Two Dimensional Area Minimizing Integral Currents are Classical Minimal Surfaces

Author(s): Sheldon Xu-Dong Chang
Source: *Journal of the American Mathematical Society*, Vol. 1, No. 4 (Oct., 1988), pp. 699-778
Published by: American Mathematical Society
Stable URL: https://www.jstor.org/stable/1990991
Accessed: 09-05-2020 20:12 UTC

JOURNAL OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 1, Number 4, October 1988

TWO DIMENSIONAL AREA MINIMIZING INTEGRAL CURRENTS ARE CLASSICAL MINIMAL SURFACES

SHELDON XU-DONG CHANG

CONTENTS

1. Introduction
2. Proof of the main regularity result
3. Terminology and basic facts
4. Assumptions and squash deformations
5. Squeeze deformation
6. Comparison surface
7. Height control
8. Order of contact
9. Order of contact revisited
10. Rate of convergence
11. Toward uniqueness
12. Separation
13. Appendix A
14. Appendix B
15. Appendix C

1. INTRODUCTION

In [A] F. Almgren proved that any area minimizing integral current on a Riemannian manifold is a smooth submanifold in the interior except a possible singular set of at most codimension two. That is so far the most general result about the regularity of generalized solutions to the problem of finding area minimizing submanifolds on a Riemannian manifold in arbitrary dimension and codimension. That result implies any two dimensional area minimizing integral current is a smooth surface except a possible 0-dimensional singular set. However, a 0-dimensional set may be very 'big.' In particular, it may not consist of just isolated points as illustrated by some Cantor type sets (not

Received by the editors June 9, 1987.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 49F20; Secondary 49F10.
The results here were presented to the Society at the annual meeting in San Antonio, 1987. The work was done with support from Princeton University.

0894-0347/88 $1 00 + $.25 per page

those obtained through the standard construction). On the other hand, all the known examples of two dimensional area minimizing currents have only isolated singular points. So the question is naturally raised whether the singular set in a two dimensional area minimizing integral current consists of just isolated points (see §12 in [Y] and the collection of open problems in [GMT]). In this paper we answer this question affirmatively. Furthermore, we give a complete description of the local structure near any interior singular point. Since it is well known that any complex curve on a Kaehler manifold is area minimizing and it may have singularities, the result obtained in this paper is optimal.

Combining this regularity result and the existence of area minimizing integral currents spanning a given boundary or representing a given integral homology class on a compact Riemannian manifold, we know now that any null homologous curve on a Riemannian manifold bounds a least area surface which is a classical minimal surface in the interior and any nontrivial two dimensional integral homology class on a compact Riemannian manifold can be represented by a finite integral linear combination of classical closed minimal surfaces which have least area and at most finitely many intersection points. As a consequence, we know the existence of closed minimal surfaces on any compact Riemannian manifold with nontrivial second integral homology group.

This result is proved by using the theory of multiple-valued functions developed in [A] to give a detailed analysis of the local structure near an interior singular point. We use two main tools. One is the branched center manifold which is a modified version of the center manifold introduced in [A]. The center manifold is very useful in the study of local structure of singular sets. In our context it eventually serves as a wedge inserted to separate the singularity. The other is the analysis of the lowest order term in the multiple-valued function approximating a two dimensional area minimizing current over a branched center manifold. To establish the existence and the uniqueness of such a lowest order term and analyze the difference between the lowest order term and the approximation function itself are the essential parts of this paper.

For a locally irreducible two dimensional area minimizing integral current, we construct two finite sequences of branched discs $\{N_l\}$, $\{Y_l\}$ near an interior singular point. The two sequences are related as follows

$$Y_1 \leftarrow N_1 \leftarrow Y_2 \leftarrow N_2 \leftarrow \cdots \leftarrow Y_k \leftarrow N_k$$

where k is bounded by the density of the integral current at the singular point, Y_1 is a small disc in the tangent plane and Y_{l+1} is a multi-sheet branched covering of N_l. To obtain this multi-sheet covering over N_l is the core of this paper. N_{l+1} is obtained through the center manifold construction and basically has the same topological structure as Y_{l+1}.

The analytical results in this paper enable us to conclude that the current coincides with the branched disc at the top level.

The present paper is organized as follows.

In §2, we present the main analytical results and show how to use them to get the regularity result. From §4 to §12 we prove Theorem B, which tells how to get Y_{i+1} from N_i if the current does not coincide with N_i. This is essential to the scheme of gradually separating the surface near a singular point. §3 gives the basic facts about geometric measure theory and multiple-valued functions, some results about the multiple-valued functions minimizing Dirichlet integrals and isothermal coordinates on a surface with a branched point. In §§4 and 5, we demonstrate how to use first variations to obtain two formulas which are familiar in the context of harmonic functions. Similar calculations appeared in [A] in a different context. The success of this work is very much due to the results in §7, where we prove with lots of work that the height integral of the multiple-valued function approximating an area minimizing current can be bounded by its Dirichlet integral. Using this we are able to define the order of contact for the multiple-valued function approximating an area minimizing current. The idea of using the order of contact to control the local behavior near the singular set was introduced in [A] where it was called frequency and proved to be a well-defined quantity for multiple-valued functions minimizing the Dirichlet integrals. In this paper we show that the order of contact is well-defined for two dimensional area minimizing currents. It is demonstrated in §§9 and 10 that the order of contact strongly controls the local behavior of the current. By applying the construction of a comparison surface generated from a multiple-valued harmonic function, we show that the lowest order term in the multiple-valued function is well defined. In §§11 and 12 it is proved that this lowest order term provides a branched disc to separate the current once more if the current does not coincide with N_i near the singular point.

In Appendix A we describe how to extend the results of F. Morgan and B. White to the case of Riemannian manifolds and we prove that a two dimensional area minimizing integral current approaches its unique tangent cone with a rate of $1 + \varepsilon$ pointwise. This is necessary for the construction of the first center manifold. In Appendices B and C we briefly describe the construction of the center manifold in [A] with necessary modifications. We use the construction in [A] to patch together a branched center manifold and express most of the area minimizing current as the graph of a multiple-valued function over this branched surface. We also describe how to get estimates on the area of the part of the current missed by the graph of this multiple-valued function and how 'centered' the branched surface is.

This paper is the revised version of the author's Ph.D. thesis at Princeton University. The author wishes to acknowledge his sincere thanks to his thesis advisor Professor Fredrick Almgren, Jr. for encouragement and many helpful discussions during the preparation of this work, especially the discussions of the construction of the center manifolds in his paper. Also the author wishes to thank Princeton University and the mathematics department for financial support during 1983–1986.

This paper is dedicated to the memory of my grandmother.

2. Proof of the main regularity result

We assume the readers have some basic knowledge of geometric measure theory. In this section we show how to use two technical results proved in this paper to prove the main regularity result.

Let T be a two dimensional (locally) area minimizing integral current on a C^5 Riemannian manifold M of dimension $2+m$ which is isometrically embedded in the Euclidean space R^{2+m+n} and $0 \in \operatorname{spt} T \sim \operatorname{spt} \partial T \subset M$.

The following is the regularity result we intend to prove in this paper.

Main regularity result. *Let M and T be as above. Then T can be decomposed into finitely many pieces of locally irreducible integral currents (cf. §3, Definition 3.1) near 0. The supports of any two such locally irreducible currents intersect only at 0.*

In a small neighborhood of 0, each such locally irreducible current is a $C^{3,\alpha}$ disc with a possible branched point at 0.

Hereafter we assume the integral current T is locally irreducible at 0. There is no loss of generality in doing so, according to the discussion after Definition 3.1. Under this assumption, we know from Appendix A that the current has a two dimensional plane with integer multiplicity as its unique tangent cone at 0, denoted by T_0 throughout this paper. We let this plane be $R^2 \times \{0\} \subset T_0M = R^{2+m} \times \{0\} \subset R^{2+m+n}$.

First we give a description of one of the two technical results proved in this paper.

Let Y be either the tangent plane $R^2 \times \{0\}$ or a $C^{3,\varepsilon}$ branched disc which is a K-sheet covering of a small neighborhood of 0 in the plane $R^2 \times \{0\}$. In the second case, we assume different branches do not intersect and approach the point 0 with a certain order. Let V be a horn neighborhood (formally called an admissible neighborhood as in Definition 3.8) of Y. Intuitively V is a thin neighborhood along each branch of Y when it is a branched covering of the tangent plane. Also V contains the support of T near 0. Then we have the following result.

Theorem A. *There is a surface N in V which is the $C^{3,\varepsilon}$ image of a disc and which has the following properties*

(a) *If Y is the tangent plane, then N is the graph of a function defined over a small neighborhood of 0 in the tangent plane. N is a branched disc if Y is, and they have the same number of sheets of covering over the plane. (Y and N almost look alike.)*

(b) *There is a multiple-valued Lipschitz function, as defined in Definition 3.2, $f_0: N_r \to Q_J(R^{2+m+n})$. For $p \in N_r$, $f_0(p)$ takes 'values' in the normal space $T_p^{\perp} N_r$ and coincides with the 'slice' of the current T in that normal space for most points p except a set $N_r^2 \subset N_r$. N_r^2 and the area of the part of the current* *missed by the graph of* f_0 *are very 'small,' which is made more precise by the assumption* (H3) *in* §4.

(c) *The surface* N *which may be a branched surface is 'highly centered' with respect to the current* T *in the following sense: Take* $p \in N$, *then the support of* $f_0(p)$ *consists of* J *points in* R^{2+m+n}. *Let* $\eta \circ f_0(p)$ *be the average of those* J *points, then the integral of* $|\eta \circ f(p)|$ *is very small, which is made more precise by the assumptions* (H1) *and* (H2) *in* §4.

The construction of center manifolds was first done by F. Almgren in [A]. It was used to carry out a very delicate construction of comparison surfaces which eventually help to bound the Hausdorff measure of the singular sets. We use the center manifolds for a different purpose. As mentioned in the Introduction, they are essentially used as wedges. Also we need center manifolds which are 'centered' all the way to the singular point under consideration. Center manifolds with this property were constructed in [A] only for area minimizing currents having at least 2-degree contact with the tangent plane, whereas in certain situations we have only $1+\varepsilon$-degree contact. We adopt the construction in [A] to get a center manifold with enough centering to meet our need. The proof of the regularity result in this paper can be simplified if the center manifold is exactly centered and there is no error term in the approximation of the area minimizing current using multiple-valued functions.

The next result is used to produce Y_{i+1} from N_i if the current does not coincide with N_i near 0 and then is used to separate the area minimizing current along the different branches of Y_{i+1}.

Let N be a surface of the type as in Theorem A, f_0 be a multiple-valued function defined on N approximating the area minimizing current in the sense of Theorem A, then

Theorem B. *If* T *does not coincide with the center manifold* N *which may be a branched surface, then the multiple-valued function* f_0 *has the following properties:*

(a) *The lowest order term of* f_0 *is a well-defined strictly multiple-valued function denoted by* g_0. *The function* g_0 *is homogeneous and minimizes the Dirichlet integral in a small neighborhood of* 0 *on* N.

(b) *The graph of* g_0 *over* N *defines a* K'*-sheet covering of the surface* N *with* $K' > 1$. *The different branches of this new branched surface* Y' *do not intersect each other except at* 0. *If the original surface* N *is a* K*-sheet covering of the tangent plane with* $K \geq 1$, *then the new surface* Y' *is a* $K \cdot K'$*-sheet covering of the tangent plane near* 0. Y' *is as smooth as* N.

(c) *The support of the area minimizing current* T *lies in a horn neighborhood* (*i.e. admissible neighborhood*) *of the new surface* Y', *where the horn neighborhood of* Y' *is a union of disjoint thin neighborhoods any one of which corresponds to one of the branches of* Y'.

Remark. The above two results can be summarized as follows: If the support of the current T is in a horn neighborhood of a surface Y which is either a plane or a branched disc, then we can insert a wedge (branched center manifold) N. If the current does not coincide with N, then we can separate the current one more time.

Proof of the main regularity result. We prove the main regularity result of this paper here by assuming Theorem A and Theorem B.

We start with $Y = Y_1$ being the tangent plane. We construct a center manifold N_1 which exists according to the estimates in Appendix A and Theorem A and which covers the tangent plane near 0 only once. If the current does not coincide with this center manifold N_1, then from Theorem B, we know the existence of a certain branched surface Y_2 which is a K_2-sheet covering of N_1 with $K_2 > 1$ and has a horn neighborhood (i.e. admissible neighborhood) containing the current T. Thus the current covers the tangent plane near 0 at least $K_1 \cdot K_2$ times.

Next we apply Theorem A to Y_2 to get a center manifold N_2. If T does not coincide with N_2 near 0, then we apply Theorem B to the current T and the center manifold N_2 to conclude the existence of a surface Y_3 such that the current lies in a horn neighborhood of this new surface Y_3. Y_3 forms a $K_1 \cdot K_2 \cdot K_3$-sheet covering of the tangent plane near the origin hence the current covers the tangent plane near 0 at least $K_1 \cdot K_2 \cdot K_3$ times.

Now we start an inductive procedure. Once we know that the current lies in some horn neighborhood of Y_n which is a branched disc covering the tangent plane $K_1 \cdots K_n$ times near 0, then we apply Theorem A to get a branched center manifold N_n. If the current coincides with N_n, then we simply stop, since the current is already the image of a disc near 0. Otherwise we apply Theorem B to get a new surface Y_{n+1}, which is a K_{n+1}-sheet covering of N_n with $K_{n+1} > 1$, hence a $K_1 \cdots K_n K_{n+1}$-sheet covering of the tangent plane near 0.

From the monotonicity of the function

$$\Theta(s) = \frac{\mathbf{M}(T \llcorner B_s^{2+m+n})}{s^2}$$

(cf. [WA] or [LS]) and the results from Appendix A, we know that $\Theta(0) = \lim_{s\to 0} \Theta(s)$ exists and is a finite positive integer. If the current T covers the tangent plane near 0 at least $K_1 \cdots K_n K_{n+1}$ times as above, where $K_2, \dots, K_{n+1} > 1$, using the fact that T is close to the tangent plane with a degree of $1 + \varepsilon$ (see Appendix A), it is easy to verify

$$\Theta(0) = \lim_{s\to 0} \Theta(s) \geq K_1 K_2 \cdots K_n K_{n+1}.$$

Hence the above inductive procedure stops after at most $\Theta(0)$ times. That means the current coincides with a certain N_n which is a $C^{3,\alpha}$ image of a disc. Thus we have the result.

Remark 1. If T is locally irreducible at 0, and $\Theta(0) = p_1 \cdots p_k$ where $\{p_i\}$ are prime numbers, then the above procedure stops after at most $k+1$ times.

Remark 2. If the density of the area minimizing current at 0 is one, then the regularity follows from the results of W. Allard in [WA]. The readers can also consult with more recent exposition by L. Simon in [LS] for this case.

3. Terminology and basic facts

We let R^h be the standard Euclidean space of dimension h.
$B^h(x, r)$, $B_r^h(x)$ both denote the ball in R^h with center x, radius r.
B_r^h denotes $B^h(0, r)$. When $h = 2$, we often omit the upper index.
We define two maps which are used often

$$\mu(r)\colon R^h \to R^h, \qquad r \in R$$
$$\mu(r)(x) = rx, \qquad x \in R^h$$

and

$$\tau(y)\colon R^h \to R^h, \qquad y \in R^h$$
$$\tau(y)(x) = x - y, \qquad x \in R^h$$

For definitions of integral currents and facts about them, the readers should consult with H. Federer's treatise *Geometric measure theory* [F]. They are generalized surfaces, including all the smooth submanifolds, polyhedral chains, analytical varieties, etc.

We continue to let T be a two dimensional area minimizing integral current on a Riemannian manifold M of dimension $2+m$ isometrically embedded in the Euclidean space of dimension $2+m+n$. Let $0 \in \operatorname{spt} T \sim \operatorname{spt} \partial T \subset M \subset R^{2+m+n}$, we assume $T_0 M = R^{2+m} \times \{0\}$ and over $B_4^{2+m} \subset T_0 M = R^{2+m}$ the manifold M can be written as the graph of a certain function denoted by ϕ which satisfies the following:

$$(1) \qquad |\phi(x)| \le |x|^2, \qquad |D\phi(x)| \le |x|, \qquad |D^i \phi(x)| \le 1$$

with $i = 1, \ldots, 5$ and $x \in B_4^{2+m}$. We assume

$$\operatorname{spt} \partial T \cap B_1^{2+m} \times R^n = \emptyset.$$

Definition 3.1. T is said to be locally irreducible at 0 if for each $1 > r > 0$, it is impossible to write $T \llcorner B_r^{2+m+n} = T_1 + T_2$ with $\operatorname{spt} T_1 \cap \operatorname{spt} T_2 = \{0\}$.

We infer from Appendix A that the density $\Theta(\|T\|, p)$ of a two dimensional integral minimizing current on a Riemannian manifold is a positive integer. If the decomposition described as in Definition 3.1 is possible, then T_1, T_2 are area minimizing as well and $\Theta(\|T\|, 0) = \Theta(\|T_1\|, 0) + \Theta(\|T_2\|, 0)$. It readily follows that in a small neighborhood of 0, T can be decomposed only into finitely many pieces of locally irreducible currents. To prove the regularity result, it is enough to prove it for the locally irreducible currents. Hereafter we assume that T is locally irreducible at 0.

From Appendix A, we know that T has a plane with positive integer density as its unique tangent cone T_0 at 0. We assume that $\operatorname{spt} T_0 = R^2 \times \{0\}$.

Multiple-valued functions. The theory of multiple-valued functions was developed in [A]. It is the most natural framework for the regularity theory in geometric measure theory and promises a lot of future development and applications in other fields. Here we introduce the basic notations and facts of multiple-valued functions. The readers are referred to [A, Chapter 1] for more details.

The space $Q_K(R^h)$ consists of all the unordered K points denoted by $\sum_{i=1}^K [\![p_i]\!]$, $p_i \in R^h$. We let $\operatorname{spt} \sum_{i=1}^K [\![p_i]\!] = \cup_{i=1}^K \{p_i\}$. $Q_K(R^h)$ is equipped with the natural metric $\mathscr{G}$ defined by

$$\mathscr{G}\left(\sum_{i=1}^K [\![p_i]\!], \sum_{i=1}^K [\![q_i]\!]\right) = \inf \left(\sum_{i=1}^K |p_i - q_{n_i}|^2\right)^{1/2}$$

where $(n_1, \dots, n_K)$ is a permutation of $(1, \dots, K)$ and the infimum is taken over all the permutations.

We let $|\sum_{i=1}^K [\![p_i]\!]|^2 = \sum_{i=1}^K |p_i|^2 = \mathscr{G}^2(\sum_{i=1}^K [\![p_i]\!], K[\![0]\!])$.

We modify the construction in [A, Chapter 1.2] as suggested by B. White in order to obtain a map

$$\xi: Q_K(R^h) \to Q_K^* \subset R^{hn(K)}$$

with $n(K) \in \mathbf{Z}^+$ such that ξ is a bilipschitz correspondence between Q_K and Q_K^* and for every $p \in Q_K(R^h)$, p has a small neighborhood in Q_K such that ξ is an equidistance map over the neighborhood. The modification mentioned is to choose the orthogonal projections $\Pi_1, \dots, \Pi_{n(K)}$ in [A, Chapter 1.2] as complete sets of coordinate projections corresponding to distinct orthonormal coordinate systems for R^h and to compose the resulting map ξ there with proper scaling to get such a ξ.

Definition 3.2. (a) f is a multiple-valued function if it is a map

$$f: U \subset R^k \to Q_K(R^h)$$

for some $K \in \mathbf{Z}^+$. Associated with f, there is an $\eta \circ f: U \to R^h$ defined as $\eta \circ f(x) = \sum_{i=1}^K p_i$ when $f(x) = \sum_{i=1}^K [\![p_i]\!]$.

(b) f is called a Lipschitz function if there is $C > 0$ such that

$$\mathscr{G}(f(x), f(y)) \le C|x - y|, \qquad x, y \in U.$$

(c) f is called affine if there are $A_1, \dots, A_K$ where each A_i is an affine map from R^k to R^h, such that

$$f(x) = \sum_{i=1}^K [\![A_i(x)]\!].$$

(d) f is called affine approximable at x_0 if there are affine maps $A_1, \dots, A_K$ from R^k to R^h such that

$$\lim_{|x-x_0|\to 0} \frac{\mathscr{G}\left(f(x), \sum_{i=1}^K [\![A_i(x)]\!]\right)}{|x-x_0|} = 0 .$$

(e) f is strongly affine approximable at x_0 if (d) holds for f at x_0 and $A_i = A_j$ if $A_i(x_0) = A_j(x_0)$.

From [A, Chapter 1.4] we know that if f is a multiple-valued Lipschitz function then it is strongly affine approximable almost every where over its domain. It implies that in the graph of f which is defined by

$$\operatorname{graph} f = \{(x, y) \in R^k \times R^h | x \in U , y \in \operatorname{spt} f(x)\}$$

for almost all $x \in U$, the tangent spaces at (x, y) and (x, y') agree if $y = y'$.

If f is affine approximable at x_0 with $\sum_{i=1}^K [\![A_i]\!]$ as its affine approximation, then obviously $f(x_0) = \sum_{i=1}^K [\![A_i(x_0)]\!]$ and $A_i(x) = A_i(x_0) + L_i(x - x_0)$ with $L_i \in \operatorname{Hom}(R^k, R^h)$.

Definition 3.3. If f is affine approximable at x_0, then

(a) $\sum_{i=1}^K [\![L_i]\!] \in Q_K(\operatorname{Hom}(R^k, R^h))$, denoted by $Df(x)$, is defined as the differential of f at x_0, we let $|Df(x_0)|^2 = \sum_{i=1}^K |L_i|^2$ where $|L|$ is the Euclidean norm of the matrix associated with any $L \in \operatorname{Hom}(R^k, R^h)$.

(b) $\sum_{i=1}^K [\![L(v)]\!]$ is defined as the derivative of f at x_0 in the direction v and is denoted by $D_v f \in Q_K(R^h)$. Let $|D_v f(x_0)|^2 = \sum_{i=1}^K |L_i(v)|^2$.

The map ξ mentioned before has the following properties

(2) $$|\xi \circ f| = |f| ,$$

(3) $$|D_v(\xi \circ f)(x)| = |D_v f| .$$

In this paper we are mostly interested in the multiple-valued functions defined over two dimensional domains. We define the following:

Definition 3.4. Assume $f : B_c^2 \subset R^2 \to Q_K(R^h)$ is affine approximable at x_0, and (r, θ) is the polar coordinates for B_c^2, we let

$$|f_r(x_0)|^2 = \sum_{i=1}^K \left| L_i \left(\frac{\partial}{\partial r}\right)\right|^2 ,$$

$$|f_\theta(x_0)|^2 = \sum_{i=1}^K \left| L_i \left(\frac{\partial}{\partial \theta}\right)\right|^2 ,$$

$$f_r \cdot f_\theta(x_0) = \sum_{i=1}^K L_i\left(\frac{\partial}{\partial r}\right) \cdot L_i\left(\frac{\partial}{\partial \theta}\right) ,$$

$$f_r(x_0) \cdot f = \sum_{i=1}^K L_i\left(\frac{\partial}{\partial r}\right) \cdot A_i(x_0) ,$$

$$|Df(x_0)|^2 = |f_r(x_0)|^2 + \frac{1}{r^2}|f_\theta(x_0)|^2 .$$

In [A, Chapter 2], the regularity of multiple-valued functions minimizing Dirichlet integrals was studied. We summarize the main results here.

Let $f: B_r^k \to Q_K(R^h)$, and f is almost everywhere affine approximable in B_r^k, we define the Dirichlet integral of f over B_r^k by

$$D(f, r) = \int_{B_r^k} |Df(x)|^2 \, d\mathscr{L}^k x$$

if the integral exists.

Following [A], we let $\mathscr{Y}$ denote all the functions on ∂B_r^k with L^2 integrable boundary derivatives.

Theorem 3.1. *Let $g: \partial B_r^k \to Q_K(R^h)$, and $g \in \mathscr{Y}$, then there is an $f: B_r^k \to Q_K(R^h)$ with $f|\partial B_r^k = g$ such that*

(a) *$D(f, r) \le D(f', r)$ for any f' with $f'|\partial B_r^k = f|\partial B_r^k = g$, and f is called minimizing the Dirichlet integral.*

(b) *f is Holder continuous and the Holder constant $\alpha = 1/K$ in case $k = 2$.*

(c) *The graph of f is a smooth k dimensional manifold in the interior except for a possible singular set of Hausdorff dimension at most of $k - 2$. In particular, when $k = 2$, the singular set is of dimension 0.*

Remark. In a related work [C], we show that in case $k = 2$, the singularity set mentioned in (c) is actually locally finite. That is much stronger than (c). That result motivates the present work.

Definition 3.5. $g: \partial B_r \to Q_K(R^h)$ is called an elementary function if and only if there is a $g_0: \partial B_r \to R^h$ continuous, such that

$$g(r, \theta) = \sum_{i=1}^{K} \left[\!\left[g_0\left(r, \frac{\theta + 2\pi i}{K}\right) \right]\!\right] .$$

Remark. Following the discussion in [F] on the structure of one dimensional integral currents, we see immediately that if $g: \partial B_r \to Q_K(R^h)$ is Lipschitz, then g can be written as the sum of elementary Lipschitz functions with positive integer multiplicities. Also following [BW], we can write any Lipschitz K-valued elementary function g over B_1 as

$$g(\theta) = \sum_{i=1}^{K} \left[\!\left[\sum_{j=0}^{\infty} \alpha_j \cos j \left(\frac{\theta + 2\pi i}{K}\right) + \beta_j \sin j \left(\frac{\theta + 2\pi i}{K}\right) \right]\!\right] .$$

Definition 3.6. $f: B_r^2 \to Q_K(R^h)$ is called simple if and only if f is continuous, $f(0) = K[\![0]\!]$, $\operatorname{card}(\operatorname{spt} f(x)) = K$ for $x \ne 0$, $K \in \mathbf{Z}^+$ and $\operatorname{graph} f \sim \{0\}$ is connected.

It is obvious that the graph of any simple function is a topological disc. For any open, connected, and simply connected $U \subset B_r^2$, $0 \notin U$, there are continuous functions $f_i: U \to R^h$, $i = 1, \dots, K$ such that $f|U = \sum_{i=1}^K [\![f_i]\!]$.

Corollary 3.1. *If $f: B_1^2 \to Q_K(R^h)$ is homogeneous of degree N, namely $f(x) = |x|^N f(x/|x|)$, and f minimizes the Dirichlet integral, then f can be decomposed as the sum of simple functions with multiplicities, the graphs of any two such functions in this decomposition intersect only at $\{0\}$. The graph of f is a union of discs which may be singular at $\{0\}$.*

Proof. We examine the link of the graph of f. It must consist of curves with constant positive integer multiplicities, and any two of them do not intersect, otherwise we have one dimensional singularity on the graph to contradict Theorem 3.1(c).

Remark. B. White characterized functions as in Corollary 3.1, he proved they are essentially holomorphic, see [C] for more details.

Multiple-valued conformal representation of branched surfaces.

Definition 3.7. A branched surface R in R^{2+h} is said to be admissible if it is the graph of a simple function $k: B_r^2 \to Q_K(R^h)$, and for any open, connected, and simply connected set $0 \notin U \subset B_r^2$, those functions $\{k_i\}$ associated with k and U as in the remark following Definition 3.6 satisfy the following: There are constants $0 < \varepsilon, C < \infty$ independent of U such that

$$|k_i(x)| \le C|x|^{1+\varepsilon},$$
$$|D^j k_i(x)| \le C|x|^{-j+1+\varepsilon},$$
$$|D^3 k_i(x) - D^3 k_i(y)| \le C(\min(|x|, |y|))^{-2+\varepsilon}|x-y|^{\varepsilon}$$

for each $x, y \in U$ and $j = 1, 2, 3$.

Remark. We allow $K = 1$ in the above definition.

Let $\overline{p}$ be the image of orthogonal projection of p into R^2. For an admissible surface R, from Definition 3.7, we know that

$$RA_r = \min\{\operatorname{dist}(p, q) | p, q \in R, p \neq q, \overline{p} = \overline{q} \in B_{2r} \sim B_{r/2}\} > 0.$$

Definition 3.8. A neighborhood V in R^{2+m} of an admissible surface R is called admissible if for some $r > 0$, there are constant, $C > 0$, $a > 0$ with

$$V = \{p \in B_r \times R^{m+n} | \operatorname{dist}(p, R) \le C|\overline{p}|^a RA_{|\overline{p}|}\}.$$

It is well known that for a surface which is the graph of a function $v: B_1^2 \to R^h$ with $|D^i v(x)| \le 1$, $i = 1, 2$, there is an $\alpha > 0$, such that any $y \in B_1^2 \times R^h$ with $\operatorname{dist}(y, \operatorname{graph} v|B_{1/2}^2) \le \alpha$ has a unique point $NP(y) \in \operatorname{graph}(v|B_{1/2}^2)$ with $\operatorname{dist}(y, NP(y)) = \operatorname{dist}(y, \operatorname{graph}(v|B_{1/2}^2))$. The set of those points is the normal neighborhood of height α.

An admissible surface R defined over B_1^2 is decomposed into a union of K pieces of surfaces over any $B^2(x, |x|/2)$, $x \neq 0$ according to the definition. Applying the above discussion to those K pieces of surfaces after normalizing them by a factor of $1/|x|$, we know that each piece in $R \cap B^2(x, |x|/4)$ admits a normal neighborhood of height $|x|$ when $|x|$ is small enough.

Lemma 3.1. *Any admissible neighborhood V of an admissible surface R has an open neighborhood near 0 which admits a nearest point retraction map NP onto $R \cap B_r^2 \times R^h \sim \{0\}$ when r is small enough.*

Proof. It follows from the above discussion and the definition of admissible neighborhood.

If an admissible surface R is a one-sheet covering of B_r^2, it is well known that it has a conformal coordinates system near 0. We want an analogy of this in case R is a multiple-sheet covering of B_r^2, i.e. $K > 1$.

Let k, B_r^2, and K be the same as in Definition 3.7. Let $s > 0$ with $s^K = r$ and

$$u: B_s^2 \to B_r^2,$$
$$x = u(y) = y^K.$$

It is obvious that there is a function $\tilde{k}: B_s \to R^h$ which is unique up to a rotation by a multiple of $2\pi/K$ such that

$$k(x) = \sum_{y \in u^{-1}(x)} [\![\tilde{k}(y)]\!] \in Q_K(R^h).$$

It is apparent that the surface R is now the image of the map $(u, \tilde{k}): B_s \to R^{2+h}$.

Using the derivatives bound on the function k as in Definition 3.7, we have that for some $C' > 0$ the following hold:

$$|\tilde{k}(y)| \leq C'|y|^{K(1+\varepsilon)}, \tag{4}$$

$$|D^i \tilde{k}(y)| \leq C'|y|^{K-i+\varepsilon K}, \qquad i = 1, 2, 3, \tag{5}$$
$$|D^3 \tilde{k}(x) - D^3 \tilde{k}(y)| \leq C'|x - y|^{\varepsilon}.$$

There is a metric defined on B_s induced by the map $(u, \tilde{k})$. This metric is singular at 0 and is C^ε on B_s, locally $C^{2,\varepsilon}$ on $B_s \sim \{0\}$. It is easy to verify that

$$g_{11}(y) = K^2|y|^{2K-2} + O(|y|^{2K-2+2\varepsilon}), \tag{6}$$
$$g_{22}(y) = K^2|y|^{2K-2} + O(|y|^{2K-2+2\varepsilon}), \tag{7}$$
$$g_{12}(y) = O(|y|^{2K-2+2\varepsilon}). \tag{8}$$

Hence the function

$$w = \frac{g_{11} - g_{22} + 2ig_{12}}{g_{11} + g_{22} + 2(g_{11}g_{22} - g_{12}^2)^{1/2}}$$

is well defined over B_s and is a C^ε function, locally $C^{2,\varepsilon}$ over $B_s \sim \{0\}$. According to Proposition 26 in [S] and the discussion there, there is a $C^{1,\varepsilon}$ map

$$v: B_t \to B_s$$

with $v(0) = 0$, $Dv(0) = 1_{R^2}$ such that the induced metric over B_t is conformal; in other words, the map $(u \circ v, \tilde{k} \circ v)$ is a conformal map. By applying the proof of Proposition 26 in [S] to $B(p, |p|/4) \subset B_t$, $p \neq 0$, we conclude that for some $0 < t' < t$, the map v satisfies further estimates:

(9) $$|D^i v(x)| \leq C|x|^{K-i+\varepsilon K}, \qquad i = 1, 2, 3, \ x \in B_{t'},$$

(10) $$|D^3 v(x) - D^3 v(y)| \leq C \min(|x|, |y|)^\varepsilon |x - y|^\varepsilon, \qquad x, y \in B_{t'}.$$

Remark. The above estimates are obtained first for the inverse map of v, then using the chain rule to get the similar estimates for v itself.

Let $\rho = t'^K$ and

$$w: B_{t'} \to B_\rho$$

with $q = w(p) = p^K$. Let

$$F_0: B_{t'} \to R^{2+h},$$
$$F: B_\rho \to Q_K(R^{2+h})$$

be defined by

(11) $$F_0(p) = (u \circ v(p), \tilde{k} \circ v(p))$$

$$F(q) = \sum_{p^K = q} [\![F_0(p)]\!]$$

(12) $$= \sum_{p^K = q} [\![u \circ v(p), \tilde{k} \circ v(p)]\!].$$

It is obvious that F has the following properties:

Lemma 3.2. (a) *For any open, connected, and simply connected* $U \subset B_\rho$, $0 \notin U$, *there are* $F_i: U \to R^{2+h}$, $i = 1, \dots, K$ *such that*

$$F|U = \sum_{i=1}^K [\![F_i]\!]$$

and F_i, $i = 1, \dots, K$, *satisfy*

(13)
$$F(0) = [\![0]\!],$$
$$|DF_i(q)| = 1 + O(|q|^\varepsilon),$$
$$|D^j F_i(q)| \leq C|q|^{-j+1+\varepsilon}, \qquad q \in B_\rho, \ j = 2, 3,$$
$$|D^3 F_j(p) - D^3 F_j(q)| \leq C(\min(|p|, |q|))^{-2+\varepsilon}|p - q|^\varepsilon, \qquad i = 1, \dots, K.$$

(b) *Each F_i is conformal. From the definition of the admissible surfaces, the graphs of F_i, $i = 1, \dots, K$, do not intersect with each other.*

(c) *If $e(x)$ is a unit normal vector field near $F_i(x)$, $x \neq 0$, then*

$$|De(x)| \leq C|x|^{-1+\varepsilon}.$$

Also for a differentiable normal vector field $\vec{n}$ near $F_i(x)$ and a unit vector $a \in T_{F_i(x)}N$, we have

$$\begin{aligned}|D_a\vec{n}(x) \cdot F_{i,s}(x)| &= |D_a(\vec{n}(x) \cdot F_{i,s}(x)) - \vec{n}(x) \cdot D_aF_{i,s}(x)| \\ &= |\vec{n}(x) \cdot D_aF_{i,s}(x)| \\ &\leq C|\vec{n}(x)||x|^{-1+\varepsilon}.\end{aligned}$$

Thus we get a multiple-valued conformal representation of the admissible surface R near 0. The only reason we use F which is multiple-valued instead of $(u, \tilde{k})$ which is single valued is that $|DF_i(q)| = 1 + O(|q|^{\varepsilon})$ so the induced metrics by F_i are similar to the Euclidean metric on B_ρ.

Multiple-valued functions over branched surfaces. Throughout this paper, we are primarily interested in multiple-valued functions over admissible surfaces. We often consider a multiple-valued function

$$f_0 \colon W \subset R \to Q_J(R^{2+h})$$

where R is an admissible surface as in Definition 3.7 with $K \geq 1$, W is an open neighborhood of $0 \in R$ and with $W \supset F(B_\rho)$ for some $\rho > 0$.

Let $U \subset B_\rho$ be open, connected, and simply connected, and let F_i, $i = 1, \dots, K$, as in Definition 3.7. If

$$f_0 \colon W \to Q_J(R^{2+h})$$

then we can define the following function :

$$f(x) = \sum_{i=1}^{K} f_0 \circ F_i(x) \in Q_{KJ}(R^{2+h}).$$

For $x \in U$, we can always write f_0 at $F_i(x) \in W$ as

$$f_0 \circ F_i(x) = \sum_{j=1}^{J} [\![f_{0j} \circ F_i(x)]\!], \qquad i = 1, \dots, K.$$

We let $f_{ij}(x) = f_{0j} \circ F_i(x)$, then we define

Definition 3.9.

$$(F + f)(x) = \sum_{i,j}^{K,J} [\![F_i(x) + f_{ij}(x)]\!]. \tag{14}$$

If f_0 is Lipschitz, then we have $f_0 \circ F_i$ almost everywhere affine approximable. Thus $F + f$ is almost everywhere affine approximable. If we let $D(f_0 \circ F_i) = \sum_{j=1}^{J} [\![L_{ij}]\!]$, then

$$Df(x) = \sum_{i=1}^{K} \sum_{j=1}^{J} [\![L_{ij}]\!] . \tag{15}$$

Hereafter we let

$$f_{ij,s} = L_{ij}\left(\tfrac{\partial}{\partial s}\right) , \tag{16}$$

$$f_{ij,\theta} = L_{ij}\left(\tfrac{\partial}{\partial \theta}\right) . \tag{17}$$

Definition 3.10. We call $f: B_\rho \to Q_{KJ}(R^{2+h})$ R-admissible if it is generated from a function $f_0: R \to Q_J(R^{2+h})$ as above.

Remark. Using the notation in (11), we can write any R-admissible function f as

$$f(q) = \sum_{p^K = q} f_0 \circ \tilde{k}(p) .$$

We adopt the following conventions for R-admissible functions: Let $x, y \in B_\rho$, $x, y \neq 0$, let U be an open, connected, and simply connected domain containing x, y, (r, θ) be the polar coordinates system on B_ρ, and let F_i, $i = 1, \dots, K$, be the same as in Definition 3.7. We define

$$|\eta \circ f(x)| = \sum_{i=1}^{K} |\eta \circ (f_0 \circ F_i(x))| ,$$

$$\mathscr{G}(f(x), f(y)) = \sum_{i=1}^{K} \mathscr{G}(f_0 \circ F_i(x), f_0 \circ F_i(y)) ,$$

$$|f|^2 = \sum_{i=1}^{K} |f_0 \circ F_i|^2 ,$$

$$|f_r(x)|^2 = \sum_{i=1}^{K} |(f_0 \circ F_i)_r(x)|^2 ,$$

$$|f_\theta(x)|^2 = \sum_{i=1}^{K} |(f_0 \circ F_i)_\theta(x)|^2 ,$$

$$(f_r, f_\theta) = \sum_{i=1}^{K} (f_0 \circ F_i)_r (f_0 \circ F_i)_\theta .$$

Definition 3.11. The Dirichlet integral for an R-admissible function f over B_s is defined by

$$D(f, s) = \int_{B_s} |Df|^2 \, d\mathscr{L}^2$$

with $|Df| = |f_r|^2 + |f_\theta|^2 / r^2$.

The height integral of f, $H(f, s)$, is defined by

$$H(f, s) = \int_{\partial B_s} |f|^2 \, d\mathscr{H}^1 .$$

An R-admissible function is said to minimize the Dirichlet integral if it is minimizing among the R-admissible functions.

Remarks. (1) The definition of Dirichlet integrals for multiple-valued functions over surfaces given here is slightly different from the one given in [A]. Our definition is more convenient for calculations. The reason that we can use the estimates involving the Dirichlet integrals in [A] will be explained in Appendix C.

(2) For any R-admissible function g generated from $g_0: W \subset R \to Q_J(R^{2+h})$, if g minimizes the Dirichlet integral then the map $F_0 = (u \circ v, \tilde{k} \circ v)$ onto a neighborhood of $0 \in R$ as introduced in (11) gives a $g_0 \circ F_0$ which minimizes the Dirichlet integral in the sense of Theorem 3.1. So all the results there apply.

If V is an admissible neighborhood of R, from the discussion after Definition 3.8, we know near 0, V admits a nearest point retraction map NP onto an open neighborhood of $0 \in R$. We define

$$V_r = NP^{-1}(F(B_r)) \tag{18}$$

which makes sense for each small $r > 0$.

Definition 3.12. A (V, r)-admissible vector field is a $C^{1,\alpha}$ map

$$X: B_1^{2+h} \to R^{2+h}$$

such that $X = 0$ outside V_r.

Associated with each (V, r)-admissible vector field X, there is a $C^{1,\alpha}$ map:

$$L(t): B_1^{2+h} \to R^{2+h},$$

$$L(t)(p) = p + tX(p), \qquad p \in B_1^{2+h}.$$

Obviously, $L(t)|(B_1^{2+h} \sim V_r)$ is the identity map and for small t, $L(t)$ is a diffeomorphism of V_r onto its image.

4. Assumptions and squash deformation

Assumptions. From now on through §12, we assume the following:

T, M, ϕ are as before.

N is an admissible surface as defined in Definition 3.7 with F the multiple conformal representation of N having estimates on the first three derivatives as in Lemma 3.2.

f is an N-admissible function associated with

$$f_0: N \to Q_J(R^{2+m+n})$$

and V is an admissible neighborhood of N which has the nearest point retraction map NP described in Lemma 3.1. V_r is defined by equation (18).

Furthermore we have the following assumptions on f_0, f, N, and T.

$$\text{(19)}\qquad \operatorname{spt}(T \llcorner V_r) \subset V_r \cup \{0\}, \qquad \operatorname{spt}\partial(T \llcorner V_r) \subset NP^{-1}(\partial B_r).$$

Let $0 \notin U \subset B_r$ be open, connected, and simply connected and F_i, $i = 1, \dots, K$, as in Lemma 3.2.

We have $B_r = C_r^1 \cup C_r^2$, $C_r^1 \cap C_r^2 = \varnothing$ such that for any $x \in U$

$$\operatorname{spt} f_0 \circ F_i(x) \subset T^{\perp}_{F_i(x)} N \cap M$$

and for any $x \in U \cap C_r^1$,

$$f_0 \circ F_i(x) = \tau_\sharp(F_i(x))\langle T, NP, F_i(x)\rangle,$$

where $\langle T, NP, , F_i(x)\rangle$ is the slice of T by $NP^{-1}(x)$, cf. [F, slice theory]. We let $T_r^1 = T \llcorner NP^{-1}(F(C_r^1))$.

The bad set C_r^2, the part of the current

$$T_r^2 = T \llcorner (V_r \cap NP^{-1}(F(C_r^2))$$

missed by the multiple-valued function f, the integral of $|\eta \circ f|$ (recalling Definition 3.2) which tells how centered N is near 0 and the Lipschitz constant of $f|B_r$ satisfy the following:

$$\text{(H1)}\qquad \int_{B_r} |\eta \circ f|\, s ds\, d\theta \le C_3\left(\int_{B_r} (|f|^2 + |f|^2 s^{-1+\varepsilon})\, s ds\, d\theta + rD(r)^{1+\gamma}\right),$$

(H2)

$$\int_{B_r} |\eta \circ f|^2 s^{-1+\varepsilon}\, s ds\, d\theta \le C_3 \int_{B_r} (|f|^2 s^{-1+\varepsilon} + |f|^2 s^{-2+2\varepsilon})\, ds\, d\theta + D^{1+\gamma}(r),$$

$$\text{(H3)}\qquad \mathbf{M}(T_r^2),\ \mathbf{M}(C_r^2) \le C_3\left(\int_{B_r} |f|^2 s^{-2+\sigma}\, s ds\, d\theta\right)^{1+\gamma} + D^{1+\gamma}(r),$$

$$\text{(H4)}\qquad \operatorname{Lip}(f|B_r) \le C_3 D^\tau(r) + C_3\left(\int_0^r H(s)\, ds\right)^\tau \le C_3 r^{2\tau}$$

with constants $0 < \varepsilon, \gamma, \sigma, \tau < 1$.

First variations. We are going to use the formula for first variations to derive two very important formulas.

First we do some calculations for first variation.

Let X be a (V, r)-admissible vector field and $L(t)$ a family of local diffeomorphisms for small t associated with X as in Definition 3.12.

Let $T \llcorner V_r$ be denoted by T_r. Using the area formula (see [F]), we get

$$\text{(20)}\qquad \mathbf{M}(L(t)_\sharp(T_r)) = \int \frac{|D_u L(t) \wedge D_v L(t)|}{|u \wedge v|}\, d\|T_r\|.$$

Here u and v span the tangent plane of T_r at p, since T_r admits a tangent plane at almost every point.

For two vectors $a, b \in R^h$, we let

$$|a \wedge b| = \left(|a|^2|b|^2 - (a \cdot b)^2\right)^{1/2} . \tag{21}$$

Let π denote the plane spanned by u, v, then the deformed unit square on this plane under the map $L(t)$ has area

$$\frac{|D_u L(t) \wedge D_v L(t)|}{|u \wedge v|} .$$

We let

$$\operatorname{div}_\pi X = \frac{d}{dt}\bigg|_{t=0} \frac{|D_u L(t) \wedge D_v L(t)|}{|u \wedge v|} . \tag{22}$$

Hereafter whenever we integrate $\operatorname{div} X$, it is assumed that div is taken with respect to the tangent planes of the relevant integral current. The tangent plane exists almost everywhere according to [F]. Hence the integrals involving divergence are well defined.

We let $\mathbf{G}_2 TM$ denote the fiber bundle of planes in TM, then there is a function

$$h: \mathbf{G}_2 TM \to T^\perp M \tag{23}$$

with $h(p, \pi)$ being the mean curvature of M restricted to the plane $\pi \in \mathbf{G}_2 T_p M$. Hereafter whenever we integrate h over an integral current, it is understood that π refers to the tangent plane of the the integral current. As before, the tangent planes are almost everywhere well defined and the integral is meaningful.

According to [A, Chapter 2], the multiple-valued function $F + f$ maps any measurable set $U \subset B_r$ with its natural orientation to an integral current in R^{2+m+n} denoted by $(F + f)_\sharp(U)$, $\operatorname{spt}(F + f)_\sharp(U) = \operatorname{graph}(F + f \mid U)$, the orientation is induced by $D(F + f)$ (see Definition 3.3).

From Definition 3.9 and the discussion after it, we know it makes sense to talk about $F_{i,s} + f_{ij,s}$, etc.

If we let

$$u_{ij} = F_{i,s} + f_{ij,s} , \tag{24}$$

$$v_{ij} = \tfrac{1}{s}(F_{i,\theta} + f_{ij,\theta}) , \tag{25}$$

we have the following version of the area formula which is particularly useful to us.

Area formula. *Let* $k: \operatorname{graph}((F + f)|B_r) \to R$ *be a measurable function, then*

$$\int k \, d\|(F + f)_\sharp B_r\| = \int_{B_r} \sum_{i=1, j=1}^{K, J} |u_{ij} \wedge v_{ij}| k\left(F_i(x) + f_{ij}(x)\right) d\mathscr{L}^2 x .$$

Proof. Let NP be the nearest point retraction map for the surface N. We define

$$l: V \to B_r$$

by $l(q) = x$ if $NP(q) \in \operatorname{spt} F(x)$. This is a well-defined function since

$$\operatorname{spt} F(x) \cap \operatorname{spt} F(y) = \varnothing$$

for $x \neq y$. Then we apply the co-area formula to l to get the desired formula after observing that

$$\|\wedge_2 Dl\|\Big|_{F_i(x)+f_{ij}(x)} = \frac{1}{|u_{ij} \wedge v_{ij}|}.$$

Lemma 4.1. (a)

$$\int \frac{|v|^2 D_u X \cdot v + |u|^2 D_v X \cdot v + (u \cdot v)(D_u X \cdot v + D_v X \cdot u)}{|u \wedge v|^2} \, d\|(F+f)_\sharp(B_r)\|$$
$$= \int X \cdot h \, d\|(F+f)_\sharp(B_r)\| + \int (\operatorname{div} X - X \cdot h) \, d\|(F+f)_\sharp(C_r^2)\|$$
$$- \int (\operatorname{div} X - X \cdot h) \, d\|T_r^2\|.$$

(b) *Let* $u = F_{i,s} + f_{ij,s}$, $v = (F_{i,\theta} + f_{ij,\theta})/s$,

$$\int \operatorname{div} X \, d\|(F+f)_\sharp(B_r)\|$$
$$= \int_{B_r} \sum_{i=1,j=1}^{K,J} \frac{(D_u X \cdot u|v|^2 + D_v X \cdot v|u|^2 - u \cdot v(D_u X \cdot v + D_v X \cdot u))}{|u \wedge v|} \, d\mathscr{L}^2.$$

Proof. By definition of div (in (22)) and direct calculations, we have

$$\operatorname{div}_\pi X = \frac{1}{|u \wedge v|^2}(|v|^2 D_u X \cdot u + |u|^2 D_v X \cdot v - u \cdot v(D_u X \cdot v + D_v x \cdot u))$$

since

$$\frac{d}{dt}\Big|_{t=0} L(t) = X.$$

If we calculate the derivative of (20), then we have

$$\frac{d}{dt}\Big|_{t=1} \mathbf{M}(L(t)_\sharp(T_r)) = \int \operatorname{div} X \, d\|T_r\|. \tag{26}$$

On the other hand we have the well-known formula for first variation, cf. [WA],

$$\frac{d}{dt}\Big|_{t=1} \mathbf{M}(L(t)_\sharp(T_r)) = \int X \cdot h \, d\|T_r\|. \tag{27}$$

Using the definition of C_r^1,

$$T_r^1 = (F+f)_\sharp(C_r^1),$$

thus

$$\begin{aligned}
\int \operatorname{div} X \, d\|(F+f)_\sharp(B_r)\|
&= \int \operatorname{div} X \, d\|T_r\| - \int \operatorname{div} X \, d\|T_r^2\| + \int \operatorname{div} X \, d\|(F+f)_\sharp(C_r^2)\| \\
&= \int X \cdot h \, d\|T_r\| - \int \operatorname{div} X \, d\|T_r^2\| + \int \operatorname{div} X \, d\|(F+f)_\sharp(C_r^2)\| \\
&= \int X \cdot h \, d\|(F+f)_\sharp(B_r)\| + \int X \cdot h \, d\|T_r^2\| - \int X \cdot h \, d\|(F+f)_\sharp(C_r^2)\| \\
&\quad + \int \operatorname{div} X \, d\|(F+f)_\sharp(C_r^2)\| - \int \operatorname{div} X \, d\|T_r^2\| .
\end{aligned}$$

Hence we have (a).

Part (b) is obtained by using the definition of $\operatorname{div} X$ and the area formula.

Squash deformation.

We use the squash deformation to get the following.

Theorem 4.1. *For almost every sufficiently small $r > 0$, we have*

$$\left| D(r) - \int_{\partial B_r} (f\,, f_r) r \, d\theta \right| \le \mathscr{E}_1(r)$$

where

$$\begin{aligned}
\mathscr{E}_1(r) = C_1 &\int_{B_r} (|\eta \circ f| s^{-1+\varepsilon} + |Df|^4 + |f|^2 s^{-2+2\varepsilon}) \, s ds \, d\theta \\
&+ C_1 \int_{\partial B_r} (\,|Df|^3 |f| + |Df||f|^2 r^{-1+\varepsilon}) r \, d\theta \\
&+ C_1 \left(\operatorname{Lip}(f|B_r) \mathbf{M}(T_r^2) + \sup_{x \in \partial B_r} |f(x)| \frac{d}{dr} \mathbf{M}(T_r^2) \right) .
\end{aligned}$$

Proof. We choose

$$\begin{aligned}
&\varphi \colon V \cup \{0\} \to R \,, \\
&\varphi(p) = 0\,, \quad \text{for } p \in V \sim V_r \,, \\
&\varphi | V_{r-\Delta r} = 1 \,, \\
&|D\varphi| \le 3(1/\Delta r) \,,
\end{aligned} \tag{28}$$

and $\varphi(q) = \varphi(NP(q))$ where NP is the same as in Lemma 3.1. Also $\varphi(F(x))$ does not depend on θ. Eventually we let φ go to the characteristic function of $NP^{-1}(N_r) \cap V$. Define

$$X = \varphi(\mathrm{id} - NP)$$

with id the identity map. Following the convention in Lemma 4.1 for vectors u, v and using the Taylor expansion, we have the following three expressions:

$$(29)\quad \frac{|v|^2}{|u\wedge v|} = 1 + \frac{1}{2\rho}\left(|f_{ij,s}|^2 - \left|\frac{f_{ij,\theta}}{s}\right|^2 + 2F_{i,s}\cdot f_{ij,s} - 2\frac{F_{i,\theta}}{s}\cdot\frac{f_{ij,\theta}}{s}\right) + C_{10}(|f|^2 s^{-2+\varepsilon} + |Df|^4),$$

$$(30)\quad \frac{|u|^2}{|u\wedge v|} = 1 + \frac{1}{2\rho}\left(\left|\frac{f_{ij,\theta}}{s}\right|^2 - |f_{ij,s}|^2 + 2\frac{F_{i,\theta}}{s}\cdot\frac{f_{ij,\theta}}{s} - 2F_{i,s}\cdot f_{ij,s}\right) + C_{10}(|f|^2 s^{-2+2\varepsilon} + |Df|^4),$$

$$(31)\quad \frac{u\cdot v}{|u\wedge v|} = \frac{1}{s\rho}(F_{i,s}\cdot f_{ij,\theta} + F_{i,\theta}\cdot f_{ij,s} + f_{ij,\theta}\cdot f_{ij,s}) + C_{10}(|f|^2 + |Df|^4)$$

where

$$\rho = |F_{i,s}|^2 = \left|\frac{F_{i,\theta}}{s}\right|^2$$

and C_{10} is a constant. Using the definitions of X, u, v, it is easy to check

$$D_uX = \varphi_s f_{ij} + \varphi f_{ij,s}, \qquad D_vX = \varphi\frac{f_{ij,\theta}}{s}.$$

So (we omit the lower indices to simplify the calculation)

$$\begin{aligned}
\frac{|v|^2 D_uX\cdot u}{|u\wedge v|} &= \frac{|v|^2}{|u\wedge v|}(\varphi f_s + \varphi_s f)\cdot(F_s + f_s)\\
&= (\varphi|f_s|^2 + \varphi F_s\cdot f_s + \varphi_s f\cdot f_s)\\
&\quad\times\left(1 + \frac{1}{2\rho}\left(\left|\frac{f_\theta}{s}\right|^2 - |f_s|^2 + 2\frac{F_\theta}{s}\cdot\frac{f_\theta}{s} - 2F_s\cdot f_s\right)\right.\\
&\qquad\left. + C_{10}(|f|^2 s{-}2 + 2\varepsilon + |Df|^2)\right)\\
&= \varphi((|f_s|^2 + F_s\cdot f_s) + C_{11}(|f|^2 s^{-2+2\varepsilon} + |Df|^4))\\
&\quad + \varphi_s(f\cdot f_s + C_{11}(|f||Df|^3 + |Df||f|^2 s^{-1+1\varepsilon})),\\
\frac{|u|^2 D_vX\cdot v}{|u\wedge v|} &= \frac{|u|^2}{|u\wedge v|}\varphi\frac{f_\theta}{s}\cdot\left(\frac{f_\theta}{s} + \frac{F_\theta}{s}\right)\\
&= \varphi\left(\left|\frac{f_\theta}{s}\right|^2 + \frac{F_\theta}{s}\cdot\frac{f_\theta}{s}\right)\\
&\quad\times\left(1 + \frac{1}{2\rho}\left(|f_r|^2 - \left|\frac{f_\theta}{s}\right|^2 + 2F_s\cdot f_s - 2\frac{F_\theta}{s}\frac{f_\theta}{s}\right)\right.\\
&\qquad\left. + C_{10}(|f|^2 s^{-2+2\varepsilon} + |Df|^4)\right)
\end{aligned}$$

(continues

$$= \varphi \left(\left| \frac{f_\theta}{s} \right|^2 + \frac{F_\theta}{s} \frac{f_\theta}{s} \right) \pm C_{11}(|f|^2 s^{-2+2\varepsilon} + |Df|^4),$$

$$\frac{u \cdot v(D_u X \cdot v + D_v X \cdot u)}{|u \wedge v|}$$
$$= \varphi \left(2 f_s \cdot \frac{f_\theta}{s} + f_s \cdot \frac{F_\theta}{s} + \frac{f_\theta}{s} F_s \right)$$
$$\times \left(\frac{1}{\rho} \left(F_s \cdot \frac{f_\theta}{s} + \frac{F_\theta}{s} \cdot f_s + f_s \cdot \frac{f_\theta}{s} \right) + C_{10}(|f|^2 s^{-2+2\varepsilon} + |Df|^4) \right)$$
$$+ \varphi_s \frac{f \cdot f_\theta}{s} \left(\frac{1}{\rho} \left(F_s \cdot \frac{f_\theta}{s} + \frac{F_\theta}{s} \cdot f_s + f_s \cdot \frac{f_\theta}{s} \right) + C_{10}(|f|^2 s^{-2+2\varepsilon} + |Df|^4) \right)$$
$$= C_{11} \varphi(|f|^2 s^{-2+2\varepsilon} + |Df|^4)$$
$$+ C_{11} \varphi_s(|f|^2 |Df| s^{-1+\varepsilon} + |f|\,|Df|^3 + |Df|\,|f|^3 s^{-2+2\varepsilon}).$$

We used Lemma 3.2 (c) in the above calculations. Putting back the subindices and summing over i, j, we get

$$\sum_{i=1, j=1}^{K, J} \frac{(D_u X \cdot u|v|^2 + D_v X \cdot v|u|^2 - u \cdot v(D_u X \cdot v + D_v X \cdot u))}{|u \wedge v|}$$
$$= \sum_{i=1, j=1}^{J, K} \varphi \left(|f_{ij,s}|^2 + \left| \frac{f_{ij,\theta}}{s} \right|^2 \right) + 3C_{11} \varphi(|f|^2 s^{-2+2\varepsilon} + |Df|^4)$$
$$+ \varphi \left(F_{i,s} \cdot f_{ij,s} + \frac{F_{i,\theta}}{s} \cdot \frac{f_{ij,\theta}}{s} \right) + \varphi_s(f_{ij} \cdot f_{ij,s})$$
$$+ C_{11} \varphi_s(|f||Df|^3 + |Df||f|^2 s^{-1+\varepsilon}).$$

After noticing that $f_{ij}(x) \in T^{\perp}_{F_i(x)} N$ and applying Lemma 3.2(c), we have

$$\int_{B_r} \sum_{i=1, j=1}^{K, J} \left(F_{i,s} \cdot f_{ij,s} + \frac{F_{i,\theta}}{s} \cdot \frac{f_{ij,\theta}}{s} \right)$$
$$= \int_{B_r} \sum_{i=1}^{K} \left(F_{i,1} \sum_{j=1}^{J} f_{ij,s} + \frac{F_{i,\theta}}{s} \sum_{j=1}^{J} \frac{f_{ij,\theta}}{s} \right) d\mathscr{L}^2$$
$$\leq \int_{B_r} \sum |\eta \circ f| s^{-1+\varepsilon} s ds\, d\theta ,$$

Also we have

$$| \overrightarrow{n} \cdot (p - q)| \leq \Gamma_1 |p - q|^2$$

for unit vector $\overrightarrow{n} \in T^{\perp}_p M$, $p, q \in M$. By the definition of h, we have $h(p) \in T^{\perp}_p M$, so

$$|h(p) \cdot (p - NP(p))| \leq \Gamma_1 |p - NP(p)|^2 .$$

We apply this to

$$\int h \cdot X \, d\|(F+f)_\sharp(C_r^1)\|$$

and obtain

$$\left| \int h \cdot X \, d\|(F+f)_\sharp(C_r^1)\| \right| \le \Gamma_1 \int_{B_r} \sum_{i=1, j=1}^{K, J} \varphi |f_{ij}|^2 |u \wedge v| \, d\mathscr{L}^2 .$$

If we let φ go to the characteristic function (same as letting Δr go to 0), then all the integrals involving φ_s become the boundary integral over ∂B_r.

As for the term involving T_r^2, using the properties of φ and

$$|(\mathrm{id} - NP)(x)| \le r$$

for $q \in \operatorname{spt} T_r$, we have

$$\int \operatorname{div} X \, d\|T_r^2\| \le C \frac{r}{\Delta r} \mathbf{M}(T_r^2 \llcorner V_r \sim V_{r-\Delta r}) + C\mathbf{M}(T_r^2) .$$

Using the slicing theory of 4.3 in [F], we conclude for almost all r,

$$\int \operatorname{div} X \, d\|T_r^2\| \le Cr \frac{d}{dr} \mathbf{M}(T_r^2) + C\mathbf{M}(T_r^2) .$$

We treat the term involving C_r^2 similarly after noticing that C_r^2 is the image of $\operatorname{spt} T_r^2$ under NP.

Combining all those results and applying Lemma 4.1, we get the theorem.

Remark. In case the function f minimizes the Dirichlet integral, then the error term $\mathscr{E}_1(r)$ is zero.

5. Squeeze deformation

A harmonic function h over B_r satisfies the following two equalities:

$$\int_{B_r} |Dh|^2 \, d\mathscr{L}^2 = \int_{\partial B_r} (h, h_r) \, d\mathscr{H}^1 , \tag{32}$$

$$\int_{\partial B_r} |h_r|^2 \, d\mathscr{H}^1 = \int_{\partial B_r} \left| \frac{h_\theta}{r} \right|^2 d\mathscr{H}^1 . \tag{33}$$

In order to calculate the derivatives of the function which defines the order of contact, we need the analogous formulas for multiple-valued functions which approximate the area minimizing currents. In §4, we already obtain the analogous formula for (32). In this section we show how to get the second one by using the so-called squeeze deformation.

Theorem 5.1. *For almost every sufficiently small $r > 0$, we have*

$$\left| \frac{r}{2} \int_{\partial B_r} \left(|f_r|^2 - \left| \frac{f_\theta}{r} \right|^2 \right) d\mathscr{H}^1 \right| \le \mathscr{E}_2(r)$$

with

$$\mathscr{E}_2(r) = C_2 \int_{B_r} (|f|^2 s^{-2+2\varepsilon} + |f||Df|s^{-1+\varepsilon} + |Df|^3)\, d\mathscr{L}^2 \tag{34}$$
$$+ C_2 \int_{B_r} |\eta \circ f| s^{-1+\varepsilon}\, d\mathscr{L}^2$$
$$+ C_2 \int_{\partial B_r} (|\eta \circ f| s^{\varepsilon} + |f|^2 s^{-1+2\varepsilon} + s|Df|^3)\, d\mathscr{H}^1 \tag{35}$$
$$+ C_2 \left(\mathbf{M}(T_r^2) + r\frac{d}{dr}\mathbf{M}(T_r^2) \right) .$$

Proof. Let $\varphi: V_r \to R$ be the same as in the proof of Theorem 4.1. Let U and F_i, $i = 1, \ldots, K$, be the same as in §4, then we can define a vector field

$$Y: N_r \to R^{2+m+n}$$

over N_r by letting $Y(p) = |x| F_{i,s}(x) \in R^{2+m+n}$ with $p = F_i(x)$. Then let

$$X: V_r \to R^{2+m+n}$$

by setting $X(q) = \varphi(q) Y(NP(q)) \in R^{2+m+n}$. Let (s, θ) be the polar coordinates over R^2. As before we keep

$$u = F_i + f_{ij,s}, \qquad v = \frac{F_{i,\theta}}{s} + \frac{f_{ij,\theta}}{s}.$$

It is easy to verify

$$D_u X = \varphi_s s F_{i,s} + \varphi F_{i,s} + \varphi s F_{i,ss}, \qquad D_v X = \varphi \frac{s F_{i,s\theta}}{s} .$$

Using (29), (30), (31), and the expressions for $D_u X$, $D_v X$, , u, v, we calculate the following (as before we omit the index i for F and i, j for f),

$$\begin{aligned}
\mathbf{I} &= \frac{|v|^2 D_u X \cdot u}{|u \wedge v|} \\
&= \frac{|v|^2}{|u \wedge v|} (\varphi_s s F_s + \varphi F_s + \varphi s F_{ss}) \cdot (F_s + f_s) \\
&= \varphi(|F_s|^2 + \frac{s}{2} F_{ss} \cdot F_s + F_s \cdot f_s + s F_s \cdot f_s) \frac{|v|^2}{|u \wedge v|} \\
&\quad + \varphi_s (s|F_s|^2 + s F_s \cdot f_s) \frac{|v|^2}{|u \wedge v|} \\
&= \varphi(|F_s|^2 + s F_{ss} \cdot F_s + F_s \cdot f_s + s F_s \cdot f_s) && (36) \\
&\quad + \varphi \frac{1}{2\rho} (|F_s|^2 + s F_{ss} \cdot F_s) \left(\left| \frac{f_\theta}{s} \right|^2 - |f_s|^2 + 2\frac{F_\theta}{s} \cdot \frac{f_\theta}{s} - 2 F_s \cdot f_s \right) && (37) \\
&\quad + C_{20} (|f||Df| s^{-1+\varepsilon} + |f|^2 s^{-2+2\varepsilon} + |Df|^3) \\
&\quad + \varphi_s (s|F_s|^2 + s F_s \cdot f_s)
\end{aligned}$$

$$+\frac{\varphi_s s|F_s|^2}{2\rho}\left(\left|\frac{f_\theta}{s}\right|^2-|f_s|^2+2\frac{F_\theta}{s}\cdot\frac{f_\theta}{s}-2F_s\cdot f_s\right)$$
$$+C_{20}(|f|^2s^{-2+2\varepsilon}+|Df|^3),$$

$$\begin{aligned}\mathrm{II}&=\frac{|u|^2D_vX\cdot v}{|u\wedge v|}\\&=\varphi F_{s\theta}\cdot\left(\frac{F_\theta}{s}+\frac{f_\theta}{s}\right)\frac{|u|^2}{|u\wedge v|}\\&=\varphi F_{s\theta}\cdot\left(\frac{F_\theta}{s}+\frac{f_\theta}{s}\right)\\&\quad\times\left(1+\frac{1}{2\rho}\left(|f_s|^2-\left|\frac{f_\theta}{s}\right|^2+2F_s\cdot f_s-2\frac{F_\theta}{s}\cdot\frac{f_\theta}{s}\right)\right.\\&\qquad\left.+C_{10}(|f|^2s^{-2+2\varepsilon}+|Df|^4)\right)\end{aligned}$$

$$=\varphi\left(F_{s\theta}\cdot\frac{F_\theta}{s}+F_{s\theta}\cdot\frac{f_\theta}{s}\right)\tag{38}$$

$$+\varphi F_{s\theta}\cdot\frac{F_\theta}{2s\rho}\left(|f_s|^2-\left|\frac{f_\theta}{s}\right|^2+2F_s\cdot f_s-2\frac{F_s}{s}\cdot\frac{f_\theta}{s}\right)\tag{39}$$
$$+C_{20}(|f|^2s^{-2+2\varepsilon}+|Df|^3),$$

$$\begin{aligned}\mathrm{III}&=\frac{u\cdot v(D_uX\cdot v+D_vX\cdot u)}{|u\wedge v|}\\&=\varphi\left(F_s\cdot\frac{F_\theta}{s}+F_{ss}\cdot F_\theta+F_s\cdot\frac{f_\theta}{s}+F_{ss}\cdot f_\theta+F_{s\theta}\cdot F_s+F_{s\theta}\cdot f_s\right)\\&\quad\times\left(\frac{1}{s\rho}(F_s\cdot f_\theta+F_\theta\cdot f_s)+f_s\cdot f_\theta+C_{10}(|f|^2s^{-2+2\varepsilon}+|Df|^4)\right)\\&\quad+\varphi_s s\left(F_s\cdot\frac{F_\theta}{s}+F_s\cdot\frac{f_\theta}{s}\right)\\&\quad\times\left(\frac{1}{s\rho}(F_s\cdot f_\theta+F_\theta\cdot f_s+f_s\cdot f_\theta)+C_{10}(|f|^2s^{-2+2\varepsilon}+|Df|^4)\right).\end{aligned}$$

Before we calculate the divergence by putting I, II, III together, we notice that F_i, $i=1,\dots,K$, are conformal maps (see Lemma 3.2), i.e.

$$|F_{i,s}|^2=\left|\frac{F_{i,\theta}}{s}\right|^2=\rho,\tag{40}$$

$$F_{i,s}\cdot F_{i,\theta}=0.\tag{41}$$

Hence we have the following identities:

$$
\begin{aligned}
F_{i,s\theta}\cdot\frac{F_{i,\theta}}{s} &= \frac{1}{2s}\frac{d|F_{i,\theta}|^2}{ds}\\
&= \frac{1}{2s}\frac{ds^2|F_{i,s}|^2}{ds}\\
&= |F_{i,s}|^2 + sF_{i,ss}\cdot F_{i,s}\,.
\end{aligned}
\tag{42}
$$

Differentiating (41), we get

$$F_{i,ss}\cdot F_{i,\theta} + F_{i,s\theta}\cdot F_{i,s} = \frac{d}{ds}F_{i,s}\cdot F_{i,\theta} = 0.$$

Using (42) we conclude that the expressions (37) in I and (39) in II are the same except with opposite signs. So we have

$$
\begin{aligned}
\mathrm{I}+\mathrm{II} = {} & \varphi(|F_s|^2 + sF_{ss}\cdot F_s + F_s\cdot f_s + sF_s\cdot f_s)\\
& + \varphi\left(F_{s\theta}\cdot\frac{F_\theta}{s} + F_\theta\cdot\frac{f_\theta}{s}\right)\\
& + C_{20}\varphi(|f||Df|s^{-1+\varepsilon} + |f|^2 s^{-2+\varepsilon} + |Df|^3)\\
& + \varphi_s(s|F_s|^2 + sF_s\cdot f_s)\\
& + \frac{\varphi_s s}{2}\left(\left|\frac{f_\theta}{s}\right|^2 - |f_s|^2 + 2\frac{F_\theta}{s}\cdot\frac{f_\theta}{s} - 2F_s\cdot f_s\right)\\
& + C_{20}\varphi_s(|f|^2 s^{-1+2\varepsilon} + s|Df|^3)\\
= {} & \varphi\left(\frac{1}{s}\frac{ds^2|F_s|^2}{ds}\right) + \varphi\left(F_s\cdot f_s + sF_{ss}\cdot f_s + F_{s\theta}\cdot\frac{f_\theta}{s}\right)\\
& + C_{20}\varphi(|f||Df|s^{-1+\varepsilon} + |f|^2 s^{-2+\varepsilon} + |Df|^3)\\
& + \frac{s}{2}\varphi_s\left(\left|\frac{f_\theta}{s}\right|^2 - |f_s|^2\right)\\
& + \varphi_s\left(s|F_s|^2 + s\frac{F_\theta}{s}\cdot\frac{f_\theta}{s}\right)\\
& + C_{20}\varphi_s(|f|^2 s^{-1+2\varepsilon} + s|Df|^3)\,.
\end{aligned}
$$

As for expression III, we apply the condition (41), its consequence (43) and Lemma 3.2(c) to get

$$
\begin{aligned}
\mathrm{III} = {} & \varphi\left(F_s\cdot\frac{f_\theta}{s} + F_{ss}\cdot f_\theta + F_{s\theta}\cdot f_s\right)\\
& \quad\times\frac{1}{\rho}\left(F_s\cdot\frac{f_\theta}{s} + \frac{F_\theta}{s}\cdot f_s + \frac{f_\theta}{s}\cdot f_s + C_{10}(|f|^2 s^{-2+2\varepsilon} + |Df|^4)\right)\\
& + \varphi_s(F_s\cdot F_\theta + F_s\cdot f_\theta)\\
& \quad\times\frac{1}{\rho}\left(F_s\cdot\frac{f_\theta}{s} + \frac{F_\theta}{s}\cdot f_s + \frac{f_\theta}{s}\cdot f_s + C_{10}(|f|^2 s^{-2+2\varepsilon} + |Df|^4)\right)
\end{aligned}
$$

$$= \varphi C_{20}(|Df||f|s^{-1+\varepsilon} + |f|^2 s^{-2+2\varepsilon} + |Df|^4) + C_{20}|\varphi_s|(|f|^2 s^{-1+2\varepsilon} + |Df|^4) .$$

We put back the subindices and integrate I + II − III over B_r to get the following

$$\int_{B_r} \sum_{i=1,j=1}^{K,J} \frac{(D_u X \cdot u|v|^2 + D_v X \cdot v|u|^2 - u \cdot v(D_u X \cdot v + D_v X \cdot u))}{|u \wedge v|} \, d\mathscr{L}^2$$

$$= \int_{B_r} J \sum_{i=1}^{K} \left(\varphi \frac{1}{s} \frac{d}{ds} s^2 |F_{i,s}|^2 + \varphi_s s |F_{i,s}|^2 \right) d\mathscr{L}^2$$

$$+ \int_{B_r} \sum_{i=1,j=1}^{K,J} \varphi \left(F_{i,s} \cdot f_{ij,s} + sF_{i,ss} \cdot f_{ij,s} + F_{i,s\theta} \cdot \frac{f_{ij,\theta}}{s} + sF_{i,ss} \cdot f_{ij,s} \right) d\mathscr{L}^2$$

$$+ C_{20} \int_{B_r} \varphi(|f|^2 s^{-2+2\varepsilon} + |f||Df|s^{-1+\varepsilon} + |Df|^3) \, d\mathscr{L}^2$$

$$+ \frac{1}{2} \int_{B_r} s\varphi_s \sum_{i=1,j=1}^{K,J} \left(\left| \frac{f_{ij,\theta}}{s} \right|^2 - |f_{ij,s}|^2 \right) d\mathscr{L}^2$$

$$+ \int_{B_r} \sum_{i=1,j=1}^{K,J} s\varphi_s \frac{F_{i,\theta}}{s} \cdot \frac{f_{ij,\theta}}{s} \, d\mathscr{L}^2$$

$$+ \int C_{20} \varphi_s (|f|^2 s^{-1+2\varepsilon} + s|Df|^3) \, d\mathscr{L}^2 .$$

Using polar coordinates (s, θ) on B_r, we have

$$\int_{B_r} \left(\varphi \frac{d}{sds} \sum_{i=1}^{K} (s^2 |F_{i,s}|^2) + \varphi_s s |F_{i,s}|^2 \right) s ds \, d\theta$$

$$= \int_0^{2\pi} \int_0^r \frac{d}{ds} \left(\varphi s^2 \sum_{i=1}^{K} |F_{i,s}|^2 \right) ds \, d\theta = 0 ,$$

because $\varphi | \partial B_r = 0$.

If we let φ go to the characteristic function of B_r as in the proof of Theorem 4.1, immediately we see that φ can be ignored in those integrals over B_r and those integrals involving φ_s become the boundary integrals over ∂B_r. We treat the following integrals as we did in the proof of Theorem 4.1 to conclude that

$$\left| \int \operatorname{div} X \, d\|(F+f)_\sharp(C_r^2)\| \right| , \qquad \left| \int \operatorname{div} X \, d\|T_r^2\| \right|$$

are less than

$$Cr \frac{d}{dr} \mathbf{M}(C_r^2) + C\mathbf{M}(C_r^2)$$

and

$$Cr \frac{d}{dr} \mathbf{M}(T_r^2) + \mathbf{M}(T_r^2)$$

with constant C respectively, since

$$|DX| \leq |D\varphi||Y| + |\varphi||DY|.$$

As for the cross terms involving f and F, we have

$$\int_{B_r} \sum_{i=1,j=1}^{K,J} (F_{i,s} \cdot f_{ij,s} + sF_{i,ss} \cdot f_{ij,s} + F_{i,s\theta} \cdot \frac{f_{ij,\theta}}{s} + sF_{i,ss} \cdot f_{ij,s})\, d\mathscr{L}^2$$
$$= \int_{B_r} \sum_{i=1}^{K} \left(F_{i,s} \cdot (\eta \circ f_i)_s + \frac{F_{i,\theta}}{s} \cdot \frac{(\eta \circ f_i)_\theta}{s} + sF_{i,ss} \cdot (\eta \circ f_i)_s \right) d\mathscr{L}^2 .$$

We estimate the first two terms as in the proof of Theorem 4.1 and for the third term we integrate by parts and apply the estimates of the third derivative of F as in the Lemma 3.2 to get

$$\left| \int_{B_r} \sum_{i=1}^{K} sF_{i,ss} \cdot (\eta \circ f_i)_s\, d\mathscr{L}^2 \right|$$
$$= \left| \int_{B_r} \sum_{i=1}^{K} \frac{d}{ds}(sF_{i,ss} \cdot (\eta \circ f_i)_s) - (\eta \circ f_i) \cdot \frac{d}{ds}(sF_{i,ss})\, d\mathscr{L}^2 \right|$$
$$= \left| \int_{\partial B_r} \sum_{i=1}^{K} sF_{i,ss}(\eta \circ f_i)\, d\mathscr{H}^1 \right|$$
$$+ \left| \int \eta \circ f_i \cdot \frac{d}{ds}(sF_{i,ss})\, d\mathscr{L}^2 \right|$$
$$\leq \int_{\partial B_r} r^{\varepsilon} |\eta f|\, d\mathscr{H}^1 + \int_{B_r} s^{-1+\varepsilon} |\eta f|\, d\mathscr{L}^2.$$

To finish the proof we need to calculate the term involving the mean curvature. Let

$$q \in \operatorname{spt}(F+f)_\sharp(C_r^1) = \operatorname{spt} T_r^1 \subset M ,$$
$$p = NP(q) \in N \subset M ,$$

by definition of X, $X(q) = X(p) \in T_pN$. We observe that at the tangent planes π_p, π_q at p, q are spanned respectively by $F_{i,s}$, $\frac{1}{s}F_{i,\theta}$ and $F_{i,s}+f_{ij,s}$, $\frac{1}{s}(F_{i,\theta} + f_{ij,\theta})$.

Assertion.

(44) $$X(q) \cdot h(q, \pi_q) = X(p)Z(p)(q-p) \pm C_{22}(|Df||f| + |f|^2)$$

with $Z: N \to \operatorname{Hom}(R^{2+m+n}, R)$ dependent only on M.

Proof of the Assertion. Let $B_p: T_pM \times T_pM \to T_p^\perp M$ be the second fundamental form of M in R^{2+m+n}, and let $p \in M$, $\pi_p \subset T_pM$ where π_p is a plane with the orthonormal frame $\{u, v\}$, then

$$h(p, \pi_p) = B_p(u,u) + B_p(v,v).$$

Let e^a, $a = 1, \dots, n$, be orthonormal normal vector fields on $M_1 \supset V_r$, let $B_p^a(\pi) = e^a \cdot h(p, \pi_p)$, then by definition of second fundamental form we have

$$B_p^a(\pi) = uDe^a(p)u + vDe^a(p)v.$$

Using Taylor expansion, derivative estimates of ϕ whose graph is M, and elementary linear algebra, we get

$$\begin{aligned}|B_p^a(\pi_p) - B_q^a(\pi_q)| &\le C(|p-q| + \|\pi_p - \pi_q\|) \\ &\le C(|f| + |Df|)\end{aligned} \tag{45}$$

with C a constant just depending on M.

Applying $X(p) = X(q)$ and $X(p) \cdot h(p, \pi_p) = 0$, we have

$$\begin{aligned} X(q) \cdot h(q, \pi_q) &\\ &= X(p) \cdot (h(q, \pi_q) - h(p, \pi_p)) \\ &= X(p) \left(\sum_a B_q^a(\pi_q) e^a(q) - B_p^a(\pi_p) e^a(p) \right) \\ &= X(p) \sum_a \left((B_q^a(\pi_q) - B_p^a(\pi_p)) e^a(q) + B_p^a(\pi_p)(e^a(q) - e^a(p)) \right). \end{aligned}$$

Using Taylor expansion of $e^a(q)$ at p,

$$e^a(q) - e^a(p) = De^a(p) \cdot (q-p) \pm C_{22}|q-p|^2.$$

Using $|q-p| \le |f|$, we have

$$X(q) \cdot h(q, \pi_q) = \left(X(p) \sum_a B_p^a De^a \right)(q-p) + C(|Df||f| + |f|^2).$$

Hence we have the assertion.

Using the assertion and noticing $p = F_i$, $q = F_i + f_{ij}$, if we sum over j, we have the following

$$\int X \cdot h \, d\|(F+f)_\sharp C_r^1\| \le \int_{C_r^1} \sum_{i=1}^K C\left(|(\eta f)_i| + |f||Df| + |f|^2\right) d\mathscr{L}^2. \tag{46}$$

Combining all those and applying Lemma 4.1, we have the result.

6. Comparison surface

In order to get useful information about the area minimizing integral current, it is important to construct 'good' comparison surfaces. Here we present comparison surfaces which are built from multiple-valued harmonic functions.

Let

$$g: B_r \to Q_{KJ}(R^{2+m+n})$$

be an N-admissible function, cf. Definition 3.10. We let

$$F + g: B_r \to Q_{KJ}(R^{2+m+n})$$

be as in Definition 3.9. So $(F+g)_\sharp(B_r)$ makes sense. We have the following results:

Lemma 6.1. *If* $(F+f)|\partial B_r = (F+g)|\partial B_r$, $\operatorname{spt}(F+g)_\sharp B_r \subset M$, *and* $D(g,r)$ *is well defined, then*

(a)
$$\mathbf{M}((F+f)_\sharp(B_r)) \le \mathbf{M}((F+g)_\sharp(B_r)) + 2\mathbf{M}((F+f)_\sharp(C_r^2)),$$

(b)
$$D(r) \le D(g,r) + \mathscr{E}_3(r)$$

with

$$D(r) = D(f,r) = \int_{B_r} |Df|^2 \, d\mathscr{L}^2,$$
$$\mathscr{E}_3(r) = C_3 \int_{B_r} (|\eta \circ f| s^{-1+\varepsilon} + |\eta \circ g| s^{-1+\varepsilon}) \, d\mathscr{L}^2$$
$$+ C_3 \int_{B_r} \left(|f|^2 s^{-2+2\varepsilon} + |Df|^3\right) \, d\mathscr{L}^2$$
$$+ 2\mathbf{M}((F+f)_\sharp(C_r^2)).$$

Proof. Since

$$T_r = T_r^1 + T_r^2 = (F+f)_\sharp(C_r^1) + T_r^2$$
$$= (F+f)_\sharp(B_r) + T_r^2 - (F+f)_\sharp C_r^2$$

using the boundary condition on $F+g$, we have

$$\partial T_r = \partial\left((F+f)_\sharp(B_r) + T_r^2 - (F+f)_\sharp C_r^2\right)$$
$$= \partial\left((F+g)_\sharp(B_r) + T_r^2 - (F+f)_\sharp(C_r^2)\right).$$

It is known that T is area minimizing and the current $(F+g)_\sharp B_r$ also has support on M, hence

$$\mathbf{M}(T_r) \le \mathbf{M}((F+g)_\sharp(B_r)) + \mathbf{M}(T_r^2) + \mathbf{M}((F+f)_\sharp(C_r^2)),$$

or

$$\mathbf{M}(T_r^1) \le \mathbf{M}((F+g)_\sharp(B_r)) + \mathbf{M}((F+f)_\sharp(C_r^2)).$$

Adding $\mathbf{M}((F+f)_\sharp(C_r^2))$ to both sides we get (a).

Proof of (b). As in §§4 and 5, we let $\rho_i = |F_{i,s}|^2 = |F_{i,\theta}/s|^2$ and

$$u_{ij} = F_{i,s} + f_{ij,s}, \qquad v_{ij} = \tfrac{1}{s}(F_{i,\theta} + f_{ij,\theta}).$$

Applying the conformal condition of F and Lemma 3.2, lipschitz condition on f, we have

$$|u_{ij} \wedge v_{ij}| = \rho_i + \frac{1}{2}\left(|f_{ij,s}|^2 + \frac{1}{s^2}|f_{ij,\theta}|^2\right)$$
$$+ F_{i,s} f_{ij,s} + \frac{1}{s^2} F_{i,\theta} f_{ij,\theta} \pm C(|f|^2 s^{-2+\varepsilon} + |Df|^3).$$

We apply the area formula in §4 to get

$$\begin{aligned}\mathbf{M}((F+f)_\sharp(B_r)) &= \int_{B_r} \sum_{i=1,j=1}^{K,J} |u_{ij} \wedge v_{ij}|\, d\mathscr{L}^2 \\ &= \int_{B_r} \left(\sum_{i=1}^{K} \rho_i + \frac{1}{2}|Df|^2 \pm |\eta \circ f| s^{-1+\varepsilon} \pm |Df|^3 \right) d\mathscr{L}^2 .\end{aligned}$$

Also we have

$$\begin{aligned}\mathbf{M}((F+g)_\sharp(B_r)) &\le \int_{B_r} \sum_{i=1,j=1}^{K,J} \left| (F_{i,s} + g_{ij,s}) \wedge \frac{1}{s}(F_{i,\theta} + g_{ij,\theta}) \right| d\mathscr{L}^2 \\ &\le \int_{B_r} \sum_{i=1,j=1}^{K,J} \frac{|F_{i,s} + g_{ij,s}|^2 + |F_{i,\theta} + g_{ij,\theta}|^2}{2}\, d\mathscr{L}^2 \\ &= \int_{B_r} \left(\sum_{i=1}^{K} \rho_i + \frac{1}{2}|Dg|^2 + |\eta \circ g| s^{-1+\varepsilon} \right) d\mathscr{L}^2 .\end{aligned}$$

Thus (b) follows from (a).

The following part is an elaborate construction of a comparison function g which has the same boundary value as f on ∂B_r, the function is built from a multiple-valued harmonic function. The construction is lengthy, since we deal with functions which are multiple-valued and have to guarantee that the image of $F+g$ is on the ambient manifold M in order to apply the area comparison lemma.

Let

$$F_0 \colon B_t \to R^{2+m+n}$$

as in (11), let

$$f_0 \colon N_r \to Q_J(R^{2+m+n})$$

be the function from which f is generated. Then

$$f(x) = \sum_{y^\wedge = x} f_0 \circ F_0(y) .$$

Over ∂B_t, $f_0 \circ F_0$ is the sum of l elementary functions with multiplicities, following the remark after Definition 3.5. In other words, there are

$$\psi_l \colon \partial B_t \to Q_{J_l}(R^{2+m+n}), \qquad l = 1 \dots, L,$$

so that

$$f_0 \circ F_0 = \sum_{l=1}^{L} c_l \psi_l$$

with $J_l, c_l \in \mathbf{Z}^+$, $\sum_{l=1}^{L} c_l J_l = J$ and

$$\psi_l(\omega) = \sum_{b=1}^{J_l} \left[\!\left[\sum_{j=0}^{\infty} \alpha_{l,j} \cos j \left(\frac{\omega + 2\pi b}{J_l} \right) + \beta_{l,j} \sin j \left(\frac{\omega + 2\pi b}{J_l} \right) \right]\!\right] .$$

Here $\alpha_{l,j}$ and $\beta_{l,j}$ are in R^{2+m+n}. We extend ψ_l over B_t to get

$$\psi_l : B_t \to Q_{J_l}(R^{2+m+n})$$

defined by

$$\psi_l(\rho, \omega) = \sum_{b=1}^{J_l} \left[\!\left[\sum_{j=0}^{\infty} \alpha_{l,j} \left(\frac{\rho}{t}\right)^{j/J_l} \cos j \left(\frac{\omega + 2\pi b}{J_l}\right) + \beta_{l,j} \left(\frac{\rho}{t}\right)^{j/J_l} \sin j \left(\frac{\omega + 2\pi b}{J_l}\right) \right]\!\right]$$

$$= \sum_{b=1}^{J_l} [\![p_{lb}(\rho, \omega)]\!] \tag{47}$$

Next we try to push the graph of ψ_l into M. Let $\overline{x} \in T_0 M$ denote the image of orthogonal projection of $x \in R^{2+m+n}$, let $\Phi = (0, \phi)$ where ϕ is associated with M as in §3. Thus we have

$$x = \overline{x} + \Phi(\overline{x}), \qquad x \in M.$$

We define

$$F_0 + g_0 : B_t \to Q_J(R^{2+m+n})$$

$$(F_0 + g_0)(\rho, \omega) = \sum_{l=1}^{L} c_l \sum_{b=1} [\![\overline{F}_0(\rho, \omega) + \overline{p}_{lb}(\rho, \omega) + \Phi(\overline{F}_0(\rho, \omega) + \overline{p}_{lb}(\rho, \omega))]\!]$$

$$= \sum_{l=1}^{L} c_l \sum_{b=1}^{J_l} [\![q_{lb}(\rho, \omega)]\!].$$

g_0 is defined by taking the difference of $F_0 + g_0$ and F_0 which is well defined. Apparently g_0 generates an N-admissible function g and $F + g$ defined by

$$(F + g)(x) = \sum_{y^K = x} (F_0 + g_0)(y)$$

whose image is supported on M and has the same boundary values as $F + f$.

For the convenience of the following calculation, we let $\overline{F}_0 + \overline{g}_0$ bear its obvious meaning.

Lemma 6.2. (a) *Let* p *be the function defined in* (47), *then*

$$D(g, r) \le (1 + Cr) D(\overline{p}, r) + CrH(r).$$

(b) *There is a constant depending only on* K, J *such that*

$$D(g, r) \le C(K, J) \cdot rD'(f, r) + rH(r).$$

(c)

$$\int_{B_r} |\eta \circ g| s^{-1+\varepsilon} \, d\mathscr{L}^2 \le \int_{\partial B_r} |\eta \circ f| r^{\varepsilon} \, d\mathscr{H}^1.$$

Proof. In the proof of this lemma, we write g_0 as g' to avoid too many lower indices.

By definition of g_0, we have

$$\begin{aligned} g'_{lb,s} &= q_{lb,s} - F_{0,s} \\ &= \overline{p}_{lb,s} + \overline{F}_{0,s} + \Phi_i(\overline{p}_{lb} + \overline{F}_0)(\overline{p}^i_{lb,s} + \overline{F}^i_{0,s}) - F_{0,s} . \end{aligned}$$

Since

$$F_0(\rho, \omega) \in M$$

we have

$$\begin{aligned} F_0(\rho, \omega) &= \overline{F}_0(\rho, \omega) + \Phi(\overline{F}_0(\rho, \omega)) \\ F_{0,s} &= \overline{F}_{0,s} + \Phi_i(\overline{F}_0)\overline{F}^i_{0,s} . \end{aligned}$$

If we use the Taylor expansion and this expression we obtain

$$\begin{aligned} g'_{lb,s} = {} & \overline{p}_{lb,s} + \Phi_i(\overline{p}_{lb} + \overline{F}_0)\overline{p}^i_{lb,s} \\ & + \Phi_{ij}(\overline{F}_0)\overline{F}^i_{0,s}\overline{p}^j_{lb} + \Phi_{ijk}(\overline{F}_0 + c\overline{p}_{lb})\overline{F}^i_{0,s}\overline{p}^j_{lb}\overline{p}^k_{lb} . \end{aligned}$$

Also we know that

$$\begin{aligned} |\Phi_l(\overline{p}_{lb} + \overline{F}_0)| &\le C|\overline{p}_{lb} + \overline{F}_0| \le Cs \\ |\Phi_{ijk}| &\le |D^3\Phi| \le C . \end{aligned}$$

Hence

$$\begin{aligned} |g'_{lb,s}|^2 = {} & |\overline{p}_{lb,s}|^2 \\ & + \left|\Phi_i(\overline{p}_{lb} + \overline{F}_0)\overline{p}^i_{lb,s} + \Phi_{ij}(\overline{F}_0)\overline{F}^i_{0,s}\overline{p}^j_{lb} + \Phi_{ijk}(\overline{F}_0 + c\overline{p}_{lb})\overline{F}^i_{0,s}\overline{p}^j_{lb}\overline{p}^k_{lb}\right|^2 \\ \le {} & |\overline{p}_{lb,s}|^2 + C(s|\overline{p}_{lb,s}|^2 + |\overline{p}_{lb}|^2 + |\overline{p}_{lb}|^2|\overline{p}_{lb,s}|^2) \\ \le {} & |\overline{p}_{lb,s}|^2 + C(s|\overline{p}_{lb,s}|^2 + |\overline{p}_{lb}|^2) . \end{aligned}$$

The last line is obtained by using the fact that $|\overline{p}_{lb}| \le s$ and $s \le 1$. Do the same calculation for $\frac{1}{s}g'_{lb,\theta}$, we have the similar estimates:

$$\left|\frac{g'_{lb,\theta}}{s}\right| \le \left|\frac{\overline{p}_{lb,\theta}}{s}\right|^2 + C\left(s\left|\frac{\overline{p}_{lb,\theta}}{s}\right|^2 + |\overline{p}_{lb}|^2\right),$$

hence

$$\int_{B_r} |Dg|^2\, s\,ds\,d\theta \le \int_{B_r} (|D\overline{p}|^2 + Cs|D\overline{p}|^2 + |\overline{p}|^2)\, s\,ds\,d\theta .$$

From the definition of $\overline{p}_{lb}(\rho , \omega)$, we calculate

$$\begin{aligned}\int_{B_r} |\overline{p}|^2 sds\,\theta &= \sum_{l=1}^{L} c_l \sum_{b=1}^{J_l}\sum_{j=0}^{\infty}\left(\frac{r^2}{2j/KJ_l+2}\right)(|\overline{\alpha}_{lj}|^2+|\overline{\beta}_{lj}|^2)\\ &\le \sum_{l=1}^{L} c_l \sum_{b=1}^{J_l}\sum_{j=0}^{\infty}\frac{r^2}{2}(|\overline{\alpha}_{lj}|^2+|\overline{\beta}_{lj}|^2)\\ &\le \frac{r}{2}\int_{\partial B_r}|\overline{g}(r,\theta)|^2 rd\theta\\ &\le \frac{r}{2}H(r)\,.\end{aligned}$$

Using the power series defining $\overline{p}$, we obtain

$$D(\overline{p},r) = \sum_{l=1}^{L} c_l \sum_{b=1}^{J_l}\sum_{j=0}^{\infty}(|\overline{\alpha}_{lj}|^2+|\overline{\beta}_{lj}|^2),$$

$$\int_0^{2\pi}\left|\frac{\overline{p}_\theta}{r}\right|^2 r\,d\theta = \frac{1}{r}\sum_{l=1}^{L} c_l \sum_{j=0}^{\infty}\frac{j^2}{KJ_l}(|\overline{\alpha}_{lj}|^2+|\overline{\beta}_{lj}|^2)\,.$$

Using the above expressions and the fact $g|\partial B_r = f|\partial B_r$ and the definition of $\overline{p}$, we conclude that there is a constant $C(K,J)$ such that

$$\begin{aligned}D(\overline{p},r) &\le C(K,J)r\int_0^{2\pi}\left|\frac{\overline{p}_\theta}{r}\right|^2 r\,d\theta\\ &\le C(K,J)r\int_0^{2\pi}\left|\frac{f_\theta}{r}\right|^2 r\,d\theta\\ &\le C(K,J)rD'(r)\,.\end{aligned}$$

Thus part (b) follows.

Using Poisson's formula for harmonic functions we get the following:

$$\eta\circ g_0(\rho,\omega) = \frac{1}{2\pi}\int_0^{2\pi}\frac{t^2-\rho^2}{|te^{iv}-\rho e^{i\omega}|^2}\eta\circ f_0(t,v)\,dv\,.$$

Using the fact that B_t is a K-sheet covering of B_r and $z = w^K$, we have

$$\begin{aligned}&\int_{B_r}|\eta\circ g(s,\theta)|s^{-1+\varepsilon}sd\,d\theta\\ &= \int_{B_t}|\eta\circ g_0(\rho,\omega)|\rho^{(-1+\varepsilon)K}\rho^{2K-1}K\,\rho d\rho\,d\omega\\ &\le \frac{1}{2\pi}\int_0^{2\pi}\left(\int_{B_t}\frac{t^2-\rho^2}{|te^{iv}-\rho e^{i\omega}|^2}\rho^{K\varepsilon+K-1}K\,d\rho\,d\omega\right)|\eta\circ f_0(t,v)|\,dv\\ &= \int_0^{2\pi}\frac{1}{K(1+\varepsilon)}t^{K\varepsilon+K}|\eta\circ f_0(t,v)|\,dv\,.\end{aligned}$$

The last equality is obtained by using the property of Poisson's integral

$$\frac{1}{2\pi}\int_0^{2\pi}\frac{t^2-\rho^2}{|te^{iv}-\rho e^{i\omega}|^2}\,d\omega = 1\,.$$

Using the fact $z = w^K$, we have

$$\begin{aligned}\int_0^{2\pi}\frac{1}{K(1+\varepsilon)}t^{K\varepsilon+K}|\eta\circ f_0(t\,,v)|\,dv &= \int_0^{2\pi}\frac{1}{1+\varepsilon}r^{\varepsilon}|\eta\circ f(r\,,\theta)|\,rd\theta\\ &= \int_{\partial B_r}|\eta\circ f|r^{\varepsilon}\,d\mathscr{H}^1\,,\end{aligned}$$

thus we have (c).

7. Height control

In this section we present one of the main estimates in this paper. We prove that the height integral introduced in §3,

$$H(r) = \int_{\partial B_r}|f|^2\,d\mathscr{H}^1$$

for the approximation function f can be bounded by its Dirichlet integral. This enables us to calculate the derivatives of the function which defines the order of contact in the next section.

The height control estimates is analogous to Poincare inequality and is proved by using the formula proved in Theorem 4.1 and the comparison surface constructed in §6.

The complication of the proof is due to the following fact: To estimate the height integral, we need to calculate its derivative and then use the formula in Theorem 4.1 to estimate that derivative in terms of the Dirichlet integral. But the formula in Theorem 4.1 is not exact and the error terms involve the height integral itself. So we need to iterate the argument several times to achieve the desired result.

Theorem 7.1. *There are constants* $C\,,r_0>0$, *so that*

$$H(r)\le CrD(r) \tag{48}$$

whenever $0<r<r_0$.

First we prove the following lemmas.

Lemma 7.1. *Let* $\varepsilon\,,\sigma$ *and* τ *be the small positive constants appearing in the hypotheses in* §4. *For* $0<\lambda=\min(\varepsilon\,,\sigma)$ *and small* $r>0$, *the following holds:*

$$\int_0^r\frac{H(s)}{s^{2-\lambda}}\le C_4\left(\frac{H(r)}{r^{1-\lambda}}+r^{\tau}D(r)\right)\,. \tag{49}$$

Proof. We calculate the left side of the inequality directly by using integral by parts,

$$\int_0^r\frac{H(s)}{s^{2-\lambda}} = \int_0^r\frac{H(s)}{s}\frac{1}{\lambda}ds^{\lambda} = \frac{H(s)}{\lambda s^{1-\lambda}} - \int_0^r\frac{s^{\lambda}}{\lambda}d\frac{H(s)}{s}\,.$$

We use the fact that for almost all r

$$\begin{aligned}\frac{d}{dr}\frac{H(s)}{s} &= \int_0^{2\pi}\frac{d}{ds}\left(\sum_{i=1,j=1}^{K,J}|f_{ij}|^2\right)d\theta\\ &= \int_0^{2\pi}2\sum_{i=1,j=1}^{K,J}f_{ij}\cdot f_{ij,s}\,d\theta\\ &= \frac{1}{s}\int_{\partial B_r}2\sum_{i=1,j=1}^{K,J}f_{ij}\cdot f_{ij,s}\,s d\theta\\ &= \frac{2}{s}\int_{\partial B_r}(f,f_r)\,d\mathscr{H}^1\\ &= \frac{1}{s}(D(s)\pm\mathscr{E}_1(s)),\end{aligned}$$

by Theorem 4.1. Thus

$$\begin{aligned}\int_0^s\frac{H(s)}{s^{2-\lambda}}ds &= \frac{H(r)}{\lambda r^{1-\lambda}}-\frac{1}{\lambda}\int_0^r\frac{1}{s^{1-\lambda}}(D(s)\pm\mathscr{E}_1(s))\,ds\\ &= \frac{H(r)}{s^{1-\lambda}}-\frac{1}{\lambda}\int_0^r\frac{D(s)}{s^{1-\lambda}}ds\pm\frac{1}{\lambda}\int_0^r\frac{\mathscr{E}_1(s)}{s^{1-\lambda}}ds\,.\end{aligned}$$

Next we calculate the integral

$$\int_0^r\frac{\mathscr{E}_1(s)}{s^{1-\lambda}}.$$

Let

$$\mathscr{E}_1(s)=\mathscr{E}_{11}(s)+\mathscr{E}_{12}(s)$$

with

$$\mathscr{E}_{11}(s)=C_1\mathbf{M}(T_s^2)+C_1\int_{B_s}(|\eta\circ f(t,\theta)|t^{-1+\varepsilon}+|Df|^3+|f|^2t^{-2+2\varepsilon})t\,dt\,d\theta\,,$$

$$\mathscr{E}_{12}(s)=C_1 s\frac{d}{ds}\mathbf{M}(T_s^2)+C_1\int_{\partial B_s}(|Df|^3|f|+|Df||f|^2s^{-1+\varepsilon})s\,d\theta$$

First we have

$$\int_0^r\frac{\mathscr{E}_{11}(s)}{s^{1-\lambda}}ds\le\mathscr{E}_{11}(r)\int_0^r\frac{1}{s^{1-\lambda}}ds=\frac{s^\lambda}{\lambda}\mathscr{E}_{11}(r)$$

since the function $\mathscr{E}_{11}$ is increasing. We estimate $\mathscr{E}_{12}$ as follows:

$$\begin{aligned}\int_0^r\frac{\mathscr{E}_{12}(s)}{s^{1-\lambda}}ds &\le C_1\int_0^r s^\lambda\frac{d}{ds}\mathbf{M}(T_s^2)\,ds\\ &\quad+C_1\int_{B_r}(|Df|^3|f|s^{-1+\lambda}+|Df||f|^2s^{-2+\varepsilon+\lambda})\,sds\,d\theta\\ &\le C_1r^\lambda\mathbf{M}(T_r^2)+C_1\int_{B_r}(C_3|Df|^3s^{\tau+\lambda}+C_3|f|^2s^{-2+\tau+\lambda})\,sds\,d\theta\\ &\le C_1C_3r^\lambda\mathscr{E}_{11}(r)\,.\end{aligned}$$

The last part is obtained by using the definition of $\mathscr{E}_{11}$. So we have

$$\int_0^r \frac{H(s)}{s^{2-\lambda}} \le \frac{H(r)}{r^{1-\lambda}} + C_1 C_3 r^\lambda \mathscr{E}_{11}(r).$$

Using (H2) and (H3) in the beginning of §4 to replace the terms in $\mathscr{E}_{11}(r)$ involving $|\eta \circ f|$ and $\mathbf{M}(T_r^2)$, we get

$$\begin{aligned}\int_0^r \frac{H(s)}{s^{2-\lambda}} \le\ & \frac{H(r)}{r^{1-\lambda}} + C_1 C_3 r^\lambda D^{1+\gamma}(r) \\ & + C_1 C_3 r^\lambda \int_{B_r} (|f|^2 s^{-2+\sigma} + |f|^2 s^{-1+\varepsilon} + |f|^2 s^{-2+2\varepsilon} + |Df|^3)\, d\mathscr{L}^2 .\end{aligned}$$

Since $\lambda \le \min(2\varepsilon, \sigma)$, we can move the terms involving $|f|$ on the right side of the inequality to the left side when r is small enough, then we have the conclusion.

Lemma 7.2. *There is a constant $C > 0$ such that for $\lambda > 0$ as in Lemma* 7.1 *and each small $r > 0$, the following holds*

$$H(r) \le 4 \int_0^r \frac{D(s)}{s}\, ds + C_4 r^{1+2\lambda} D(r) . \tag{50}$$

Proof. We keep the notation in the proof of Lemma 7.1. Using (H2) and (H2), we obtain

$$\begin{aligned}\int_0^r & \frac{\mathscr{E}_{11}(s)}{s}\, ds \\ & \le C_1 \int_0^r \frac{1}{s} \left(\mathbf{M}(T_s^2) + \int_{B_s} (|\eta \circ f| t^{-1+\varepsilon} + |Df|^3 + |f|^2 t^{-2+2\varepsilon})\, d\mathscr{L}^2 \right) ds \\ & \le C_1 C_3 \int_0^r \frac{1}{s} \left(\int_{B_s} |f|^2 t^{-2+\sigma}\, d\mathscr{L}^2 + D^{1+\gamma}(s) \right) ds \\ & \quad + C_1 C_3 \int_0^r \left(\frac{1}{s} \int_{B_s} (|f|^2 t^{-2+2\varepsilon} + |f|^2 t^{-1+\varepsilon})\, d\mathscr{L}^2 + \frac{1}{s} D^{1+\gamma}(s) \right) ds \\ & \quad + C_1 C_3 \int_0^r \left(\frac{1}{s} \int_{B_s} (|Df|^3 + |f|^2 t^{-2+2\varepsilon})\, d\mathscr{L}^2 \right) ds \\ & \le 3 C_1 C_3 \int_0^r \left(\frac{1}{s} \int_{B_s} |f|^2 (t^{-2+\sigma} + t^{-2+2\varepsilon} + t^{-1+\varepsilon})\, d\mathscr{L}^2 \right) ds \\ & \quad + 3 C_1 C_3 \int_0^r \frac{1}{s} (C_3 s^{2\tau} D(s) + D^{1+\gamma}(s))\, ds \\ & \le C_4 \int_0^r \frac{1}{s} \left(\frac{H(s)}{s^{1-\lambda}} + s^\lambda D(s) \right) ds \\ & \quad + \int_0^r \frac{1}{s} (s^{2\tau} D(s) + D^{1+\gamma}(s))\, ds .\end{aligned}$$

If we apply Lemma 7.1 again, we get

$$\int_0^r \frac{\mathscr{E}_{11}(s)}{s}\, ds \le C_4' \left(\frac{H(r)}{r^{1-\lambda}} + r^\lambda D(r) \right) . \tag{51}$$

Using (H3) again, we estimate

$$\begin{aligned}
\int_0^r \frac{\mathscr{E}_{12}(s)}{s}\,ds \;&\le\; C_1\int_0^r \frac{d}{ds}\mathrm{M}(T_s^2)\,ds\\
&\quad + C_1\int_{B_r}(|Df|^3|f|s^{-1} + |Df||f|^2 s^{-2+\varepsilon})\,d\mathscr{L}^2\\
&\le\; C_1C_3\left(D^{1+\gamma} + \int_{B_r}|f|^2 s^{-2+\sigma}\,d\mathscr{L}^2\right)\\
&\quad + C_1\int_{B_r}(s^{2\tau}|Df|^2 + s^{2\tau}|f|^2 s^{-2+\varepsilon})\,d\mathscr{L}^2\\
&\le\; C_1C_3C_4\left(\frac{H(r)}{s^{1-\lambda}} + (r^{2\tau} + r^{2\gamma})D(r)\right).
\end{aligned}$$

Thus we have

$$\begin{aligned}
H(r) &= r\int_0^r \left(\frac{H(s)}{s}\right)'\,ds\\
&= 2r\int_0^r \frac{1}{s}(D(s) + \mathscr{E}_1(s))\,ds\\
&= 2r\int_0^r \frac{D(s)}{s}\,ds + r\int_0^r \frac{\mathscr{E}_{11}(s) + \mathscr{E}_{12}(s)}{s}\,ds\\
&\le 2r\int_0^r \frac{D(s)}{s}\,ds + C_4'(r^\lambda H(r) + (r^{1+2\tau} + r^{1+2\gamma})D(r))\\
&\le 2r\int_0^r \frac{D(s)}{s}\,ds + C_4'(r^\lambda H(r) + 2r^{1+\lambda}D(r)).
\end{aligned}$$

When $C_4' r^\lambda \le \frac{1}{2}$, we have the desired result.

Proof of Theorem 7.1. Using Lemma 6.1(b), Lemma 6.2(a), we obtain

$$\begin{aligned}
H(r) &\le 4r\int_0^r \frac{D(s)}{s}\,ds + C_5 r^{1+\lambda}D(r)\\
&\le 4r\int_0^r \frac{D(g,s) + \mathscr{E}_3(s)}{s}\,ds + C_5 r^{1+\lambda}D(r)\\
&\le 4r\int_0^r \frac{C(K,J)sD'(s) + sH(s) + \mathscr{E}_3(s)}{s}\,ds + C_5 r^{1+\lambda}D(r)\\
&\le 4C(K,J)rD(r) + 4\int_0^r \frac{\mathscr{E}_3(s) + sH(s)}{s}\,ds + C_5 r^{1+\lambda}D(r).
\end{aligned}$$

It is easy to see

$$\begin{aligned}
\mathscr{E}_3(s) &\le \mathscr{E}_1(s) + \int_{B_r}|\eta\circ g|s^{-1+\varepsilon}\,s\,ds\,d\theta\\
&\le \mathscr{E}_1(s) + C\int_{\partial B_r}|\eta\circ f|s^\varepsilon s\,d\theta,
\end{aligned}$$

using the definitions of $\mathscr{E}_1, \mathscr{E}_3$ and applying Lemma 6.2(c). So

$$r\int_0^r \frac{\mathscr{E}_1(s)+\mathscr{E}_3(s)}{s}\,ds \;\le\; 2r\int_0^r\left(\frac{\mathscr{E}_1(s)}{s}+\int_{\partial B_s}|\eta\circ f|s^{-1+\varepsilon}s\,d\theta\right)ds$$

We already know

$$r\int_o^r \frac{\mathscr{E}_1(s)}{s}\,ds \le C_4(r^\lambda H(r)+r^{1+\lambda}D(r)).$$

Using (H3) and Lemma 7.1, we get

$$\begin{aligned} r\int_{B_r}|\eta\circ f|s^{-1+\varepsilon}\,d\mathscr{L}^2 \;&\le\; C_3 r\left(\int_{B_r}|f|^2 s^{-2+2\varepsilon}s ds\,d\theta + D^{1+\gamma}(r)\right)\\ &\le\; C_3C_4(r^\lambda H(r)+r^{1+\tau}D(r)) + C_3D^{1+\gamma}(r)\,. \end{aligned}$$

So we have

$$\begin{aligned} r\int_0^r \frac{\mathscr{E}_1(s)+\mathscr{E}_3(s)}{s}\,ds \;&\le\; 2r\int_0^r\left(\frac{\mathscr{E}_1(s)}{s}+\int_{\partial B_s}|\eta\circ f|s^{-1+\varepsilon}s\,d\theta\right)ds\\ &\le\; 2rC_4(r^\lambda H(r)+r^{1+\lambda}D(r)) + 2r\int_0^r|\eta\circ f|s^{-1+\varepsilon}\,d\mathscr{L}^2\\ &\le\; C_5(r^\lambda H(r)+r^{1+\lambda}D(r))\,. \end{aligned}$$

By letting $0<r<1$ and using Lemma 7.1, we have

$$r\int_0^r H(s)\,ds \le r\int_0^r \frac{H(s)}{s^{2-\lambda}}\,ds \le C_4(r^\lambda H(r)+r^\tau D(r)).$$

So

$$H(r)\le C_5\left(rD(r)+r^\lambda H(r)+3r^{1+u}D(r)\right).$$

If $C_5r\lambda\le\frac{1}{2}$, then

$$H(r)\le C_6 rD(r).$$

8. Order of contact

In this section we define the order of contact, or rather the inverse of it, between the branched center manifold and the approximation function. The bulk of the text is devoted to prove it is a well-defined quantity. In the next section we shall prove the order of contact is not ∞, unless the current coincides with the center manifold N. To start with, we define

$$K(r)=\frac{H(r)}{rD(r)}.$$

The major result in this section is

Theorem 8.1. $\lim_{r\to 0} K(r)$ *exists.*

Remark. In case f is a harmonic function, $\lim_{r\to 0} K(r)$ is the inverse of the order of the first nonzero term appearing in the power series.

It is proved by establishing a differential inequality to prove a certain monotonicity property of $K(r)$.

We need the following lemmas in the proof of the theorem.

Lemma 8.1.

$$D(s)\int_{\partial B_s}(f\ ,f_s)\,d\mathscr{H}^1 \le H(s)\int_{\partial B_s}|f_s|^2\,d\mathscr{H}^1 + D(s)\mathscr{E}_1(s)\,.$$

Proof.

$$\begin{aligned}
&D(s)\int_{\partial B_s}(f\ ,f_s)\,d\mathscr{H}^1\\
&\le D(s)\left(\int \partial B_s|f|^2\,d\mathscr{H}^1\right)^{1/2}(|f_s|^2\,d\mathscr{H}^1)^{1/2}\\
&=\frac{1}{2}H(s)\int_{\partial B_s}|f_s|^2\,d\mathscr{H}^1+\frac{1}{2}D^2(s)\\
&\le\frac{1}{2}H(s)\int_{\partial B_s}|f_s|^2\,s d\theta+\frac{1}{2}D(s)\left(\int_{\partial B_s}(f\ ,f_s)\,d\mathscr{H}^1+\mathscr{E}_1(s)\right),
\end{aligned}$$

by using Theorem 4.1. So

$$D(s)\int_{\partial B_s}(f\ ,f_s)\,d\mathscr{H}^1 \le H(s)\int_{\partial B_s}|f_s|^2\,d\mathscr{H}^1 + D(s)\mathscr{E}_1(s)\,.$$

Applying the height control estimates, we can simplify the four hypotheses in §4 as follows. Using

$$H(s)\le \Gamma s D(s)$$

we estimate

$$\begin{aligned}
\int_{B_r}|f|^2 s^{-2+\lambda} &= \int_0^r\int_0^{2\pi}|f|^2 s^{-2+\lambda} s\,ds\,d\theta\\
&= \int_0^r H(s)s^{-2+\lambda}\,ds\\
&\le \int_0^r \Gamma s D(s) s^{-2+\lambda}\,ds\\
&\le \int_0^r \Gamma D(s) s^{-1+\lambda}\,ds\\
&\le \Gamma D(r)\int_0^r s^{-1+\lambda}\,ds\\
&\le \Gamma D(r)\frac{r^\lambda}{\lambda}
\end{aligned}$$

for any $\lambda > 0$. As a consequence of this, we have the following simplified version of (H2), (H3), and (H4):

$$\int_{B_r} |\eta \circ f| s^{-1+\varepsilon} s ds\, d\theta \leq C_3' \frac{r^{2\varepsilon}}{2\varepsilon} D(r) + C_3' D^{1+\gamma}(r), \tag{52}$$

$$\mathbf{M}(T_r^2), \mathbf{M}(C_r^2) \leq C_3' D^{1+\gamma}(r), \tag{53}$$

$$\operatorname{Lip}(f|_{B_r}) \leq C_3' D^{\tau}(r). \tag{54}$$

Lemma 8.2. *Let* $0 < \lambda < \lambda_0 = \frac{1}{2}\min(\varepsilon, \gamma, \sigma, \tau)$, *where* $\varepsilon, \gamma, \sigma, \tau$ *appear in Hypotheses* (H1), (H2), (H3) *and* (H4) *in* §4. *Then*

(a)

$$\mathscr{A}_\lambda(r) = \int_0^r \frac{\mathscr{E}_1(s)}{s^{1+\lambda} D(s)}\, ds \leq C r^{\lambda_0 - \lambda}, \tag{55}$$

(b)

$$\mathscr{B}_\lambda(r) = \int_0^r \frac{\mathscr{E}_2}{s^{1+\lambda} D(s)}\, ds \leq C r^{\lambda_0 - \lambda}. \tag{56}$$

We denote $\mathscr{A}_0(s), \mathscr{B}_0(s)$ *by* $\mathscr{A}(s)$, $\mathscr{B}(s)$ *respectively.*

Remark. Part (a) is similar to the estimates in §7, but we are unable to prove this until we have the height control estimates.

Proof. Let

$$\mathscr{E}_1(s) = \mathscr{E}_{11}(s) + \mathscr{E}_{12}(s)$$

with

$$\begin{aligned}\mathscr{E}_{11}(s) &= C_1 \mathbf{M}(T_s^2) + C_1 \int_{\partial B_s} |Df||f|^2 s^{-1+\varepsilon} s\, d\theta \\ &\quad + C_1 \int_{B_s} (|\eta \circ f(s, \theta)| s^{-1+\varepsilon} + |Df|^3 + |f|^2 s^{-2+2\varepsilon}) s ds\, d\theta \\ \mathscr{E}_{12}(s) &= C_1 s \frac{d}{ds} \mathbf{M}(T_s^2) + C_1 \int_{\partial B_s} |Df|^3 |f| s\, d\theta.\end{aligned}$$

As a consequence of (52), (53), and (54), we have the following

$$\mathscr{E}_{11}(s) \leq C_1' \left(D^{1+\gamma}(s) + s^{2\varepsilon} D(s) \right). \tag{57}$$

Thus

$$\begin{aligned}\int_0^r \frac{\mathscr{E}_{11}(s)}{s^{1+\lambda} D(s)}\, ds &\leq \int_0^r C_1' \left(\frac{D^{\gamma}(s)}{s^{1+\lambda}} + s^{\tau - 1 - \lambda} \right) ds \\ &\leq C_1'' \int_0^r \left(s^{2\gamma - 1 - \lambda} + s^{\tau - 1 - \lambda} \right) ds \\ &\leq C 2 r^{(\lambda_0 - \lambda)}\end{aligned} \tag{58}$$

by the assumption on λ_0.

As to the term involving $\mathscr{E}_{12}$, first we estimate $|f(t,\theta)| \le \operatorname{Lip}(f|B_t)t \le t$ and $|Df(t,\theta)| \le D^\tau(s)$, $0 < t < s$. Thus

$$\int_0^s \frac{\mathscr{E}_{12}(t)}{t}\,dt \le C_1'\left(D^{1+\gamma}(s) + D^{1+\tau}(s)\right). \tag{59}$$

Using (59) and integration by parts, we have

$$\begin{aligned}
&\int_0^r \frac{\mathscr{E}_{12}(s)}{s^{1+\lambda}D(s)}\,ds \\
&= (r^\lambda D(r))^{-1}\int_0^r \frac{\mathscr{E}_{12}(s)}{s}\,ds \\
&\quad + \int_0^r \left(\frac{1}{sD(s)}\right)'\int_0^s t^{-1}\mathscr{E}_{12}(t)\,dt\,ds \\
&\le (r^\lambda D(r))^{-1}C_1'\left(D^{1+\gamma}(r) + D^{1+\tau}(r)\right) \\
&\quad + \int_0^r \left(\frac{\lambda}{s^{1+\lambda}D(s)} + \frac{D'(s)}{s^\lambda D^2(s)}\right)(D^{1+\gamma}(s) + D^{1+\tau}(s))\,ds \\
&= \int_0^r \left(\lambda\frac{D^\gamma(s) + D^\tau(s)}{s^\lambda} + \frac{D^{\gamma/2}(s)}{s^\lambda}\frac{D'(s)}{D^{1-\gamma/2}(s)} + \frac{D^{\tau/2}(s)}{s^\lambda}\frac{D'(s)}{D^{1-\tau/2}}\right)ds \\
&\le C_2 D^{\lambda_0/2}(r)
\end{aligned}$$

by using the fact that $D^c(s) \le Cs^{2c}$ which is a consequence of (54) and the assumption on λ_0.

As to part (b), we let

$$\mathscr{E}_2(r) = C\mathscr{E}_{21}(r) + \mathscr{E}_{22}(r)$$

with

$$\begin{aligned}
\mathscr{E}_{21} &= \int_{B_r}(|f|^2 s^{-2+2\varepsilon} + |f||Df|s^{-1+\varepsilon} + |Df|^3)\,d\mathscr{L}^2 \\
&\quad + C_2\int_{B_r}|\eta\circ f|s^{-1+\varepsilon}\,d\mathscr{L}^2 + C_2\int_{\partial B_r}|f|^2 r^{-1+2\varepsilon}\,d\mathscr{H}^1 \\
&\quad + C_2\mathbf{M}(T_r^2), \\
\mathscr{E}_{22} &= C_2\int_{\partial B_r}(|\eta\circ f|s^\varepsilon + s|Df|^3)\,d\mathscr{H}^1 + C_2 r\frac{d}{dr}\mathbf{M}(T_r^2).
\end{aligned}$$

We notice most terms in $\mathscr{E}_2$ are the same as in $\mathscr{E}_1$ and we indicate how to get the others. The additional terms in $\mathscr{E}_{21}$ but not in $\mathscr{E}_{11}$ are

$$A(r) = \int_{B_r}|f||Df|s^{-1+\varepsilon}\,d\mathscr{L}^2 + \int_{\partial B_r}|f|^2 s^{-1+2\varepsilon}\,\mathscr{H}^1$$

and the following estimates hold for $A(s)$ by using Cauchy-Schwarz inequality and the height control estimates Theorem 7.1,

$$\begin{aligned} A(s) &\le \left(\int_{B_s} |f|^2 t^{-2+2\varepsilon}\, d\mathscr{L}^2\right)^{1/2} \left(\int_{B_s} |Df|^2\, d\mathscr{L}^2\right)^{1/2} + \int_{\partial B_s} |f|^2 s^{-1+\varepsilon}\, d\mathscr{H}^1 \\ &= \left(\int_0^s H(t) t^{-2+2\varepsilon}\, dt\right)^{1/2} D^{1/2}(s) + H(s) s^{-1+\varepsilon} \\ &\le C\left(\left(\int_0^s D(t) t^{-1+2\varepsilon}\, dt\right)^{1/2} D^{1/2}(s) + D(s) s^{\varepsilon}\right) \qquad (60) \\ &\le C' s^{\varepsilon} D(s)\,. \end{aligned}$$

Using (60) then

$$\begin{aligned} \int_0^r \frac{A(s)}{s^{1+\lambda} D(s)}\, ds &\le \int_0^r s^{\varepsilon-\lambda-1}\, ds \\ &\le C r^{\lambda_0-\lambda}\,. \qquad (61) \end{aligned}$$

The additional term appearing in $\mathscr{E}_{22}$ not in $\mathscr{E}_{12}$ is the following

$$B(s) = \int_{\partial B_s} |\eta \circ f| s^{\varepsilon}\, d\mathscr{H}^1\,. \qquad (62)$$

We integrate by parts to get

$$\begin{aligned} \int_0^r \frac{B(s)}{s^{1+\lambda} D(s)} &\le (r^{\lambda} D(r))^{-1} \int_0^r s^{-1} B(s)\, ds \\ &\quad - \int_0^r \left(\frac{1}{s^{\lambda} D(s)}\right)' \int_0^s t^{-1} B(t)\, dt\, ds \qquad (63) \\ &= J_1 + J_2\,. \qquad (64) \end{aligned}$$

We use (62), (H2) in §4 and height control estimates to get

$$\begin{aligned} \int_0^s t^{-1} B(t)\, dt &\le \int_{B_s} |\eta \circ f| t^{-1+\varepsilon}\, d\mathscr{L}^2 \\ &\le \int_{B_s} |f|^2 t^{-2+2\varepsilon}\, d\mathscr{L}^2 + D^{1+\gamma}(s)\,. \qquad (65) \end{aligned}$$

It will not work if one tries to put this estimate directly back into (63), instead we do the following:

Using the height control estimates, we see

$$|J_1| \le C r^{\lambda_0-\lambda}\,.$$

As for J_2,

$$\begin{aligned} |J_2| &\le \int_0^r \left(\frac{\lambda}{s^{1+\lambda} D(s)} + \frac{D'(s)}{s^{\lambda} D^2(s)}\right) \left(D^{1+\gamma}(s) + \int_{B_s} |f|^2 t^{-2+2\varepsilon}\, d\mathscr{L}^2\right) ds \\ &\le C D^{\gamma/2}(r) + J_3\,. \end{aligned}$$

Use the fact that $((s^\lambda D(s))^{-1})'$ is negative, since the function is decreasing to get

$$\begin{aligned} J_3 &= -\int_0^r \left(\frac{1}{s^\lambda D(s)}\right)' \int_{B_s} |f|^2 t^{-2+2\varepsilon}\, d\mathscr{L}^2\, ds \\ &= -\frac{1}{s^\lambda D(r)} \int_{B_r} |f|^2 t^{-2+2\varepsilon}\, d\mathscr{L}^2 \\ &\quad + \int_0^r \left(\frac{1}{s^\lambda D(s)} \int_{\partial B_s} |f|^2 s^{-2+2\varepsilon}\, d\mathscr{L}^2\right) ds \\ &\le \int_0^r \left(\frac{1}{s^\lambda D(s)} \int_{\partial B_s} |f|^2 s^{-2+2\varepsilon}\, d\mathscr{L}^2\right) ds \\ &\le \int_0^r \frac{1}{s^\lambda D(s)} C D(s) s^{-1+2\varepsilon}\, ds \\ &\le Cr^\varepsilon . \end{aligned}$$

Combining the estimates for J_1, J_2, J_3 we get the estimates for (64), thus we have the estimates for $\mathscr{E}_{22}$ to conclude (b).

Proof of Theorem 8.1. We prove the theorem by calculating the derivative of $K(r)$ to establish a differential inequality. For almost all r,

$$\begin{aligned} K'(s) &= \frac{1}{D^2(s)}\left(\left(\frac{H(s)}{s}\right)' D(s) - \frac{H(s)}{s} D'(s)\right) \\ &= D^{-2}\left(2\int_0^{2\pi} (f\,,f_s)\, d\theta D(s) - \frac{H(s)}{s} D'(s)\right) \\ &= \frac{1}{sD^2(s)}\left(2\int (f\,,f_s)\, s d\theta D(s)\right. \\ &\quad \left. -2H(s)\int |f_s|^2 s d\theta + H(s)\int_{\partial B_s}\left(|f_s|^2 - \left|\frac{f_\theta}{s}\right|^2\right)\right) \\ &\le \frac{1}{sD(s)^2}\left(2\mathscr{E}_1(s)D(s) + 2H(s)\frac{\mathscr{E}_2(s)}{s}\right) . \end{aligned}$$

The last inequality is obtained by using Lemma 8.1 and the squeeze deformation Theorem 5.1. So

$$K'(s) \le \frac{2\mathscr{E}_1(s)}{sD(s)} + 2K(s)\cdot \frac{\mathscr{E}_2(s)}{sD(s)}$$

or we can write this as

$$\left(K(s)e^{-2\int_0^s \mathscr{E}_2(t)/tD(t)\,dt}\right)' \le e^{-2\int_0^s \mathscr{E}_2(t)/tD(t)\,dt}\mathscr{E}_1(s)/sD(s) \tag{66}$$

which is the same as

$$(67) \qquad (K(s)e^{-2\mathscr{B}(s)})' \leq 2e^{-2\mathscr{B}(s)} \frac{\mathscr{E}_1(s)}{sD(s)} \leq 4\frac{\mathscr{E}_1(s)}{sD(s)}$$

by using the definitions of $\mathscr{A}$, $\mathscr{B}$ and their properties in Lemma 8.2. If we integrate this inequality, then for $0 < s < r$,

$$K(r)e^{-2\mathscr{B}(r)} - K(s)e^{-2\mathscr{B}(s)} \leq 4\int_s^r \frac{\mathscr{E}_1(s)}{sD(s)}\, ds = 4\mathscr{A}(r) - 4\mathscr{A}(s)$$

which is the same as

$$(68) \qquad K(r)e^{-2\mathscr{B}(r)} - 4\mathscr{A}(r) \leq K(s)e^{-2\mathscr{B}(s)} - 4\mathscr{A}(s)\,.$$

Hence the limit $\lim_{s\to 0}(K(s)e^{-2\mathscr{B}(s)} - 4\mathscr{A}(s))$ exists. As a consequence of Lemma 8.2, we know that

$$\lim_{s\to 0}\mathscr{A}(s) = 0\,, \qquad \lim_{s\to 0}\mathscr{B}(s) = 0\,,$$

so the limit

$$\lim_{s\to 0} K(s)$$

exists. From height control estimates Theorem 7.1, we conclude that

$$\lim_{s\to 0} K(s) \neq \infty\,.$$

In the next section we show that if

$$\lim_{s\to 0} K(s) = 0\,,$$

then the current T coincides with the center manifold N near 0.

9. Order of contact revisited

In §8 we proved that the order of contact of the multiple-valued approximation function to the area minimizing current with the center manifold is well defined.

Here we show that the behavior of this multiple-valued function is somewhat similar to the analytical function: The order of contact $N = K^{-1} = \lim_{r\to 0} K^{-1}(r)$ is finite unless the current coincides with the center manifold itself. In that case the regularity problem is already settled since the center manifold is a $C^{3,\varepsilon}$ branched disk.

We prove the following lemma first.

Lemma 9.1. *Let* $0 < \lambda \leq \frac{1}{2}\min(\gamma\,,\tau)$ *with* $\gamma\,,\tau$ *appearing in* §4. *If* $K = 0$, *then*

$$K(r) \leq \Gamma_2 D^{\lambda}(r)\,.$$

Proof. It is done by using the monotonicity result in §8. There we proved

$$K(r)e^{-2\mathscr{B}(r)} - 4\mathscr{A}(r) \leq K(s)e^{-2\mathscr{B}(s)} - 4\mathscr{A}(s)$$

for $r > s > 0$. If we let $s \to 0$, then

$$\lim_{s \to 0} K(s) = K = 0\,, \qquad \lim_{s \to 0} \mathscr{A}(s) = 0\,.$$

So

$$K(r) e^{-2\mathscr{B}(r)} \leq 4\mathscr{A}(r)\,.$$

Using our knowledge of $\mathscr{B}$ from Lemma 8.2, we conclude that

$$K(r) \leq 8\mathscr{A}(r) \tag{69}$$

when r is small enough. Thus it is enough to prove

$$\mathscr{A}(r) \leq \Gamma_2 D^{\lambda}(r)\,.$$

Notice that $\mathscr{A}(s)$ involves the height integral itself. Small $K(s)$ leads to small $\mathscr{E}_1$ hence small $\mathscr{A}(s)$, thus makes $K(s)$ smaller, thanks to (69).

Now we verify this idea by analyzing the error term $\mathscr{E}_1(s)$ from the squash deformation in detail. Applying (H2) and (H3) in §4 to $\mathscr{E}_1(s)$ defined in Theorem 4.1, we obtain

$$\begin{aligned}
\mathscr{A}(r) &= \int_0^r \frac{\mathscr{E}_1(s)}{sD(s)}\,ds \\
&\leq \int_0^r \frac{1}{sD(s)} \left(\int_{B_s} |f|^2 t^{-2+2\varepsilon}\, d\mathscr{L}^2 + D^{1+\tau}(s) \right) \\
&\quad + \int_0^r \frac{1}{sD(s)} \left(\int_{\partial B_s} (|Df|^3 |f| + |Df||f|^2 s^{-1+\varepsilon}) s\, d\theta + s \frac{d}{ds} \mathbf{M}(T_s^2) \right) \\
&\leq \int_0^r \frac{1}{sD(s)} \left(\int_0^s K(t) D(t) t^{-1+2\varepsilon} + D^{1+\tau}(s) \right) ds \\
&\quad + \int_0^r \frac{1}{sD(s)} \left(\int_{\partial B_s} (|Df|^3 |f| + |Df||f|^2 s^{-1+\varepsilon}) s\, d\theta + s \frac{d}{ds} \mathbf{M}(T_s^2) \right) ds \\
&\leq \int_0^r \frac{1}{sD(s)} \left(\mathscr{A}(s) D(s) \int_0^s t^{-1+2\varepsilon}\, dt + D^{1+\tau}(s) \right) ds \\
&\quad + \int_0^r \left(D^{-1+\tau}(s) D'(s) + K(s) s^{\varepsilon} + \frac{1}{D(s)} \frac{d}{ds} \mathbf{M}(T_s^2) \right) ds \\
&\leq \frac{r^{2\varepsilon}}{2\varepsilon} \mathscr{A}(r) + D^{\tau/2}(r) + \mathscr{A}(r) \frac{r^{\varepsilon}}{\varepsilon} + \int_0^r \frac{1}{D(s)} \frac{d}{ds} \mathbf{M}(T_s^2)\, ds\,.
\end{aligned}$$

We estimate the last term in the last inequality by using integration by parts and (52),

$$\begin{aligned}
\int_0^r \frac{1}{D(s)} \frac{d}{ds} \mathbf{M}(T_s^2)\, ds &= \frac{\mathbf{M}(T_r^2)}{D(r)} + \int_0^r \mathbf{M}(T_s^2) \frac{D'(s)}{D^2(s)}\, ds \\
&\leq C_3' D^{\gamma}(r) + \int_0^r D^{-1+\gamma} D'(s)\, ds \\
&= C_3' \left(D^{\gamma}(r) + \frac{D^{\gamma}(r)}{\gamma} \right)\,.
\end{aligned}$$

Putting back this estimate, then

$$\mathscr{A}(r) \le \mathscr{A}(r)\left(\frac{r^{2\varepsilon}}{2\varepsilon}+\frac{r^{\varepsilon}}{\varepsilon}\right)+D^{\tau/2}(r)+C_3'\left(D^{\gamma}(r)+\frac{D^{\gamma}(r)}{\gamma}\right).$$

When $(r^{2\varepsilon}/2\varepsilon + r^{\varepsilon}/\varepsilon) \le 1/2$, we move the first term on the right to the left side and the lemma follows. Now we are ready to prove the theorem.

Theorem 9.1.

$$K = \lim_{s\to 0} K(s) \neq 0$$

unless the current T_r *coincides with* N_r *for small* $r > 0$.

Proof. We use the formula obtained from the squash deformation to show there is a contradiction if $K = 0$ and $D(r) > 0$ for all small $r > 0$.

$$\begin{aligned}
D(s) &\le \int_{\partial B_s}(f\,,f_s)\,d\mathscr{H}^1+\mathscr{E}_1(s)\\
&\le \left(\int_{\partial B_s}|f|^2 s\,d\theta\right)^{1/2}\left(\int_{\partial B_s}|f_s|^2 s\,d\theta\right)^{1/2}+\mathscr{E}_1(s)\\
&\le (H(s))^{1/2}(D'(s))^{1/2}+\mathscr{E}_1(s)\\
&\le D^{\lambda/2}(s)sD'(s)+D^{-\lambda/2}\frac{H(s)}{s}+\mathscr{E}_1(s)
\end{aligned}$$

by Cauchy-Schwarz inequality. Using

$$K(s) \le \Gamma_2 D^{\lambda}$$

which is equivalent to

$$H(s) \le sD^{1+\lambda}(s)\ ,$$

we have

$$D(s) \le \Gamma_2 D^{1+\lambda/2}(s)+D^{\lambda/2}(s)sD'(s)+\mathscr{E}_1(s).$$

When s is small, move the first term on the right to the left, then for $C > 0$

$$\frac{1}{s} \le C\frac{D'(s)}{D^{1-\lambda/2}(s)}+C\frac{\mathscr{E}_1(s)}{sD(s)}.$$

We integrate this inequality, then the following holds for $0 < s < r$ for small r,

$$\log\left(\frac{r}{s}\right) \le \frac{2}{\lambda}\left(D^{\lambda/2}(r)-D^{\lambda/2}(s)\right)+\mathscr{A}(r)-\mathscr{A}(s)$$

(recalling the definition of $\mathscr{A}(s)$ in Lemma 8.2). Let s go to 0, then we have a finite number on the right and positive ∞ on the left, an obvious contradiction.

So we have $D(r) = 0$ for small $r > 0$, assuming $K = 0$. From the height control estimates we conclude that

$$H(r) = 0$$

for each small $r > 0$. So the part of the current T_r^1 which is in the graph of $F + f$ lies on the center manifold. Using (H3) in §4, we conclude that the whole current T_r lies on the center manifold for small r. Since ∂T_r is supported in $NP^{-1}(\partial B_r)$, we conclude that $T_r = k[\![N_r]\!]$ where $[\![N_r]\!]$ is the current associated with N_r.

10. Rate of convergence

From §9, we know that the function

$$N(r) = K^{-1}(r)$$

is a well-defined function for sufficiently small $r > 0$, and

$$N = \lim_{r \to 0} N(r) \neq 0$$

unless T_r coincides with N_r, in which case we need not to do anything further. So we shall assume that $N < \infty$. The fact that $N > 0$ is already known following the work on height control estimates.

The result in this section is essential to this paper. We prove the following estimates for the rate of convergence for the function $N(s)$ which defines the order of contact.

Theorem 10.1. *Let $0 < \lambda < \frac{1}{2}\min(\varepsilon, \gamma, \sigma, \tau)$, then there is a constant C such that*

$$|N(s) - N| \leq Cs^{\lambda}$$

for each small $s > 0$.

In the next section we shall use this estimate to prove that the normalized sequence of the approximation function restricted to ∂B_s uniquely converges to a multiple-valued homogeneous function. The estimate is obtained using the comparison surface constructed in §6.

There we show that for any r small, we have a g with the same boundary values as f and its Dirichlet integral can be estimated by

$$D(r, g) \leq (1 + Cr)D(r, \overline{p}) + CrH(r).$$

Here the function

$$\overline{p} : B_r \to Q_{KJ}(R^{2+m} \times \{0\})$$

is the multiple-valued harmonic function with the boundary value $\overline{f}$. Using the expression in §6 defining the function $\overline{p}$, we obtain

$$\int_{B_r} |D\overline{p}|^2 \, d\mathscr{L}^2 = \sum_{l=1}^{L} c_l \sum_{j=0}^{\infty} j(|\overline{\alpha}_{lj}|^2 + |\overline{\beta}_{lj}|^2) \tag{70}$$

and

$$\int_{\partial B_r} \left|\frac{\overline{p}_\theta}{r}\right|^2 r\,d\theta = \frac{1}{r}\sum_{l=1}^{L} c_l \sum_{j=0}^{\infty} \frac{j^2}{KJ_l}(|\overline{\alpha}_{lj}|^2 + |\overline{\beta}_{lj}|^2), \tag{71}$$

$$H(\overline{p}, r) = \int_{\partial B_r} |f|^2 r\,d\theta$$

$$= r\sum_{l=1}^{L} c_l \sum_{j=0}^{\infty} KJ_l(|\overline{\alpha}_{lj}|^2 + |\overline{\beta}_{lj}|^2). \tag{72}$$

We first prove the following elementary algebraic fact:

Lemma 10.1 *There is $\lambda_0 > 0$, just depending on $\kappa > 0$, such that for any $0 < \lambda < \lambda_0$ the following holds:*

$$\frac{j^2}{KJ_l} - \kappa^2 KJ_l \geq (2\kappa + \lambda)(j - \kappa KJ_l)$$

for $j \in \mathbf{Z}^+ \cup \{0\}$.

Proof. It is easy to check that there is $\lambda_0 > 0$, for any $0 < \lambda < \lambda_0$, and for any

$$j \in \mathbf{Z}^+ \cup \{0\}$$

we have either

$$j = \kappa KJ_l \qquad \text{or} \qquad |j - \kappa KJ_l| \geq \lambda KJ_l. \tag{73}$$

If $j \leq \kappa KJ_l$, then

$$j + \kappa KJ_l \leq 2\kappa KJ_l \leq (2\kappa + \lambda)KJ_l$$

and

$$j - \kappa KJ_l \leq 0$$

so

$$(j + \kappa KJ_l)(j - \kappa KJ_l) \geq (2\kappa + \lambda)KJ_l(j - \kappa KJ_l),$$

$$\left(\frac{j^2}{KJ_l} - \kappa^2 KJ_l\right) \geq (2\kappa + \lambda)(j - \kappa KJ_l).$$

If $j > \kappa KJ_l$ then using (73), we know that

$$j + \kappa KJ_l \geq 2\kappa KJ_l + \lambda KJ_l$$

so

$$(j + \kappa KJ_l)(j - \kappa KJ_l) \geq (2\kappa KJ_l + \lambda KJ_l)(j - \kappa KJ_l)$$

$$\left(\frac{j^2}{KJ_l} - \kappa^2 KJ_l\right) \geq (2\kappa + \lambda)(j - \kappa KJ_l).$$

Combining those two cases, we have the desired result.

This inequality enables us to get the following:

Lemma 10.2 *For the function* p *constructed in* §6, *we have that for any* $\kappa > 0$, *there is* $\lambda_0 > 0$ *such that for* $0 < \lambda \leq \lambda_0$ *the following holds:*

$$D(\overline{p}, s) \leq \frac{1}{(2\kappa + \lambda)} \left(s \int_{\partial B_s} \left| \frac{\overline{p}_\theta}{s} \right|^2 s\, d\theta + \kappa(\kappa + \lambda) \frac{H(\overline{p}, s)}{s} \right) .$$

Proof. Using the expressions (70), (71), (72) and applying Lemma 10.1 to every coefficient in those expressions, we obtain

$$s \int_{\partial B_s} \left| \frac{\overline{p}_\theta}{s} \right|^2 s\, d\theta - \kappa^2 \frac{H(\overline{p}, s)}{s} \geq (2\kappa + \lambda) \left(D(\overline{p}, s) - \kappa \frac{H(\overline{p}, s)}{s} \right)$$

which is equivalent to the Lemma 10.2.

Proof of Theorem 10.1. Since $g|_{\partial B_s} = f|_{\partial B_s}$,

$$\int_{\partial B_s} \left| \frac{g_\theta}{s} \right|^2 s\, d\theta = \int_{\partial B_s} \left| \frac{f_\theta}{s} \right|^2 s\, d\theta .$$

From the definition of g in §6 (6.2), we see immediately that

$$\int_{\partial B_s} \left| \frac{\overline{p}_\theta}{s} \right|^2 d\mathscr{H}^1 \leq \int_{\partial B_s} \left| \frac{g_\theta}{s} \right|^2 d\mathscr{H}^1,$$

$$H(\overline{p}, s) \leq H(g, s),$$

since $\overline{p}$ is the projection of g into the tangent space T_0M. Using Lemma 6.1(b), Lemma 6.2(a), Lemma 10.2 by setting $\kappa = N$ and choosing $\lambda < \frac{1}{2}\min(\varepsilon, \gamma, \sigma, \tau)$, we have

$$\begin{aligned} D(s) &\leq D(g, r) + \mathscr{E}_3(s) \\ &\leq (1 + Cs) D(\overline{p}, s) + CsH(s) + \mathscr{E}_3(s) \\ &\leq \frac{(1 + Cs)}{2N + \lambda} \left(s \int_{\partial B_s} \left| \frac{f_\theta}{s} \right|^2 s\, d\theta + (N + \lambda) N \frac{H(s)}{s} \right) + CsH(s) + \mathscr{E}_3(s) \\ &\leq \frac{(1 + Cs)}{2N + \lambda} \left(\frac{s}{2} D'(s) + \frac{s}{2} \int_{\partial B_s} \left(\left| \frac{f_\theta}{s} \right|^2 - |f_s|^2 \right) s\, d\theta + (N + \lambda) N \frac{H(s)}{s} \right) \\ &\quad + CsH(s) + \mathscr{E}_3(s) \qquad (74) \\ &\leq \frac{(1 + Cs)}{2N + \lambda} \left(\frac{s}{2} D'(s) + (N + \lambda) N \frac{H(s)}{s} \right) + CsH(s) + \mathscr{E}_3(s) + \mathscr{E}_2(s) \end{aligned}$$

the last inequality is obtained by using Theorem 5.1. By rearranging this inequality, we have

$$\begin{aligned} \frac{sD'(s)}{2} &\geq (2N + \lambda) D(s) - N(N + \lambda) \frac{H(s)}{s} \\ &\quad - C \left(sH(s) + \mathscr{E}_2(s) + \mathscr{E}_3(s) + sD(s) \right) . \qquad (75) \end{aligned}$$

Let

$$\mathscr{E}_4(s) = C\left(sH(s) + sD(s) + \mathscr{E}_2(s) + \mathscr{E}_3(s)\right) . \tag{76}$$

Multiplying by $H(s)/s$ on both sides of (75), we have

$$\frac{H(s)D'(s)}{2} \geq (2N+\lambda)D(s)\frac{H(s)}{s} - (N+\lambda)N\frac{H^2(s)}{s^2} - \frac{H(s)}{s}\mathscr{E}_4(s)$$

so

$$\begin{aligned} D^2(s) - \frac{H(s)D'(s)}{2} &\leq D^2(s) - (2N+\lambda)D(s)\frac{H(s)}{s} \\ &\quad + (N+\lambda)N\frac{H^2(s)}{s^2} + \frac{H(s)}{s}\mathscr{E}_4(s) . \end{aligned} \tag{77}$$

Using Lemma 8.1 and (77), the derivatives of $K(s)$ can be estimated by

$$\begin{aligned} K'(s) &= \frac{2\int_{\partial B_s}(f\ , f_s)s\,d\theta D(s) - H(s)D'(s)}{sD^2(s)} \\ &= 2\frac{D^2(s) - H(s)D'(s)/2}{sD^2(s)} + \frac{\mathscr{E}_1(s)}{sD(s)} \\ &\leq \frac{1}{sD^2(s)}\left(D^2(s) - (2N+\lambda)D(s)\frac{H(s)}{s} + (N+\lambda)N\frac{H^2(s)}{s^2}\right) \\ &\quad + \frac{1}{sD^2(s)}\frac{H(s)}{s}\mathscr{E}_4(s) + \frac{\mathscr{E}_1(s)}{sD(s)} . \end{aligned} \tag{78}$$

The first term on the right can be rewritten as

$$\begin{aligned} &\frac{1}{sD^2(s)}\left(D^2(s) - (2N+\lambda)D(s)\frac{H(s)}{s} + (N+\lambda)N\frac{H^2(s)}{s^2}\right) \\ &\quad = \frac{1}{s}\left(1 - (2N+\lambda)\frac{H(s)}{sD(s)} + (N+\lambda)N\left(\frac{H(s)}{sD(s)}\right)^2\right) \\ &\quad = \frac{1}{s}N\left(K(s) - \frac{1}{N}\right)((N+\lambda)K(s) - 1) \end{aligned}$$

by definition of $K(s)$. Putting this back into (78),

$$\begin{aligned} K'(s) &\leq \frac{1}{s}N\left(K(s) - \frac{1}{N}\right)((N+\lambda)K(s) - 1) \\ &\quad + \frac{1}{sD^2(s)}\frac{H(s)}{s}\mathscr{E}_4(s) + \frac{1}{sD(s)}\mathscr{E}_1(s) . \end{aligned} \tag{79}$$

Using the monotonicity of $K(s)e^{-2\mathscr{B}(s)} - 4\mathscr{A}$ established in §8 and Lemma 8.2, we have

$$K(s) - K - 4\mathscr{A}(s) - 4\mathscr{B}(s) \leq 0 \tag{80}$$

for s small enough. Since $(N+\lambda)K(s)$ goes to $(N+\lambda)K = 1+\lambda K$, there is a $\xi > 0$ such that

$$(N+\lambda)K(s) - 1 \geq (1-\xi)\lambda K > 0. \tag{81}$$

As a consequence of (80) and (81), we have

$$\begin{aligned} &((N+\lambda)K(s)-1)\,(K(s)-K-4\mathscr{A}(s)-\mathscr{B}(s)) \\ &\leq (1-\xi)\lambda K\,(K(s)-K-4\mathscr{A}(s)-4\mathscr{B}(s)) \end{aligned} \tag{82}$$

which is the same as

$$\begin{aligned} &\left(K(s)-\frac{1}{N}\right)((N+\lambda)K(s)-1) \\ &\leq K\lambda(1-\xi)\left(K(s)-\frac{1}{N}\right)+C\mathscr{A}(s)+C\mathscr{B}(s)\,. \end{aligned} \tag{83}$$

Using (83) to bound the left side of (79), we have the following differential inequality:

$$\begin{aligned} (K(s)-K)' &= K'(s) \\ &\leq \frac{1}{s}K\lambda(1-\xi)N\left(K(s)-\frac{1}{N}\right) \\ &\quad + K(s)\frac{1}{sD(s)}\mathscr{E}_4(s)+\frac{1}{sD(s)}\mathscr{E}_1(s)+C\frac{\mathscr{A}(s)}{s}+C\frac{\mathscr{B}(s)}{s}\,. \end{aligned} \tag{84}$$

Using $KN = 1$, the fact that $K(s) \leq C$, and dividing both sides by $1/s^{\lambda(1-\xi)}$ we finally have

$$\begin{aligned} \frac{d}{ds}\left(\frac{K(s)-K}{s^{\lambda(1-\xi)}}\right) &\leq \frac{C\mathscr{E}_4(s)+\mathscr{E}_1(s)}{s^{1+\lambda}D(s)} \\ &\quad + C\frac{\mathscr{A}(s)+\mathscr{B}(s)}{s^{1+\lambda}}\,. \end{aligned} \tag{85}$$

Using Lemma 8.2 and the choice of λ, we know that

$$\frac{\mathscr{A}(s)}{s^{1+\lambda}}, \qquad \frac{\mathscr{B}(s)}{s^{1+\lambda}}, \qquad \frac{\mathscr{E}_1(s)}{s^{1+\lambda}D(s)}$$

are integrable over $(0, r)$ for small r. As for $\mathscr{E}_4(s)$, by definition

$$\mathscr{E}_4(s) = C\left(sH(s)+sD(s)+\mathscr{E}_2(s)+\mathscr{E}_3(s)\right)\,. \tag{86}$$

Obviously

$$\frac{sH(s)+sD(s)+\mathscr{E}_2(s)}{s^{1+\lambda}D(s)}$$

is integrable as a consequence of the height control estimates Theorem 7.1 and Lemma 8.2. From the definition of $\mathscr{E}_3(s)$ in Lemma 6.1 and using Lemma 6.2 (c) we see immediately that $\mathscr{E}_3(s)$ can be bounded by $\mathscr{E}_1(s)+\mathscr{E}_2(s)$. Thus

$$\frac{\mathscr{E}_3(s)}{s^{1+\lambda}D(s)}$$

is integrable over $(0, r)$.

We integrate (85) over (s, r) for small $0 < s < r$ to obtain

$$\frac{K(r) - K}{r^{\lambda(1-\xi)}} \leq \frac{K(s) - K}{s^{\lambda(1-\xi)}} + C \tag{87}$$

which leads to

$$K(s) - K \geq C' s^{\lambda(1-\xi)}$$

where the constant C' depends on $r > 0$, but we fix this r. On the other hand, following the monotonicity we have

$$K(s) - K \leq 4\mathscr{A}(s) + 4\mathscr{B}(s) \leq Cs^{\lambda}$$

by Lemma 8.2. Hence we have

$$|K(s) - K| \leq Cs^{\lambda/2}$$

which is equivalent to the statement in the theorem, since $N(s) = K^{-1}(s)$ and $N \neq \infty$.

11. Toward uniqueness

In this section we prove one of the main results of this paper, that is the uniqueness of the limit of normalized sequences defined over the branched center manifold. This is the key step in separating the current near any singular point. The other fact we need is the pointwise convergence of the sequence which is proved in the next section.

First we have the following theorem, which is a consequence of the theorem in the last section and which gives more evidence of the niceness of the multiple-valued function approximating area minimizing current.

Let $\varepsilon, \gamma, \sigma, \tau$ be the small positive constants appearing in §4, and let N be the order of contact defined in §9.

Theorem 11.1. *There exist positive constants h_0, d_0, such that for some constants $C > 0$, $\lambda \leq \frac{1}{3}\min(\varepsilon, \gamma, \sigma, \tau)$ and for each small $s > 0$, the following hold:*

$$\left|\frac{H(s)}{s^{2N+1}} - h_0\right| \leq Cs^{\lambda}, \qquad \left|\frac{D(s)}{s^{2N}} - d_0\right| \leq Cs^{\lambda}.$$

In particular, both

$$\lim_{s \to 0} \frac{H(s)}{s^{2N+1}}, \qquad \lim_{s \to 0} \frac{D(s)}{s^{2N}}$$

exist.

Proof. It is proved by calculating the derivative of $\log(H(s)/s^{2N+1})$,

$$\begin{aligned}\frac{d}{ds}\log\left(\frac{H(s)}{s^{2N+1}}\right) &= \frac{d}{ds}\left(\log\left(\frac{H(s)}{s}\right)-\log(s^{2N})\right)\\ &= \frac{2}{H(s)}\int_{B_s}(f(s,\theta),f_s(s,\theta))s\,d\theta-\frac{2N}{s}\\ &= \frac{2D(s)}{H(s)}+\frac{\mathscr{E}_1(s)}{H(s)}-\frac{2N}{s}\\ &= \frac{2}{s}\left(\frac{sD(s)}{H(s)}-N\right)+\frac{sD(s)}{H(s)}\frac{\mathscr{E}_1(s)}{sD(s)}\\ &= \frac{2}{s}(N(s)-N)+\frac{sD(s)}{H(s)}\frac{\mathscr{E}_1(s)}{sD(s)}.\end{aligned}$$

Using Theorem 10.1, estimates on $\mathscr{A}(s)$ in Lemma 8.2, and the fact that $N(s)=sD(s)/H(s)\le 2N$ for small s, we have

$$\left|\frac{d}{dt}\log\left(\frac{H(t)}{t^{2N+1}}\right)\right| \le \frac{2C}{t^{1-\lambda}}+2N\frac{\mathscr{E}_1(t)}{tD(t)}.$$

We integrate over (s,r) for small $r,s>0$ to get

$$\left|\int_s^r\frac{d}{dt}\log\left(\frac{H(t)}{2N+1}\right)dt\right|\le\frac{2C}{\lambda}(r^\lambda-s^\lambda)+2N(\mathscr{A}(r)-\mathscr{A}(s)),\tag{88}$$

$$\left|\log\left(\frac{H(r)}{r^{2N+1}}\frac{s^{2N+1}}{H(s)}\right)\right|\le\frac{2C}{\lambda}(r^\lambda-s^\lambda)+2N(\mathscr{A}(r)-\mathscr{A}(s)).\tag{89}$$

From this we see that

$$\frac{H(s)}{s^{2N+1}}e^{-2Cs^\lambda/\lambda-2N\mathscr{A}(s)}$$

is a decreasing function and

$$\frac{H(s)}{s^{2N+1}}e^{2Cs^\lambda/\lambda+2N\mathscr{A}(s)}$$

is an increasing function, so

$$\lim_{s\to0}\frac{H(s)}{s^{2N+1}}=h_0\ne0$$

unless $H(s)=0$ for small $s>0$. Using the fact $\mathscr{A}(s)\le s^\lambda$ and the monotonicity of the above two functions, we have

$$h_0\le\frac{H(s)}{s^{2N+1}}e^{2Cs^\lambda/\lambda+\mathscr{A}(s)}\le\frac{H(s)}{s^{2N+1}}(1+C's^\lambda).$$

Also

$$h_0\ge\frac{H(s)}{s^{2N+1}}e^{-2Cs^\lambda/\lambda-\mathscr{A}(s)}\ge\frac{H(s)}{s^{2N+1}}(1-C's^\lambda),$$

so we have

$$\left|\frac{H(s)}{s^{2N+1}}-h_0\right|\le C's^\lambda.$$

Since

$$\frac{D(s)}{s^{2N}} = \frac{sD(s)}{H(s)} \frac{H(s)}{s^{2N+1}} = N(s) \frac{H(s)}{s^{2N+1}} ,$$

$$\left| \frac{H(s)}{s^{2N+1}} - h_0 \right| \leq C' s^{\lambda}$$

and

$$|N(s) - N| \leq C s^{\lambda} ,$$

the following holds

$$\left| \frac{D(s)}{s^{2N}} - Nh_0 \right| \leq C'' s^{\lambda} ,$$

in particular

$$\lim_{s \to 0} \frac{D(s)}{s^{2N}} = d_0$$

exists.

Using Theorem 11.1, we can prove the following:

Theorem 11.2. (1) *There is a multiple-valued function* g *defined over* ∂B_1 *such that for small* $s > 0$,

$$\int_0^{2\pi} \mathcal{G}^2 \left(\frac{f(s, \theta)}{s^N}, g(\theta) \right) d\theta \leq C s^{\lambda}$$

with constants $C > 0$, $\lambda > 0$, *where* λ *is the same as in Theorem* 11.1.

(2) $g(s, \theta) = s^N g(\theta)$ *is a strictly multiple-valued* N*-admissible function.*

(3) *The function* g *minimizes the Dirichlet integral among all the* N*-admissible functions* (*see Definition* 3.10).

graph$(F + g)$ *is a multiple-sheet covering of the branched center manifold* $N = \text{graph}\, F$.

Remark. The third part relies on the pointwise convergence of the sequence to g, which is proved in the next section without relying on this part of the theorem (also see the end of the following proof).

Proof. Let $\mathcal{G}$ denote the distance function on the space $Q_{KJ}(R^{2+m+n})$ as defined in [A], the map

$$\xi : Q_K(R^{2+m+n}) \to R^{(2+m+n)n(J)}$$

mentioned in §3 (modified version of the map ξ defined in [A, Chapter 1.12]) is the isometric map between this multi-valued space and a subset of the vector space of high dimension. For some of the properties of ξ one can look up §3 or [A] for more details. The following calculations adopt the conventions for

N-admissible functions as mentioned in §3 after Definition 3.10.

$$\begin{aligned}
&\int_0^{2\pi} \mathcal{G}^2\left(\frac{f(s,\theta)}{s^N}, \frac{f(r,\theta)}{r^N}\right) d\theta \\
&\quad \le C_\xi \int_0^{2\pi} \left|\frac{\xi\circ f(s,\theta)}{s^N} - \frac{\xi\circ f(r,\theta)}{r^N}\right|^2 d\theta \\
&\quad \le \int_0^{2\pi} \left(\int_s^r \left|\left(\frac{\xi\circ f(t,\theta)}{t^N}\right)'\right| dt\right)^2 d\theta \\
&\quad \le \int_0^{2\pi} (r-s) \left(\int_s^r \left|\left(\frac{\xi\circ f(t,\theta)}{t^N}\right)'\right|^2 dt\right) d\theta \\
&\quad = \int_0^{2\pi} (r-s) \left(\int_s^r \left|\frac{(\xi\circ f)_t(t,\theta)}{t^N} - N\frac{\xi\circ f(t,\theta)}{t^{N+1}}\right|^2 dt\right) d\theta \\
&\quad = \int_0^{2\pi} (r-s) \left(\int_s^r \left(\frac{|(\xi\circ f)_t(t,\theta)|^2}{t^{2N}}\right.\right. \\
&\qquad\qquad \left.\left. - 2N\frac{(\xi\circ f(s,\theta), (\xi\circ f)_t(t,\theta))}{t^{2N+1}} + N^2\frac{|\xi\circ f(t,\theta)|^2}{t^{2N+2}}\right) dt\right) d\theta \\
&= \int_0^{2\pi} (r-s) \left(\int_s^r \left(\frac{|f_t(t,\theta)|^2}{t^{2N}}\right.\right. \\
&\qquad\qquad \left.\left. - 2N\frac{(f(t,\theta), f_t(t,\theta))}{t^{2N+1}} + N^2\frac{|f(t,\theta)|^2}{t^{2N+2}}\right) dt\right) d\theta .
\end{aligned}$$

The last equality is obtained by using the properties of the map ξ mentioned before Definition 3.2. We change the order of the integration to get

$$\begin{aligned}
&\int_0^{2\pi} \mathcal{G}^2\left(\frac{f(s,\theta)}{s^N}, \frac{f(r,\theta)}{r^N}\right) d\theta \\
&\le (r-s)\int_s^r \left(\int_0^{2\pi} \left(\frac{|f_t(t,\theta)|^2}{t^{2N}} - 2N\frac{(f(t,\theta), f_t(t,\theta))}{t^{2N+1}} + N^2\frac{|f(t,\theta)|^2}{t^{2N+2}}\right) d\theta\right) dt
\end{aligned}$$

$$= (r-s)\int_s^r \left(\frac{D'(t)}{2t^{2N+1}} - 2N\frac{D(t)}{t^{2N+2}} + N^2\frac{H(t)}{t^{2N+3}}\right) dt \tag{90}$$

$$+(r-s)\int_s^r \left(\frac{\mathcal{E}_2(t)}{t^{2N+2}} + 2N\frac{\mathcal{E}_1(t)}{t^{2N+2}}\right) dt . \tag{91}$$

The first part in the last inequality can expressed as

$$\begin{aligned}
&(r-s)\int_s^r \left(\frac{1}{2t}\left(\frac{D(t)}{t^{2N}}\right)' - N^2\frac{D(t)}{t^{2N+2}}\left(\frac{1}{N} - \frac{H(t)}{tD(t)}\right)\right) dt \\
&\qquad = (r-s)\int_s^r \left(\frac{1}{2t}\left(\frac{D(t)}{t^{2N}}\right)' - N^2\frac{D(t)}{t^{2N+2}}(K - K(t))\right) dt .
\end{aligned}$$

If we let $s = \frac{1}{2}r$ and use Theorems 10.1 and 11.1, then

$$\begin{aligned}(r-s)\int_s^r \left(\frac{1}{2t}\left(\frac{D(t)}{t^{2N}}\right)' - N^2\frac{D(t)}{t^{2N+2}}(K-K(t))\right) dt &\le \frac{r}{2}\int_{r/2}^r \left(\frac{1}{r}\left(\frac{D(t)}{t^{2N}}\right)' + \frac{4CN^2 d_0}{t^2}t^\lambda\right) dt \\ &\le \frac{1}{2}\left|\frac{D(r)}{r^{2N}} - \frac{D(s)}{s^{2N}}\right| + C' r^\lambda \\ &\le C'' r^\lambda ,\end{aligned}$$

by using Theorem 11.1 again for the first term involving Dirichlet integral. On the other hand we already know from Lemma 8.2 that

$$\begin{aligned}(r-s)\int_s^r \left(\frac{\mathscr{E}_1(t)+\mathscr{E}_2(t)}{t^{2N+2}}\right) dt &= (r-s)\int_s^r \frac{1}{t}\left(\frac{\mathscr{E}_1(t)+\mathscr{E}_2(t)}{tD(t)}\frac{D(t)}{t^{2N}}\right) dt \\ &\le 4d_0 r\int_0^r \left(\frac{\mathscr{E}_1(t)+\mathscr{E}_2(t)}{tD(t)}\right) dt \\ &\le 4d_0 r^\lambda .\end{aligned}$$

Putting the above two estimates back into (90) and (91) we get

$$\frac{r}{2}\int_{r/2}^r \left(\frac{1}{2t}\left(\frac{D(t)}{t^{2N}}\right)' - N^2\frac{D(t)}{t^{2N+1}}(K-K(t)) + \frac{\mathscr{E}_1(t)+\mathscr{E}_2(t)}{t^{2N+2}}\right) dt \le C_1 r^\lambda ,$$

which implies

$$\left(\int_0^{2\pi} \mathscr{G}^2\left(\frac{f(r,\theta)}{r^N}, \frac{f(r/2,\theta)}{(r/2)^N}\right) d\theta\right)^{1/2} \le C_1' r^{\lambda/2} .$$

By replacing r by $r/2^k$, we obtain

$$\left(\int_0^{2\pi} \mathscr{G}^2\left(\frac{f(r/2^k,\theta)}{(r/2^k)^N}, \frac{f(r/2^{k+1},\theta)}{(r/2^{k+1})^N}\right) d\theta\right)^{1/2} \le C_1'(\frac{r}{2^k})^{\lambda/2} .$$

Hence

$$\begin{aligned}\left(\int_0^{2\pi} \mathscr{G}^2\left(\frac{f(r,\theta)}{r^N}\right), \frac{f(r/2^{k+1},\theta)}{(r/2^{k+1})^N}) d\theta\right)^{1/2} &\le C_1' \sum_{i=0}^{k-1}\left(\frac{r}{2^i}\right)^{\lambda/2} \\ &\le C_1'' r^{\lambda/2} . \end{aligned} \tag{92}$$

This is enough to conclude that the sequence $\{f(r,\theta)/r^N\}$ for small r is a Cauchy sequence. Therefore there is a unique limit

$$g: \partial B_1 \to Q_{KJ}(R^{2+m+n})$$

and

$$\lim_{s\to 0}\int_0^{2\pi} \mathscr{G}^2\left(\frac{f(s,\theta)}{s^N}, g(\theta)\right) d\theta \le Cs^\lambda . \tag{93}$$

By passing to the K-sheet covering disc of B_r where f_0 is defined, it is easy to see that the sequence $\{f_0(\rho^K, \omega)/\rho^{KN}\}$ converges to some function g_0, we conclude that the function

$$g: B_r \to Q_{KJ}(R^{2+m+n}),$$
$$g(s, \theta) = s^N g(\theta)$$

is an N-admissible function. Using the properties of L^2 convergence which are valid for multiple-valued functions as well, by applying the map ξ mentioned before, we have

$$\int_0^{2\pi} |g(\theta)|^2 \, d\theta = \lim_{s\to 0} \int_0^{2\pi} \left| \frac{f(s, \theta)}{s^N} \right|^2 d\theta = \lim_{s \to 0} \frac{H(s)}{s^{2N+1}} = h_0 \neq 0 .$$

So the function g is not a trivial function. Furthermore, the function is a strictly multiple-valued. This can be proved by the following argument. Using the centering hypothesis (H1) in §4, we have

$$\int_{B_r} |\eta \circ f(s, \theta)| \, sds \, d\theta \le C_1 \int_{B_r} |f(s, \theta)|^2 s^{-1+\varepsilon} \, sds \, d\theta + C_1 r D^{1+\gamma}(r) .$$

From this we get

$$\begin{aligned}
&\frac{r^{N+2}}{N+2} \int_0^{2\pi} |\eta \circ g(\theta)| \, d\theta \\
&= \int_{B_r} s^N |\eta \circ g(s, \theta)| sds \, d\theta \\
&\le \int_{B_r} |\eta \circ f(s, \theta)| sds \, d\theta \\
&\quad + \int_{B_r} |\eta \circ f(s, \theta) - s^N \eta \circ g(\theta)| sds \, d\theta \\
&\le C_1 \int_{B_r} |f(s, \theta)|^2 s^{-1+\varepsilon} sds \, d\theta + C_1 r D^{1+\gamma}(r) \\
&\quad + C_1 \int_{B_r} \mathscr{G}(f(s, \theta), s^N g(\theta)) sds \, d\theta \\
&= C_1 \int_0^r H(s) s^{-1+\varepsilon} \, ds + 2C_1 r (d_0 r^{2N})^{1+\gamma} \\
&\quad + \sqrt{\pi} r \left(\int_{B_r} \mathscr{G}^2(f(s, \theta), s^N g(\theta)) sds \, d\theta \right)^{1/2} \\
&\le 2C_1 h_0 \frac{r^{2N+1+\varepsilon}}{2N+1+\varepsilon} + 2C_1 r (d_0 r^{2N})^{1+\gamma} \\
&\quad + \sqrt{\pi} r \left(\int_0^r s^{2N+1} \left(\int_0^{2\pi} \mathscr{G}^2 \left(\frac{f(s, \theta)}{s^N}, g(\theta) \right) d\theta \right) \right)^{1/2} \\
&\le 2C_1 h_0 \frac{r^{2N+1+\varepsilon}}{2N+1+\varepsilon} + 2C_1 r (d_0 r^{2N})^{1+\gamma} + \sqrt{\pi} r \left(C r^{\lambda} \frac{r^{2N+2}}{2N+2} \right)^{1/2} .
\end{aligned}$$

The above holds only when

$$\int_0^{2\pi} |\eta \circ g| \, d\theta = 0 .$$

Since we already know that the function g is not trivial, we conclude that g is strictly multiple-valued and its lifting to the K-sheet covering disc is also strictly multiple-valued following the definition of $|\eta \circ g|$ for N-admissible functions (see Definition 3.10).

Thus the $\operatorname{graph}(F + g)$ defines a strictly multiple-sheet covering of $\operatorname{graph} F$.

This function g actually takes values in $Q_{KJ}(R^{2+m} \times \{0\})$ for the following reason.

As in §6,

$$p = \overline{p} + \Phi(\overline{p}), \qquad p \in M \cap B_4^{2+m} .$$

So we have

$$F_i(x) + f_{ij}(x) = \overline{F}_i(x) + \overline{f}_{ij}(x) + \Phi(\overline{F}_i(x) + \overline{f}_{ij}(x)) ,$$
$$F_i(x) = \overline{F}_i(x) + \Phi(\overline{F}_i(x)).$$

By omitting the lower indices, we can write

$$\begin{aligned} f &= \overline{f} + \Phi(\overline{F} + \overline{f}) - \Phi(\overline{F}) \\ &= \overline{f} + \Phi_i(\overline{F})\overline{f}^i + c|f|^2 \\ &= \overline{f} + R . \end{aligned}$$

The last part is obtained by using the Taylor expansion of Φ. Using the assumption on ϕ and knowledge of f, we have

$$|\Phi_i(\overline{F}(s, \theta))| \le cs , \qquad |f(s, \theta)| \le cs .$$

Thus

$$\int_0^{2\pi} \left| \frac{R(s, \theta)}{s^N} \right|^2 d\theta \le (cs)^2 \int_0^{2\pi} \left| \frac{f(s, \theta)}{s^N} \right|^2 d\theta \le 2h_0(cs)^2$$

with h_0 in Theorem 11.1. So the projection of the sequence of the functions $\{f(s, \theta)/s^N\}$ in the normal direction of T_0M go to 0 with the rate of $O(s^2)$. So the limit function g takes values only in $T_0M = R^{2+m} \times \{0\} \subset R^{2+m+n}$.

As to the fact that g minimizes the Dirichlet integral, it can be verified by using the area comparison Lemma 6.1 and the pointwise convergence proved in the next section. The idea is to show that if there is an N-admissible function h which has less Dirichlet integral, then the savings in the Dirichlet integral can be translated into the savings of area to contradict the assumed fact that T_r is area minimizing. The pointwise convergence theorem in §12 guarantees that there is a integral current S so that

$$S = (F + f)_\sharp(B_r) - (F + g)_\sharp(B_r)$$

with $\mathbf{M}(S) \le Cr^{N+a}$, $a > 0$.

The readers may also consult with [A, Chapters 2 and 3] for additional information.

12. Separation

From the previous section we know that the normalized sequence has a well-defined limit and the convergence has a certain rate. Here we prove that the convergence is actually pointwise. This result is crucial when one tries to separate the current. It is well known that the L^2 convergence does not imply the pointwise convergence, but with a certain extra condition, it is true. In particular, we have the following:

Theorem 12.1. *Suppose $h : B_{r_0} \sim \{0\} \to R^K$ is locally Lipschitz and that there exist constants $a > 0$, $b > 0$, $c > 0$ such that for small each $r > 0$ the following holds:*

$$\int_{B_r} |h(s, \theta)|^2 s ds\, d\theta \le c r^{2+4a}, \tag{94}$$

$$\mathrm{Lip}(h|_{B_{r_0} - B_r}) \le c r^{-b}, \tag{95}$$

$$\int_{B(x, |x|^{1+a})} |Dh(s, \theta)|^2 s ds\, d\theta \le c|x|^a. \tag{96}$$

Then $|h(s, \theta)| \le s^{a/4}$.

Proof. Suppose that $|x|$ is very small and $|h(x)| > |x|^{a/4}$. Let

$$S = \{y \in B(x, |x|^{1+a}) : |h(y)| > \tfrac{1}{2}|x|^{a/4}\}.$$

We estimate

$$\tfrac{1}{4}|x|^{a/2}\mathscr{L}^2(S) \le \int_{B(x, |x|^{1+a})} |h|^2\, d\mathscr{L}^2 \le \int_{B(0, 2x)} |h|^2\, d\mathscr{L}^2 \le c|2x|^{2+4a}$$

by (96). So

$$\mathscr{L}^2(S) \le 2^{4+4a} c|2x|^{2+3a}.$$

Also

$$\mathscr{L}^2(B(x, |x|^{1+a})) = \pi|x|^{2+2a}$$

which implies

$$\frac{\mathscr{L}^2(S)}{\mathscr{L}^2(B(x, |x|^{1+a}))} \le c 2^{3+a}|x|^a. \tag{97}$$

Let $\Omega \subset S^1 \subset R^2$ denotes all $\omega \in S^1$ for which there is $|x|^{1+a} > \rho(\omega) > 0$ with

$$|h(x + \rho(\omega)\omega)| \le \tfrac{1}{2}|x|^{a/4}.$$

Applying the estimates on S we conclude that $\mathscr{H}^1(\Omega) > \pi$. Using Rademacher's theorem we obtain for almost every $\omega \in \Omega$ that

$$\begin{aligned}
\tfrac{1}{2}|x|^{a/4} &\le \left|h(x) - h(x+\rho(\omega))\right| \\
&\le \int_0^{\rho(\omega)} |Dh(x+t\omega)|\,dt \\
&\le \int_0^{|x|^{1+2a+b}} |Dh(x+t\omega)|\,dt + \int_{|x|^{1+2a+b}}^{|x|^{1+a}} |h(x+t\omega)|\,dt \\
&\le c\left|\frac{x}{2}\right|^{-b} \cdot |x|^{1+2a+b} + \left(\int_{|x|^{1+2a+b}}^{|x|^{1+a}} \frac{1}{t}\,dt\right)^{1/2} \\
&\qquad \times \left(\int_{|x|^{1+2a+b}}^{|x|^{1+a}} |Dh(x+t\omega)|^2 t\,dt\right)^{1/2}.
\end{aligned}$$

The last inequality is obtained by using the condition on Lipschitz constant (95) and the Cauchy-Schwarz inequality. Since the above inequality is true for almost every $\omega \in \Omega$, we integrate over Ω and get

$$\begin{aligned}
\frac{\pi}{2}|x|^{a/4} &\le 2^{b+1} c\pi |x|^{1+2a} \\
&\quad + \left(\log\left(\frac{|x|^{1+a}}{|x|^{1+2a+b}}\right)\right)^{1/2} \int_\Omega \left(\int_{|x|^{1+2a+b}}^{|x|^{1+a}} |Dh(x+t\omega)|^2\, tdt\right)^{1/2} d\omega \\
&\le c\pi|x|^{2a} + \left((-k-a)\log(|x|)\cdot c|x|^a\right)^{1/2}.
\end{aligned}$$

If $|x|$ is small enough, then

$$0 \le -(k+a)\log(|x|)|x|^{a/4} \le 1$$

hence

$$\frac{\pi}{2}|x|^{a/4} \le 2c\pi|x|^{2a} + \left(c|x|^{3a/4}\right)^{1/2}.$$

This is impossible for small $|x|$.

With this we can prove the following:

Theorem 12.2. *There is a constant $C > 0$. For a fixed $0 < \lambda < \frac{1}{4}\min(\varepsilon, \gamma, \sigma, \tau)$ and for each small s we have*

$$\mathscr{G}\left(\frac{f(s,\theta)}{s^N}, g(\theta)\right) \le C|s|^{\lambda/16}$$

Proof. We let

$$h(s,\theta) = \frac{\xi \circ f(s,\theta)}{s^N} - \xi \circ g(\theta).$$

From Theorem 11.2, and the Lipschitz condition on f we can see that conditions (94) and (95) can be satisfied with $k = N$. We need to check the third condition. We let $a = \lambda/4$ and let

$$U = B(0, |x| + |x|^{1+a}) \sim B(0, |x| - |x|^{1+a}) \supset B(x, |x|^{1+a}).$$

Then

$$\begin{aligned}
&\int_{B(x, |x|^{1+a})} |Dh|^2 \, d\mathscr{L}^2 \\
&\le \int_{B(x, |x|^{1+a})} \left(\left| D\frac{\xi \circ f}{s^N} \right|^2 + |D\xi \circ g|^2 \right) dA \\
&\le \int_U \left(\left| D\frac{f(s, \theta)}{s^N} \right|^2 + |Dg(\theta)|^2 \right) s \, ds \, d\theta \\
&\le \int_U \left(\frac{|f_s|^2}{s^{2N}} - 2N\frac{(f_s, f)}{s^{2N+1}} + N^2 \frac{|f|^2}{s^{2N+2}} + \frac{1}{s^2}\frac{|f_\theta|^2}{s^{2N}} \right) s ds \, d\theta \\
&\quad + \int_U \frac{1}{s^2} |g_\theta|^2 s ds \, d\theta \\
&= \int_{|x|-|x|^{1+a}}^{|x|+|x|^{1+a}} \left(\left(\frac{D(s)}{s^{2N}} \right)' + N^2 \frac{H(s)}{s^{2N+2}} + 2N \frac{\mathscr{E}_1(s)}{s^{2N+1}} \right) ds \\
&\quad + \int_{|x|-|x|^{1+a}}^{|x|+|x|^{1+a}} \frac{1}{s} ds \int_0^{2\pi} |g_\theta|^2 \, d\theta \\
&\le 2c_2(|x| + |x|^{1+a})^\lambda + 2h_0 \log \left(\frac{|x| + |x|^{1+a}}{|x| - |x|^{1+a}} \right) \\
&\quad + \int_0^{2\pi} |g_\theta|^2 \, d\theta \log \left(\frac{|x| + |x|^{1+a}}{|x| - |x|^{1+a}} \right) .
\end{aligned}$$

Since

$$\log \left(\frac{|x| + |x|^{1+a}}{|x| - |x|^{1+a}} \right) \le \log(1 + 3|x|^a) \le 4|x|^a$$

we get

$$\int_{B(x, |x|^{1+a})} |Dh|^2 \, d\mathscr{L}^2 \ \le \ c|x|^a.$$

Thus we have

$$\mathscr{G} \left(\frac{f(s, \theta)}{s^N}, g(\theta) \right) \ \le \ cs^{a/4}.$$

After replacing a by $\lambda/4$ we have the desired result.

Let

$$Y_r' = \{p \in R^{2+m+n} | \exists x \in B_r, \ p \in \operatorname{spt}(F + g)(x)\}$$

and

$$W_a(Y_r') = \{p \in R^{2+m+n} | \exists q \in Y_r', \ \operatorname{dist}(p, q) = dist(p, Y_r') \le |p|^a\}.$$

Theorem 12.3. (1) *For* $0 < a = \lambda/20$ *and each small* $r > 0$, *we have* $\operatorname{spt} T_r \subset W_{N+a}(Y'_r)$.

(2) *The surface defined by* $\operatorname{graph}(F+g)$ *is a branched disc as smooth as the center manifold* $N = \operatorname{graph} F$ *itself and is a multiple-sheet covering of* N.

Proof. It is enough to show that $\operatorname{spt} T_r \sim \operatorname{spt} T_{r/2} \subset W_{N+a}$ for all small $r > 0$. If $p \in \operatorname{spt} T_r^1$ where T_r^1 is defined in §4, Assumptions, then $p = F_i(x) + f_{ij}(x)$ for some i, j and $x \in B_r$. From Theorem 12.2, we know that

$$\operatorname{dist}(\operatorname{spt}(F(x)+f(x)), \operatorname{spt}(F(x)+g(x)) \le Cx^{N+\lambda/16}.$$

Thus $p \in W_{N+\lambda/17}$ for small r. When $p \in \operatorname{spt} T_r^2$, if $p \notin W_{N+\lambda/20}$, then the support of $T \llcorner B(p, |p|^{N+\lambda/20})$ lies outside of $W_{N+\lambda/17}$ hence a part of T_r^2.

From monotonicity of area, we have a constant $C(M)$ depending only on M such that

$$\mathbf{M}(T \llcorner B(p, |p|^{N+\lambda/20}) \ge C(M)|p|^{2N+\lambda/10} \ge Cr^{2N+\lambda/10}$$

since $p \in \operatorname{spt} T_r \sim \operatorname{spt} T_{r/2}$. But from the estimates of T_r^2 in (52), Theorem 11.1, and the assumption on λ, we have

$$\mathbf{M}(T_r^2) \le D^{1+\tau}(r) < Cr^{2N+2\lambda}$$

to contradict the above inequality when r is small enough.

This enables us to conclude that T_r separates along the different branches of V'_r. The support of the current goes to the surface V'_r very fast.

Since we proved Theorem 12.2, (3) in Theorem 11.2, is now valid. Thus g minimizes the Dirichlet integral. Using Theorem 12.3, Corollary 3.1 and the assumption that T is locally irreducible, we see immediately $g_0|\partial B_t$ must be a single elementary function with possible multiplicity. Also from the proof of Theorem 11.2, we know that g_0 is a strictly multiple-valued function. Thus $F+g$ is a branched disc which is an admissible surface and covers N more than one time. From Theorem 12.2 we know T is in an admissible neighborhood of $\operatorname{graph}(F+g)$. Using the regularity of multiple-valued functions minimizing the Dirichlet integral, we know g is smooth except at the branched point 0. Since F satisfies the estimates as in Lemma 3.2 and comes from F_0 which is $C^{3,\varepsilon}$, so $\operatorname{graph}(F)$ is the $C^{3,\varepsilon}$ image of a disc. Thus $(F+g)B_r$ is the $C^{3,\varepsilon}$ image of a disc.

13. Appendix A

In order to prove the regularity result we assumed the existence of the branched center manifold and several conditions on the approximation function (see Assumptions in §4). In the remaining part of this paper we verify those hypotheses.

The discussion in this section is only for two dimensional currents. In [FM], F. Morgan proved that any area minimizing cone in R^k is a union of planes with multiplicities and any two of such planes intersect only at the origin. Using this, B. White proved that the area-minimizing current in R^k has unique tangent cone at any interior point.

We are going to generalize those two results to area minimizing currents on Riemannian manifolds and give a pointwise convergence estimate.

First we embed the manifold M into R^{2+m+n}, where $2+m$ is the dimension of M. Assume $0 \in \operatorname{spt} T \sim \operatorname{spt} \partial T \subset M$ and $T_0 M = R^{2+m} \times \{0\}$. If $\{\lambda_i\}$ is a sequence of positive numbers such that $\lambda_i \to 0$, then according to [F], there is a subsequence, let us say λ_i itself, such that $\mu(1/\lambda_i)_\sharp(T \llcorner B_{\lambda_i}^{2+m+n})$ converges to a cone T_0 in R^{2+m+n} under the flat norm. We are going to give necessary refinement in the following theorem so that we can apply the results of F. Morgan and B. White.

Theorem 13.1. *The cone T_0 obtained above is area minimizing in the tangent space $T_0 M$. Hence F. Morgan's result applies to T_0 and*

$$T_0 = \sum_{j=1} n_j T_j$$

with $n_j \in \mathbf{Z}^+$, planes $\{T_j\}$ and $T_j \cap T_{j'} = \{0\}$.

The tangent cone T_0 is unique in the sense that it is independent of the choice of the sequence λ_i.

Proof. Let ϕ be as in §3. We define

$$\phi_\lambda : B_1^{2+m} \to R^n, \qquad \Phi_\lambda : B_1^{2+m} \to R^{2+m+n}$$

by

$$\phi_\lambda = \mu(1/\lambda) \circ \phi \circ \mu(\lambda) : B_1^{2+m+n} \to R^n$$

and

$$\Phi_\lambda(x) = (x, \phi_\lambda(x)).$$

Obviously if

$$|\phi(x)| \le |x|^2, \qquad |D\phi(x)| \le |x|$$

for $x \in B^{2+m}(0, 1)$, then

$$|\phi_\lambda(x)| \le \lambda^2 |x|^2, \qquad |D\phi_\lambda(x)| \le \lambda |x|$$

and

$$\operatorname{graph}(\phi_\lambda) = \operatorname{image}(\Phi_\lambda) = M_\lambda = \mu(1/\lambda)_\sharp(M \cap B_\lambda^{2+m} \times R^n).$$

The proof of the fact that the cone T_0 lies in the tangent space $T_0 M$ is easy. Since

$$\begin{aligned} \operatorname{spt} \mu_\sharp(1/\lambda)(T \llcorner B_\lambda^{2+m+n}) &\subset M_\lambda \\ &= \operatorname{graph}(\phi_\lambda) \\ &\subset B^{2+m}(0, 1) \times B^n(0, \lambda^2), \end{aligned}$$

so when $\lambda \to 0$, we have

$$\operatorname{spt}(T_0) \subset B^{2+m}(0,1) \times \{0\}.$$

We denote the projection map onto $T_0 M \subset R^{2+m+n}$ by P. Assume that the cone T_0 is not area minimizing in $B_1^{2+m} \subset T_0 M = R^{2+m} \times \{0\}$, then there is an integral current C such that

$$\partial C = \partial T_0$$

and

$$\mathbf{M}(C) \le \mathbf{M}(T_0) - a$$

for some $a > 0$. Since the sequence of the current

$$T_{\lambda_i} = \mu_\sharp(1/\lambda_i)(T \llcorner B^{2+m+n}(0, \lambda_i))$$

converges to T_0 in the flat norm so we have a sequence of integral currents $\{S_{\lambda_i}\}$ such that

$$\partial T_{\lambda_i} = \partial(T_0 + S_{\lambda_i}) = \partial(C + S_{\lambda_i}),$$

with $\mathbf{M}(S_{\lambda_i}) \le a/4$. Since the induced maps $\Phi_{\lambda\sharp}$ and $P_\sharp$ on the space of integral currents commute with the boundary operator ∂ and the induced map $\Phi_{\lambda_i\sharp} P_\sharp$ is identity on the space of integral currents supported on M_{λ_i}, we have

$$\partial\left(\Phi_{\lambda_i\sharp} P_\sharp(C + S_{\lambda_i})\right) = \partial T_{\lambda_i}.$$

Also the support of $\Phi_{\lambda_i\sharp} P_\sharp(C - S_{\lambda_i})$ lies on the manifold M_{λ_i}. Using the fact T_{λ_i} is area minimizing in M_{λ_i}, we obtain the following comparison on the areas,

$$\begin{aligned}
\mathbf{M}(T_{\lambda_i}) &\le \mathbf{M}(\Phi_{\lambda_i,\sharp} P_\sharp(C - S_{\lambda_i})) \\
&\le (1 + \lambda_i)\mathbf{M}(P_\sharp(C - S_{\lambda_i})) \\
&\le (1 + \lambda_i)\mathbf{M}(C - S_{\lambda_i}) \\
&\le (1 + \lambda_i)(\mathbf{M}(C) + \mathbf{M}(S_{\lambda_i})) \\
&\le (1 + \lambda_i)(\mathbf{M}(T_0) - a + a/4).
\end{aligned}$$

Using the lower semicontinuity of the mass (or area) with respect to the flat norm convergence and the estimates above, we have

$$\mathbf{M}(T_0) \le \varliminf_{\lambda_i \to 0} \mathbf{M}(T_{\lambda_i}) \le \mathbf{M}(T_0) - 3a/4.$$

This is impossible. So T_0 is an area minimizing cone in the Euclidean space $R^{2+m} \times \{0\}$ and we can apply the result of F. Morgan [FM] to conclude

$$T_0 = \sum_{j=1} n_j S_j$$

where the support of S_i is a plane and $\operatorname{spt} S_i \cap \operatorname{spt} S_j = 0$.

The uniqueness follows from [BW] with a little modification. For the comparison surface F constructed in [BW], we use the induced map $\Phi_{\lambda\sharp}$ on the space of integral currents from Φ to map the surface F into the manifold M_λ. Then we have the following estimates

$$\mathbf{M}(\Phi_{\lambda\sharp}F) \leq (1+\lambda)\mathbf{M}(F) .$$

When λ is small enough, it follows from White's estimates that the epiperimetric inequality holds for area minimizing currents on a Riemannian manifold. Hence we have the uniqueness result.

Next we show that the sequence $\{T_\lambda\}$ converges to T_0 with a certain rate.

Theorem 13.2. *There are constants $C, \varepsilon > 0$ such that for each small $\lambda > 0$ the following hold:*

$$\mathscr{F}(T_\lambda, T_0) \leq C\lambda^\varepsilon ,$$

$$\operatorname{dist}\left(\operatorname{spt}(T \llcorner B^{2+m+n}(0,\lambda)), \operatorname{spt} T_0\right) \leq C\lambda^{1/4\varepsilon}$$

with $\mathscr{F}$ the flat norm and $T_\lambda = \mu(\frac{1}{\lambda})_\sharp(T \llcorner B_\lambda^{2+m+n})$.

Proof. According to [BW], there are constants $c, \varepsilon > 0$ such that for small λ,

$$\mathbf{M}(\partial T_\lambda) - \mathbf{M}(\partial T_0) \leq c\lambda^\varepsilon .$$

Also he proved that for each small $s > 0$, the following estimates hold:

$$\lim_{k\to\infty} \mathbf{M}(R_\sharp T_{s, s/2^k}) \leq cs^\varepsilon \tag{98}$$

where $R: B_1^{2+m+n} \sim \{0\} \to \partial B_1^{2+m+n}$ is defined by $R(x) = x/|x|$ and $T_{s, s/2^k}$ denotes the integral current obtained by restricting the current T to the annulus region $B^{2+m+n}(0, s) \sim B^{2+m+n}(0, s/2^k)$.

To show the first statement of the theorem, we need to demonstrate that there are currents A and B where $A \in \mathscr{R}_3(R^{2+m+n})$, $B \in \mathscr{R}_2(R^{2+m+n})$, such that

$$T_\lambda - T_0 = B + \partial A$$

with

$$\mathbf{M}(A) , \mathbf{M}(B) \leq c\lambda^\varepsilon .$$

The construction of A and B. . Let $Z_\delta = [\delta, 1] \times T_\lambda$ be the three dimensional current in $R \times R^{2+m+n}$ as defined in [F], let d be the map $d: R \times R^{2+m+n} \to R$ such that $d(t, x) = t - |x|$, also $Y_a = \{(t, x) : t - |x| \geq a\}$ where $a > 0$ is as small as we like. We denote the annulus region $B^{2+m+n}(0, s) \sim B^{2+m+n}(0, t)$ by $A(t, s)$. We define

$$l: R \times R^{2+m+n} \sim \{0\} \to R^{2+m+n}$$

such that $l_a(t, x) = (t-a)(x/|x|)$. Also let

$$X_a = ([\delta, 1] \times (T_\lambda \llcorner A(\delta, 1)) \llcorner Y_a .$$

It is obvious that

$$\begin{aligned}\partial X_a &= \Big(\partial([\delta, 1] \times (T_\lambda \llcorner A(\delta, 1)))\Big) \llcorner Y_a \\ &\quad + \langle [\delta, 1] \times (T_\lambda \llcorner A(\delta, 1)), d, a\rangle \\ &= \{1\} \times T_\lambda \llcorner A(\delta, 1-a) \\ &\quad -[\delta + a, 1] \times \partial(T_\lambda \llcorner A(0, a)) \llcorner Y_a + \langle Z_\delta, d, a\rangle .\end{aligned}$$

Thus

$$\begin{aligned}\partial l_{a\sharp} X_a &= R_\sharp T_\lambda \llcorner A(\delta, 1-a) \\ &\quad - l_{a\sharp}([\delta + a, 1] \times \partial T_\lambda \llcorner A(0, \delta)) + l_{a\sharp}\langle Z_\delta, d_a, a\rangle .\end{aligned}$$

It is easy to see that

$$l_{a\sharp}\langle Z_\delta , d_a , a\rangle = T_\lambda \llcorner A(\delta , 1)$$

since $l_a(t , x) = (t-a)(x/|x|)$ and $t - a = |x|$ for $(t , x) \in \operatorname{spt}(\langle Z_\delta , d , a\rangle)$. If we let a go to 0, then $R_\sharp T_\lambda \llcorner A(\delta , 1-a)$ go to $R_\sharp T_\lambda \llcorner A(\delta , 1)$, and

$$l_{a\sharp}\Big([\delta + a , 1] \times \partial(T_\lambda \llcorner A(0 , a))\Big)$$

go to $T_0 \llcorner A(\delta , 1)$, hence we have

$$\partial l_\sharp X = R_\sharp T_\lambda \llcorner A(\delta , 1) - T_0 \llcorner A(\delta , 1) + T_\lambda \llcorner A(\delta , 1) .$$

So when δ goes to 0,

$$\partial l_\sharp X = \lim_{\delta\to 0} R_\sharp T_\lambda \llcorner A(\delta , 1) - T_0 + T_\lambda .$$

We let $A = l_\sharp X_0$ and $B = \lim_{\delta\to 0} R_\sharp T_\lambda \llcorner A(\delta , 1)$, so

(99) $$A = B - T_0 + T_\lambda .$$

From (98), we know that the B is well defined, and

$$\mathbf{M}(B) \le c\lambda^\varepsilon .$$

Next we estimate $\mathbf{M}(A)$. We still let $Z = (0 , 1]\times T_\lambda$. First we notice that if $\vec{Z}$ is the tangent space for Z at $(t , x) \in \operatorname{spt} Z$, then Z is spanned orthogonally by $\partial/\partial t$ and $\vec{T}$ which is the tangent plane for the current T_λ at $x \in \operatorname{spt} T_\lambda$. Also it is easy to check that the image of $\partial/\partial t$ under the differential of l is perpendicular to the image of $\vec{T}$, since one is in the radial direction and the other is in the spherical direction. Furthermore at (t , x),

$$dl\left(\frac{\partial}{\partial t}\right) = \frac{x}{|x|} ,$$

$$dl(\vec{T}) = t\, dR(\vec{T})$$

where R is the radial projection defined earlier. Hence the Jacobian of l at $\vec{Z}$ is

$$\langle \wedge_3 l , \vec{Z}(t , x)\rangle = t\langle \wedge_2 R , \vec{T}\rangle ,$$

and

$$\begin{aligned}\mathbf{M}(A) &= \int \langle \wedge_3 l, \overrightarrow{Z}(t,x)\rangle \, d\|Z \llcorner Y_0\|(t,x) \\ &\leq \int_0^1 \left(\int t\langle \wedge_2 R, \overrightarrow{T}(x)\rangle \, d\|T_\lambda\|(x) \right) dt \leq c\lambda^\varepsilon .\end{aligned}$$

The last inequality is obtained again by using (98) again. This proves the first part of the theorem. In order to prove the second part, we let $V_b = \{x : \operatorname{dist}(x, \operatorname{spt} T_0) \leq b|x|^{1+\varepsilon/2}\}$. Using elementary slicing theory and the estimates on the area of the current A, we observe that for any small $\lambda > 0$, there is a $b_\lambda \in [1/2, 1]$ such that

$$\mathscr{F}(T_\lambda \llcorner V_{b_\lambda}, T_0) \leq c'\lambda^{\varepsilon/2}.$$

Thus the sequence $T_\lambda \llcorner V_{b_\lambda}$ converges to T_0 under the $\mathscr{F}$ norm. If there is $p \in \operatorname{spt} T_\lambda \cap A(1/2, 1)$ such that

$$\operatorname{dist}(p, \operatorname{spt} T_0) \geq \lambda^{\varepsilon/4},$$

then according to monotonicity of area minimizing currents,

$$\mathbf{M}(T_\lambda \llcorner B(p, \tfrac{1}{2}\lambda^{\varepsilon/4})) \geq C(M)\lambda^{\varepsilon/2}$$

with $C(M)$ depending just on the ambient manifold M. Hence we have

$$\begin{aligned}\mathbf{M}(T_\lambda) &\geq \mathbf{M}(T_\lambda \llcorner V_{b_\lambda}) + \mathbf{M}(T_\lambda \llcorner B(p, \tfrac{1}{2}\lambda^{\varepsilon/4}) \\ &\geq \mathbf{M}(T_0) + C(M)\lambda^{\varepsilon/2}. \qquad (100)\end{aligned}$$

Similar to the proof of the last theorem, we use B_λ constructed from (99) (where we did not use subindices) to construct

$$\Phi_{\lambda\sharp} P_\sharp (T_0 - B_\lambda)$$

with Φ_λ, P the same as in the proof of Theorem 13.1. We observe that this current is supported on M_λ and have the same boundary as T_λ, hence

$$\begin{aligned}\mathbf{M}(T_\lambda) &\leq \mathbf{M}(\Psi_{\lambda\sharp} P_\sharp (T_0 - B_\lambda)) \\ &\leq (1+\lambda)(\mathbf{M}(T_0) + \mathbf{M}(B_\lambda)) \\ &\leq (1+\lambda)(\mathbf{M}(T_0) + c'\lambda^\varepsilon)\end{aligned}$$

which is not compatible with (100) when λ is sufficiently small. Thus we have

$$\operatorname{dist}(\operatorname{spt} T_0, \operatorname{spt} T_\lambda \cap A(1/2, 1)) \leq \lambda^{\varepsilon/4}.$$

This is enough to conclude the second half of the theorem after scaling the space by λ.

Remark. If we estimate more carefully, we can prove the rate of convergence is actually $\varepsilon/2$.

Once we know that the tangent cone of any area minimizing current on a Riemannian manifold is unique and the tangent cone is simply the union of planes with multiplicity, we conclude that

$$\Theta(\|T\|, 0) = \mathbf{M}(T_0) = \lim_{\to 0} \frac{\mathbf{M}(T \llcorner B_r^{2+m+n})}{r^2} = \lim_{r \to 0} \Theta(\|T\|, r)$$

is a positive integer by using the lower semi-continuity of mass under the flat norm convergence and the fact that T_λ is area minimizing on M_λ.

14. Appendix B

In this section, we present some important results of F. Almgren, in particular the result of representing area minimizing integral currents by multiple-valued functions under certain assumptions on the local properties of the currents. We include those which are necessary for the construction of the center manifold. All these results are proved in [A] which is certainly one of the longest papers ever written in mathematics.

Before stating those results, we fix some notations.

Let ϕ^*, M^*, and T^* be as before with the obvious meanings assigned. We have the following bounds on the function ϕ^*,

$$\begin{aligned} &|\phi^*(x)| \leq l^2|x|^2, \\ (101)\qquad &|D\phi^*(x)| \leq l|x|, \\ &|D^j\phi^*(x)| \leq l^{j-1}, \qquad j = 1, \dots, 5. \end{aligned}$$

$B(\pi, p, t)$ is the disc of radius t with center p in the 2-plane π passing p.

$B(\pi, p, t, h)$ is the cylinder with base $B(\pi, p, t)$ and height h in the normal direction of π, here we allow $h = \infty$.

$K(x, r)$ denotes the square in $R^2 \times \{0\} \subset T_0M \subset R^{2+m+n}$ with sides parallel to a fixed coordinate system, center $x \in R^2$, and side length r.

$R_i = R_0/2^i$.

P_π denotes the orthogonal projection of R^{2+m+n} onto the plane π. We sometime use π_0 to denote the plane $R^2 \times \{0\}$ which is the same as $\operatorname{spt} T_0$.

There exist constants $\theta_0, b_0 > 0$ such that if we have a function

$$g: B(\pi, p, 3R_k) \to R_\pi^{m+n}$$

with $|\pi - \pi_0| \leq \theta_0$ and $|g| \leq b_0 R_k$, $|Dg| \leq b_0$, we can always find

$$g^*: K(x, 2R_k) \to R^{m+n}$$

such that

$$\operatorname{graph}(g^*) = K(x, 2R_k) \times R^{m+n} \cap \operatorname{graph}(g).$$

We define the excess of a two dimensional current S with respect to a two plane $\pi \in \mathbf{G}_2^{2+m+n}$ by

$$EXC(S, \pi) = \int |P_\pi - P_{\vec{S}}|^2 \, d\|S\|.$$

Theorem 14.1. *Let* $S = T^* \llcorner B(\pi, p, 4r) \times R^{m+n}$, *and* δ, β *are certain small constants. For* $r > 0$, *assume* S *satisfies the following:*

(1) $\operatorname{spt} S \subset B(\pi, p, 4r, l^\beta r^{1+\beta})$, $\operatorname{spt} \partial S \subset \partial B(\pi, p, 4r) \times R^{m+n}$.

(2) *The excess of* S *with respect to the plane* π *satisfies*

$$EXC(S, \pi) \le l^{2-\delta} r^{4-\delta}$$

(3) $P_{\pi\sharp} S = JB(\pi, p, 4r)$,

(4) $\mathbf{M}(S) - J\mathbf{M}(B(\pi, p, 4r)) \le c$.

Then there is a multiple-valued function

$$f: B(\pi, p, 3r) \to Q_J(R^{m+n})$$

and constants C, σ, $\gamma > 0$ *such that the following are true:*

(a) $B = B(\pi, p, 3r)$ *can be decomposed into* C^1, C^2, $B = C^1 \cup C^2$, $C^1 \cap C^2 = \emptyset$. *Over* C^1, *the graph of* f *coincides with the slice of the integral current in the normal direction, i.e.* $f(x) = \langle S, P, x\rangle$.

(b) *The set* C^2 *and* $S^2 = S \llcorner B^2 \times R^{m+n}$ *satisfy :*

$$\mathscr{L}^2(C^2), \mathbf{M}(S^2) \le C l^{2+\sigma} r^{4+\sigma}.$$

(c) *The function* f *has the following bound on its Lipschitz constant*

$$\operatorname{Lip}(f) \le (lr)^\gamma.$$

This is one of the major results in [A]. The readers are referred to Chapter 3 in [A] for the entire proof.

For the construction of the center manifold mentioned before, the following theorem is important:

Theorem 14.2. *Let* f *be the function obtained in Theorem* 14.1, *then there is a single valued function* $g: B(\pi, p, 3r) \to R^{m+n}$ *such that*

(a) *There is a constant* $C > 0$ *such that for* $x \in B(\pi, p3r)$,

$$|g(x)| \le C l^\beta r^{1+\beta},$$

$$|Dg(x)| \le C l^{1-\delta/2} r^{1-\delta/2}. \tag{102}$$

(b) *For any measurable set* $U \subset B(\pi, p, 3r)$,

$$\int_U |\eta \circ f - g| \, d\mathscr{L}^2 \le l \int_U |f|^2 \, d\mathscr{L}^2 + C l^\beta r^{2+\sigma}.$$

(c) *With the assumption on* π *as before, one can define* $g^*: K(x, 2r) \subset T_0 M \to R^{m+n}$ *such that* $\operatorname{graph}(g^*) = K(x, 2r) \times R^{m+n} \cap \operatorname{graph}(g)$ *where* $x = P(p)$. *Also we define*

$$H: K(x, 2r) \to R^{m+n}, \qquad H(x) = \phi^*(P_{\pi_0}(g^*(x))).$$

One will see in the next section that this is the mean tool in the construction of the center manifold. As to the higher derivative bounds we know the following result from [A]:

Lemma 14.1. *Let* $S_k = T \llcorner B(\pi, p, 4R_k) \times R^{m+n}$ *, if* S_k *satisfies the same hypothesis as in the Theorem* 14.1 *with* r *there replaced by* R_k *, for* $k = 1, \dots, N$ *, let* g_k *be the function in Theorem* 14.2 *corresponding to* f_k *, then there are the following bounds on the functions and the derivatives:*

$$
\begin{aligned}
|g_k(x)| &\leq Cl^\beta R_k^{1+\beta}, \\
|Dg_k(x)| &\leq Cl^\beta R_k^{1-\delta/2}, \\
|D^i g_k(x)| &\leq Cl^\beta
\end{aligned}
$$

for $x \in B(\pi, p, 3R_k)$, $i = 1, \dots, 4$, *and* $k = 1, \dots, N$.

Remark. This lemma is the key step in [A] to prove the $C^{3,\alpha}$ condition on the center manifold. The function g is essentially a harmonic function plus a correction term related to the mean curvature of the ambient manifold. To bound the high derivatives of harmonic functions over small domains is impossible without using additional information. The idea here is that if g is associated with f which approximates a small part of an area minimizing integral current, then it is possible to bound the high derivatives as we can always control the second derivatives of a harmonic function defined over a large domain.

The following lemma gives estimates of the differences between two functions which are obtained as in Theorem 14.2 and are defined over nearby squares.

Lemma 14.2. *Let* S_k, f_k, g_k *be the same as in Lemma* 14.1, *and let*

$$
S_k' = T^* \llcorner B(\pi', p', 4R_k) \times R^{m+n}, \qquad k = 1, \dots, N'
$$

with $|N - N'| \leq 1$, *and they satisfy the same conditions on the height of the support and the excess. Let* f_k' *and* g_k' *related to them as in Lemma* 14.1.

Let $x = P_{\pi_0}(p)$, $x' = P_\pi(p')$.

Assume $|x - x'| \leq 2R_N$ *and* $|\pi - \pi'| \leq l^\beta$.

Let

$$
\begin{aligned}
g^* &: K(x, 2R_N) \to R^{m+n}, \\
g'^* &: K(x', 2R_{N'}) \to R^{m+n}
\end{aligned}
$$

be associated with g_N *and* $g_{N'}$ *as in Theorem* 14.2, *then the following hold:*

$$
\begin{aligned}
|g^*(x) - g'^*(x)| &\leq Cl^\beta R_N^{3+\sigma}, \\
|D(g^*(x) - g'^*(x)| &\leq Cl^\beta R_N^{2+\sigma}, \\
|D^2(g^*(x) - g'^*(x))| &\leq Cl^\beta R_N^{1+\sigma}, \\
|D^3(g^*(x) - g'^*(x))| &\leq Cl^\beta R_N^{\sigma}, \\
|D^4(g^*(x) - g'^*(x))| &\leq Cl^\beta R_N^{\sigma}.
\end{aligned}
$$

The proof of this lemma is in 4.17 in [A] with slight modification, since we have a stronger conditions on the height and excess estimates. After going through the machinery there one can readily verify that we have a slightly stronger estimate on those bounds. Namely in the estimates, the small constant l is present which is necessary for our purpose later on.

15. Appendix C

From §4 to §12, we described how to get Y_i from N_{i-1} if the current does not coincide with N_{i-1} near 0. In this section we give a description of how to get N_i from Y_i for $i = 1, \dots, k$. That will complete the constructions of two sequences of branched discs as promised in the Introduction.

Let T, M, etc. be the same as in §3. There it is shown that we can always work just with locally irreducible current. From Appendix A, we know that the support of the tangent cone of T at 0 is a plane which we call the tangent plane as we do for smooth surfaces and is identified with $R^2 \times \{0\} \subset T_0 M$.

We assume from now on that Y is either the tangent plane itself or is an admissible surface defined over a neighborhood of 0 in the tangent plane. Let W be an admissible neighborhood of Y in R^{2+m+n}, then over any simply connected region L in $B(0, r_0)$ excluding 0, Y is the union of Q pieces of surfaces and W is the union of Q corresponding neighborhoods, hence the current T restricted over L is the sum of Q currents which do not intersect each other.

Theorem 15.1. *Assume* $\operatorname{spt}(T \llcorner B_{r_0} \times R^{m+n}) \subset W \cup \{0\}$ *for some* $r_0 > 0$, *then there is a* (*branched*) *center manifold* N *which is an admissible surface in the sense of Definition* 3.7. *There is* $0 < r_1 < r_0$ *such that*

$$N \cap B_{r_1} \times R^{m+n} \subset W \cup \{0\}.$$

Also there is a function

$$f_0 \colon N \to Q_J(R^{m+n}).$$

If we put the conformal structure on N *as described in* §3, *and let* N_r, B_r, F, *and* f *associated with* f_0 *as in Definition* 3.10, *then we have that all the assumptions in* §4 *are satisfied, in particular hypotheses* (H1), (H2), (H3), *and* (H4) *are satisfied.*

Remark. This is essentially Theorem A in §1.

The construction of the center manifold N here is a modified version of the construction in [A]. Since the original construction is very long (more than 700 pages), it is impossible to write down the whole construction here. Instead, we give a very brief description and point out the necessary modification. We refer the readers to [A] to find how the estimates are obtained.

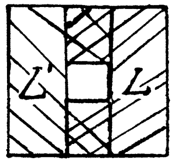

FIGURE 1

The description of the construction. The construction of the branched center manifold involves the following steps:

(1) First write $K(0,1) = \bigcup_\alpha A_\alpha$, $\alpha \in \mathbf{Z}^+$ where

$$A_\alpha = K\left(0, \frac{1}{2^\alpha}\right) \sim K\left(0, \frac{1}{2^{\alpha+2}}\right),$$

with $\alpha \in \mathbf{Z}^+$. Then let $A_\alpha = L \cup L'$ as shown in Figure 1. Over the simply connected region and connected region L, Y is decomposed into Q pieces and the admissible neighborhood W is also decomposed into Q disjoint neighborhoods, hence the current is decomposed into Q pieces whose supports are disjoint.

(2) Construct a center manifold for each piece as similarly done in [A], we shall give more details on the construction later.

(3) Estimate the derivatives of the center manifold.

(4) Patch together those center manifolds over L and L' to get a connected surface which is Q sheet coverings of A_α.

(5) Patch together those surfaces obtained from step 3 over different A_α to get a single branched surface over $K(0,1)$.

(6) Verify those hypotheses (H1), (H2), (H3), and (H4).

Next we describe how to achieve step 2 in the above description. We shall explain that step 4 and step 5 can be done without too much work. Step 3 follows from the estimates in [A] with some modifications. Step 6 is achieved by applying the estimates in [A] to the center manifold(s) built over the connected and simply connected regions L (L') and then patch the results together to get the more global estimates.

Throughout this section we let

$$\lambda = 1/2^\alpha. \tag{103}$$

Let $L^2 = L \cap K(0, \frac{1}{2})$, $L^1 = L \sim L^2$ be the inner half and outer half of L respectively.

Take A_α as in step 1. We normalize it by dilating the whole space by a factor of $\lambda^{-1} = 2^\alpha$.

Let

$$M_\alpha = \mu(\lambda)_\sharp M \cap B(0,3) \times R^{m+n},$$

$$\phi_\alpha = \mu(\lambda^{-1}) \circ \phi \circ \mu(\lambda) : B^{2+m}(0,3) \subset T_0 M \to R^n$$

such that $\operatorname{graph}(\phi_\alpha) = M_\alpha$ and ϕ_α satisfies the following estimates:

$$\begin{aligned} |\phi_\alpha(x)| &\le \lambda^2|x|^2, \\ |D\phi_\alpha(x)| &\le \lambda|x|, \\ |D^i\phi_\alpha(x)| &\le \lambda^{i-1}, \qquad i = 2, 3, 4, 5. \end{aligned}$$

Let $l = \lambda^a = (1/2^\alpha)^a$ for a small constant $a > 0$. Let $\overline{p}$ be the image of the orthogonal projection of $p \in R^{2+m+n}$ into R^2. We say

$$T(p, \pi, k) = T_\alpha \llcorner B(\pi, p, 4R_k, +\infty)$$

satisfying

(1) HT condition if

$$\operatorname{spt} T(p, \pi, k) \subset B(\pi, p, 4R_k, c_p l^\beta R_k^{1+\beta}).$$

(2) EX condition if $EXC(T(p, \pi, k), \pi) \le c'_p l^{2-\delta} R_k^{4-\delta}$
where $c_p = c'_p = 1$ if $\overline{p} \in L^1$ and $c_p = 2^{(a+1)\beta}$, $c'_p = 2^{(a+1)(2-\delta)}$ if $\overline{p} \in L^2$.

In the construction of the center manifold, the main ingredient is the following $\sum$-procedure.

The $\sum$-procedure. Let $H: K(x, 2R_k) \to R^{m+n}$, $\operatorname{graph}(H) \subset M_\alpha$, let $p \in M_\alpha$ and π be a certain plane. Assume that the current $T(p, \pi, k)$ satisfies the HT and EX conditions.

We let $K(x, R_k)$ be the union of four small squares of length R_{k+1}, denoted by $K(x^j, R_{k+1})$, $j = 1, 2, 3, 4$.

Let $p^i = (x^i, H(x^i))$ and π^i be the tangent plane of $\operatorname{graph}(H)$ at p^i.

Assume the Theorem 14.1 is applicable to the currents $T(p^i, \pi^i, k+1)$, $i = 1, \dots, 4$, over $B(\pi^i, p^i, 3R_{k+1})$, then from Theorem 14.2, we get

$$g^i: B(\pi^i, p^i, 3R_{k+1}) \to R^{m+n}_{\pi^i},$$

hence we get

$$g^{i*}, H^i: K(x^i, 2R_{k+1}) \to R^{m+n}.$$

Thus we get four functions H^i whose graphs lie in the ambient manifold M_α.

The construction of the center manifold as mentioned in step 2. In the following construction, if Theorems 14.1 and 14.2 are applicable to $T(\pi, p, k)$, then they are automatically applicable to the four induced currents. This follows from estimates in 4.28 of [A] by using the derivative estimates.

Part 1. As mentioned earlier, we dilate the whole space by $\lambda = 2^\alpha$ when working with A_α. Under this dilation, we have T_α and M_α. Let $K(1) \sim K(1/2^2)$, which is the image of A_α under dilation, be $L \cup L'$ (here we do not use extra notations for the dilated ones). Over L, $Y_\alpha \llcorner L \times R^{m+n} = \bigcup_{i=1}^Q Y^i$, in case Y is a branched surface. The integral current T_α is decomposed into Q pieces

$\{T^i\}$, since T is in an admissible neighborhood of the branched surface Y, and we take one of them. If the original Y is the tangent plane T_0, we take just the restriction of T_k above the region L.

In case Y is the tangent plane, we have the estimates

$$\operatorname{dist}(\operatorname{spt} T_\alpha, \operatorname{spt} T_0) \le \lambda^\varepsilon, \qquad \varepsilon > 0,$$

which implies that

$$EXC(T_\alpha, \pi_0) \le \lambda^{2\varepsilon}$$

(see [LS]). In case Y is a branched surface,

$$\operatorname{dist}(\operatorname{spt} T^i, Y^i) \le \lambda^{N-1+\varepsilon}.$$

The first assertion follows from the estimates in Appendix A, the second case from §12.

Also it follows from there that the function $w: L \to R^{m+n}$ such that $\operatorname{graph}(w) = Y^i$ has the following bounds:

$$|w(x)| \le \lambda^\varepsilon,$$
$$|D^j w(x)| \le \lambda^\varepsilon, \qquad j = 1, \dots, 3. \tag{104}$$

Part 2. Let $\mathscr{K}_1$ be a set of squares of side length R_0 such that

$$L = \bigcup_{K \in \mathscr{K}_1} K,$$

and K, K' in this decomposition do not intersect each other in the interior. This gives L the structure of a chess board.

Part 3. We apply Theorems 14.1 and 14.2 to the area minimizing current $T(x, \pi_0, 4R_0)$. $T(x, \pi_0, 4R_0)$ satisfies the HT and EX conditions because of the estimates in Part 1 on the distance between the current and the tangent plane. Applying Theorem 14.2, we get a function $H: K(x, 2R_0) \to R^{m+n}$ with $\operatorname{graph}(H) \subset M_\alpha$ for each $K \in \mathscr{K}_1$.

Recall $L^2 = L \cap K(0, \frac{1}{2})$, $L^1 = L \sim L^2$ are the inner half and outer half of L respectively.

We apply the $\sum$-procedure to all the squares in $\mathscr{K}_1 \subset L^1$ to obtain

$$\mathscr{K}_2' = \{H(K) | K \in \mathscr{K}_2\}$$

where $\mathscr{K}_2$ is the collection of squares obtained by dividing every K in $\mathscr{K}_1$ into four small squares of side length R_2.

For $K \in \mathscr{K}_1$, $K \subset L^2$, we divide K into four small squares K^i, $i = 1, \dots, 4$, as before but instead of applying Theorems 14.1 and 14.2 to $T(p^i, \pi^i, 4R_2)$, we apply those theorems to $T(p^i, \pi_0, 4R_2)$. So we obtain $\mathscr{K}_2'' = \{H(K) | K \in \mathscr{K}_2, K \subset L^2\}$ where $H(K)$ corresponds to $T(p^i, \pi_0, 4R_2)$ as in Theorem 14.2.

This difference in the construction is used later to patch the constructed surfaces over different A_α.

Let $\mathscr{K}_2 = \mathscr{K}_2' \cup \mathscr{K}_2''$. Let $\mathscr{L}_2$ be the set of those K in $\mathscr{K}_2$ such that $T(K)$ fails to satisfy either the HT or EX condition.

Part 4. Let $K \in \mathscr{J}_2 = \mathscr{K}_2 \sim \mathscr{L}_2$, then $T(K) = T(p, \pi, R_2)$ satisfies the condition in the $\sum$-procedure, hence we can apply it to $T(K)$ and get $K = \bigcup K^i$, p^i and the functions $\{H(K^i)\}$ corresponding to $\{T(p^i, \pi^i, 4R_3)\}$.

Let $\mathscr{L}_3'$ be those $K \in \mathscr{K}_3$, $(K - \partial K) \cap L = \varnothing$, $L \in \mathscr{L}_2$ such that $T(K)$ fails either the EX condition or the HT condition and let $\mathscr{L}_3''$ be the set of squares in $\mathscr{K}_3$ which is not a subset of any squares in $\mathscr{L}_2$ but is the neighbor of one of $\mathscr{L}_2$. Then finally let $\mathscr{L}_3 = \mathscr{L}_3' \cup \mathscr{L}_3''$. We can apply the $\sum$-procedure for any squares in $\mathscr{J}_3 = \mathscr{K}_3 \sim \mathscr{L}_3 \sim \mathscr{S}L_3$ where

$$\mathscr{S}L_3 = \{K | K \in \mathscr{K}_3, K \subset L \in \mathscr{L}_1 \cap \mathscr{L}_2\}.$$

Part 5. In general we have the following inductive procedure: Given $\mathscr{L}_i$, $\mathscr{K}_i$, $\mathscr{J}_i$, we subdivide all the squares in $\mathscr{K}_i$ to get a collection of squares of side length R_{i+1} to get $\mathscr{K}_{i+1}$ which gives an even finer chessboard structure for L.

For any $K \in \mathscr{J}_i$, we apply the $\sum$-procedure. Let $\mathscr{P}J_{i+1}$ consists of all the smaller squares generated from the $\sum$-procedure. Let $\mathscr{L}_{i+1}'$ be the set consisting of all the elements in $\mathscr{P}J_{i+1}$ such that $T(K) \in \mathscr{P}J_{i+1}$ fails to satisfy either the HT or EX conditions, and let $\mathscr{L}_{i+1}''$ be the subset of $\mathscr{P}J_{i+1}$ consisting of all the small squares which sit next to some larger square in $\mathscr{L}_i$. $\mathscr{L}_{i+1}$ is the union of $\mathscr{L}_{i+1}'$ and $\mathscr{L}_{i+1}''$. Let

$$\mathscr{J}_{i+1} = \mathscr{P}J_{i+1} \sim \mathscr{L}_{i+1}.$$

So the current $T(K)$ associated with any square K in $\mathscr{J}_{i+1}$ which is of side length R_{i+1} will satisfy both HT and EX conditions, so we apply the $\sum$-procedure again to them and continue the inductive procedure.

Part 6. For any $K \in \mathscr{K}_i$, if it is in $\mathscr{J}_i$ then there is the function $H(K)$ associated with it. If $K = K(x, R_i)$ is a subset of some $K' \in \mathscr{L}_j$, $l < i$, then we simply define

$$H(K) = H(K')|_{K(x, 2R_l)} : \to R^{m+n},$$

so in any case we have a function whose graph lies in M_α associated with it.

Let $\psi(K) : K(x, 2R_l) \to R^+ \cup \{0\}$ with

$$\psi(K)\Big|_K = 1, \qquad \psi(K)\Big|_{\partial K(x, 2R_l)} = 0,$$

$$|D\psi(K)| \le 3R_l^{-l}, \qquad i = 1, \dots, 4.$$

We define $G^i = \sum_{K \in \mathscr{K}_i} \psi(K) H(K)$. From Chapter 4.19 in [A] we know that the limit of G^i exists when i goes to infinity. Furthermore G has the following bounds:

$$|G(x)| \le Cl^a, \qquad |D^l(x)| \le Cl^a.$$

The proof of those two facts are in 4.19 of [A], except here we have a small factor l^a on the right side. It is there because we have a better estimate for the various $H(K)$ since we start with the EX and HT conditions which involve this factor. If we go through all the proofs in [A], this factor will be maintained. The only exception is 4.17 in [A]. If we change the hypothesis $|x-x'| \leq \frac{5}{16}S_N$ $\|1-\theta\| \leq \theta_{4.16}$ by $|x-x'| \leq R_N$, $|1-\theta| \leq l^a$, the same proof goes through.

This finishes steps 2 and 3 in the description of the main construction.

Part 7. The constructions over L and L' agree over a slightly reduced domain in $L \cap L'$, because both are very local constructions. If $K \in \mathscr{K}_1$ lies in $L \cap L'$ we know that the construction over it is the same. No matter whether we work over L or L', the subsequent constructions are the same inside K. So both constructions agree over K. Using the definitions of G^i and $\psi(K)$, we know that if $K \subset L \cap L'$ with

$$\operatorname{dist}(K, \partial(L \cap L')) \leq R_0$$

then the function G^i agrees with the corresponding function constructed over L'. Hence we can patch those surfaces built for different pieces T_i and over L, L' together naturally to get a multi-sheet covering on the annulus region $K(0,1) \sim K(0, \frac{1}{4})$.

Using the nearest point retraction of the ambient manifold M_α similar to 4.30 of [A], we project this multi-sheet covering of $K(0,1) \sim K(0, \frac{1}{4})$ into M_α to get the partly finished center manifold. We denote the surface by N_α^*.

Part 8. Let G_α, $G_{\alpha+1}$ be the surfaces constructed as above over A_α and $A_{\alpha+1}$, then the construction of $\mu(1/2)_\sharp(G_{\alpha+1})$ agrees with G_α over a slightly reduced domain of $\mu(2^\alpha)A_\alpha \cap \mu(1/2)\mu(2^{\alpha+1})$ because the modification in Part 2 for $\mathscr{K}_2$ and the deliberate difference of stopping conditions for squares in the outer half L^1 of L and inner half L^2. This adjustment makes both constructions identical for squares in $\mu(2^\alpha)A_\alpha \cap \mu(1/2)\mu(2^{\alpha+1})A_{\alpha+1}$. Hence as in Part 7, the functions constructed over different annulus regions agree after suitable dilation. Thus those partial center manifolds $\{N_\alpha^*\}$ patch naturally after the dilations. Hence we have N.

Part 9. To complete the proof of Theorem A, we need to verify that the surface constructed stays inside the admissible neighborhood W so that different branches of the surface N^* will not intersect with each other near 0. Before proving it, we observe that if the surface is exactly centered, then N^* automatically stays inside the neighborhood W.

Let $\mathscr{L}_{\mathrm{HT}}$ and $\mathscr{L}_{\mathrm{EX}}$ be the collection of squares K such that $T(K)$ fails to satisfy the HT and EX conditions respectively.

Lemma 15.1. *There is a constant* $C > 0$, *such that for* $K \in \mathscr{L}_{\mathrm{HT}} \cup \mathscr{L}_{\mathrm{EX}}$ *we have*

$$R_K \leq C\lambda^{N-1}.$$

Remark. This lemma implies that if the current coincides with Y, then the surface N coincides with Y as well near 0.

Proof. Let $K' \supset K$, then $K' \notin \mathscr{L}_{\mathrm{HT}} \cup \mathscr{L}_{\mathrm{EX}}$. Let

$$B(K') = B(\pi(K'), p(K'), 4R_{K'}, cl^{\beta} R_{K'}^{1+\beta})$$

and $B_{\infty}(K')$ be the cylinder with the same base and direction as the above one except with ∞ as its height.

Let

$$B(K) = B(\pi(K), p(K), 4R_K, cl^{\beta} R_K^{1+\beta})$$

and $B_{\infty}(K)$ as above.

Let $p + \pi$ denote the plane parallel to π and passing p.

Then

$$\operatorname{spt} T(K') \subset B(K'), \tag{105}$$

$$EXC(T(K'), \pi(K')) \leq cl^{2-\delta} R_{K'}^{4-\delta}. \tag{106}$$

Since

$$\operatorname{dist}(\operatorname{spt} T(K'), Y \cap B(K')) \leq \lambda^{N-1+\varepsilon},$$

and Y satisfies (104), we have

$$\operatorname{dist}(\operatorname{spt} T(K'), (q + \pi') \cap B_{\infty}(K')) \leq \lambda^{N-1+\varepsilon} + c\lambda^{\varepsilon} R_{K'}^2 \tag{107}$$

with $\pi' = T_q Y$, $q \in Y$. This implies

$$\operatorname{dist}(\operatorname{graph} g(K'), (q + \pi') \cap B_{\infty}(K')) \leq \lambda^{N-1+\varepsilon} + c\lambda^{\varepsilon} R_{K'}^2$$

with $g(K')$ from Theorem 14.2, since $g(K')$ is almost the average of $T(K')$ near $\pi(K')$. The above inequality also holds if we replace the function $g(K')$ by the function $H(K')$, following the estimates in 4.28 in [A].

Using the estimates on the second derivative of $H(K')$ and the fact that $p(K) \in \operatorname{graph} H(K')$, we have

$$\begin{aligned} &\operatorname{dist}((\pi(K) + p(K)) \cap B_{\infty}(K'), (q + \pi') \cap B_{\infty}(K')) \\ &\quad \leq 4(\lambda^{N-1+\varepsilon} + c\lambda^{\varepsilon} R_{K'}^2). \end{aligned} \tag{108}$$

Combining (107) and (108), we have

$$\operatorname{dist}(\operatorname{spt} T(K), (\pi(K) + p(K)) \cap B_{\infty}(K)) \leq 8(\lambda^{N-1+\varepsilon} + 4\lambda^{\varepsilon} R_K^2). \tag{109}$$

If $K \in \mathscr{L}_{\mathrm{HT}}$, then immediately

$$\begin{aligned} l^{\beta} R_K^{1+\beta} &\leq \operatorname{dist}(\operatorname{spt} T(K), (\pi(K) + p(K)) \cap B_{\infty}(K)) \\ &\leq 8(\lambda^{N-1+\varepsilon} + 4\lambda^{\varepsilon} R_K^2). \end{aligned} \tag{110}$$

Hence the assertion follows. If $K \in \mathscr{L}_{\mathrm{EX}}$, then

$$EXC(T(K), \pi(K)) \geq c_p l^{2-\delta} R^{4-\delta}.$$

On the other hand (109) implies that

$$EXC(T(K), \pi(K)) \le 8R_K^2(\lambda^{N-1+\varepsilon} + 4\lambda^{\varepsilon} R_K^2). \tag{111}$$

Recalling the definitions of l, λ, we conclude that

$$R_K \le C\lambda^{N-1}.$$

This concludes the lemma.

Now we are ready to prove that $N \cap B_r \times R^{m+n} \subset W \cup \{0\}$.

From the lemma we conclude that $R_K \le C\lambda^{N-1}$ for $K \in \mathscr{L}$ since any $K \in \mathscr{L}$ is either in $\mathscr{L}_{\mathrm{HT}} \cap \mathscr{L}_{\mathrm{EX}}$ or is stopped by the neighboring condition. Using the same estimates as that of 4.33 (24) in [A], we obtain for any $p \in N \cap (K \times R^{m+n})$, $K \in \mathscr{L}_i$ and $q \in NP^{-1}(p)$ the following holds

$$|p - q| \le l^{\beta} R_l^{1+\beta}, \tag{112}$$

$$\begin{aligned} \operatorname{dist}(p, Y \cap (K \times R^{m+n})) &\le \operatorname{dist}(p, q) + \operatorname{dist}(q, Y) \\ &\le l^{\beta} R_K^{1+\beta} + \lambda^{N-1} \\ &\le l^{\beta} \lambda^{(N-1)(1+\beta)} + \lambda^{N-1+\varepsilon}. \end{aligned} \tag{113}$$

Thus we have N in W when λ is small.

Part 10. From 4.26, 4.33 in [A], the multiple-valued function defined over surfaces as constructed from Part 1 through Part 6 is constructed and is proved to have numerous properties. In particular, the function agrees with the slice of the relevant area minimizing current by the nearest point retraction associated with the center manifold for the most part. In our situation, we need to patch the functions in [A] over different pieces of the center manifold together to get a multiple-valued function defined over the whole N. Since those small pieces of center manifolds coincide over their common domains (maybe a slightly reduced one), the nearest point retractions for different center manifolds agree with each other over the common domain. This makes it possible to patch together the functions defined for those small pieces of center manifolds. The only difference is that we extend the multiple-valued function after patching them together, which makes no difference for the estimates.

As for the hypotheses (H1) through (H4), it follows from the similar estimates in 4.33, 5.11, 5.12, and 5.14 in [A]. The necessary modification we need is to apply the estimates in [A] for the surfaces built over different regions, then to assemble them together after suitable dilation. The proper scaling gives the term $r^{-1+\varepsilon}$ in (H1) and $r^{-2+\varepsilon}$ in (H2).

The definition of Dirichlet integral is slightly different from the original definition in [A]. Our definition makes it easy to calculate the integral, especially for those multiple-valued harmonic functions. In order to use the estimates there to get (H1) through (H4), we notice

$$\int_{W'} |D^* f|^2 \, d\mathscr{L}^2 \le \int_{W'} |Df|^2 \mathscr{L}^2 + \int_{W'} |f|^2 \mathscr{L}^2.$$

Here the left side is the Dirichlet integral used in [A]. This implies that we can replace the Dirichlet integrals in the estimates by ours after adding an additional term of the L^2 integral of the function.

REFERENCES

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

[A] F. J. Almgren, **Q** *valued functions minimizing Dirichlet's integral and the regularity of area minimizing rectifiable currents up to codimension two*, preprint.

[C] S. Chang, *The structure of the singularity of multiple-valued functions minimizing Dirichlet integral over two dimensional domain*, preprint.

[F] H. Federer, *Geometric measure theory*, Springer-Verlag, 1969.

[FM] F. Morgan, *On the singular structure of two-dimensional area minimizing surfaces in* R^n, Trans. Amer. Math. Soc. **276** (1983), 137-143.

[LS] L. Simon, *Lectures on geometric measure theory*, Proc. Centre for Mathematical Analysis, vol. 3, Australian National University, 1984.

[LS1] ——, *Asymptotics for a class of nonlinear evolution equations with application to geometric problems*, Ann. of Math. (2) **118** (1983), 525-571.

[S] M. Spivak, *A comprehensive introduction to differential geometry*, vol. 4, Chapter 9, Addendum 1. *Isothermal coordinates*, Publish or Perish, Boston, 1975.

[BW] B. White, *Tangent cones to two-dimensional area minimizing integral currents are unique*, Duke. Math. J. **50** (1983), 143-160.

[Y] S. T. Yau, *Survey on partial differential equations in differential geometry*, Seminar on Differential Geometry (S. T. Yau, ed.), vol. 102, Ann. of Math. Studies.

[GMT] *Geometric measure theory and the calculus of variations*, Proc. Sympos. Pure Math., vol. 44, Amer. Math. Soc., Providence, RI, 1985.

ABSTRACT. Geometric measure theory guarantees the existence of area minimizing integral currents spanning a given boundary or representing a given integral homology class on a compact Riemannian manifold. We study the regularity of such generalized surfaces. We prove that in case the dimension of the area minimizing integral currents is two, then they are classical minimal surfaces. Among the consequences of this regularity result, we know now that any two dimensional integral homology class on a compact Riemannian manifold can be represented by a finite integral linear combination of classical closed minimal surfaces that have only finitely many intersection points.

The result is proved by using the theory of multiple-valued functions developed by F. Almgren in [A]. We extend many important estimates in his paper and extend his construction of center manifolds. We use the branched center manifolds and lowest order term in the multiple-valued functions approximating the area minimizing currents to construct two sequences of branched surfaces near an interior singular point to separate the nearby singularity gradually. The analysis developed in this paper enables us to conclude the generalized surface must coincide with one of the branched surfaces.

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY, CAMBRIDGE, MASSACHUSETTS 02138