2+2\beta_2}$. We refer the reader to [31, Proposition 3.1] (see also [72, Section 5.5]) for the precise statement.

In the second case, namely when (e) above holds) we would like to say that $\int_{\Omega} |DN|^2$ is at least $C\ell(L)^{m+2-2\delta_2}$ for any region Ω which is a ball in \mathcal{M} of radius $c\ell(L)$, sufficiently close to \mathcal{L} . This is true but much more subtle and illustrated through a principle which, inspired by a pioneering paper of Rivière, cf. [63], we call the *splitting before tilting* phenomenon.

Recall that, if H is the father of L , the excess in \mathbf{B}_H is smaller than $C_e \mathbf{m}_0 \ell(H)^{2-2\delta_2}$. If the parameter M_0 is chosen very large, \mathbf{B}_L and \mathbf{B}_H are almost concentric and the radius of \mathbf{B}_L is twice the radius of \mathbf{B}_H . On the other hand we know that, in \mathbf{B}_L , the current T can be approximated by a Dir-minimizing Q -valued map. If the latter were too close to a multiple copy (with multiplicity Q) of a classical single valued harmonic function, then the argument illustrated in Section 4.1 tells us that the excess in \mathbf{B}_L should be almost $\frac{1}{4}$ of that in \mathbf{B}_H , cf. Remark 4.3. But this is not the case because the ratio is instead at least $2^{-2+2\delta_2}$: although small, the parameter δ_2 makes here a big difference! Thus in \mathbf{B}_L T is close to a “non-classical”

Dir-minimizer $f = \sum_i \llbracket f_i \rrbracket$, more precisely we can certainly assume that the Dirichlet energy of $\tilde{f} = \sum_i \llbracket f_i - \boldsymbol{\eta} \circ f_i \rrbracket$ is a non-negligible fraction of the Dirichlet energy of f . For such maps we have the important property that their energy must be sufficiently large on any region. Passing now to the “curvilinear coordinates” we can infer the same conclusion for $\int_\Omega |DN|^2$ whenever Ω is a region of a suitable size sufficiently close to \mathcal{L} . For the precise statement we refer the reader to [31, Proposition 3.4] (see also [72, Section 5.5]).

13. The frequency function again

We are now ready to discuss the proof of Theorem 3.4. We assume by contradiction that the theorem is false and, recalling Proposition 10.2, we fix an area minimizing current T of dimension m , a plane π_0 (which without loss of generality we assume to be $\mathbb{R}^m \times \{0\}$) an integer $Q > 0$ and a sequence of radii $r_k \downarrow 0$ with the following properties:

- The excess $\mathbf{E}(T_{0,r_k}, \mathbf{C}_8(\pi_0, 0))$ converges to 0 as $k \uparrow \infty$;
- The point 0 is singular, $\Theta(T, 0) = Q$ and $\mathcal{H}_\infty^{m-2+\alpha}(\mathbf{B}_1 \cap \text{Sing}_Q(T_{0,r_k})) \geq \eta > 0$.

13.1. A sequence of center manifolds. We wish to approximate the current with an \mathcal{M} -normal approximation N over a center manifold \mathcal{M} . A first attempt could be the following: for some r sufficiently large, the excess $\mathbf{E}(T_{0,r}, \mathbf{C}_r(\pi_0, 0))$ will be sufficiently small (and, in case T is area minimizing in a Riemannian submanifold Σ , $\Sigma_{0,r}$ will be sufficiently flat). Assuming without loss of generality that $r = 1$, we can then construct a center manifold \mathcal{M} and an \mathcal{M} -normal approximation. However, we have no guarantee that this approximation is accurate at very small scales around 0. This would certainly be the case if 0 belonged to the contact set $\mathbf{\Gamma}$ described in Section 11.1, but of course it might be the case that 0 belongs to some cube $L \in \mathscr{W}$ where the refining procedure has stopped. If this is the case, then at a certain small scale around 0 the graph of N might have a completely different behavior than T .

If we set $t_1 := 1$, we can then distinguish two situations. In the first one $0 \in \mathbf{\Gamma}$ and we are thus satisfied with our center manifold and our approximating map, which we denote by \mathcal{M}_1 and N_1 . In the second one at some distance s_1 from the origin we encounter the first cube of size sufficiently large compared to s_1 , say $\bar{c}_s s_1$ for some specified constant \bar{c}_s (this number turns out to be $1/(16\sqrt{m})$ in our proof, cf. [32, Section 2]: its precise value is however not important, as long as it is a fixed parameter which does not depend upon those which enter in the construction of the center manifold). We can assume that s_1 is (much) smaller than t_1 : we just need to set the parameter N_0 in the construction of the center manifold accordingly large.

At the scale s_1 the pair (\mathcal{M}_1, N_1) , i.e. the center manifold and the corresponding approximation, is not anymore serving our purposes. Now consider T_{0,s_1} : the latter current might or might not have sufficiently small excess to

construct another center manifold. In the first case we set $t_2 = s_1$, otherwise we let t_2 be the first radius smaller than s_1 at which the excess goes below the desired threshold. We then construct the pair (\mathcal{M}_2, N_2) taking T_{0,t_2} as reference area minimizing current.

The procedure above delivers a sequence, or a finite number, of intervals $]s_i, t_i[$ which we call *intervals of flattening*. If they are finitely many then the endpoint of one of them is $s_{l_0} = 0$. It turns out that, for k sufficiently large, each radius r_k belongs to one of such intervals, which will be denoted by $I_{l(k)}$. For each interval of flattening I_l we have a center manifold \mathcal{M}_l and an \mathcal{M}_l -normal approximation N_l . Similarly, all the relevant objects needed in their constructions, like the family \mathscr{W}^l of cubes in the corresponding Whitney decompositions, will be indexed with l .

13.2. The frequency function. We have now gained a sequence of center manifolds $\mathcal{M}_{l(k)}$ and of approximations $N_{l(k)}$. Let us rescale the center manifolds by a factor $r_k/t_{l(k)}$, so to gain a sequence of manifolds $\bar{\mathcal{M}}_k$ and maps \bar{N}_k at “scale 1”, which should give rather good approximations of $T_k := T_{0,r_k}$. We fix for convenience a “central point” for each center manifold \mathcal{M}_l : it will be the point p_l lying in the plane $\{0\} \times \mathbb{R}^n$. Correspondingly $\bar{p}_k := p_{l(k)}t_{l(k)}/r_k$ is the central point of $\bar{\mathcal{M}}_k$. The geodesic balls with center q and radius ρ will be denoted by $\mathcal{B}_\rho(q)$ in any of these manifolds. We next normalize further the maps \bar{N}_k dividing them by their “ L^2 norm”, namely by

$$(69) \quad \mathbf{h}_k := \left(\int_{\mathcal{B}_1(\bar{p}_{l(k)})} |\bar{N}_k|^2 \right)^{1/2},$$

i.e. we set

$$(70) \quad N_k^b := \frac{\bar{N}_k}{\mathbf{h}_k}.$$

Observe that \mathbf{h}_k must be positive: by the discussion in Section 12.3, it can vanish only if $\mathscr{W}^{l(k)}$ is empty. However in the latter case T_{0,r_k} would coincide with Q copies of the (smooth) manifold $\bar{\mathcal{M}}_k$, which cannot be the case because 0 is a singular point.

We wish to take a limit for (a suitable subsequence of) N_k^b : since $\bar{\mathcal{M}}_k$ “flattens” (i.e. converges to π_0) we hope that the limit N_∞^b is a Q -valued map which has L^2 norm equal to 1, because the convergence is strong in L^2 . Such strong convergence will be achieved if we could prove that the rescaled maps have bounded Dirichlet energy, namely if we had an inequality of type

$$(71) \quad \int_{\mathcal{B}_1(\bar{p}_k)} |DN_k^b|^2 \leq C,$$

for some constant C which does not depend on k . In turn this bound corresponds to a “reverse Sobolev” inequality for the $N_{l(k)}$, i.e.

$$\int_{\mathcal{B}_r(p_{l(k)})} |DN_{l(k)}|^2 \leq Cr^{-2} \int_{\mathcal{B}_r(p_{l(k)})} |N_{l(k)}|^2.$$

Recalling Section 8.3 this is precisely the type of information that the monotonicity of the frequency function yields for the blow-ups of the Dir-minimizing Q -valued functions. In the case at hand we know that $N_{l(k)}$ must be close to a Dir-minimizing function because it approximates well an area minimizing current. We can therefore hope that the frequency function of $N_{l(k)}$ is almost monotone.

We thus consider the Lipschitz (piecewise linear) function $\phi : [0 + \infty[\rightarrow [0, 1]$ given by

$$\phi(r) := \begin{cases} 1 & \text{for } r \in [0, \frac{1}{2}], \\ 2 - 2r & \text{for } r \in]\frac{1}{2}, 1], \\ 0 & \text{for } r \in]1, +\infty[. \end{cases}$$

For every interval of flattening $I_j =]s_j, t_j]$, let N_j be the normal approximation of T_j .

DEFINITION 13.1 (Frequency functions, cf. [32, Definition 3.1]). For every $r \in]0, 3]$ we define:

$$\mathbf{D}_j(r) := \int_{\mathcal{M}^j} \phi\left(\frac{d_j(p)}{r}\right) |DN_j|^2(p)$$

and

$$\mathbf{H}_j(r) := - \int_{\mathcal{M}^j} \phi'\left(\frac{d_j(p)}{r}\right) \frac{|N_j|^2(p)}{d_j(p)} dp,$$

where $d_j(p)$ is the geodesic distance on \mathcal{M}_j between p and p_j . Finally we set

$$\mathbf{I}_j(r) := \frac{r\mathbf{D}_j(r)}{\mathbf{H}_j(r)}.$$

The following is the main analytical estimate of the proof of Theorem 3.4, which allows us to exclude infinite order of contact among the different sheets of a minimizing current.

THEOREM 13.2 (Main frequency estimate, cf. [32, Theorem 3.2]). *There exists a geometric constant C_0 such that, for every $[a, b] \subset [\frac{s_j}{t_j}, 1]$ with $\mathbf{H}_j|_{[a,b]} > 0$, we have*

$$(72) \quad \mathbf{I}_j(a) \leq C_0(1 + \mathbf{I}_j(b)).$$

Obviously if we can find an upper bound for $\mathbf{I}_j(1)$ independent of j , we would gain that the frequency function is always bounded and this would give the desired reverse Sobolev inequality. If the intervals of flattening are finite, then there is nothing to prove. What happens however if the intervals are not finite? Consider some “starting point” t_j . It essentially can be of two kinds:

- $t_j = s_{j-1}$; in this case the center manifold \mathcal{M}_{j-1} had ceased to be “good” at scale s_{j-1} and we encounter a cube $L \in \mathcal{W}^{j-1}$ which is of size s_{j-1}/t_{j-1} where the refining procedure has stopped. It is the first scale at which we encounter a large cube close to 0: for this very precise reason it is not difficult to show that it must necessarily be a cube where either the excess or the height condition fails. But then the discussion in Section 12.3 implies that the approximating map N_{j-1} is well separated at this scale: this means that the denominator of the frequency function \mathbf{I}_{j-1} is not too small compared to the numerator. It remains to transfer this information to the frequency function \mathbf{I}_j : note however that at the scale which interests us there is a large overlap between the graphs of (suitable rescalings of) N_j and N_{j-1} since they both approximate the same piece of the same current.
- $t_j < s_{j-1}$; in this case the excess of T_{0,s_j} is not small enough to construct a center manifold. t_j is then the first radius at which the excess goes below the desired threshold. As such we also know that $\mathbf{E}(T, \mathbf{B}_{t_j})$ cannot be *too* small, i.e. the current cannot be too close to a flat plain at the scale t_j . But then we can expect that the denominator $\mathbf{H}_j(1)$ is not too small compared to the numerator $\mathbf{D}_j(1)$.

This discussion can be made rigorous and we refer the reader to [32, Theorem 5.1] for the details.

13.3. The monotonicity of the frequency function. We close this section with an idea of how Theorem 13.2 is proved. Recalling the proofs of Theorem 8.4 and Theorem 8.6, the monotonicity of the corresponding quantity for Dir-minimizing functions rely on two identities which correspond to “internal” and “external” variations, cf. Proposition 8.2. In turn both variations can be interpreted as suitable deformations of the graphs of the functions. We then know that:

- The variations of the Dirichlet energy for N_j is close to the variation of the mass of its graph;
- The variation of the mass of the graph of N_j is close to the variation of the mass of T_{0,t_j} , which in turn is 0 because it is area minimizing (this must be suitably adjusted if the minimizing property is inside some ambient manifold Σ).

We thus can write down identities which correspond to those of Proposition 8.2 but will include several error terms. We wish to control these error terms, hoping to derive, for instance, an inequality of the form

$$\frac{d}{dr} \mathbf{I}_j(r) \geq -E(r) \mathbf{I}_j(r),$$

which in turn gives the monotonicity of $\mathbf{I}_j(r)e^{\int_0^r E(\tau)d\tau}$. In fact the inequality which we derive is somewhat weaker, but still good enough to show Theorem 13.2.

In order to control the error terms, we localize them in each Whitney region \mathcal{L} of \mathcal{M}_j , since the “contact set” Γ thus not contribute to them. We then use the estimates in Theorem 12.2 to bound such error terms with suitable powers of $\ell(L)$. In turn we bound such powers with $\int_{\Omega} |N_j^2|$ or with $\int_{\Omega} |DN_j|^2$, following the discussion in Section 12.3, where Ω is a suitable domain close to \mathcal{L} . Observe the following crucial fact: we have some freedom in the choice of the domain Ω for each cube \mathcal{L} , because Ω just needs to have essentially comparable size and be at a comparable distance. This freedom is fundamental, since the error terms must be controlled finally at the very same scale r , namely an inequality of the form

$$\frac{d}{dr}\mathbf{I}_j(r) \geq -E(r)\mathbf{I}_j(s) \quad \text{for some } s(r) > r ,$$

is completely useless for our purposes, even if $s(r)$ is only slightly larger than r . For a more careful description on how these difficulties are overcome we refer the reader to [72, Sections 5.7, 6.1 and 6.2].

14. The persistence of singularities

We have now finally gained our “blown-up” map N_{∞}^b and we know that it is a Q -valued map on $B_1(0, \pi_0)$. Note that the estimate (68) (and the lower bounds discussed in Section 12.3) deliver the extra information that $\boldsymbol{\eta} \circ N_{\infty}^b \equiv 0$. This also helps us in the Taylor expansion of the area functional to conclude that N_{∞}^b is Dir-minimizing, cf. Theorem 12.1 (we are of course ignoring the complications given by the ambient Riemannian manifold Σ). We now wish to succeed where the strategy outlined in Section 10 failed, i.e. in showing that the blown-up map N_{∞}^b must “remember” the singularities of the rescaled currents. Note however that we just need to show that the lower bound

$$\mathcal{H}_{\infty}^{m-2+\alpha}(\mathbf{B}_1 \cap \text{Sing}_Q(T_{0,r_k})) \geq \eta > 0$$

induces a similar lower bound on the Hausdorff measure of the set of points p where $N_{\infty}^b(p) = Q \llbracket \boldsymbol{\eta} \circ N_{\infty}^b(p) \rrbracket$. Indeed, from such a lower bound and Theorem 7.9 we would conclude that N_{∞}^b is a classical single valued harmonic function counted Q times. On the other hand $\boldsymbol{\eta} \circ N_{\infty}^b \equiv 0$ would then imply $N_{\infty}^b \equiv Q \llbracket 0 \rrbracket$. This would finally contradict what we concluded from the previous section, namely that

$$\int |N_{\infty}^b|^2 = 1 .$$

Hence, consider the closed sets $D_Q(T_{0,r_k} \cap \bar{\mathbf{B}}_1)$ and let Γ be their Hausdorff limit (after extraction of a subsequence). We wish to show that most points in Γ are points of “multiplicity Q ” for N_{∞}^b , or briefly Q -points for N_{∞}^b .

Recall that N_∞^b is continuous (because Dir-minimizing). Thus a statement analogous to Theorem 10.3 would guarantee that *all* points of Γ are Q -points for N_∞^b . However such a theorem is not available at present. We rather show that the subset of points of Γ which fail to be Q -points of N_∞^b must be a set of $\mathcal{H}^{m-2+\alpha}$ -measure, which however is enough for our purposes. We finish therefore this survey by sketching the argument.

Indeed we will follow a slightly different strategy. Summarizing what achieved so far, N_∞^b is a nontrivial Dir-minimizing map which has $\boldsymbol{\eta} \circ N_\infty^b \equiv 0$. Thus the set of points of multiplicity Q for N_∞^b coincides with the closed set of p where $N_\infty^b(p) = Q \llbracket 0 \rrbracket$. Such set must have $\mathcal{H}^{m-2+\alpha}$ measure zero. We can then identify a closed set Λ for which, at the same time, we have

- For some positive ϑ

$$\sum_i |N_{\infty,i}^b - \boldsymbol{\eta} \circ N_{\infty,i}^b|^2 \geq 2\vartheta \quad \text{everywhere on } \Lambda$$

- Λ is the Hausdorff limit of closed subsets Λ_k of $D_Q(T_{0,r_k})$ with $\mathcal{H}^{m-2+\alpha}$ measure larger than $\frac{\eta}{2}$.

By the Hölder continuity of N_∞^b there must be a fixed constant $\bar{\sigma}$ such that

$$\int_{B_\sigma(q)} \sum_i |N_{\infty,i}^b - \boldsymbol{\eta} \circ N_{\infty,i}^b|^2 \geq \vartheta > 0 \quad \forall q \in \Lambda, \forall \sigma < \bar{\sigma}.$$

We now fix a positive $\sigma < \bar{\sigma}$, whose choice will be specified only at the very end.

By L^2 convergence, for k large we inherit the inequality

$$(73) \quad \int_{B_\sigma(q)} \sum_i |\bar{N}_{k,i} - \boldsymbol{\eta} \circ \bar{N}_k|^2 \geq \frac{\vartheta}{2} \mathbf{h}_k^2 > 0 \quad \forall q \in \mathbf{p}_{\bar{\mathcal{M}}_k}(\Lambda_k),$$

where $\mathbf{p}_{\bar{\mathcal{M}}_k}$ denotes the orthogonal projection onto $\bar{\mathcal{M}}_k$. Now observe that, for k large enough, it is also true that for any $q \in \mathbf{p}_{\bar{\mathcal{M}}_k}(\Lambda_k)$ there is a point $p \in D_Q(T_{0,t_{l(k)}})$ which is in the proximity of q , at a scale much smaller than σ .

A very favorable situation is when q belongs to the contact set where T_{0,r_k} coincides with $Q \llbracket \bar{\mathcal{M}}_k \rrbracket$ (which in particular implies that q itself belongs to $D_Q(T_{0,t_{l(k)}})$): in this case

$$\lim_{\rho \downarrow 0} \int_{B_\rho(q)} \sum_i |\bar{N}_{k,i} - \boldsymbol{\eta} \circ \bar{N}_k|^2 = 0$$

On the other hand, even if this is not the case, it is possible to argue that for some radius $\rho \ll \sigma$ the integral above goes below the threshold $\frac{\vartheta}{4} \mathbf{h}_k^2$, namely

$$(74) \quad \int_{B_\rho(q)} \sum_i |\bar{N}_{k,i} - \boldsymbol{\eta} \circ \bar{N}_k|^2 \leq \frac{\vartheta}{4} \mathbf{h}_k^2.$$

Now, the reader familiar with Morrey spaces will realize that (73) and (74) force the existence of some intermediate radius $t(q) \in]\rho, \sigma[$ with

$$(75) \quad \frac{c_0 \vartheta}{\sigma^\alpha} \mathbf{h}_k^2 \leq \frac{1}{t(q)^{m-2+\alpha}}, \int_{\mathcal{B}_{t(q)}(q)} |D\bar{N}_k|^2,$$

where c_0 is a universal constant.

For each $p \in \Lambda_k$ let $t(p)$ be $t(\mathbf{p}_{\bar{\mathcal{M}}_k}(p))$. Now using an elementary covering argument we can cover Λ_k with balls $\mathbf{B}_{10t(p_i)}(p_i)$ in such a way that the balls $\mathcal{B}_{2t(p_i)}(p_i)$ are disjoint. Being $\bar{\mathcal{M}}_k$ almost flat it is not difficult to see that even the balls $\mathcal{B}_{t(q_i)}(q_i) := \mathcal{B}_{t(p_i)}(\mathbf{p}_{\bar{\mathcal{M}}_k}(p_i))$ must be disjoint. Since Λ_k has $\mathcal{H}^{m-2+\alpha}$ measure larger than $\eta/2$, we achieve

$$(76) \quad \begin{aligned} \frac{\eta}{2} &\leq \sum_i \omega_{m-2+\alpha} (10t(q_i))^{m-2+\alpha} \leq \sum_i \frac{C\sigma^\alpha}{\mathbf{h}_k^2} \int_{\mathcal{B}_{t(q_i)}(q_i)} |D\bar{N}_k|^2 \\ &\leq \frac{C\sigma^\alpha}{\mathbf{h}_k^2} \int_{\mathcal{B}_1(\bar{p}_k)} |D\bar{N}_k|^2 \leq C\sigma^\alpha. \end{aligned}$$

In the last inequality we have used that \mathbf{h}_k^2 , the L^2 norm of \bar{N}_k , controls the Dirichlet energy, recall (69), (70) and (71). Although the constant C in (76) depends upon the sequence of rescalings, the current T and several other parameters, it does not depend upon σ and η . Thus for a suitable choice of σ we reach a contradiction.

References

- [1] W. K. Allard. On the first variation of a varifold. *Ann. Math. (2)*, 95:417–491, 1972.
- [2] F. J. Almgren, Jr. Existence and regularity almost everywhere of solutions to elliptic variational problems among surfaces of varying topological type and singularity structure. *Ann. Math. (2)*, 87:321–391, 1968.
- [3] F. J. Almgren, Jr. Some interior regularity theorems for minimal surfaces and an extension of Bernstein’s theorem. *Ann. Math. (2)*, 84:277–292, 1966.
- [4] F. J. Almgren, Jr. Existence and regularity almost everywhere of solutions to elliptic variational problems with constraints. *Mem. Am. Math. Soc.*, 4(165):viii+199, 1976.
- [5] F. J. Almgren, Jr. *Almgren’s Big Regularity Paper*. World Scientific Monograph Series in Mathematics, vol. 1. World Scientific Publishing Co. Inc., 2000.
- [6] L. Ambrosio. Metric space valued functions of bounded variation. *Ann. Sc. Norm. Super. Pisa, Cl. Sci. (4)*, 17(3):439–478, 1990.
- [7] L. Ambrosio, C. De Lellis, and T. Schmidt. Partial regularity for area minimizing currents in Hilbert spaces. To appear in *J. Reine Angew. Math.*, 2013.
- [8] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.
- [9] L. Ambrosio and B. Kirchheim. Currents in metric spaces. *Acta Math.*, 185(1):1–80, 2000.
- [10] E. Bombieri. Regularity theory for almost minimal currents. *Arch. Ration. Mech. Anal.*, 78(2):99–130, 1982.
- [11] E. Bombieri, E. De Giorgi, and E. Giusti. Minimal cones and the Bernstein problem. *Invent. Math.*, 7:243–268, 1969.
- [12] R. Caccioppoli. Misura e integrazione sugli insiemi dimensionalmente orientati. *Atti Accad. Naz. Lincei, Rend. Cl. Sci. Fis. Mat. Nat. (8)*, 12:3–11, 1952.

- [13] S. X. Chang. Two-dimensional area minimizing integral currents are classical minimal surfaces. *J. Am. Math. Soc.*, 1(4):699–778, 1988.
- [14] J. Cheeger, R. Haslhofer, and A. Naber. Quantitative stratification and the regularity of mean curvature flow. *Geom. Funct. Anal.*, 23(3):828–847, 2013.
- [15] J. Cheeger and A. Naber. Lower bounds on Ricci curvature and quantitative behavior of singular sets. *Invent. Math.*, 191(2):321–339, 2013.
- [16] J. Cheeger and A. Naber. Quantitative stratification and the regularity of harmonic maps and minimal currents. *Commun. Pure Appl. Math.*, 66(6):965–990, 2013.
- [17] J. Cheeger, A. Naber, and D. Valtorta. Critical sets of elliptic equations. *Commun. Pure Appl. Math.*, 68(2):173–209, 2015.
- [18] G. David. Should we solve Plateau’s problem again? In C. Fefferman, A. D. Ionescu, D. H. Phong, S. Wainger, editors, *Advances in Analysis: The Legacy of Elias M. Stein*. Princeton University Press, 2014.
- [19] E. De Giorgi. Su una teoria generale della misura $(r - 1)$ -dimensionale in uno spazio ad r dimensioni. *Ann. Mat. Pura Appl. (4)*, 36:191–213, 1954.
- [20] E. De Giorgi. Nuovi teoremi relativi alle misure $(r - 1)$ -dimensionali in uno spazio ad r dimensioni. *Ric. Mat.*, 4:95–113, 1955.
- [21] E. De Giorgi. *Frontiere Orientate di Misura Minima*. Seminario di Matematica della Scuola Normale Superiore di Pisa, 1960–61. Editrice Tecnico Scientifica, Pisa, 1961.
- [22] E. De Giorgi. Una estensione del teorema di Bernstein. *Ann. Sc. Norm. Super. Pisa, Cl. Sci. (3)*, 19:79–85, 1965.
- [23] E. De Giorgi. *Selected Papers*. L. Ambrosio, G. Dal Maso, M. Forti, M. Miranda and S. Spagnolo, editors. Springer Collected Works in Mathematics. Springer, Heidelberg, 2013.
- [24] C. De Lellis. Allard’s interior regularity theorem: an invitation to stationary varifolds. Lecture notes available at http://www.math.uzh.ch/fileadmin/user/delellis/publikation/allard_24.pdf, 2012.
- [25] C. De Lellis. The size of the singular set of area-minimizing currents. *Preprint*, 2015.
- [26] C. De Lellis, F. Ghiraldin, and F. Maggi. A direct approach to Plateau’s problem. To appear in *J. Eur. Math. Soc.*, 2014.
- [27] C. De Lellis, C. R. Grisanti, and P. Tilli. Regular selections for multiple-valued functions. *Ann. Mat. Pura Appl. (4)*, 183(1):79–95, 2004.
- [28] C. De Lellis and E. Spadaro. Center manifold: A case study. *Discrete Contin. Dyn. Syst.*, 31(4):1249–1272, 2011.
- [29] C. De Lellis and E. Spadaro. Q -valued functions revisited. *Mem. Am. Math. Soc.*, 211(991):vi+79, 2011.
- [30] C. De Lellis and E. Spadaro. Multiple valued functions and integral currents. To appear in *Ann. Sc. Norm. Super. Pisa, Cl. Sci. (5)*, 2013.
- [31] C. De Lellis and E. Spadaro. Regularity of area-minimizing currents II: Center manifold. To appear in *Ann. Math. (2)*, 2013.
- [32] C. De Lellis and E. Spadaro. Regularity of area-minimizing currents III: Blow-up. To appear in *Ann. Math. (2)*, 2013.
- [33] C. De Lellis and E. Spadaro. Regularity of area minimizing currents I: Gradient L^p estimates. *Geom. Funct. Anal.*, 24(6):1831–1884, 2014.
- [34] C. De Lellis, E. Spadaro, and L. Spolaor. Regularity theory for 2-dimensional almost minimal currents I: Lipschitz approximation. *Preprint*, 2015.
- [35] C. De Lellis, E. Spadaro, and L. Spolaor. Regularity theory for 2-dimensional almost minimal currents II: Center manifold. *Preprint*, 2015.
- [36] C. De Lellis, E. Spadaro, and L. Spolaor. Regularity theory for 2-dimensional almost minimal currents III: blowup. *Preprint*, 2015.
- [37] C. De Lellis, E. Spadaro, and L. Spolaor. Uniqueness of tangent cones for 2-dimensional almost minimizing currents. *Preprint*, 2015.

- [38] G. De Philippis, A. De Rosa, and F. Ghiraldin. A direct approach to Plateau's problem in any codimension. *Preprint*, 2015.
- [39] G. De Philippis and E. Paolini. A short proof of the minimality of Simons cone. *Rend. Semin. Mat. Univ. Padova*, 121:233–241, 2009.
- [40] G. De Rham. *Variétés Différentiables. Formes, Courants, Formes Harmoniques*. Actualités Sci. Ind., vol. 1222, Publ. Inst. Math. Univ. Nancago III. Hermann et Cie, Paris, 1955.
- [41] U. Dierkes, S. Hildebrandt, and F. Sauvigny. *Minimal Surfaces*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 339. Second edition. With assistance and contributions by A. Küster and R. Jakob. Springer, Heidelberg, 2010.
- [42] J. Douglas. Solution of the problem of Plateau. *Trans. Am. Math. Soc.*, 33(1):263–321, 1931.
- [43] L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.
- [44] H. Federer. Some theorems on integral currents. *Trans. Am. Math. Soc.*, 117:43–67, 1965.
- [45] H. Federer. *Geometric Measure Theory*. Die Grundlehren der mathematischen Wissenschaften, vol. 153. Springer-Verlag New York Inc., New York, 1969.
- [46] H. Federer. The singular sets of area minimizing rectifiable currents with codimension one and of area minimizing flat chains modulo two with arbitrary codimension. *Bull. Am. Math. Soc.*, 76:767–771, 1970.
- [47] H. Federer and W. H. Fleming. Normal and integral currents. *Ann. Math. (2)*, 72:458–520, 1960.
- [48] W. H. Fleming. On the oriented Plateau problem. *Rend. Circ. Mat. Palermo (2)*, 11:69–90, 1962.
- [49] M. Focardi, A. Marchese, and E. Spadaro. Improved estimate of the singular set of dir-minimizing q -valued functions via an abstract regularity result. To appear in *J. Funct. Anal.*, 2013.
- [50] N. Garofalo and F. Lin. Monotonicity properties of variational integrals, A_p weights and unique continuation. *Indiana Univ. Math. J.*, 35(2):245–268, 1986.
- [51] N. Garofalo and F. Lin. Unique continuation for elliptic operators: A geometric-variational approach. *Commun. Pure Appl. Math.*, 40(3):347–366, 1987.
- [52] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Classics in Mathematics. Reprint of the 1998 edition. Springer-Verlag, Berlin, 2001.
- [53] E. Giusti. *Minimal Surfaces and Functions of Bounded Variation*. Monographs in Mathematics, vol 80. Birkhäuser Verlag, Basel, 1984.
- [54] J. Harrison and H. Pugh. Existence and soap film regularity of solutions to Plateau's problem. *Preprint*, 2013. [arXiv:1310.0508](https://arxiv.org/abs/1310.0508).
- [55] R. Harvey and H. B. Lawson Jr. Calibrated geometries. *Acta Math.*, 148:47–157, 1982.
- [56] R. L. Jerrard and H. M. Soner. Functions of bounded higher variation. *Indiana Univ. Math. J.*, 51(3):645–677, 2002.
- [57] A. Naber and D. Valtorta. The singular structure and regularity of stationary and minimizing varifolds. *Preprint*, 2015.
- [58] T. Radó. On Plateau's problem. *Ann. Math. (2)*, 31(3):457–469, 1930.
- [59] E. R. Reifenberg. Solution of the Plateau Problem for m -dimensional surfaces of varying topological type. *Acta Math.*, 104:1–92, 1960.
- [60] E. R. Reifenberg. On the analyticity of minimal surfaces. *Ann. Math. (2)*, 80:15–21, 1964.
- [61] Yu. G. Reshetnyak. Sobolev classes of functions with values in a metric space. *Sib. Mat. Zh.*, 38(3):657–675, iii–iv, 1997.

- [62] Yu. G. Reshetnyak. Sobolev classes of functions with values in a metric space. II. *Sib. Mat. Zh.*, 45(4):855–870, 2004.
- [63] T. Rivière. A lower-epiperimetric inequality for area-minimizing surfaces. *Commun. Pure Appl. Math.*, 57(12):1673–1685, 2004.
- [64] R. Schoen and L. Simon. Regularity of stable minimal hypersurfaces. *Commun. Pure Appl. Math.*, 34(6):741–797, 1981.
- [65] R. Schoen and L. Simon. A new proof of the regularity theorem for rectifiable currents which minimize parametric elliptic functionals. *Indiana Univ. Math. J.*, 31(3):415–434, 1982.
- [66] R. Schoen, L. Simon, and S. T. Yau. Curvature estimates for minimal hypersurfaces. *Acta Math.*, 134(3-4):275–288, 1975.
- [67] L. Simon. Asymptotics for a class of nonlinear evolution equations, with applications to geometric problems. *Ann. Math. (2)*, 118(3):525–571, 1983.
- [68] L. Simon. *Lectures on Geometric Measure Theory*. Proceedings of the Centre for Mathematical Analysis, Australian National University, vol. 3. Australian National University Centre for Mathematical Analysis, 1983.
- [69] L. Simon. Rectifiability of the singular sets of multiplicity 1 minimal surfaces and energy minimizing maps. In *Surveys in Differential Geometry, Vol. II (Cambridge, MA, 1993)*, pp. 246–305. Int. Press, Cambridge, MA, 1995.
- [70] L. Simon. *Theorems on Regularity and Singularity of Energy Minimizing Maps*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 1996. Based on lecture notes by Norbert Hungerbühler.
- [71] J. Simons. Minimal varieties in riemannian manifolds. *Ann. Math. (2)*, 88:62–105, 1968.
- [72] E. Spadaro. Regularity of higher codimension area minimizing integral currents. In L. Ambrosio, editor, *Geometric Measure Theory and Real Analysis*, pp. 131–192, Proceedings of the ERC school in Pisa. Edizioni SNS, 2014.
- [73] E. N. Spadaro. Complex varieties and higher integrability of Dir-minimizing Q -valued functions. *Manuscr. Math.*, 132(3-4):415–429, 2010.
- [74] E. M. Stein and G. Weiss. *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton Mathematical Series, vol. 32. Princeton University Press, Princeton, N.J., 1971.
- [75] D. Triscari. Sulle singolarità delle frontiere orientate di misura minima. *Ann. Sc. Norm. Super. Pisa, Cl. Sci. (3)*, 17:349–371, 1963.
- [76] C. Villani. *Topics in Optimal Transportation*. Graduate Studies in Mathematics, vol. 58. American Mathematical Society, Providence, RI, 2003.
- [77] B. White. Tangent cones to two-dimensional area-minimizing integral currents are unique. *Duke Math. J.*, 50(1):143–160, 1983.
- [78] B. White. The least area bounded by multiples of a curve. *Proc. Am. Math. Soc.*, 90(2):230–232, 1984.
- [79] B. White. Stratification of minimal surfaces, mean curvature flows, and harmonic maps. *J. Reine Angew. Math.*, 488:1–35, 1997.
- [80] N. Wickramasekera. A general regularity theory for stable codimension 1 integral varifolds. *Ann. Math. (2)*, 179(3):843–1007, 2014.
- [81] W. Wirtinger. Eine Determinantenidentität und ihre Anwendung auf analytische Gebilde in euklidischer und Hermitescher Maßbestimmung. *Monatshefte Math. Phys.*, 44(1):343–365, 1936.
- [82] L. C. Young. Some extremal questions for simplicial complexes. V. The relative area of a Klein bottle. *Rend. Circ. Mat. Palermo (2)*, 12:257–274, 1963.