SECOND ORDER RECTIFIABILITY OF INTEGRAL VARIFOLDS OF LOCALLY BOUNDED FIRST VARIATION

ULRICH MENNE

ABSTRACT. In this work it is shown that for every integral n varifold in an open subset U of $\mathbb{R}^{n+m}$, $n, m \in \mathbb{N}$, of locally bounded first variation there exists a countable collection C of n dimensional submanifolds of U of class $\mathcal{C}^2$ such that $\mu(U \sim \bigcup C) = 0$ and for each member M of C

$$\vec{\mathbf{H}}_\mu(x) = \vec{\mathbf{H}}_M(x) \quad \text{for } \mu \text{ almost all } x \in M.$$

CONTENTS

INTRODUCTION

First, some definitions will be recalled. Suppose throughout the introduction that $n, m \in \mathbb{N}$ and U is an open subset of $\mathbb{R}^{n+m}$. Using [Sim83, Theorem 11.8] as a definition, μ is an integral n varifold in U if and only if μ is a Radon measure on U and for μ almost all $x \in U$ there exists an approximate tangent plane $T_x\mu \in G(n+m, n)$ with multiplicity $\theta^n(\mu, x) \in \mathbb{N}$ of μ at x, $G(n+m, n)$ denoting the set of n dimensional, unoriented planes in $\mathbb{R}^{n+m}$. The distributional first variation of mass of μ equals

$$(\delta\mu)(\eta) = \textstyle\int \operatorname{div}_\mu \eta \, \mathrm{d}\mu \quad \text{whenever } \eta \in C^1_{\mathrm{c}}(U, \mathbb{R}^{n+m})$$

where $\operatorname{div}_\mu \eta(x)$ is the trace of $D\eta(x)$ with respect to $T_x\mu$. $\|\delta\mu\|$ denotes the total variation measure associated to $\delta\mu$ and μ is said to be of locally bounded first variation if and only if $\|\delta\mu\|$ is a Radon measure, in this case the generalised mean curvature vector $\vec{\mathbf{H}}_\mu(x) \in \mathbb{R}^{n+m}$ can be defined by the requirement

$$\vec{\mathbf{H}}_\mu(x) \bullet v = -\lim_{\varrho \downarrow 0} \frac{(\delta\mu)(\chi_{B_\varrho(x)} v)}{\mu(B_\varrho(x))} \quad \text{for } v \in \mathbb{R}^{n+m}$$

whenever this limit exists for $x \in U$; here $\bullet$ denotes the usual inner product on $\mathbb{R}^{n+m}$. The mean curvature vector $\vec{\mathbf{H}}_\mu(x)$ is perpendicular to $T_x\mu$ at μ almost all

Date: August 28, 2008.
2000 *Mathematics Subject Classification.* Primary 49Q15; Secondary 35J60.
The author acknowledges financial support via the DFG Forschergruppe 469.
AEI publication number. AEI-2008-065.

$x \in U$, as shown by Brakke in [Bra78, 5.8]. From the above definition of an integral n varifold μ one obtains that μ almost all of U is covered by a countable collection of n dimensional submanifolds of U of class $\mathcal{C}^1$. This concept is extended to higher orders of differentiability by adapting a definition of Anzellotti and Serapioni in [AS94] as follows: A rectifiable n varifold μ in U is called countably rectifiable of class $\mathcal{C}^{k,\alpha}$ $[\mathcal{C}^k]$, $k \in \mathbb{N}$, $0 < \alpha \leq 1$, if and only if there exists a countable collection of n dimensional submanifolds of U of class $\mathcal{C}^{k,\alpha}$ $[\mathcal{C}^k]$ covering μ almost all of U. Throughout the introduction this will be abbreviated to $\mathcal{C}^{k,\alpha}$ $[\mathcal{C}^k]$ rectifiability. Note that $\mathcal{C}^{k,1}$ rectifiability and $\mathcal{C}^{k+1}$ rectifiability agree by [Fed69, 3.1.15]. An integral n varifold μ in U of locally bounded first variation which is $\mathcal{C}^2$ rectifiable is said to satisfy (L) if and only if

$$(L) \qquad \vec{\mathbf{H}}_\mu(x) = \vec{\mathbf{H}}_M(x) \quad \text{for } \mu \text{ almost all } x \in M$$

whenever M is an n dimensional submanifold of U of class $\mathcal{C}^2$.

Suppose for the rest of the introduction that μ is an integral n varifold in U. The following two questions will addressed in this work.

Question 1. Which assumptions on $\delta\mu$ imply $\mathcal{C}^{1,\alpha}$ rectifiability of μ, $0 < \alpha \leq 1$?

Question 2. Suppose μ is of locally bounded first variation and μ is $\mathcal{C}^2$ rectifiable. Which conditions on $\delta\mu$ imply (L)?

Among the possible conditions of $\delta\mu$ there are the following integrability conditions. μ is said to satisfy (H_p), $1 \leq p \leq \infty$, if and only if it is of locally bounded first variation, $\vec{\mathbf{H}}_\mu \in L^p_{\mathrm{loc}}(\mu, \mathbb{R}^{n+m})$, and, in case $p > 1$, satisfies

$$(H_p) \qquad (\delta\mu)(\eta) = -\textstyle\int \vec{\mathbf{H}}_\mu \bullet \eta \,\mathrm{d}\mu \quad \text{whenever } \eta \in C^1_{\mathrm{c}}(U, \mathbb{R}^{n+m}).$$

In order to state the related results, the tilt-excess and the height-excess of μ are defined by

$$\operatorname{tiltex}_\mu(x, \varrho, T) = \varrho^{-n} \textstyle\int_{B_\varrho(x)} |T_\xi\mu - T|^2 \,\mathrm{d}\mu(\xi),$$
$$\operatorname{heightex}_\mu(x, \varrho, T) = \varrho^{-n-2} \textstyle\int_{B_\varrho(x)} \operatorname{dist}(\xi - x, T)^2 \,\mathrm{d}\mu(\xi)$$

whenever $x \in \mathbb{R}^{n+m}$, $0 < \varrho < \infty$, $B_\varrho(x) \subset U$, $T \in G(n+m, n)$; here $S \in G(n+m, n)$ is identified with the orthogonal projection of $\mathbb{R}^{n+m}$ onto S and $|\cdot|$ denotes the norm induced by the usual inner product on $\operatorname{Hom}(\mathbb{R}^{n+m}, \mathbb{R}^{n+m})$. Of basic importance is the following theorem due to Brakke.

Theorem 5.7 in [Bra78]. *If μ is satisfies (H_1), then*

$$\operatorname{tiltex}_\mu(x, \varrho, T_x\mu) = o_x(\varrho), \ \operatorname{heightex}_\mu(x, \varrho, T_x\mu) = o_x(\varrho) \quad \text{as } \varrho \downarrow 0$$

for μ almost all $x \in U$.

If (H_1) is replaced by (H_2), then $o_x(\varrho)$ can be replaced by $O_x(\varrho^{2-\varepsilon})$ for every $\varepsilon > 0$. Using the following lemma which is an adaption of [Sch04b, Appendix A], one infers that (H_1) implies $\mathcal{C}^{1,1/2}$ rectifiability and (H_2) implies $\mathcal{C}^{1,\alpha}$ rectifiability for every $0 < \alpha < 1$.

Lemma. *If $0 < \alpha \leq 1$ and μ satisfies* $\operatorname{heightex}_\mu(x, \varrho, T_x\mu) = O_x(\varrho^{2\alpha})$ *for μ almost all $x \in U$, then μ is $\mathcal{C}^{1,\alpha}$ rectifiable.*

In codimension 1 under the condition (H_p), $p > n$, $p \geq 2$, the above questions were completely answered by Schätzle extending earlier results in [Sch01]:

Theorems 4.1, 5.1, 6.1 in [Sch04a]. *If $m = 1$ and μ satisfies (H_p) for some p with $p > n$, $p \geq 2$, then*

$$\operatorname{tiltex}_\mu(x, \varrho, T_x\mu) = O_x(\varrho^2), \ \operatorname{heightex}_\mu(x, \varrho, T_x\mu) = O_x(\varrho^2) \quad \text{as } \varrho \downarrow 0$$

for μ almost all $x \in U$, and μ is $\mathcal{C}^2$ rectifiable and satisfies (L).

In fact, inspecting the proof, one notices that the decay rates imply (L) under weaker conditions on $\delta\mu$. Moreover, Schätzle showed in arbitrary codimension the following equivalence:

Theorems 3.1, 4.1 in [Sch04b]**.** *Suppose μ satisfies (H_2).*
Then the following two conditions are equivalent:
(1) *μ is $\mathcal{C}^2$ rectifiable.*
(2) *For μ almost all $x \in U$*

$$\operatorname{tiltex}_\mu(x,\varrho,T_x\mu) = O_x(\varrho^2), \ \operatorname{heightex}_\mu(x,\varrho,T_x\mu) = O_x(\varrho^2) \quad \text{as } \varrho \downarrow 0.$$

In this case μ satisfies (L).

The two conditions are not equivalent if μ is merely required to satisfy (H_p) for some p with $1 \leq p < \frac{2n}{2+n}$. In fact, in [Men08b, 1.5] the existence of a $\mathcal{C}^2$ rectifiable integral n varifold satisfying (H_p) not having quadratic decay of neither tilt-excess nor height-excess was confirmed.

In case $n = 1$ an answer to the second question was obtained by Leonardi and Masnou:

Theorem 2.1 in [LM08]**.** *If $n = 1$, μ satisfies (H_1) and is $\mathcal{C}^2$ rectifiable, then μ satisfies (L).*

Moreover, in the same paper a partial extension of this result to the case $n \geq 2$ has been obtained by assuming additionally that for $\mathcal{H}^n$ almost all x with $\theta^n(\mu,x) \geq 1$ there exists an n dimensional submanifold M of class $\mathcal{C}^2$ of U and $Q \in \mathbb{N}$ such that $x \in M$ and $\theta^n(\mu,y) = Q$ for $\mathcal{H}^n$ almost all $y \in M$.

In the present work the condition (H_1) is shown to be sufficient to obtain an affirmative answer to both questions (with $\alpha = 1$):

Theorem 3.7. *If μ satisfies (H_1), then μ is $\mathcal{C}^2$ rectifiable and satisfies (L).*

Using the previous theorem of Schätzle, one directly obtains:

Corollary. *If μ satisfies (H_2), then for μ almost all $x \in U$*

$$\operatorname{tiltex}_\mu(x,\varrho,T_x\mu) = O_x(\varrho^2), \ \operatorname{heightex}_\mu(x,\varrho,T_x\mu) = O_x(\varrho^2) \quad \text{as } \varrho \downarrow 0.$$

Also, using the Sobolev Poincaré type inequality relating tilt and height quantities (cf. [Men08c, 2.9, 2.10]), one obtains:

Corollary. *If μ satisfies (H_2), $q = \infty$ if $n = 1$, $1 \leq q < \infty$ if $n = 2$, $q = \frac{2n}{n-2}$ if $n > 2$, then for μ almost all $x \in U$*

$$\limsup_{\varrho \downarrow 0} \varrho^{-2-n/q} \| \operatorname{dist}(\cdot - x, T_x\mu) \|_{L^q(\mu \llcorner B_\varrho(x))} < \infty.$$

In case $n > 2$, the exponent q cannot be replaced by any larger number as it is shown by [Men08b, 1.5]. It is not known to the author if one can take $q = \infty$ in case $n = 2$.

The next parts of the introduction will describe in an informal style the main ideas of proof whereas the more technical issues will be explained in the body of the text. The basic strategy is to cover μ by a countable set of suitably rotated graphs of Lipschitzian functions satisfying a partial differential equation ensuring $\mathcal{C}^2$ rectifiability.

The starting point to do so is the work of Brakke. The above mentioned decay rates of tilt-excess and height-excess, i.e.

$$\operatorname{tiltex}_\mu(x,\varrho,T_x\mu) = o_x(\varrho), \ \operatorname{heightex}_\mu(x,\varrho,T_x\mu) = o_x(\varrho) \quad \text{as } \varrho \downarrow 0$$

for μ almost all $x \in U$, will be crucial, despite the fact that they only correspond to $\mathcal{C}^{1,1/2}$ rectifibility via the above mentioned lemma. One reason for this is that they allow μ to be approximated near μ almost all $x \in U$ for each $0 < \varrho < \infty$ by a Lipschitzian multivalued function f_ϱ in the ball $B_\varrho(x)$ such that the scale invariant measure of the error sets, where the approximation fails, decays like $o_x(\varrho)$ as $\varrho \downarrow 0$; one order higher than a generic set with n density 0 at x would do. This implies that these sets do not affect the limit $\varrho \downarrow 0$ obtained as long as the derivative of the test functions to be rescaled is bounded in L^∞, a fact that has been used by Brakke in his proof of the perpendicularity of mean curvature. Taking into account that the Lipschitz constant of the approximating function can be prescribed to be arbitrarily small (cp. [Sch04a, (D.9)]), one is led, after some calculations, to consider the following model case for the Laplace operator.

Suppose $u : \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitzian, A is $\mathcal{L}^n$ measurable and to each $a \in A$, $0 < \varrho < \infty$ there corresponds a harmonic function $v_{a,\varrho} : B_\varrho(a) \to \mathbb{R}^m$ such that

$$\limsup_{\varrho \downarrow 0} \varrho^{-2-n/p} \| u - v_{a,\varrho} \|_{L^p(B_\varrho(a), \mathbb{R}^m)} < \infty$$

with $1 \leq p < \infty$, $p < n/(n-1)$ if $n > 1$; the approximating functions $v_{a,\varrho}$ being constructable in a straightforward way using $u|\partial B_\varrho(a)$ as boundary values. If $v_{a,\varrho}$ were affine functions, this would immediately yield second order differentiability in L^p by [CZ61, Theorem 5] implying $\mathcal{C}^2$ rectifiability of u by [CZ61, Theorem 9] (see also [Zie89, 3.6–8]); here a function $v : B \to \mathbb{R}^m$, $B \subset \mathbb{R}^n$, is called $\mathcal{C}^2$ rectifiable if and only if there exists a sequence of functions $v_i : \mathbb{R}^n \to \mathbb{R}^m$ of class $\mathcal{C}^2$ such that $\mathcal{L}^n\,(B \sim \{x : v(x) = v_i(x) \text{ for some } i\}) = 0$. Despite the fact that harmonic functions are themselves smooth and satisfy well known a priori estimates, there can exist points $a \in A$ such that there do not exist affine functions $P_{a,\varrho} : \mathbb{R}^n \to \mathbb{R}^m$ with

$$\limsup_{\varrho \downarrow 0} \varrho^{-2-n/p} \| u - P_{a,\varrho} \|_{L^p(B_\varrho(a), \mathbb{R}^m)} < \infty$$

if $n > 1$.[1] To circumvent this difficulty, one considers closed sets A_k, $k \in \mathbb{N}$, such that

$$\varrho^{-n/p} \| u - v_{a,\varrho} \|_{L^p(B_\varrho(a), \mathbb{R}^m)} \leq k \varrho^2 \quad \text{whenever } 0 < \varrho < 1/k$$

and constructs functions $v_k : \mathbb{R}^n \to \mathbb{R}^m$ which agree with u on A_k and satisfy $\Delta v_k \in L^r(\mathcal{L}^n \llcorner V_k, \mathbb{R}^m)$ for some $1 < r < \infty$ and some open neighbourhood V_k of A_k, hence $v_k | V_k \in W^{2,r}_{\mathrm{loc}}(V_k, \mathbb{R}^m)$ and infers second order differentiability of v_k in L^r from [Reš68], hence $\mathcal{C}^2$ rectifiability of $v_k | V_k$ and $u|A$. The functions v_k are constructed by use of the partition of unity with estimates in [Fed69, 3.1.13] from the functions $v_{a,\varrho}$ only using classical local Calderón-Zygmund type a priori estimates, [GT01, 9.4–5].

In reducing the general case to a slight extension of the model case, the following three aspects are important.

(1) The result in the model case is of $\mathcal{L}^n$ almost everywhere type.
(2) The approximating function is multivalued.
(3) The right hand side of the equation has to be estimated in terms of a norm corresponding to a "divergence of a Radon measure", more precisely in a scale invariant norm on the dual of $C^1_c(B_\varrho(a), \mathbb{R}^m)$, as will be explained in 2.14.

Concerning (1), the Lipschitzian approximation of Brakke has to be extended to construct a *single* multivalued function in a ball $B_r(x_0)$ together with estimates an *every* ball $B_\varrho(x)$ contained in $B_r(x_0)$ such that x is an element of a "good" set having full n density at x_0. This greatly contributes to the complexity of the

[1] If $n = 2$, this may be seen by considering the behaviour of the continuous function $u : \mathbb{C} \to \mathbb{R}$ such that $u(re^{i\varphi}) = r^2 (\log r) \cos(2\varphi)$ for $0 < r < \infty$, $\varphi \in \mathbb{R}$ at 0, noting $\Delta u(re^{i\varphi}) = 4\cos(2\varphi)$ for $0 < r < \infty$, $\varphi \in \mathbb{R}$.

estimates involved. Having to consider all points x in a set of positive measure simultanuously, it also rules out the possibility to assume $Du(x) = 0$ for the points x examined which is often useful in order to approximate the minimal surface operator by the Laplace operator. Concerning (2), one can again use Brakke's tilt-excess estimates to control the error occuring when passing to the average of the multivalued function. Concerning (3), first note that for u, as its Lipschitz constant is small, the Euler Lagrange differential operator L_F corresponding to the nonparametric area integrand F and the operator L_G corresponding to a perturbation G of the Dirichlet integrand yield the same distribution. It is now easy to construct functions $v_{a,\varrho}$ with $L_G(v_{a,\varrho}) = 0$ and boundary values $u|\partial B_\varrho(a)$. Also, once an estimate of $\|u - v_{a,\varrho}\|_{L^p(B_\varrho(a),\mathbb{R}^m)}$ is established, the arguments from the model case carry over from Δ to L_G by rather straightforward perturbation arguments. The difficulty lies in establishing the mentioned L^p estimate for $u - v_{a,\varrho}$. Since in the linear case, i.e. $\Delta w = \operatorname{div} f$, $w : \mathbb{R}^n \to \mathbb{R}^m$ and $f : \mathbb{R}^n \to \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ functions of class $\mathcal{C}^\infty$ with compact support, $\|Dw\|_{L^1(\mathcal{L}^n,\operatorname{Hom}(\mathbb{R}^n,\mathbb{R}^m))}$ cannot be controlled by $\|f\|_{L^1(\mathcal{L}^n,\operatorname{Hom}(\mathbb{R}^n,\mathbb{R}^m))}$, a perturbation argument to pass from Δ to L_G seems to be impossible. Instead of this the construction of u is examined to conclude that u inherites decays of a tilt quantity, i.e.

$$\lim_{\varrho \downarrow 0} \varrho^{-n-1} \textstyle\int_{B_\varrho(a)} |Du(x) - Du(a)|^2 \, \mathrm{d}\mathcal{L}^n x = 0$$

for $\mathcal{L}^n$ almost all a in a relevant set A from the tilt-excess decay of the varifold μ. These estimates are used to explicitely estimate the difference to suitably chosen linear differential operators close to Δ and thereby resolving the problem.

Concerning the ideas of proof, it should be finally remarked that the second order differentiability in L^p in fact obtained for u does not imply directly that these differentials satisfy the equation in differentiated form. However, using once more the tilt estimates, one obtains estimates for comparison functions $w_{a,\varrho}$ with $L_G(w_{a,\varrho})$ given by suitable constants and boundary values given by $u|\partial B_\varrho(a)$ as before to establish the desired relation via a blow up argument. This relation then yields the condition (L) for μ.

The notation is taken from [Sim83, GT01, Fed69, Alm00]. With the exception of standard notation like $\mathbb{N}$ denoting the positive integers, $\mathbb{R}$ denoting the real numbers, $B_\varrho(x)$ and $\bar{B}_\varrho(x)$ denoting respectively the open and closed balls centered at x with radius ϱ, as well as Lebesgue and Sobolev spaces, which have already been used without warning, all notation is introduced or recalled before first occurance. Definitions will be denoted by '='. To simplify verification, in case a statement asserts the existence of a constant, small (ε) or large (Γ), depending on certain parameters this number will be referred to by using the number of the statement as index and what is supposed to replace the parameters in the order of their appearance given in brackets, for example $\varepsilon_{1.6}(n, m, q, 2p)$.

The work is organised as follows. Section 1 provides a criterion for second order differentiability in Lebesgue spaces. In Section 2 the approximation by multivalued functions is constructed The results of these two sections are combined in Section 3 to prove the main theorem. In Appendix A Almgren's notation for Q valued, i.e. multivalued, functions is summarised for the convenience of the reader. Finally, in Appendix B a result about Lebesgue points for distributions is included which clarifies some arguments in the main body of the text.

Acknowledgements. The author offers his thanks to Professor Reiner Schätzle for introducing him to elliptic partial differential equations and the questions of geometric measure theory this paper deals with. The author also would like to thank Professor Tom Ilmanen for inviting him to the ETH Zürich in 2006 and several extensive discussions.

1. A criterion for second order differentiability in Lebesgue spaces

In 1.1–1.8 the situation considered is introduced and modifications and simple applications of standard results are given in the precise form needed to prove the main lemma in 1.9. The criterion is then proved in 1.11.

1.1. In this section and partly also in Section 3 the concept of weakly differentiable functions, in particular Sobolev spaces, as introduced e.g. in [GT01] and the concept of distributions as introduced e.g. in [Fed69] will be used. In doing so, the following conventions and abbreviations will be employed.

Suppose $n, m \in \mathbb{N}$. Since some of the results are pointwise, no identification of functions agreeing almost everywhere will be used; instead the conventions will be employed that for any i times weakly differentiable function $u : U \to \mathbb{R}^m$, U an open subset of $\mathbb{R}^n$, $i \in \mathbb{N}_0$, the weak i-th derivative will be denoted by $D^i u$,

$$a \in \operatorname{dmn} D^i u \quad \text{if and only if} \quad \lim_{r \downarrow 0} (\omega_n r^n)^{-1} \textstyle\int_{B_r(a)} D^i u \, \mathrm{d}\mathcal{L}^n \text{ exists}$$

and in this case $D^i u(a)$ equals the limit in question.[2]

Moreover, the following abbreviations will be convenient. For $i \in \mathbb{N}$, and vector spaces V and W denote by $\odot^i(V, W)$ for $i \in \mathbb{N}$ the set of all symmetric multilinear maps from the i fold product of V into W, $\odot^0(V, W) = W$ and let $\odot^i V = \odot^i(V, \mathbb{R})$ for $i \in \mathbb{N}_0$. Suppose $a \in \mathbb{R}^n$, $0 < r < \infty$, H is a finite dimensional Hilbert space, and $f : B_r(a) \to H$ is $\mathcal{L}^n \llcorner B_r(a)$ measurable. Then

$$|f|_{p;a,r} = \|f\|_{L^p(\mathcal{L}^n \llcorner B_r(a), H)} \quad \text{for } 1 \le p \le \infty.$$

Note, concerning the case $f = D^i u$ for some $u : B_r(a) \to \mathbb{R}^m$, $i \in \mathbb{N}$, the Hilbert space norm is given by $i! \, |\phi|^2 = \sum_{s \in \mathcal{S}(n,i)} |\phi(e_{s(1)}, \ldots, e_{s(i)})|^2$ for $\phi \in \odot^i(\mathbb{R}^n, \mathbb{R}^m)$ where $e_1, \ldots, e_n$ denotes an orthonormal base of $\mathbb{R}^n$ and $\mathcal{S}(n, i)$ is the set of all maps from $\{1, \ldots, i\}$ into $\{1, \ldots, n\}$, see [Fed69, 1.10.5]. Suppose U is an open subset of $\mathbb{R}^n$, $a \in \mathbb{R}^n$, $0 < r < \infty$, $B_r(a) \subset U$, and $T \in \mathcal{D}'(U, \mathbb{R}^m)$. Then

$$|T|_{-1,p;a,r} = \sup\{T(\theta) : \theta \in \mathcal{D}(U, \mathbb{R}^m), \operatorname{spt}\theta \subset B_r(a), |D\theta|_{p';a,r} \le 1\}$$

where p' denotes the conjugate exponent to p. If $|T|_{-1,p;a,r} < \infty$, then, in case $p > 1$, T induces an element of

$$\big(W_0^{1,p'}(B_r(a), \mathbb{R}^m)\big)^*$$

by unique continuous extension of $T | W_0^{1,p'}(B_r(a), \mathbb{R}^m)$. In case $p = 1$, a similar assertions holds with $W_0^{1,\infty}(B_r(a), \mathbb{R}^m)$ replaced by its subspace of functions $u : B_r(a) \to \mathbb{R}^m$ whose extension to $\mathbb{R}^n$ by 0 is of class $\mathcal{C}^1$.

1.2. Suppose $n, m \in \mathbb{N}$, the bilinear form $\Upsilon \in \odot^2 \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ is defined by

$$\Upsilon(\sigma, \tau) = \sigma \bullet \tau \quad \text{for } \sigma, \tau \in \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m),$$

$\bullet$ denoting the inner product on $\operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m)$, $F : \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m) \to \mathbb{R}$ is of class $\mathcal{C}^2$, $0 \le \varepsilon < \infty$, and

$$\|D^2 F(\sigma) - \Upsilon\| \le \varepsilon \quad \text{whenever } \sigma \in \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m);$$

here $\|\Psi\|$ denotes for a bilinear form Ψ on $\operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ the smallest nonnegative number M such that $\Psi(\sigma, \tau) \le M|\sigma||\tau|$ for $\sigma, \tau \in \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m)$. $\operatorname{Lip} D^2 F$ will also be computed with respect to $\|\cdot\|$ on $\odot^2 \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ and $|\cdot|$ on $\operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m)$.

[2] $\operatorname{dmn} f$ denotes the domain of the function f and $\omega_n = \mathcal{L}^n(B_1^n(0))$.

To each such F there corresponds the Euler Lagrange differential operator L_F which associates to every $u \in W^{1,1}(U, \mathbb{R}^m)$ for some open subset U of $\mathbb{R}^n$ a distribution $L_F(u)$ in $\mathcal{D}'(U, \mathbb{R}^m)$ defined by

$$L_F(u)(\theta) = -\textstyle\int_U \langle D\theta(x), DF(Du(x)) \rangle \, d\mathcal{L}^n x \quad \text{for } \theta \in \mathcal{D}(U, \mathbb{R}^m);$$

here $\langle v, \psi \rangle := \psi(v)$ for a linear map $\psi : V \to \mathbb{R}$ and $v \in V$. There also occurs the linear function $C_F(\sigma) : \odot^2(\mathbb{R}^n, \mathbb{R}^m) \to \mathbb{R}^m$ which for $\sigma \in \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ is given by

$$\langle \phi, C_F(\sigma) \rangle = \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^m \big\langle (X_i \upsilon_j, X_k \upsilon_l), D^2 F(\sigma) \big\rangle \, (\phi(e_i, e_k) \bullet \upsilon_j) \upsilon_l$$

for $\phi \in \odot^2(\mathbb{R}^n, \mathbb{R}^m)$ where $e_1, \ldots, e_n$ and $X_1, \ldots, X_n$ are dual orthonormal bases of $\mathbb{R}^n$ and $\odot^1 \mathbb{R}^n$, $X\upsilon$ maps $w \in \mathbb{R}^n$ onto $X(w)\upsilon \in \mathbb{R}^m$ for $X \in \odot^1 \mathbb{R}^n$, $\upsilon \in \mathbb{R}^m$, and $\upsilon_1, \ldots, \upsilon_m$ form an orthonormal base of $\mathbb{R}^m$. Hence one obtains by partial integration for $u \in W^{2,1}(U, \mathbb{R}^m)$, $\theta \in \mathcal{D}(U, \mathbb{R}^m)$

$$L_F(u)(\theta) = \textstyle\int_U \theta(x) \bullet \big\langle D^2 u(x), C_F(Du(x)) \big\rangle \, d\mathcal{L}^n x.$$

Similarly, one defines $S : \odot^2(\mathbb{R}^n, \mathbb{R}^m) \to \mathbb{R}^m$ corresponding to the Dirichlet integrand (and therefore to Υ) and obtains $\langle \phi, S \rangle = \sum_{i=1}^n \sum_{k=1}^n \phi(e_i, e_k)$ with $e_1, \ldots, e_n$, and ϕ as in the definition of $C_F(\sigma)$. One may check that with $\kappa = 2^{1/2} nm$

$$|C_F(\sigma)| \leq \kappa \|D^2 F(\sigma)\|, \quad |C_F(\sigma) - S| \leq \kappa \varepsilon,$$
$$|C_F(\sigma) - C_F(\tau)| \leq \kappa \|D^2 F(\sigma) - D^2 F(\tau)\|$$

for $\sigma, \tau \in \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ where $|\cdot|$ denotes the norm associated to the inner product on $\operatorname{Hom}(\odot^2(\mathbb{R}^n, \mathbb{R}^m), \mathbb{R}^m)$.

1.3. **Theorem.** *Suppose $n, m \in \mathbb{N}$, and $1 < q < \infty$, $1 < p < \infty$.*

Then there exists a positive, finite number ε with the following property.

If Υ is as in 1.2, $a \in \mathbb{R}^n$, $0 < r < \infty$,

$$A : \mathbf{B}_r(a) \to \odot^2 \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m) \text{ is } \mathcal{L}^n \llcorner \mathbf{B}_r(a) \text{ measurable},$$
$$\|A(x) - \Upsilon\| \leq \varepsilon \quad \text{whenever } x \in \mathbf{B}_r(a),$$

and $u \in W^{1,q}(\mathbf{B}_r(a), \mathbb{R}^m)$, $T \in \mathcal{D}'(\mathbf{B}_r(a), \mathbb{R}^m)$ satisfy

$$-\textstyle\int_{\mathbf{B}_r(a)} \langle (D\theta(x), Du(x)), A(x) \rangle \, d\mathcal{L}^n x = T(\theta) \quad \text{for } \theta \in \mathcal{D}(\mathbf{B}_r(a), \mathbb{R}^m),$$

then

$$|Du|_{p;a,r/2} \leq \Gamma \big(r^{-n-1+n/p} |u|_{1;a,r} + |T|_{-1,p;a,r} \big)$$

where Γ is a positive, finite number depending only on n, m, and p.

1.4. **Theorem.** *Suppose $n, m \in \mathbb{N}$, and $1 < p < \infty$.*

Then there exists a positive, finite number ε with the following property.

If S is as in 1.2, $a \in \mathbb{R}^n$, $0 < r < \infty$,

$$B : \mathbf{B}_r(a) \to \operatorname{Hom}(\odot^2(\mathbb{R}^n, \mathbb{R}^m), \mathbb{R}^m) \text{ is } \mathcal{L}^n \llcorner \mathbf{B}_r(a) \text{ measurable},$$
$$|B(x) - S| \leq \varepsilon \quad \text{whenever } x \in \mathbf{B}_r(a),$$

and $u \in W^{2,p}(\mathbf{B}_r(a), \mathbb{R}^m)$, $f \in L^p(\mathcal{L}^n \llcorner \mathbf{B}_r(a), \mathbb{R}^m)$ satisfy

$$\big\langle D^2 u(x), B(x) \big\rangle = f(x) \quad \text{for } \mathcal{L}^n \text{ almost all } x \in \mathbf{B}_r(a),$$

then

$$|D^2 u|_{p;a,r/2} \leq \Gamma \big(r^{-2-n+n/p} |u|_{1;a,r} + |f|_{p;a,r} \big)$$

where Γ is a positive, finite number depending only on n, m, and p.

Proof of 1.3 and 1.4. Using some standard modifications, the techniques described in [GT01, 9.4–5] apply. □

1.5. *Remark.* Using the elementary solution constructed in [Fed69, 5.2.13], one verifies with essentially the same proof that 1.3 and 1.4 remain valid if Υ is only required to be strongly elliptic as defined in [Fed69, 5.2.11] provided the constants ε, Γ are allowed to depend additionally on an ellipticity bound of Υ and a bound for $\|\Upsilon\|$. In fact, this remark holds for all results of the present section.

1.6. **Lemma.** *Suppose $n, m \in \mathbb{N}$, and $1 < q < \infty$, $1 < p < \infty$.*

Then there exists a positive, finite number ε with the following property.

If F is related to ε as in 1.2, $a \in \mathbb{R}^n$, $0 < r < \infty$, $u \in W^{1,q}(B_r(a), \mathbb{R}^m)$, and $f \in L^p(\mathcal{L}^n \llcorner B_r(a), \mathbb{R}^m)$ satisfy

$$L_F(u)(\theta) = \textstyle\int_{B_r(a)} \theta(x) \bullet f(x) \,\mathrm{d}\mathcal{L}^n x \quad \text{whenever } \theta \in \mathcal{D}(B_r(a), \mathbb{R}^m),$$

then u is twice weakly differentiable and for every affine function $P : \mathbb{R}^n \to \mathbb{R}^m$ there holds

$$|D^2 u|_{p;a,r/2} \leq \Gamma\big(r^{-2-n+n/p} |u - P|_{1;a,r} + |f|_{p;a,r}\big)$$

where Γ is a positive, finite number depending only on n, m, and p.

Proof. The assertion may be obtained from 1.3 using difference quotients. □

1.7. **Lemma.** *Suppose $n, m \in \mathbb{N}$, and $1 < q \leq p < \infty$.*

Then there exists a positive, finite number ε with the following property.

If F is related to ε as in 1.2, $\operatorname{Lip} D^2 F < \infty$, $a \in \mathbb{R}^n$, $0 < r < \infty$, and $u_i \in W^{1,q}(B_r(a), \mathbb{R}^m)$ with $i \in \{1, 2\}$ satisfy $L_F(u_i) = 0$, then u_i are twice weakly differentiable and for every affine function $P : \mathbb{R}^n \to \mathbb{R}^m$ there holds

$$\begin{aligned} r^{-n/p+1} |D^2(u_2 - u_1)|_{p;a,r/2} \leq \Gamma\big(& r^{-n-1} |u_2 - u_1|_{1;a,r} \\ & + (r^{-n-1} |u_1 - P|_{1;a,r}) \operatorname{Lip}(D^2 F)(r^{-n-1} |u_2 - u_1|_{1;a,r})\big) \end{aligned}$$

where Γ is a positive, finite number depending only on n, m, and p.

Proof. Using an elementary covering argument, it is enough to prove the assertion with $|D^2(u_2 - u_1)|_{p;a,r/2}$ replaced by $|D^2(u_2 - u_1)|_{p;a,r/4}$. For this purpose let $\kappa = 2^{1/2} nm$,

$$\begin{gathered} \varepsilon = \inf\{\varepsilon_{1.6}(n, m, q, 2p), \varepsilon_{1.4}(n, m, p)/\kappa, \varepsilon_{1.3}(n, m, q, 2p)\}, \\ \Gamma_1 = \Gamma_{1.6}(n, m, 2p), \quad \Gamma_2 = \Gamma_{1.4}(n, m, p), \\ \Gamma_3 = \Gamma_{1.3}(n, m, 2p), \quad \Gamma = \Gamma_2 \sup\{2^{2+n-n/p}, \kappa \Gamma_1 \Gamma_3\}. \end{gathered}$$

Suppose F, a, r, and u_i satisfy the hypotheses with ε and that $P : \mathbb{R}^n \to \mathbb{R}^m$ is an affine function. In order to show that they satisfy the modified conclusions with Γ, it will be assumed $a = 0$ and $r = 1$. Abbreviate $\Lambda = \operatorname{Lip}(D^2 F)$.

By 1.6 the functions u_i are twice weakly differentiable with

$$|D^2 u_i|_{2p;0,1/2} \leq \Gamma_1 |u_i - P|_{1;0,1} \quad \text{for } i \in \{1, 2\}$$

and one obtains from 1.2 for $\mathcal{L}^n$ almost all $x \in B_1^n(0)$

$$\begin{gathered} \big\langle D^2 u_i(x), C_F(Du_i(x)) \big\rangle = 0 \quad \text{for } i \in \{1, 2\}, \\ \big\langle D^2(u_2 - u_1)(x), C_F(Du_2(x)) \big\rangle = \big\langle D^2 u_1(x), C_F(Du_1(x)) - C_F(Du_2(x)) \big\rangle. \end{gathered}$$

Therefore by 1.4, 1.2 and Hölder's inequality

$$\begin{aligned} |D^2(u_2 - u_1)|_{a,1/4;p} \leq \Gamma_2\big(& 2^{2+m-m/p} |u_2 - u_1|_{0,1/2;1} \\ & + \kappa \Lambda |D^2 u_1|_{2p;0,1/2} |D(u_2 - u_1)|_{2p;0,1/2}\big). \end{aligned}$$

To estimate $|D(u_2 - u_1)|_{2p;0,1/2}$, one computes

$$-\!\!\!\!\!\!\int_{B_1^n(0)} \langle (D\theta(x), D(u_2 - u_1)(x)), A(x)\rangle \, \mathrm{d}\mathcal{L}^n x = 0 \quad \text{for } \theta \in \mathcal{D}(B_1^n(0), \mathbb{R}^m),$$
$$\text{where } A(x) = \textstyle\int_0^1 D^2 F(tDu_2(x) + (1-t)Du_1(x)) \, \mathrm{d}\mathcal{L}^1 t,$$

and obtains from 1.3

$$|D(u_2 - u_1)|_{2p;0,1/2} \leq \Gamma_3 |u_2 - u_1|_{1;0,1}. \qquad \square$$

1.8. **Lemma.** *Suppose H is a Hilbert space with $\dim H = N < \infty$, $k, l \in \mathbb{N}_0$, $l \geq k$, $F : H \to \mathbb{R}$ is of class l, $a \in H$, $0 < \delta < \infty$, and*

$$s = \sup\{\|D^k F(x) - D^k F(a)\| : x \in \bar{B}_r(a)\}.$$

Then there exists $G : H \to \mathbb{R}$ of class l such that

$$D^i G(x) = D^i F(x) \quad \textit{for } x \in \bar{B}_{\delta/2}(a),\ i = 0, \ldots, k,$$
$$\|D^k G(x) - D^k F(a)\| \leq \Gamma s \quad \textit{for } x \in H,$$

$G|H \sim \bar{B}_\delta(a)$ is the restriction of a polynomial function of degree at most k

where Γ is a positive, finite number depending only on N and k.

Proof. Choosing $\varphi \in \mathcal{E}(\mathbb{R}, \mathbb{R})$ with $0 \leq \varphi \leq 1$,

$$\{t : -\infty < t \leq 1/2\} \subset \operatorname{Int}\{t : \varphi(t) = 1\}, \quad \{t : 1 \leq t < \infty\} \subset \operatorname{Int}\{t : \varphi(t) = 0\}$$

one defines $P : H \to \mathbb{R}$, $G : H \to \mathbb{R}$ by

$$P(x) = \textstyle\sum_{i=0}^k \left\langle (x-a)^i / i!, D^i F(a) \right\rangle,$$
$$G(x) = P(x) + \varphi(|x - a|/\delta)(F(x) - P(x))$$

for $x \in H$ and readily estimates $\|D^k G(x) - D^k F(a)\|$ be means of Taylor's formula (cf. [Fed69, 3.1.11]).[3] $\square$

1.9. **Lemma.** *Suppose $n, m \in \mathbb{N}$, $1 \leq p \leq r < \infty$, and $1 < q < \infty$.*

Then there exist a positive, finite number ε, a positive, finite number Γ_1 depending only on n and p, and a positive, finite number Γ_2 depending only on n, m, p, and r with the following property.

If F is related to ε as in 1.2, $\operatorname{Lip} D^2 F < \infty$, $j \in \{0, 1\}$, A is a closed subset of $\mathbb{R}^n$, $u : \mathbb{R}^n \cap \{x : \operatorname{dist}(x, A) < 1\} \to \mathbb{R}^m$ is j times weakly differentiable, $0 \leq \gamma < \infty$, and if for each $a \in A$, $0 < \varrho \leq 1$ there are $v_{a,\varrho} \in W^{1,q}(B_\varrho(a), \mathbb{R}^m)$ and an affine function $P_{a,\varrho} : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$L_F(v_{a,\varrho}) = 0,$$
$$\textstyle\sum_{i=0}^j \varrho^{-n/p+i} |D^i(u - v_{a,\varrho})|_{p;a,\varrho} \leq \gamma \varrho^2, \quad \varrho^{-n/p} |u - P_{a,\varrho}|_{p;a,\varrho} \leq \gamma\varrho$$

then there exists a twice weakly differentiable function $v : \mathbb{R}^n \cap \{x : \operatorname{dist}(x, A) < \frac{1}{36}\} \to \mathbb{R}^m$ with

$$\textstyle\sum_{i=0}^j \varrho^{-n/p+i} |D^i(u - v)|_{p;a,\varrho} \leq \Gamma_1 \gamma \varrho^2,$$
$$\varrho^{-n/r} |D^2 v|_{r;a,\varrho} \leq \Gamma_2 \big(\gamma(1 + \operatorname{Lip}(D^2 F)\gamma)^2 + \varrho^{-n-2} |u - P_{a,2\varrho}|_{1;a,2\varrho}\big)$$

whenever $a \in A$, $0 < \varrho \leq \frac{1}{36}$.

[3] $\operatorname{Int} A$ denotes the interior of A.

Proof. Assume $r \geq q$ and define

$$\varepsilon = \min\{1, \varepsilon_{1.6}(n,m,q,2r), \varepsilon_{1.7}(n,m,q,2r), \varepsilon_{1.6}(n,m,q,r)\}.$$

Suppose F, j, A, u, and γ satisfy the hypotheses with ε and abbreviate $\Lambda = \operatorname{Lip}(D^2 F)$.

By 1.6 and Hölder's inequality

$$\textstyle\sum_{i=0}^{j} |D^i v_{a,\varrho}|_{2r;a,1/2} < \infty, \quad \sum_{i=0}^{j} |D^i u|_{p;a,1/2} < \infty$$

whenever $a \in A$. Therefore taking limits (for example by use of an interpolation inequality similar to [Mor66, Lemma 6.2.2] and weak compactness properties of Sobolev spaces [Mor66, Theorem 3.2.4(e)]) the conclusion can be deduced from the following assertion: *There exist a positive, finite number Γ_1 depending only on n and p, and a positive, finite number Γ_2 depending only on n, m, p and r such that for every $0 < \delta \leq \frac{1}{18}$ there exists a function $v : \mathbb{R}^n \to \mathbb{R}^m$ whose restriction to $\left\{x \in \mathbb{R}^n : \operatorname{dist}(x, A) < \frac{1}{18}\right\}$ is twice weakly differentiable satisfying*

$$\textstyle\sum_{i=0}^{j} \varrho^{-n/p+i} |D^i(u-v)|_{p;a,\varrho} \leq \Gamma_1 \gamma \varrho^2,$$
$$(\varrho/2)^{-n/r} |D^2 v|_{r;a,\varrho/2} \leq \Gamma_2 \big(\gamma(1+\Lambda\gamma)^2 + (\varrho/2)^{-n-2} |u - P_{a,\varrho}|_{1;a,\varrho}\big)$$

whenever $a \in A$, $\delta \leq \varrho \leq \frac{1}{18}$.

Assume $A \neq \emptyset$, let $\Phi = \{\mathbb{R}^n \sim A\} \cup \{B_\delta(a) : a \in A\}$, note $\bigcup \Phi = \mathbb{R}^n$, define

$$h(x) = \tfrac{1}{20} \sup\{\min\{1, \operatorname{dist}(x, \mathbb{R}^n \sim U)\} : U \in \Phi\} \quad \text{for } x \in \mathbb{R}^n,$$

and apply [Fed69, 3.1.13] to obtain a countable subset S of $\mathbb{R}^n$ and functions $\varphi_s : \mathbb{R}^n \to \{t : 0 \leq t \leq 1\}$ of class $\mathcal{C}^\infty$ corresponding to $s \in S$ such that with $S_x = \left\{s \in S : \bar{B}_{10h(x)}(x) \cap \bar{B}_{10h(s)}(s) \neq \emptyset\right\}$ for $x \in \mathbb{R}^n$ and a sequence V_i of positive, finite numbers depending only on n there holds

$$\#S_x \leq (129)^n, \quad \operatorname{spt} \varphi_s \subset \bar{B}_{10h(s)}(s) \quad \text{for } s \in S,$$
$$1/3 \leq h(x)/h(s) \leq 3 \quad \text{for } s \in S_x, \quad |D^i \varphi_s(x)| \leq V_i (h(x))^{-i} \quad \text{for } s \in S,\ i \in \mathbb{N},$$
$$\sum_{s \in S} \varphi_s(y) = \sum_{s \in S_x} \varphi_s(y) = 1, \quad \sum_{s \in S} D^i \varphi_s(y) = \sum_{s \in S_x} D^i \varphi_s(y) = 0 \quad \text{for } i \in \mathbb{N}$$

whenever $x \in \mathbb{R}^n$, $y \in \bar{B}_{10h(x)}(x)$. Note for $x \in \mathbb{R}^n$, $y \in \bar{B}_{10h(x)}(x)$, $s \in S$, $i \in \mathbb{N}$

$$|D^i \varphi_s(y)| \leq V_i (h(y))^{-i} \leq (20)^i V_i (10h(x))^{-i},$$

because $h(x) - h(y) \leq \frac{1}{20}|x-y| \leq \frac{1}{2} h(x)$. Choose $\xi : S \to A$ such that

$$|\xi(s) - s| = \operatorname{dist}(s, A) \quad \text{whenever } s \in S.$$

Note $20h(x) \leq \max\{\operatorname{dist}(x, A), \delta\}$ for $x \in \mathbb{R}^n$ and observe

$$\bar{B}_{20h(x)}(x) \subset \bar{B}_{120h(s)}(\xi(s)), \quad 120h(s) \leq 1$$

whenever $x \in \mathbb{R}^n$, $\operatorname{dist}(x, A) \leq \frac{1}{18}$, $s \in S_x$, because

$$|x - s| \leq 10h(x) + 10h(s) \leq 40h(x) \leq 2\max\{\operatorname{dist}(x, A), \delta\} \leq 1/9,$$
$$|s - \xi(s)| = \operatorname{dist}(s, A) \leq |x - s| + \operatorname{dist}(x, A) \leq 1/6,$$
$$|x - \xi(s)| \leq |x - s| + |s - \xi(s)| \leq 40h(s) + 20h(s) = 60h(s),$$
$$|x - \xi(s)| + 20h(x) \leq 120h(s) \leq 360h(x) \leq 1.$$

Define $R = \bigcup\{S_x : x \in \mathbb{R}^n \text{ and } \operatorname{dist}(x, A) \leq \frac{1}{18}\}$,

$$v_s = v_{\xi(s),120h(s)} \quad \text{and} \quad P_s = P_{\xi(s),120h(s)} \quad \text{for } s \in R$$

and, denoting by v_s' the extension of v_s to $\mathbb{R}^n$ by 0, $v : \mathbb{R}^n \to \mathbb{R}^m$ by

$$v(x) = \sum_{s \in R} \varphi_s(x) v_s'(x) \quad \text{whenever } x \in \mathbb{R}^n.$$

Suppose for the rest of the proof $x \in \mathbb{R}^n$ with $\operatorname{dist}(x, A) \le \frac{1}{18}$ and observe

$$v(y) = \sum_{s \in S_x} \varphi_s(y) v_s(y) \quad \text{whenever } y \in \bar{B}_{10h(x)}(x).$$

The asserted weak differentiability is a consequence of 1.6.

One estimates

$$\begin{aligned} &|D^i(u - v_s)|_{p;x,20h(x)} \le |D^i(u - v_s)|_{p;s,120h(s)} \\ &\le \gamma(120h(s))^{n/p+2-i} \le (18)^{n/p+2} \gamma(20h(x))^{n/p+2-i} \end{aligned}$$

for $i \in \{0, j\}$, $s \in S_x$, hence by Hölder's inequality

$$\text{(I)} \quad \begin{aligned} &(20h(x))^{-n} |u - v_s|_{1;x,20h(x)} \\ &\le \omega_n^{1-1/p} \textstyle\sum_{i=0}^j (20h(x))^{-n/p+i} |D^i(u - v_s)|_{p;x,20h(x)} \le 2c_1 \gamma (20h(x))^2 \end{aligned}$$

for $s \in S_x$ where $c_1 = \omega_n^{1-1/p}(18)^{n/p+2}$. Also

$$\begin{aligned} (20h(x))^{-n} |u - P_s|_{1;x,20h(x)} &\le \omega_n^{1-1/p} (20h(x))^{-n/p} |u - P_s|_{p;\xi(s),120h(s)} \\ &\le c_1 \gamma (20h(x)), \end{aligned}$$

$$\text{(II)} \quad (20h(x))^{-n} |v_s - P_s|_{1;x,20h(x)} \le 3c_1 \gamma (20h(x))$$

for $x \in S_x$. Using

$$v(y) - u(y) = \sum_{s \in S_x} \varphi_s(y)(v_s(y) - u(y)) \quad \text{whenever } y \in \bar{B}_{10h(x)}(x)$$

and the Leibnitz formula, one obtains from (I)

$$\textstyle\sum_{i=0}^j (10h(x))^{-n/p+i} |D^i(u - v)|_{p;x,10h(x)} \le c_2 \gamma (10h(x))^2$$

where $c_2 = \omega_n^{1/p-1} 8c_1 2^{n/p}(1 + 20V_1)(129)^n$.

In case $x \in \bar{B}_\varrho(a)$ for some $a \in A$, $\delta \le \varrho \le \frac{1}{18}$,

$$20h(x) \le \max\{\operatorname{dist}(x, A), \delta\} \le \varrho, \quad \bar{B}_{20h(x)}(x) \subset \bar{B}_{2\varrho}(a)$$

and Vitali's covering theorem yields a countable subset T of $\bar{B}_\varrho(a)$ such that

$$\{\bar{B}_{2h(t)}(t) : t \in T\} \text{ is disjointed}, \quad \bar{B}_\varrho(a) \subset \bigcup \{\bar{B}_{10h(t)}(t) : t \in T\}$$

and one estimates for $i \in \{0, j\}$

$$\begin{aligned} &|D^i(u - v)|_{p;a,\varrho}^p \\ &\le \textstyle\sum_{t \in T} |D^i(u - v)|_{p;t,10h(t)}^p \\ &\le (c_2\gamma)^p \textstyle\sum_{t \in T} (10h(t))^{n+(2-i)p} \\ &= (5^{n/p+2-i} c_2 \gamma)^p \omega_n^{-1-(2-i)p/n} \textstyle\sum_{t \in T} \mathcal{L}^n(\bar{B}_{2h(t)}(t))^{1+(2-i)p/n} \\ &\le (5^{n/p+2-i} c_2 \gamma)^p \omega_n^{-1-(2-i)p/n} \mathcal{L}^n(\bar{B}_{2\varrho}(a))^{1+(2-i)p/n} \\ &= \big((10)^{n/p+2-i} c_2 \gamma\big)^p \varrho^{n+(2-i)p}. \end{aligned}$$

Therefore one obtains for $a \in A$, $\delta \le \varrho \le \frac{1}{18}$, $i \in \{0, j\}$

$$\text{(III)} \quad \varrho^{-n/p+i} |D^i(u - v)|_{p;a,\varrho} \le (10)^{n/p+2} c_2 \gamma \varrho^2$$

and one may take $\Gamma_1 = 2(10)^{n/p+2} c_2$ in the first estimate of the assertion.

According to 1.6 the functions v_s are twice weakly differentiable and satisfy for $s \in S_x$

$$(20h(x))^{-n/(2r)+2} |D^2 v_s|_{2r;x,10h(x)} \leq \Gamma_3 (20h(x))^{-n} |v_s - P_s|_{1;x,20h(x)}$$

where $\Gamma_3 = \Gamma_{1.6}(n,m,2r)$. Combining this with (II) yields

$$\text{(IV)} \qquad (10h(x))^{-n/(2r)+2} |D^2 v_s|_{2r;x,10h(x)} \leq 2^{n/(2r)} 3c_1 \gamma (10h(x))$$

for $s \in S_x$.

Using 1.7, one obtains for $s, t \in S_x$

$$\begin{aligned}(20h(x))^{-n/(2r)+1} |D^2(v_s - v_t)|_{2r;x,10h(x)} &\leq \Gamma_4 \big((20h(x))^{-n-1} |v_s - v_t|_{1;x,20h(x)} \\ &+ \Lambda((20h(x))^{-n-1} |v_s - P_s|_{1;x,20h(x)})((20h(x))^{-n-1} |v_s - v_t|_{1;x,20h(x)}) \big)\end{aligned}$$

where $\Gamma_4 = \Gamma_{1.7}(n,m,2r)$. Since

$$(20h(x))^{-n} |v_s - v_t|_{1;x,20h(x)} \leq 4c_1 \gamma (20h(x))^2$$

by (I), one estimates using (II)

$$(10h(x))^{-n/(2r)} |D^2(v_s - v_t)|_{2r;x,10h(x)} \leq c_3 \gamma (1 + \Lambda\gamma)$$

where $c_3 = \Gamma_4(4c_1 + 3c_1 \max\{4c_1, 1\})$. Using an interpolation inequality (which may be proved similarly to [Mor66, Lemma 6.2.2]), one infers with a positive, finite number Γ_5 depending only n, m, and r

$$\begin{aligned}&\textstyle\sum_{i=0}^{2} (10h(x))^{-n/(2r)+i} |D^i(v_s - v_t)|_{2r;x,10h(x)} \\ &\leq \Gamma_5 \big((10h(x))^{-n/(2r)+2} |D^2(v_s - v_t)|_{2r;x,10h(x)} + (10h(x))^{-n} |v_s - v_t|_{1;x,10h(x)} \big) \\ &\leq \Gamma_5 \big(c_3(1 + \Lambda\gamma) + 2^{n+4} c_1 \big) \gamma (10h(x))^2.\end{aligned}$$

Together this implies for $s, t \in S_x$

$$\textstyle\sum_{i=0}^{2} (10h(x))^{-n/(2r)+i} |D^i(v_s - v_t)|_{2r;x,10h(x)} \leq c_4 \gamma (1 + \Lambda\gamma)(10h(x))^2$$

where $c_4 = \Gamma_5(c_3 + 2^{n+4} c_1)$. Noting $(v - v_s)(y) = \sum_{t \in S_x} \varphi_t(y)(v_t - v_s)(y)$ for $s \in S_x$, $y \in B_{10h(x)}(x)$, one infers using the Leibnitz formula

$$\text{(V)} \qquad (10h(x))^{-n/(2r)+i} |D^i(v - v_s)|_{2r;x,10h(x)} \leq c_5 \gamma (1 + \Lambda\gamma)(10h(x))^2$$

for $s \in S_x$, $i \in \{0,1,2\}$ where $c_5 = 2(1 + 20V_1 + 400V_2) c_4 (129)^n$.

Using 1.2, one defines

$$f(y) = \big\langle D^2 v(y), C_F(Dv(y)) \big\rangle$$

whenever $y \in B_{10h(z)}(z)$ for some $z \in \mathbb{R}^n$ with $\operatorname{dist}(z, A) \leq \frac{1}{18}$ and computes for $s \in S_x$

$$f(y) = \big\langle D^2 v_s(y), C_F(Dv(y)) - C_F(Dv_s(y)) \big\rangle + \big\langle D^2(v - v_s)(y), C_F(Dv(y)) \big\rangle$$

for $\mathcal{L}^n$ almost all $y \in B_{10h(x)}(x)$. Hölder's inequality implies

$$\begin{aligned}|f|_{r;x,10h(x)} &\leq \kappa \Lambda(|D(v - v_s)|_{2r;x,10h(x)} |D^2 v_s|_{2r;x,10h(x)} \\ &+ \kappa (M+1) \omega_n^{1/(2r)} (10h(x))^{n/(2r)} |D^2(v - v_s)|_{2r;x,10h(x)},\end{aligned}$$

hence by (IV) and (V)

$$(10h(x))^{-n/r} |f|_{r;x,10h(x)} \leq c_6 \gamma (1 + \Lambda\gamma)^2$$

where $c_6 = \kappa \omega_n^{1/(2r)} (\max\{c_5 \omega_n^{-1/(2r)}, 1\} 2^{n/(2r)} 3c_1 + (M+1) c_5)$. Similarly but simpler as in the deduction of (III), one obtains for $\delta \leq \varrho \leq \frac{1}{18}$, $a \in A$

$$|f|_{r;a,\varrho} \leq c_6 (10)^{n/r} \gamma (1 + \Lambda\gamma)^2 \varrho^{n/r}$$

and thus, using 1.6 with $\Gamma_6 = \Gamma_{1.6}(n,m,r)$ and (III),

$$\begin{aligned}\varrho^{-n/r}|D^2 v|_{r;a,\varrho/2} &\leq \Gamma_6\big(\varrho^{-n-2}(|u-v|_{1;a,\varrho} + |u-P_{a,\varrho}|_{1;a,\varrho}) + \varrho^{-n/r}|f|_{r;a,\varrho}\big) \\ &\leq c_7\big(\gamma(1+\Lambda\gamma)^2 + \varrho^{-n-2}|u-P_{a,\varrho}|_{1;a,\varrho}\big)\end{aligned}$$

where $c_7 = \Gamma_6(\omega_n^{1-1/p}(10)^{n/p+2}c_2 + c_6(10)^{n/r}+1)$. Therefore one may take $\Gamma_2 = 2^{n/r}c_7$ in the second estimate of the assertion and the proof is completed. □

1.10. *Remark.* In fact, by [CZ61, Theorem 10 (ii)] (see also [Zie89, Lemma 3.7.2]), or by [Men08b, 3.1]

$$\lim_{\varrho\downarrow 0} \varrho^{-2}\textstyle\sum_{i=0}^{j}\varrho^{-n/p+i}|D^i(u-v)|_{p;a,\varrho} = 0$$

for $\mathcal{L}^n$ almost all $a \in A$. Now, Rešetnyak's result in [Reš68] applied to v yields that for $\mathcal{L}^n$ almost all $a \in A$ there exists a polynomial function $Q_a : \mathbb{R}^n \to \mathbb{R}^m$ of degree at most 2 such that

$$\limsup_{\varrho\downarrow 0} \varrho^{-2}\textstyle\sum_{i=0}^{j}\varrho^{-n/p+i}|D^i(u-Q_a)|_{p;a,\varrho} = 0.$$

1.11. **Theorem.** *Suppose $n, m \in \mathbb{N}$, $1 \leq p < \infty$, and $1 < q < \infty$.*

Then there exists a positive, finite number ε with the following property.

If F is related to ε as in 1.2, $\operatorname{Lip} D^2F < \infty$, *$U$ is an open subset of $\mathbb{R}^n$, $j \in \{0,1\}$, $u : U \to \mathbb{R}^m$ is weakly differentiable,*

$$h(a,r) = \inf\Big\{\textstyle\sum_{i=0}^{j} r^{-n/p+i}|D^i(u-v)|_{p;a,r} : v \in W^{1,q}(B_r(a),\mathbb{R}^m) \text{ and } L_F(v) = 0\Big\}$$

whenever $B_r(a) \subset U$ for some $a \in \mathbb{R}^n$, $0 < r < \infty$, and if A denotes the set of all $a \in U$ such that

$$\limsup_{r\downarrow 0} r^{-2}h(a,r) < \infty,$$

then A is a Borel set and for $\mathcal{L}^n$ almost all $a \in A$ there exists a polynomial function $Q_a : \mathbb{R}^n \to \mathbb{R}^m$ with degree at most 2 such that

$$\lim_{r\downarrow 0} r^{-2}\textstyle\sum_{i=0}^{j} r^{-n/p+i}|D^i(u-Q_a)|_{p;a,r} = 0.$$

Proof. In view of 1.6 one may assume $q \geq p$. Let $\varepsilon = \varepsilon_{1.9}(n,m,p,p,q)$. Suppose F, U, j, and u satisfy the hypotheses with ε. Define the open set V by

$$V = \Big\{x \in U : \textstyle\sum_{i=0}^{j}|D^i u|_{p;x,r} < \infty \text{ for some } 0 < r < \operatorname{dist}(x, \mathbb{R}^n \sim U)\Big\}$$

and note $A \subset V$. Denote by D the set of all $v \in W^{1,q}(B_1^n(0),\mathbb{R}^m)$ such that $L_F(v) = 0$ and define

$$W = \{(a,r) \in V \times \mathbb{R} : 0 < r < \operatorname{dist}(a, \mathbb{R}^n \sim V)\}$$

and the continuous map $T : W \to W^{1,1}(B_1^n(0),\mathbb{R}^m)$ by

$$T(a,r)(x) = r^{-1}u(a+rx) \quad \text{whenever } (a,r) \in W,\ x \in B_1^n(0).$$

Since $D \neq \emptyset$ and

$$h(a,r) = r\inf\big\{\textstyle\sum_{i=0}^{j}|D^i(T(a,r)-v)|_{p;0,1} : v \in D\big\} \quad \text{for } (a,r) \in W,$$

h is continuous. Therefore A is a Borel set. Similarly, denoting by D' the set of all affine functions mapping $\mathbb{R}^n$ into $\mathbb{R}^m$ one defines a continuous map $h' : W \to \mathbb{R}$ by

$$h'(a,r) = r\inf\{|T(a,r)-w|_{1;0,1} : w \in D'\} \quad \text{for } (a,r) \in W.$$

By [Reš68]

$$\limsup_{\varrho \downarrow 0} \varrho^{-1} h'(a,\varrho) < \infty \quad \text{for } \mathcal{L}^n \text{ almost all } a \in U.$$

Define

$$C_k = \{x \in V : \operatorname{dist}(x, \mathbb{R}^n \sim V) \geq 1/k\},$$
$$A_k = \big\{a \in C_k : h(a,r) \leq kr^2 \text{ and } h'(a,r) \leq kr \text{ for } 0 < r < 1/k\big\}$$

for $k \in \mathbb{N}$ and observe that the sets A_k are closed and

$$\mathcal{L}^n(A \sim \bigcup\{A_k : k \in \mathbb{N}\}) = 0.$$

Finally, the conclusion is obtained by applying (for each $k \in \mathbb{N}$) 1.9 in conjunction with 1.10 to rescaled versions of u, A_k and a suitable number γ. □

1.12. *Remark.* Instead of using [Reš68], one can also use the functions v occuring in the definition of $h(a,r)$ in a way reminiscent of the familiar harmonic approximation procedure to deduce

$$\limsup_{\varrho \downarrow 0} \varrho^{-1} h'(a,\varrho) < \infty \quad \text{whenever } a \in A.$$

Therefore u could have been required to be merely j times weakly differentiable.

1.13. *Remark.* This theorem generalises even in the case of the Laplace operator similar criterions (see [CZ61, Theorem 5], [Zie89, 3.8.1]) where the functions in the definition of $h(a,r)$ are required to be affine. Also note that in case $n > 1$

$$\limsup_{\varrho \downarrow 0} \varrho^{-2} h(a,\varrho) < \infty$$

does not imply the existence of a function $v \in W^{1,q}(B_r(a), \mathbb{R}^m)$ for some $0 < r < \infty$ with $L_F(v) = 0$ such that

$$\limsup_{\varrho \downarrow 0} \varrho^{-2} \sum_{i=0}^{j} r^{i-n/p} |u - v|_{p;a,\varrho} < \infty;$$

in fact this is a consequence of the example given in the Introduction, because harmonic functions are of class $\mathcal{C}^2$.

2. Approximation of integral varifolds of locally bounded first variation by Q valued functions

In this section Brakke's Lipschitz approximation [Bra78, 5.4] is reexamined along the lines of [Men08c, 1.14] to construct the covering of the varifold by suitably rotated graphs of Lipschitzian functions satisfying certain additional properties in 2.12. Before doing so, some facts about universally measurable sets and a multilayer monotonicity are recalled in 2.1–2.6.

2.1. **Definition.** A subset of a topological space is called *universally measurable* if and only if it is measurable with respect to every Borel measure on that space.

A function between topological spaces is *universally measurable* if and only if every preimage of an open set is universally measurable.

2.2. *Remark.* The corresponding definition for measures defined on Borel families (σ algebras) can found for example in [CV77, III.21].

2.3. *Remark.* If $f : X \to Y$ is a Borel function and A is a universally measurable subset of Y, then $f^{-1}(A)$ is universally measurable as may be verified with the help of [Fed69, 2.1.2].

2.4. *Remark.* The universally measurable sets form a Borel family (σ algebra).

2.5. **Lemma.** *Suppose X is a complete, separable metric space, Y is a Hausdorff topological space, $f : X \to Y$ is continuous, B is a Borel subset of X, and $g : B \to \{t : 0 \leq t \leq \infty\}$ is a Borel function.*

Then $h : Y \to \{t : 0 \leq t \leq \infty\}$ defined by

$$h(y) = \sum_{f^{-1}(\{y\})} g \quad \text{whenever } y \in Y$$

is universally measurable.

Proof. [Fed69, 2.10.10, 2.3.1 (6)] may be adapted by use of [Fed69, 2.2.13, 2.3.3] to obtain the conclusion. □

2.6. **Lemma** (Multilayer monotonicity with variable offset, cf. [Men08c, 1.6])**.** *Suppose $n, m, Q \in \mathbb{N}$, $0 \leq M < \infty$, $\delta > 0$, and $0 \leq s < 1$.*

Then there exists a positive, finite number ε with the following property.

If $X \subset \mathbb{R}^{n+m}$, $T \in G(n+m, n)$, $0 \leq d < \infty$, $0 < r < \infty$, $0 < t < \infty$, $f : X \to \mathbb{R}^{n+m}$,

$$|T(y-x)| \leq s|y-x|, \quad |T(f(y)-f(x))| \leq s|f(y)-f(x)|,$$
$$f(x) - x \in \bar{B}^{n+m}_d(0) \cap T, \quad d \leq Mt, \quad d + t \leq r$$

for $x, y \in X$, μ is an integral n varifold in $\bigcup\{B_r(x) : x \in X\}$ with locally bounded first variation,

$$\textstyle\sum_{x \in X} \theta^n_*(\mu, x) \geq Q - 1 + \delta, \quad \mu(B_r(x)) \leq M\omega_n r^n \quad \text{for } x \in X \cap \operatorname{spt}\mu,$$

and whenever $0 < \varrho < r$, $x \in X \cap \operatorname{spt}\mu$

$$\|\delta\mu\|(\bar{B}_\varrho(x)) \leq \varepsilon\, \mu(\bar{B}_\varrho(x))^{1-1/n}, \quad \textstyle\int_{\bar{B}_\varrho(x)} |T_\xi\mu - T| \,\mathrm{d}\mu(\xi) \leq \varepsilon\, \mu(\bar{B}_\varrho(x)),$$

then

$$\mu\big(\bigcup\big\{B_t(f(x)) \cap \{y : |T(y-x)| > s|y-x|\} : x \in X\big\}\big) \geq (Q-\delta)\omega_n t^n.$$

2.7. **Lemma.** *Suppose X, Y are normed, finite dimensional vector spaces, $f : X \to Y$ is of class $\mathcal{C}^1$, $a \in X$, $0 < r < \infty$, $Q \in \mathbb{N}$, $x_i \in \bar{B}_r(a)$ for $i = 1, \ldots, Q$, and $\lambda = \operatorname{Lip}(Df|\bar{B}_r(a))$.*

Then

$$\left| \frac{1}{Q}\sum_{i=1}^{Q} f(x_i) - f\left(\frac{1}{Q}\sum_{i=1}^{Q} x_i\right)\right| \leq \lambda r^2.$$

Proof. Let $P : X \to Y$ by defined by $P(x) = f(a) + \langle x - a, Df(a)\rangle$ for $x \in X$. Then for $x \in \bar{B}_r(a)$

$$|f(x) - P(x)| = \big|\big\langle x-a, \textstyle\int_0^1 Df(a + t(x-a)) - Df(a)\,\mathrm{d}\mathcal{L}^1 t\big\rangle\big| \leq (\lambda/2)r^2.$$

Since $\frac{1}{Q}\sum_{i=1}^Q P(x_i) = P(Q^{-1}\sum_{i=1}^Q x_i)$, this implies the conclusion. □

2.8. **Definition.** Whenever $k, l \in \mathbb{N}$, $k \geq l$ the set of orthogonal projections $\pi : \mathbb{R}^k \to \mathbb{R}^l$ will be denoted by $O^*(k, l)$.

2.9. Whenever $n, m \in \mathbb{N}$, and $T \in G(n+m, n)$ there exist $\pi \in O^*(n+m, n)$, $\sigma \in O^*(n+m, m)$ such that $T = \operatorname{im}\pi^*$ and $\pi \circ \sigma^* = 0$,[4] hence

$$T = \pi^* \circ \pi, \quad T^\perp = \sigma^* \circ \sigma, \quad \mathbb{1}_{\mathbb{R}^{n+m}} = \pi^* \circ \pi + \sigma^* \circ \sigma.$$

[4] $\operatorname{im} f$ denotes the image of a map f.

Whenever $a \in \mathbb{R}^{n+m}$, $0 < r < \infty$, $0 < h \leq \infty$ the *closed cuboid* $C(T, a, r, h)$ is defined by

$$C(T, a, r, h) = \left\{ x \in \mathbb{R}^{n+m} : |T(x-a)| \leq r \text{ and } |T^\perp(x-a)| \leq h \right\}$$
$$= \left\{ x \in \mathbb{R}^{n+m} : |\pi(x-a)| \leq r \text{ and } |\sigma(x-a)| \leq h \right\}.$$

This definition extends Allard's definition in [All72, 8.10] where $h = \infty$.

2.10. **Lemma** (Approximation by Q valued functions)**.** *Suppose $n, m, Q \in \mathbb{N}$, $0 < L < \infty$, $1 \leq M < \infty$, and $0 < \delta_i \leq 1$ for $i \in \{1, 2, 3, 4\}$.*

Then there exists a positive, finite number ε with the following property.

If a, r, h, T, π, and σ are as in 2.9, $h > 2\delta_4 r$,

$$U = \left\{ x \in \mathbb{R}^{n+m} : \operatorname{dist}(x, C(T, a, r, h)) < 2r \right\},$$

μ is an integral n varifold in U with locally bounded first variation,

$$(Q - 1 + \delta_1)\omega_n r^n \leq \mu(C(T, a, r, h)) \leq (Q + 1 - \delta_2)\omega_n r^n,$$
$$\mu(C(T, a, r, h + \delta_4 r) \sim C(T, a, r, h - 2\delta_4 r)) \leq (1 - \delta_3)\omega_n r^n,$$
$$\mu(U) \leq M\omega_n r^n,$$

$0 < \varepsilon_1 \leq \varepsilon$, B denotes the set of all $x \in C(T, a, r, h)$ with $\theta^{*n}(\mu, x) > 0$ such that

$$\text{either} \quad \|\delta\mu\|(\bar{B}_\varrho(x)) > \varepsilon_1\, \mu(\bar{B}_\varrho(x))^{1-1/n} \quad \text{for some } 0 < \varrho < 2r,$$
$$\text{or} \quad \textstyle\int_{\bar{B}_\varrho(x)} |T_\xi \mu - T| \, d\mu(\xi) > \varepsilon_1\, \mu(\bar{B}_\varrho(x)) \quad \text{for some } 0 < \varrho < 2r,$$

$A = C(T, a, r, h) \sim B$, $A(y) = \{x \in A : \pi(x) = y\}$ for $y \in \mathbb{R}^n$, Y is the set of all $y \in \bar{B}_r(\pi(a))$ such that

$$\textstyle\sum_{x \in A(y)} \theta^n(\mu, x) = Q \quad \text{and} \quad \theta^n(\mu, x) \in \mathbb{N}_0 \text{ for } x \in A(y),$$

Z is the set of all $z \in \bar{B}_r(\pi(a))$ such that

$$\textstyle\sum_{x \in A(z)} \theta^n(\mu, x) \leq Q - 1 \quad \text{and} \quad \theta^n(\mu, x) \in \mathbb{N}_0 \text{ for } x \in A(z),$$

and $N = \bar{B}_r(\pi(a)) \sim (Y \cup Z)$, then the following eight statements hold:

(1) *Y and Z are universally measurable, and $\mathcal{L}^n(N) = 0$.*

(2) *A and B are Borel sets and*

$$\sigma(A \cap \operatorname{spt} \mu) \subset \bar{B}_{h - \delta_4 r}(\sigma(a)).$$

(3) *$\pi(\{x \in A : \theta^n(\mu, x) = Q\}) \subset Y$.*

(4) *A function $f : Y \to Q_Q(\mathbb{R}^m)$ is uniquely characerised by the requirement*

$$\theta^n(\mu, x) = \theta^0(\|f(y)\|, \sigma(x)) \quad \text{whenever } y \in Y \text{ and } x \in A(y).$$

(5) *The function f defined in* (4) *is Lipschitzian with* $\operatorname{Lip} f \leq L$.

(6) *Defining f as in* (4) *and $G = \{x \in \mathbb{R}^{n+m} : \sigma(x) \in \operatorname{spt} f(\pi(x))\}$, for $\mathcal{L}^n$ almost all $y \in Y$ the following is true:*

   (a) *f is approximately strongly affinely approximable at y.*

   (b) *$T_x \mu$ is mapped onto* $\operatorname{Tan}\left(\operatorname{graph}_Q \operatorname{ap} Af(y), (y, \sigma(x))\right)$ *by the isometry $\pi \bowtie \sigma : \mathbb{R}^{n+m} \to \mathbb{R}^n \times \mathbb{R}^m$ whenever $x \in G$ with $\pi(x) = y$ (see A.3).*[5]

(7) *If $b \in A$, $\theta^n(\mu, b) = Q$, $0 < \varrho \leq r - |T(b - a)|$,*

$$B_{b,\varrho} = C(T, b, \varrho, \delta_4 \varrho) \cap B,$$
$$C_{b,\varrho} = \bar{B}_\varrho(\pi(b)) \sim (Y \sim \pi(B_{b,\varrho})),$$
$$D_{b,\varrho} = C(T, b, \varrho, \delta_4 \varrho) \cap \pi^{-1}(C_{b,\varrho}),$$

[5] Here $\operatorname{Tan}(S, a)$ denotes the closed tangent cone of S at a in the sense of [Fed69, 3.1.21].

then $B_{b,\varrho}$ *is a Borel set,* $C_{b,\varrho}$*,* $D_{b,\varrho}$ *are universally measurable and there holds*

$$\mathcal{L}^n(C_{b,\varrho}) + \mu(D_{b,\varrho}) \leq \Gamma_{(7)}\, \mu(B_{b,\varrho})$$

with $\Gamma_{(7)} = 3 + 2Q + (12Q+6)5^n$.

(8) *If* b, ϱ, $C_{b,\varrho}$, $D_{b,\varrho}$ *are as in* (7)*,* $g : \mathbb{R}^n \to \mathbb{R}^m$ *is a Lipschitzian extension of* $\boldsymbol{\eta}_Q \circ f$*,* $\tau \in \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m)$*,* $\theta \in \mathcal{D}(\mathbb{R}^n, \mathbb{R}^m)$*,* $\psi \in \mathcal{D}(\mathbb{R}^m, \mathbb{R})$*,*

$$\operatorname{spt}\theta \subset B_\varrho(\pi(b)), \quad \operatorname{spt}\psi \subset B_{\delta_4\varrho}(\sigma(b)),$$
$$\bar{B}_{(\delta_4/2)\varrho}(\sigma(b)) \subset \operatorname{Int}\{z : \psi(z) = 1\},$$

and $F : \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m) \to \mathbb{R}$ *denotes the nonparametric area integrand, then*

$$\begin{aligned}&\big|Q{\textstyle\int_{B_\varrho(\pi(b))}} \langle D\theta(x), DF(Dg(x))\rangle \, \mathrm{d}\mathcal{L}^n x - (\delta\mu)((\psi\circ\sigma)\cdot(\sigma^*\circ\theta\circ\pi))\big| \\ &\leq \gamma_1 Q n^{1/2} \operatorname{Lip} g {\textstyle\int_{C_{b,\varrho}}} |D\theta| \, \mathrm{d}\mathcal{L}^n + \gamma_2 Q {\textstyle\int_{E_{b,\varrho} \sim C_{b,\varrho}}} |D\theta(x)| t(x,\tau)^2 \, \mathrm{d}\mathcal{L}^n x \\ &\qquad + n^{1/2} {\textstyle\int_{D_{b,\varrho}}} |D((\psi\circ\sigma)\cdot(\sigma^*\circ\theta\circ\pi))| \, \mathrm{d}\mu\end{aligned}$$

where

$$\gamma_1 = \sup \|D^2F\|(\bar{B}_{n^{1/2}\operatorname{Lip} g}(0)),$$
$$\gamma_2 = \operatorname{Lip}\big(D^2F | \bar{B}_{n^{1/2}(L+2\|\tau\|)}(0)\big),$$
$$E_{b,\varrho} = \bar{B}_\varrho(\pi(b)) \cap \big\{y \in Y : \theta^0(\|f(y)\|, g(y)) \neq Q\big\}$$

and $t(x,\tau)$ *is the supremum of all numbers*

$$|\tau_i - \tau|$$

corresponding to all $z_1, \ldots, z_Q \in \mathbb{R}^m$*,* $\tau_1, \ldots, \tau_Q \in \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ *such that*

$$\operatorname{ap} Af(x)(v) = \sum_{i=1}^{Q} [\![z_i + \langle v, \tau_i\rangle]\!] \quad \textit{whenever } v \in \mathbb{R}^n.$$

Proof of (1)–(6). The existence of a number ε with $1 - n\varepsilon^2 \geq 1/2$ such that (1)–(6) are true is essentially proved in [Men08c, 1.14 (1) (2) (7)]; the sets Y and Z are defined as in the proof cited, their universal measurability follows from 2.4 and 2.5, $\mathcal{L}^n(N) = 0$ occurs in the last paragraph of the proof of [(1) and (2), loc. cit.], and (3) is a consequence of a slight modification of the third paragraph of the proof of [(1) and (2), loc. cit.].[6] □

Choice of constants. One can assume $2L \leq \delta_4$. Let ε_0 be a positive, finite number such that ε_0 in place of ε has the property asserted when the last two statements are omitted.

Choose $0 < s < 1$ close to 1 satisfying

$$(s^{-2} - 1)^{1/2} \leq \min\{\delta_4, L\}$$

and define $\varepsilon > 0$ so small that

$$\varepsilon \leq \min\{\varepsilon_0, \varepsilon_{2.6}(n, m, Q, M, 1/4, s)\}, \quad (1 - n\varepsilon^2)(Q - 1/4) \geq Q - 1/2.$$

Clearly, ε_1 satisfies the same inequalities as ε and one can assume $a = 0$, and $r = 1$. □

[6]For $x \in A$ with $\theta^n(\mu, x) = Q$ one defines $\delta = \inf\{\delta_2/2, (2n\gamma_n)^{-n}/\omega_n\}$, $X = A(\pi(x))$ and applies 2.6 (noting [Men08b, 2.5]) with Q, d, r, t, and f replaced by $Q+1$, 1, 2, 1, and $T|X$ to obtain $\sum_{\xi \in A(\pi(x))} \theta^n_*(\mu, \xi) < Q + \delta$ provided $\varepsilon \leq \varepsilon_{2.6}(n, m, Q+1, M, \delta, s)$ and $(s^{-2}-1)^{1/2} \leq \delta_4/2$, hence [Men08b, 2.5] implies (3).

Proof of (7). $\pi(B_{b,\varrho})$ is a universally measurable set by [Fed69, 2.2.13], hence $C_{b,\varrho}$, $D_{b,\varrho}$ are universally measurable sets by 2.3, 2.4. (2) shows

$$\delta_4\varrho + |T^\perp(b)| \leq h, \quad C(T,b,\varrho,\delta_4\varrho) \subset C(T,0,1,h).$$

Let ν denote the Radon measure characterised by

$$\nu(X) = \textstyle\int_X J^\mu T(\xi)\,\mathrm{d}\mu(\xi)$$

whenever X is a Borel subset of U,[7] and note

$$|T_x\mu - T| \leq \varepsilon \quad \text{for } \mu \text{ almost all } x \in A,$$

hence $1 - J^\mu T(x) \leq 1 - (J^\mu T(x))^2 \leq n\varepsilon^2$ for those x. Therefore

$$(1-n\varepsilon^2)\,\mu \llcorner A \leq \nu \llcorner A.$$

This implies the *coarea estimate*

$$\begin{aligned}(1-n\varepsilon^2)\,&\mu\big(C(T,b,\varrho,\delta_4\varrho)\cap\pi^{-1}(W)\big)\\ &\leq \mu\big(B_{b,\varrho}\cap\pi^{-1}(W)\big) + Q\mathcal{L}^n(Y\cap W) + (Q-1)\mathcal{L}^n(Z\cap W)\end{aligned}$$

for every subset W of $\mathbb{R}^n$; in fact the estimate holds for every Borel set by [Fed69, 3.2.22 (3)] and $\pi_\#(\mu\llcorner B)$ is a Radon measure by [Fed69, 2.2.17]. In particular, taking $W = \bar{B}_\varrho(\pi(b))$ yields

$$(1-n\varepsilon^2)\mu(C(T,b,\varrho,\delta_4\varrho)) \leq \mu(B_{b,\varrho}) + Q\omega_n\varrho^n,$$

thus one can assume, since $8Q+6 \leq \Gamma_{(7)}$, that

$$\mu(B_{b,\varrho}) \leq \tfrac{1}{4}\omega_n\varrho^n.$$

Next, it will be shown that this assumption implies

$$\mathcal{L}^n(Y\cap\bar{B}_\varrho(\pi(b))) > 0.$$

Verifying, since $(s^{-2}-1)^{1/2} \leq \delta_4$, that

$$\big\{\xi \in B_\varrho(b) : |T(\xi-b)| > s|\xi-b|\big\} \subset C(T,b,\varrho,\delta_4\varrho),$$

2.6 may be applied with

$$\begin{gathered}\delta,\ X,\ d,\ r,\ t,\text{ and } f \text{ replaced by,}\\ 1/4,\ \{b\},\ 0,\ 2,\ \varrho,\text{ and } \mathbb{1}_{\{b\}}\end{gathered}$$

to obtain

$$\mu(C(T,b,\varrho,\delta_4\varrho)) \geq (Q-1/4)\omega_n\varrho^n.$$

Hence by the coarea estimate with $W = \bar{B}_\varrho(\pi(b))$ it follows

$$\begin{gathered}\begin{aligned}&(Q-1/2)\omega_n\varrho^n\\ &\leq \mu(B_{b,\varrho}) + Q\mathcal{L}^n(Y\cap\bar{B}_\varrho(\pi(b))) + (Q-1)\mathcal{L}^n(Z\cap\bar{B}_\varrho(\pi(b)))\\ &\leq (Q-1/2)\omega_n\varrho^n + \mathcal{L}^n(Y\cap\bar{B}_\varrho(\pi(b))) - \tfrac{1}{4}\mathcal{L}^n(Z\cap\bar{B}_\varrho(\pi(b))),\end{aligned}\\ \mathcal{L}^n(Z\cap\bar{B}_\varrho(\pi(b))) \leq 4\,\mathcal{L}^n(Y\cap\bar{B}_\varrho(\pi(b))), \quad \mathcal{L}^n(Y\cap\bar{B}_\varrho(\pi(b))) > 0.\end{gathered}$$

Next, in order to estimate $\mathcal{L}^n(Z\cap\bar{B}_\varrho(\pi(b)))$, the following assertion will be proved. *If $z \in Z\cap\bar{B}_\varrho(\pi(b))$ and $\theta^n(\mathcal{L}^n\llcorner\mathbb{R}^n\sim Z, z) = 0$, then there exist $\zeta \in \mathbb{R}^n$ and $0 < t < \infty$ with*

$$z \in \bar{B}_t(\zeta) \subset \bar{B}_\varrho(\pi(b)), \quad \mathcal{L}^n(\bar{B}_{5t}(\zeta)) \leq 6\cdot 5^n\,\mu\big(B_{b,\varrho}\cap\pi^{-1}(\bar{B}_t(\zeta))\big).$$

[7]Here $J^\mu T(\xi)$ denotes the Jacobian of T with respect to μ at ξ which can be expressed as $\|\Lambda_n(T|T_\xi\mu)\|$, cf. [Fed69, 3.2.22].

Since $\mathcal{L}^n(Y \cap \bar{B}_\varrho(\pi(b))) > 0$, some element $\bar{B}_t(\zeta)$ of the family of balls

$$\{\bar{B}_{\theta\varrho}((1-\theta)z + \theta\pi(b)) : 0 < \theta \leq 1\}$$

will satisfy

$$z \in \bar{B}_t(\zeta) \subset \bar{B}_\varrho(\pi(b)), \quad 0 < \mathcal{L}^n(Y \cap \bar{B}_t(\zeta)) \leq \tfrac{1}{2}\mathcal{L}^n(Z \cap \bar{B}_t(\zeta)).$$

Hence there exists $y \in Y \cap B_t(\zeta)$. Noting[8] for $\xi \in A(y)$ with $\theta^n(\mu, \xi) > 0$

$$\begin{gathered} B_t(\eta_{\pi^*(y-\zeta),1}(\xi)) \subset \pi^{-1}(\bar{B}_t(\zeta)), \\ |T^\perp(\xi - b)| \leq L|T(\xi - b)| \leq L\varrho \quad \text{by (5)}, \\ (s^{-2} - 1)^{1/2}|T(\kappa - \xi)| \leq L2t \leq 2L\varrho \leq \delta_4\varrho \quad \text{for } \kappa \in \pi^{-1}(\bar{B}_t(\zeta)), \end{gathered}$$

the inclusion

$$\big\{\kappa \in B_t(\eta_{\pi^*(y-\zeta),1}(\xi)) : |T(\kappa - \xi)| > s|\kappa - \xi|\big\} \subset C(T, b, \varrho, \delta_4\varrho) \cap \pi^{-1}(\bar{B}_t(\zeta))$$

is valid and 2.6 can be applied with

$$\begin{gathered} \delta,\ X,\ d,\ r, \text{ and } f \text{ replaced by} \\ 1/4,\ \{\xi \in A(y) : \theta^n(\mu, \xi) > 0\},\ t,\ 2, \text{ and } \eta_{\pi^*(y-\zeta),1} | \{\xi \in A(y) : \theta^n(\mu, \xi) > 0\} \end{gathered}$$

to obtain

$$(Q - 1/4)\omega_n t^n \leq \mu\big(C(T, b, \varrho, \delta_4\varrho) \cap \pi^{-1}(\bar{B}_t(\zeta))\big).$$

The coarea estimate with $W = \bar{B}_t(\zeta)$ now implies

$$\begin{aligned} &(Q - 1/2)\omega_n t^n \\ &\leq \mu\big(B_{b,\varrho} \cap \pi^{-1}(\bar{B}_t(\zeta))\big) + Q\mathcal{L}^n(Y \cap \bar{B}_t(\zeta)) + (Q-1)\mathcal{L}^n(Z \cap \bar{B}_t(\zeta)) \\ &= \mu\big(B_{b,\varrho} \cap \pi^{-1}[\bar{B}_t(\zeta)]\big) + (Q - 1/2)\omega_n t^n \\ &\quad + \tfrac{1}{2}\mathcal{L}^n(Y \cap \bar{B}_t(\zeta)) - \tfrac{1}{2}\mathcal{L}^n(Z \cap \bar{B}_t(\zeta)), \end{aligned}$$

hence

$$\tfrac{2}{3}\mathcal{L}^n(\bar{B}_t(\zeta)) \leq \mathcal{L}^n(Z \cap \bar{B}_t(\zeta)) \leq 4\,\mu\big(B_{b,\varrho} \cap \pi^{-1}(\bar{B}_t(\zeta))\big)$$

and the assertion follows.

$\mathcal{L}^n$ almost all $z \in Z \cap \bar{B}_\varrho(\pi(b))$ satisfy the assumption of the last assertion (cf. [Fed69, 2.9.11]) and Vitali's covering theorem (cf. [Fed69, 2.8.5]) implies

$$\mathcal{L}^n(Z \cap \bar{B}_\varrho(\pi(b))) \leq 6 \cdot 5^n \mu(B_{b,\varrho}).$$

Clearly,

$$\mathcal{L}^n(\pi(B_{b,\varrho})) \leq \mathcal{H}^n(B_{b,\varrho}) \leq \mu(B_{b,\varrho}).$$

Since $C_{b,\varrho} \sim N \subset (Z \cap \bar{B}_\varrho(\pi(b))) \cup \pi(B_{b,\varrho})$, it follows

$$\mathcal{L}^n(C_{b,\varrho}) \leq (1 + 6 \cdot 5^n)\mu(B_{b,\varrho}).$$

Finally, applying the coarea estimate with $W = C_{b,\varrho}$ yields

$$(1 - n\varepsilon^2)\mu(D_{b,\varrho}) \leq \mu(B_{b,\varrho}) + Q\mathcal{L}^n(C_{b,\varrho}) \leq (1 + Q + 6Q \cdot 5^n)\mu(B_{b,\varrho}). \qquad \square$$

[8] Recall from [Sim83] that the functions $\eta_{a,r} : \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$ are given by $\eta_{a,r}(x) = r^{-1}(x-a)$ for $a, x \in \mathbb{R}^{n+m}$, $0 < r < \infty$.

Proof of (8). Let I, f_i be associated to f as in A.6, and define $B_i = \operatorname{dmn} f_i$ for $i \in I$ and G as in (6). Note by (3) (5) (4), since $L \leq \delta_4/2$,

$$\begin{gathered} G \cap \pi^{-1}(\bar{B}_\varrho(\pi(b)) \sim C_{b,\varrho} = G \cap C(T,b,\varrho,(\delta_4/2)\varrho) \sim \pi^{-1}(C_{b,\varrho}), \\ \pi(B_{b,\varrho}) \subset C_{b,\varrho}, \quad \mu\big(C(T,b,\varrho,\delta_4\varrho) \sim (G \cup \pi^{-1}(C_{b,\varrho}))\big) = 0. \end{gathered}$$

Therefore one computes using A.6 and recalling that $C_{b,\varrho}$, $D_{b,\varrho}$, and, by 2.3, also $\pi^{-1}(C_{b,\varrho})$ are universally measurable

$$\begin{aligned} &\sum_{i \in I} \textstyle\int_{B_i \cap \bar{B}_\varrho(\pi(b)) \sim C_{b,\varrho}} \langle D\theta(x), DF(\operatorname{ap} Df_i(x))\rangle \,\mathrm{d}\mathcal{L}^n x \\ &\qquad = \delta\big(\mu \llcorner G \cap \pi^{-1}(\bar{B}_\varrho(\pi(b)) \sim C_{b,\varrho}\big)(\sigma^* \circ \theta \circ \pi) \\ &\qquad = \delta\big(\mu \llcorner G \cap C(T,b,\varrho,(\delta_4/2)\varrho) \sim \pi^{-1}(C_{b,\varrho})\big)((\psi \circ \sigma) \cdot (\sigma^* \circ \theta \circ \pi)) \\ &\qquad = \delta\big(\mu \llcorner C(T,b,\varrho,\delta_4\varrho) \sim \pi^{-1}(C_{b,\varrho})\big)((\psi \circ \sigma) \cdot (\sigma^* \circ \theta \circ \pi)) \\ &\qquad = (\delta\mu)((\psi \circ \sigma) \cdot (\sigma^* \circ \theta \circ \pi)) - \delta(\mu \llcorner D_{b,\varrho})((\psi \circ \sigma) \cdot (\sigma^* \circ \theta \circ \pi)), \end{aligned}$$

hence

$$\begin{aligned} &Q \textstyle\int_{B_\varrho(\pi(b))} \langle D\theta(x), DF(Dg(x))\rangle \,\mathrm{d}\mathcal{L}^n x - (\delta\mu)((\psi \circ \sigma) \cdot (\sigma^* \circ \theta \circ \pi)) \\ &= Q \textstyle\int_{C_{b,\varrho}} \langle D\theta(x), DF(Dg(x))\rangle \,\mathrm{d}\mathcal{L}^n x \\ &\quad + Q\Big(\textstyle\int_{\bar{B}_\varrho(\pi(b)) \sim C_{b,\varrho}} \langle D\theta(x), DF(Dg(x))\rangle \,\mathrm{d}\mathcal{L}^n x \\ &\qquad - \dfrac{1}{Q} \displaystyle\sum_{i \in I} \textstyle\int_{B_i \cap \bar{B}_\varrho(\pi(b)) \sim C_{b,\varrho}} \langle D\theta(x), DF(\operatorname{ap} Df_i(x))\rangle \,\mathrm{d}\mathcal{L}^n x\Big) \\ &\quad - \delta(\mu \llcorner D_{b,\varrho})((\psi \circ \sigma) \cdot (\sigma^* \circ \theta \circ \pi)). \end{aligned}$$

The first summand may be estimated using

$$DF(0) = 0, \quad \|DF(\alpha)\| \leq \gamma_1 |\alpha| \leq \gamma_1 n^{1/2} \operatorname{Lip} g$$

for $\alpha \in \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ with $\|\alpha\| \leq \operatorname{Lip} g$. The second summand can be treated noting

$$Dg(x) = \frac{1}{Q} \sum_{i \in I(x)} \operatorname{ap} Df_i(x) \quad \text{where } I(x) = \{i \in I : x \in \operatorname{dmn} \operatorname{ap} Df_i\}$$

for $\mathcal{L}^n$ almost all $x \in \bar{B}_\varrho(\pi(b)) \sim C_{b,\varrho}$ and applying 2.7 with

$$\begin{gathered} X,\ Y,\ f,\ a,\ r, \text{ and } \{x_1, \ldots, x_Q\} \\ \text{replaced by } \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m),\ \operatorname{Hom}(\operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m), \mathbb{R}),\ DF,\ \tau,\ t(x,\tau), \\ \text{and } \{\operatorname{ap} Df_i(x) : i \in I(x)\} \end{gathered}$$

for $\mathcal{L}^n$ almost all $x \in E_{b,\varrho} \sim C_{b,\varrho}$. Finally, the third summand is estimated by use of

$$|S \bullet \beta| \leq n^{1/2} |\beta| \quad \text{for } S \in G(n+m, n),\ \beta \in \operatorname{Hom}(\mathbb{R}^{n+m}, \mathbb{R}^{n+m}). \qquad \square$$

2.11. *Remark.* Concerning measurability, note that $\mathcal{L}^n$ measurability of W does not imply μ measurability of $\pi^{-1}(W)$ but only ν measurability. An example is provided by taking $m = 1$, $n > 1$, W to be a $\mathcal{H}^{n-1}$ nonmeasurable subset of $S = \{x \in \mathbb{R}^n : |x| = 1\}$ and $\mu = \mathcal{H}^n \llcorner \pi^{-1}(S)$ as may be verified by use of [Fed69, 2.2.4, 2.6.2, 3.2.23]. In the case $W = C_{b,\varrho}$ this difficulty could also have been resolved by making use of $\pi^{-1}(Y \sim \pi(B_{b,\varrho})) \cap B_{b,\varrho} = \emptyset$.

2.12. **Lemma.** *Suppose $n, m \in \mathbb{N}$, U is an open subset of $\mathbb{R}^{n+m}$, μ is an integral n varifold of locally bounded first variation in U, and $0 < L < \infty$.*

Then there exists a countable, disjointed family H of μ measurable sets covering μ almost all of U such that for each member Z of H there exist

$$g : \mathbb{R}^n \to \mathbb{R}^m, \quad G : \mathbb{R}^n \to \mathbb{R}^m, \quad A \subset \mathbb{R}^n, \quad Q \in \mathbb{N},$$
$$\pi \in O^*(n,m), \quad \sigma \in O^*(n, n-m), \quad T \in \mathcal{D}'(\mathbb{R}^n, \mathbb{R}^m)$$

with the following five properties:

(1) *$\sigma \circ \pi^* = 0$, $G = \pi^* + \sigma^* \circ g$, and $G(A) = Z$.*
(2) *$\operatorname{Lip} g \leq L$.*
(3) *A is an $\mathcal{L}^n$ measurable subset of $\operatorname{dmn} Dg$.*[9]
(4) *$-\int \langle D\theta(x), DF(Dg(x))\rangle \, \mathrm{d}\mathcal{L}^n x = T(\theta)$ whenever $\theta \in \mathcal{D}(\mathbb{R}^n, \mathbb{R}^m)$ where $F : \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m) \to \mathbb{R}$ denotes the nonparametric area integrand.*
(5) *Whenever $y \in A$ there holds*

$$\theta^n(\mu, G(y)) = Q, \quad \operatorname{im} DG(y) = T_{G(y)}\mu,$$
$$\lim_{\varrho \downarrow 0} \varrho^{-n-1} \int_{B_\varrho(y)} |Dg(x) - Dg(y)|^2 \, \mathrm{d}\mathcal{L}^n x = 0,$$
$$\lim_{\varrho \downarrow 0} \varrho^{-n-1} |T - T_y|_{-1,1;y,\varrho} = 0$$

where $T_y \in \mathcal{D}'(\mathbb{R}^n, \mathbb{R}^m)$ is defined by

$$T_y(\theta) = \int F(Dg(y)) \sigma(\vec{\mathbf{H}}_\mu(G(y))) \bullet \theta(x) \, \mathrm{d}\mathcal{L}^n x$$

whenever $\theta \in \mathcal{D}(\mathbb{R}^n, \mathbb{R}^m)$.

Proof. Observe that if some μ measurable set Z has the properties listed in the conclusion so does every μ measurable subset of Z. Therefore, in order to prove the assertion, it is enough to show that for μ almost all $a \in U$ there exists a μ measurable set Z having the above mentioned properties and additionally satisfies $\theta^{*n}(\mu \llcorner Z, a) > 0$; in fact one can then take a maximal, disjointed family H of such Z (hence $\mu(Z) > 0$) and note H is countable and $\theta^n(\mu \llcorner \bigcup H, a) = 0$ for $\mathcal{H}^n$ almost all $a \in U \sim \bigcup H$ by [Fed69, 2.10.19 (4)] so that $\mu(U \sim \bigcup H) > 0$ would contradict the maximality of H.

Assume $L \leq 1/8$. Fix $Q \in \mathbb{N}$. Define

$$\delta_1 = \delta_2 = \delta_3 = 1/2, \quad \delta_4 = 1/4, \quad \alpha = 1/2, \quad q = 2, \quad M = 6^m Q,$$
$$\varepsilon = \varepsilon_{2.10}(n, m, Q, L, M, \delta_1, \delta_2, \delta_3, \delta_4), \quad \varepsilon_1 = \varepsilon,$$

and $S : \operatorname{dmn} T_\mu \to \operatorname{Hom}(\mathbb{R}^{n+m}, \mathbb{R}^{n+m})$ by

$$S(x) = T_x\mu \quad \text{whenever } x \in \operatorname{dmn} T_\mu.$$

For $i \in \mathbb{N}$ let C_i denote the set of all $x \in \operatorname{spt} \mu$ such that either $\bar{B}_{1/i}(x) \not\subset U$ or

$$\|\delta\mu\|(\bar{B}_\varrho(x)) > \varepsilon \, \mu(\bar{B}_\varrho(x))^{1-1/n} \quad \text{for some } 0 < \varrho < 1/i,$$

let $D_i(b)$ for $b \in \operatorname{dmn} T_\mu$ denote the set of all $x \in U$ such that either $\bar{B}_{1/i}(x) \not\subset U$ or

$$\int_{\bar{B}_\varrho(x)} |S(\xi) - S(b)|^2 \, \mathrm{d}\mu(\xi) > (\varepsilon^2/4) \, \mu(\bar{B}_\varrho(x)) \quad \text{for some } 0 < \varrho < 1/i$$

and define X_i for $i \in \mathbb{N}$ by

$$X_i = \left\{ x \in U : \theta^{n^2/(n-1)}(\mu \llcorner C_i, x) = 0 \right\} \quad \text{if } n > 1,$$
$$X_i = U \sim \overline{C_i} \quad \text{if } n = 1,$$

[9] In contrast to 1.1, Dg here denotes the classical derivative.

as well as Y_i for $i \in \mathbb{N}$ by

$$Y_i = U \cap \{ b : \theta^{n+\alpha q}(\mu \llcorner D_i(b), b) = 0 \}.$$

Note $X_i \subset X_{i+1}$, $Y_i \subset Y_{i+1}$ for $i \in \mathbb{N}$. X_i are Borel sets by [Men08b, 2.9]. Y_i are μ measurable sets by [Men08b, 3.7 (2)]. Moreover,

$$\mu\big(U \sim \textstyle\bigcup\{X_i : i \in \mathbb{N}\}\big) = 0$$

by [Men08b, 2.5], [Men08b, 2.9, 2.10] and

$$\mu\big(U \sim \textstyle\bigcup\{Y_i : i \in \mathbb{N}\}\big) = 0$$

by [Men08b, 3.7 (2)] and Brakke's estimate of $\operatorname{tiltex}_\mu$ in [Bra78, 5.5, 5.7].

The conclusion will be shown at a point a such that for some $i \in \mathbb{N}$

$$\begin{gathered} \theta^n(\mu, a) = Q, \quad a \in \operatorname{dmn} T_\mu, \quad \bar{B}_{2/i}(a) \subset U, \\ a \in X_i \cap Y_i, \quad \theta^n(\mu \llcorner U \sim X_i, a) = 0, \quad \theta^n(\mu \llcorner U \sim Y_i, a) = 0, \\ \theta^n(\mu, \cdot) \text{ and } S \text{ are approximately continuous at } a \text{ with respect to } \mu. \end{gathered}$$

μ almost all $a \in U$ with $\theta^n(\mu, a) = Q$ satisfy this conditions by the preceding remarks and [Fed69, 2.9.11, 2.9.13]. Fix such a, i, abbreviate $T = T_a \mu$, and choose $\sigma \in O^*(n, n-m)$, $\pi \in O^*(n, m)$ such that $\sigma \circ \pi^* = 0$ and $\operatorname{im} \pi^* = T$. Moreover, choose $0 < 6r < 1/i$ such that

$$\begin{gathered} (Q - 1/2)\omega_n r^n \le \mu(C(T, a, r, r)) \le (Q + 1/2)\omega_n r^n, \\ \mu(C(T, a, r, 5r/4) \sim C(T, a, r, r/2)) \le (1/2)\omega_n r^n, \\ \mu(\{x \in \mathbb{R}^n : \operatorname{dist}(x, C(T, a, r, r)) < 2r\}) \le \mu(\bar{B}_{2^{1/2}(3r)}(a)) \le M \omega_n r^n. \end{gathered}$$

Now apply 2.10 with $h = r$ to obtain B, A, Y, f, G, $B_{b,\varrho}$, $C_{b,\varrho}$, $D_{b,\varrho}$, and $E_{b,\varrho}$ with the properties listed there and use Kirszbraun's theorem (cf. [Fed69, 2.10.43]) to extend $\boldsymbol{\eta}_Q \circ f$ to a function $g : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\operatorname{Lip} g = \operatorname{Lip}(\boldsymbol{\eta}_Q \circ f) \le \operatorname{Lip} f \le L.$$

Define

$$\begin{aligned} W_0 &= B_r(a) \cap X_i \cap Y_i \cap \{b : |S(b) - S(a)| \le \varepsilon/2\}, \\ W &= A \cap \{b : \theta^n(\mu, b) = Q\} \cap W_0. \end{aligned}$$

Next, *it will be shown*

$$B \subset C_i \cup D_i(b) \quad \textit{whenever } b \in W_0.$$

If $x \in B$, then $x \in C(T, a, r, r)$, $\theta^{*n}(\mu, x) > 0$ and

$$\begin{aligned} &\text{either} \quad \|\delta\mu\|(\bar{B}_s(x)) > \varepsilon\, \mu(\bar{B}_s(x))^{1-1/n} \quad \text{for some } 0 < s < 2r, \\ &\text{or} \quad \textstyle\int_{\bar{B}_s(x)} |S(\xi) - S(a)| \, d\mu(\xi) > \varepsilon\, \mu(\bar{B}_s(x)) \quad \text{for some } 0 < s < 2r. \end{aligned}$$

In the first case, this implies $x \in C_i$, in the second case,

$$\begin{aligned} \varepsilon\, \mu(\bar{B}_s(x)) &< \textstyle\int_{\bar{B}_s(x)} |S(\xi) - S(a)| \, d\mu(\xi) \\ &\le \textstyle\int_{\bar{B}_s(x)} |S(\xi) - S(b)| \, d\mu(\xi) + |S(b) - S(a)|\, \mu(\bar{B}_s(x)), \\ (\varepsilon/2)\, \mu(\bar{B}_s(x)) &< \textstyle\int_{\bar{B}_s(x)} |S(\xi) - S(b)| \, d\mu(\xi) \\ &\le \mu(\bar{B}_s(x))^{1/2} \big(\textstyle\int_{\bar{B}_s(x)} |S(\xi) - S(b)|^2 \, d\mu(\xi) \big)^{1/2}, \end{aligned}$$

hence $x \in D_i(b)$, and the claim is proved. It implies the estimate

$$\lim_{\varrho \downarrow 0} \varrho^{-n-1} \mu(B_{b,\varrho}) = 0 \quad \text{for } b \in W_0$$

which will be central to the remaining arguments (here the definition $B_{b,\varrho} = C(T, b, \varrho, \varrho/4) \cap B$ is extended to $b \in \mathbb{R}^{n+m}$, $0 < \varrho < \infty$). A simple consequence is, since $a \in W_0$, that $\theta^n(\mu \llcorner B, a) = 0$, hence

$$\theta^n(\mu \llcorner U \sim W, a) = 0, \quad \theta^n(\mu \llcorner W, a) = Q.$$

The proof will be concluded by showing the existence of a set N with $\mu(N) = 0$ such that $Z = W \sim N$ has the desired properties. Define $T \in \mathcal{D}'(\mathbb{R}^n, \mathbb{R}^m)$ by

$$T(\theta) = -\textstyle\int \langle D\theta(x), DF(Dg(x)\rangle \,\mathrm{d}\mathcal{L}^n x \quad \text{for } \theta \in \mathcal{D}(\mathbb{R}^n, \mathbb{R}^m)$$

where $F : \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m) \to \mathbb{R}$ denotes the nonparametric area integrand. Note $W \subset G$ and $\pi(W) \subset Y$ by 2.10 (3) (4) (6). Consider $b \in W$ such that $\pi(b) \in \operatorname{dmn} Dg$, $\operatorname{im} DG(\pi(b)) = T_b\mu$ and

$$\lim_{\varrho \downarrow 0} \varrho^{-n-1} \textstyle\int_{C(T,b,\varrho,\delta_4\varrho)} |S(\xi) - S(b)|^2 \,\mathrm{d}\mu(\xi) = 0.$$

These conditions are satisfied by μ almost all $b \in W$ by 2.10 (3) (5) (6a) (6b), [Fed69, 3.1.5, 3.2.17], and Brakke's estimate of $\operatorname{tiltex}_\mu$ in [Bra78, 5.5, 5.7]. Therefore it remains to verify the last two statements of (5) for μ almost all such b, i.e.

$$\lim_{\varrho \downarrow 0} \varrho^{-n-1} \textstyle\int_{B_\varrho(y)} |Dg(x) - Dg(y)|^2 \,\mathrm{d}\mathcal{L}^n x = 0,$$
$$\lim_{\varrho \downarrow 0} \varrho^{-n-1} |T - T_y|_{-1,1;y,\varrho} = 0$$

where T_y is defined as in (5).

For this purpose choose f_i, I as in A.4. First, observe that for every $0 < \gamma < \infty$ for $\mathcal{L}^n$ almost all $y \in Y$ with

$$\sup\{|\operatorname{ap} Df_i(y) - Dg(b)|^2 : y \in \operatorname{dmn} \operatorname{ap} Df_i, i \in I\} > \gamma$$

there exists $\xi \in G$ such that

$$\pi(\xi) = y, \quad |S(\xi) - S(b)|^2 > c\gamma$$

with $c = (1+L^2)(1-(2L)^2)^{-1}m$ by A.4, 2.10 (6) and estimates on tilted planes, see e.g. [All72, 8.9 (5)]. Since $L \leq 1/8$ this implies by 2.10 (5) that $\xi \in C(T, b, \varrho, \delta_4\varrho)$, hence

$$B_\varrho(\pi(b)) \cap \{y \in Y : t(y, Dg(b))^2 > \gamma\}$$

is $\mathcal{H}^n$ almost contained in

$$\pi\big(\{\xi \in C(T, b, \varrho, \delta_4\varrho) : |S(\xi) - S(b)|^2 > c\gamma\}\big)$$

for $0 < \varrho < r - |b-a|$, hence one obtains the *tilt estimate*

$$\textstyle\int_{Y \cap B_\varrho(\pi(b))} t(y, Dg(b))^2 \,\mathrm{d}\mathcal{L}^n y \leq c^{-1} \textstyle\int_{C(T,b,\varrho,\delta_4\varrho)} |S(\xi) - S(b)|^2 \,\mathrm{d}\mu(\xi).$$

Since, by $B_\varrho(\pi(b)) \sim Y \subset C_{b,\varrho}$,

$$\begin{aligned} &\textstyle\int_{B_\varrho(\pi(b))} |Dg(x) - Dg(b)|^2 \,\mathrm{d}\mathcal{L}^n x \\ &\leq \textstyle\int_{Y \cap B_\varrho(\pi(b))} t(x, Dg(b))^2 \,\mathrm{d}\mathcal{L}^n x + 4mL^2 \mu(C_{b,\varrho}), \end{aligned}$$

the first of the two remaining statements follows from 2.10 (7).

To prove the last remaining statement, suppose that $\theta \in \mathcal{D}(\mathbb{R}^n, \mathbb{R}^m)$ and $\psi \in \mathcal{D}(\mathbb{R}^m, \mathbb{R})$ satisfy

$$\operatorname{spt} \theta \subset B_1^n(0), \quad |D\theta|_{\infty;0,1} \leq 1,$$
$$\operatorname{spt} \psi \subset B_{1/4}^m(0), \quad \bar{B}_{1/8}^m(0) \subset \operatorname{Int}\{z : \psi(z) = 1\}, \quad 0 \leq \psi \leq 1.$$

Moreover, let

$$\theta_{b,\varrho} = \varrho^{-n} \theta \circ \eta_{\pi(b),\varrho}, \quad \psi_{b,\varrho} = \psi \circ \eta_{\sigma(b),\varrho}$$

for $0 < \varrho < r - |b - a|$ and such θ, ψ, and define

$$\gamma_1 = \sup \|D^2 F\|(\bar{B}_{n^{1/2}L}(0)), \quad \gamma_2 = \operatorname{Lip}\big(D^2 F | \bar{B}_{3n^{1/2}L}(0)\big).$$

Apply 2.10 (8) with $\tau = Dg(\pi(b))$

$$\begin{aligned}
&\big| Q {\textstyle\int_{B_\varrho(\pi(b))}} \langle D\theta_{b,\varrho}(x), DF(Dg(x))\rangle \,\mathrm{d}\mathcal{L}^n x - (\delta\mu)((\psi_{b,\varrho} \circ \sigma) \cdot (\sigma^* \circ \theta_{b,\varrho} \circ \pi))\big| \\
&\quad \leq \gamma_1 Q n^{1/2} L {\textstyle\int_{C_{b,\varrho}}} |D\theta_{b,\varrho}| \,\mathrm{d}\mathcal{L}^n + \gamma_2 Q {\textstyle\int_{E_{b,\varrho} \sim C_{b,\varrho}}} |D\theta_{b,\varrho}(x)| t(x, Dg(b))^2 \,\mathrm{d}\mathcal{L}^n x \\
&\qquad + n^{1/2} {\textstyle\int_{D_{b,\varrho}}} |D((\psi_{b,\varrho} \circ \sigma) \cdot (\sigma^* \circ \theta_{b,\varrho} \circ \pi))| \,\mathrm{d}\mu.
\end{aligned}$$

The first and the third summand on the right hand side may be estimated by use of 2.10 (7) as follows

$$\begin{aligned}
{\textstyle\int_{C_{b,\varrho}}} |D\theta_{b,\varrho}| \,\mathrm{d}\mathcal{L}^n &\leq \varrho^{-n-1} \mathcal{L}^n(C_{b,\varrho}) \leq \Gamma \varrho^{-n-1} \mu(B_{b,\varrho}), \\
{\textstyle\int_{D_{b,\varrho}}} |D((\psi_{b,\varrho} \circ \sigma) \cdot (\sigma^* \circ \theta_{b,\varrho} \circ \pi))| \,\mathrm{d}\mu &\leq \varrho^{-n-1}(1 + |D\psi|_{\infty;0,1}) \mu(D_{b,\varrho}) \\
&\leq \Gamma \varrho^{-n-1}(1 + |D\psi|_{\infty;0,1}) \mu(B_{b,\varrho})
\end{aligned}$$

where $\Gamma = \Gamma_{2.10\,(7)}(Q, n)$. To estimate the remaining summand, one computes

$$\textstyle\int_{E_{b,\varrho} \sim C_{b,\varrho}} |D\theta_{b,\varrho}(x)| t(x, Dg(b))^2 \,\mathrm{d}\mathcal{L}^n x \leq \varrho^{-1-n} \int_{Y \cap B_\varrho(\pi(b))} t(x, Dg(b))^2 \,\mathrm{d}\mathcal{L}^n x$$

and uses the tilt estimate. Therefore one infers that the supremum of all numbers

$$\big| Q \textstyle\int_{B_\varrho(\pi(b))} \langle D\theta_{b,\varrho}(x), DF(Dg(x))\rangle \,\mathrm{d}\mathcal{L}^n x - (\delta\mu)((\psi_{b,\varrho} \circ \sigma) \cdot (\sigma^* \circ \theta_{b,\varrho} \circ \pi))\big|$$

corresponding to $\theta \in \mathcal{D}(\mathbb{R}^n, \mathbb{R}^m)$ such that $\operatorname{spt} \theta \subset B^n_1(0)$, $|D\theta|_{\infty;0,1} \leq 1$ tends to 0 as $\varrho \downarrow 0$. Moreover, for every such θ

$$\begin{aligned}
|(\delta\mu)((\psi_{b,\varrho} \circ \sigma) \cdot (\sigma^* \circ \theta_{b,\varrho} \circ \pi))| &\leq \|\delta\mu\|(C(T, b, \varrho, \varrho/4)) \varrho^{-n} |\theta|_{\infty;0,1} \\
&\leq \|\delta\mu\|(C(T, b, \varrho, \varrho/4)) \varrho^{-n},
\end{aligned}$$

hence

$$\limsup_{\varrho \downarrow 0} \varrho^{-n-1} |T|_{-1,1;\pi(b),\varrho} < \infty$$

for μ almost all $b \in W$ by [Fed69, 2.9.5]. Since also, noting

$$\begin{gathered}
(\psi_{b,\varrho} \circ \sigma) \cdot (\sigma^* \circ \theta_{b,\varrho} \circ \pi) = \varrho^{-n} \big((\psi \circ \sigma) \cdot (\sigma^* \circ \theta \circ \pi)\big) \circ \eta_{b,\varrho}, \\
C(T, 0, 1, \infty) \cap T_b\mu \subset C(T, 0, 1, 1/8)
\end{gathered}$$

by the restriction imposed on L,

$$\begin{aligned}
&\lim_{\varrho \downarrow 0} (\delta\mu)((\psi_{b,\varrho} \circ \sigma) \cdot (\sigma^* \circ \theta_{b,\varrho} \circ \pi)) \\
&\qquad = -Q \textstyle\int \vec{\mathbf{H}}_\mu(b) \bullet (\psi \circ \sigma)(x) (\sigma^* \circ \theta \circ \pi)(x) \,\mathrm{d}(\mathcal{H}^n \llcorner T_b\mu)(x) \\
&\qquad = -Q \textstyle\int F(Dg(\pi(b))) \sigma\big(\vec{\mathbf{H}}_\mu(b)\big) \bullet \theta(x) \,\mathrm{d}\mathcal{L}^n x
\end{aligned}$$

for μ almost all $b \in W$ as may be verified by use of [Fed69, 2.9.9, 2.9.10], one infers the conclusion from B.2. $\square$

2.13. *Remark.* From Brakke's perpendicularity of mean curvature, see [Bra78, 5.8], one infers by an elementary calculation that

$$\vec{\mathbf{H}}_\mu(G(y)) = (\sigma^* - \pi^* \circ (Dg(y))^*)(\sigma(\vec{\mathbf{H}}_\mu(G(y))))$$

for $\mathcal{L}^n$ almost all $y \in A$.

2.14. *Remark.* Since $J^\mu T$ needs not be bounded from below on $D_{b,\varrho}$ by a positive function, the use of $|D\theta|_{\infty;0,1}$ instead of $|D\theta|_{p;0,1}$ for some $1 \leq p < \infty$ in the estimation of $\int_{D_{b,\varrho}} |D((\psi_{b,\varrho} \circ \sigma) \cdot (\sigma^* \circ \theta_{b,\varrho} \circ \pi))| \,\mathrm{d}\mu$ seems to be inevitable. The resulting complications will be resolved in 3.1–3.5.

3. Proof of the main theorem

The crucial estimate which allows to combine the preceding two sections in order to prove 3.5 and hence the main theorem 3.7 is given in 3.4. For this purpose the precise form of some standard estimates needed is given in 3.1 and 3.3.

3.1. **Lemma.** *Suppose $n, m \in \mathbb{N}$, $1 \leq p < \infty$, and $p < n/(n-1)$ if $n > 1$.*

Then there exists a positive, finite number ε with the following property.

If Υ is as in 1.2, $\Psi \in \odot^2 \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ with $\|\Psi - \Upsilon\| \leq \varepsilon$, $a \in \mathbb{R}^n$, $0 < r < \infty$, and $u \in W_0^{1,2}(B_r(a), \mathbb{R}^m)$, $T \in \mathcal{D}'(B_r(a), \mathbb{R}^m)$ satisfy

$$-\textstyle\int_{B_r(a)} \langle (D\theta(x), Du(x)), \Psi \rangle \, \mathrm{d}\mathcal{L}^n x = T(\theta) \quad \text{whenever } \theta \in \mathcal{D}(B_r(a), \mathbb{R}^m),$$

then

$$r^{-1-n/p} |u|_{p;a,r} \leq \Gamma r^{-n} |T|_{-1,1;a,r}$$

where Γ is a positive, finite number depending only on n, m, and p.

Proof. The estimate is considered to be classical and can be shown, for example, as follows.

Assuming $p > 1$, one deduces from [GT01, Lemma 9.17] in conjunction with a perturbation argument that, for a suitable number ε, L^p theory is available for the differential operator associated to Ψ, the asserted estimate then being provable by a duality argument. □

3.2. *Remark.* L^p theory is availabe for a much wider class of elliptic differential operators, see [ADN59, ADN64]. However, the smallness condition on $\|\Psi - \Upsilon\|$ makes it possible to refer to more elementary methods.

3.3. **Lemma.** *Suppose $n, m \in \mathbb{N}$, $0 < c \leq M < \infty$,*

$$F : \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m) \to \mathbb{R} \text{ is of class } \mathcal{C}^2,$$

$$\|D^2 F(\sigma)\| \leq M, \quad \big\langle (\tau, \tau), D^2 F(\sigma) \big\rangle \geq c|\tau|^2 \qquad \text{for } \sigma, \tau \in \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m),$$

$a \in \mathbb{R}^n$, $0 < r < \infty$, and $u, v \in W^{1,2}(B_r(a), \mathbb{R}^m)$ with

$$u - v \in W_0^{1,2}(B_r(a), \mathbb{R}^m).$$

Then for every affine function $P : \mathbb{R}^n \to \mathbb{R}^m$

$$|D(v-u)|_{2;a,r} \leq c^{-1} \big(M |D(u-P)|_{2;a,r} + |L_F(v)|_{-1,2;a,r} \big)$$

where L_F is defined as in 1.2.

Proof. Compute for $\theta \in \mathcal{D}(B_r(a), \mathbb{R}^m)$

$$\begin{aligned} L_F(v)(\theta) &= -\textstyle\int_{B_r(a)} \langle D\theta(x), DF(Dv(x)) - DF(DP(x)) \rangle \, \mathrm{d}\mathcal{L}^n x \\ &= -\textstyle\int_{B_r(a)} \langle (D\theta(x), D(v-P)(x)), A(x) \rangle \, \mathrm{d}\mathcal{L}^n x \\ &\text{where } A(x) = \textstyle\int_0^1 D^2 F(tDv(x) + (1-t) DP(x)) \, \mathrm{d}\mathcal{L}^1 t. \end{aligned}$$

This implies for $\theta \in \mathcal{D}(B_r(a), \mathbb{R}^m)$

$$\begin{aligned} &\textstyle\int_{B_r(a)} \langle (D\theta(x), D(v-u)(x)), A(x) \rangle \, \mathrm{d}\mathcal{L}^n x \\ &= -\textstyle\int_{B_r(a)} \langle (D\theta(x), D(u-P)(x)), A(x) \rangle \, \mathrm{d}\mathcal{L}^n x - L_F(v)(\theta). \end{aligned}$$

Letting θ approximate $v - u$ in $W^{1,2}(B_r(a), \mathbb{R}^m)$, one obtains

$$c(|D(v-u)|_{2;a,r})^2 \leq \big(M |D(u-P)|_{2;a,r} + |L_F(v)|_{-1,2;a,r} \big) |D(v-u)|_{2;a,r}. \qquad \square$$

3.4. **Lemma.** *Suppose $n, m \in \mathbb{N}$, $1 \leq p < \infty$, $p < n/(n-1)$ if $n > 1$, F is related to $\varepsilon = \varepsilon_{3.1}(n,m,p)$ as in 1.2, $\operatorname{Lip} D^2 F < \infty$, $a \in \mathbb{R}^n$, $0 < r < \infty$, and $u, v \in W^{1,2}(B_r(a), \mathbb{R}^m)$ with $u - v \in W_0^{1,2}(B_r(a), \mathbb{R}^m)$.*

Then for every affine function $P : \mathbb{R}^n \to \mathbb{R}^m$

$$r^{-1-n/p} |v - u|_{p;a,r} \leq \Gamma r^{-n} \big(|L_F(v) - L_F(u)|_{-1,1;a,r} \\ + \operatorname{Lip}(D^2 F)(|D(u-P)|_{2;a,r} + |D(v-P)|_{2;a,r})^2 \big)$$

where Γ is a positive, finite number depending only on n, m, and p.

Proof. Let $\Lambda = \operatorname{Lip} D^2 F$, choose $\sigma \in \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ such that $DP(x) = \sigma$ for $x \in \mathbb{R}^n$, and define $\Psi = D^2 F(\sigma)$, $T = L_F(v) - L_F(u)$, the $\mathcal{L}^n \llcorner B_r(a)$ measurable function $A : B_r(a) \to \odot^2 \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ by

$$A(x) = \textstyle\int_0^1 D^2 F(tDv(x) + (1-t)Du(x)) - D^2 F(\sigma) \, \mathrm{d}\mathcal{L}^1 t$$

whenever $x \in B_r(a)$, and $S \in \mathcal{D}'(B_r(a), \mathbb{R}^m)$ by

$$S(\theta) = \textstyle\int_{B_r(a)} \langle (D\theta(x), D(v-u)(x)), A(x) \rangle \, \mathrm{d}\mathcal{L}^n x + T(\theta)$$

whenever $\theta \in \mathcal{D}(B_r(a), \mathbb{R}^m)$. One computes

$$DF(Dv(x)) - DF(Du(x)) \\ = \Big\langle D(v-u)(x), \textstyle\int_0^1 DDF(tDv(x) + (1-t)Du(x)) \, \mathrm{d}\mathcal{L}^1 t \Big\rangle$$

for $\mathcal{L}^n$ almost all $x \in B_r(a)$ and infers

$$S(\theta) = -\textstyle\int_{B_r(a)} \langle (D\theta(x), D(v-u)(x)), \Psi \rangle \, \mathrm{d}\mathcal{L}^n x$$

whenever $\theta \in \mathcal{D}(B_r(a), \mathbb{R}^m)$, hence by 3.1

$$r^{-1-n/p} |v-u|_{p;a,r} \leq \Gamma_1 r^{-n} |S|_{-1,1;a,r}$$

where $\Gamma_1 = \Gamma_{3.1}(n,m,p)$. It remains to estimate $|S|_{-1,1;a,r}$ by use of the definition of S. One estimates

$$\begin{aligned} \|A(x)\| &\leq \textstyle\int_0^1 \|D^2 F(tDv(x) + (1-t)Du(x)) - D^2 F(t\sigma + (1-t)\sigma)\| \, \mathrm{d}\mathcal{L}^1 t \\ &\leq \Lambda \textstyle\int_0^1 t|D(v-P)(x)| + (1-t)|D(u-P)(x)| \, \mathrm{d}\mathcal{L}^1 t \\ &= \Lambda(|D(v-P)(x)| + |D(u-P)(x)|)/2 \end{aligned}$$

for $\mathcal{L}^n$ almost all $x \in B_r(a)$. Finally,

$$|S|_{-1,1;a,r} \leq |T|_{-1,1;a,r} + \Lambda/2 \textstyle\int_{B_r(a)} (|D(u-P)(x)| + |D(v-P)(x)|)^2 \, \mathrm{d}\mathcal{L}^n x. \quad \square$$

3.5. **Lemma.** *Suppose $n, m \in \mathbb{N}$, $1 \leq p < \infty$, and $p < n/(n-1)$ if $n > 1$.*

Then there exists a positive, finite number ε with the following property.

If F is related to ε as in 1.2, $\operatorname{Lip} D^2 F < \infty$, U is an open subset of $\mathbb{R}^{n+m}$, $u : U \to \mathbb{R}^m$ is weakly differentiable, A denotes the set of all $a \in \operatorname{dmn} Du$ such that

$$\limsup_{r \downarrow 0} r^{-n-1} \textstyle\int_{B_r(a)} |Du(x) - Du(a)|^2 \, \mathrm{d}\mathcal{L}^n x < \infty,$$

B denotes the set of all $a \in \operatorname{dmn} Du$ such that

$$\lim_{r \downarrow 0} r^{-n-1} \textstyle\int_{B_r(a)} |Du(x) - Du(a)|^2 \, \mathrm{d}\mathcal{L}^n x = 0,$$

and C denotes the set of all $a \in U$ such that

$$\limsup_{r \downarrow 0} r^{-n-1} |L_F(u)|_{-1,1;a,r} < \infty,$$

then A, B, and C are Borel sets and the following two statements hold:

(1) *For $\mathcal{L}^n$ almost all $a \in A \cap C$ there exists a polynomial function $Q_a : \mathbb{R}^n \to \mathbb{R}^m$ of degree at most 2 such that*

$$\lim_{r \downarrow 0} r^{-2-n/p} |u - Q_a|_{p;a,r} = 0.$$

(2) *For $\mathcal{L}^n$ almost all $a \in B \cap C$ the polynomial function Q_a of part (1) and the (constant) distribution T_a of B.2 are related by*

$$T_a(\theta) = \textstyle\int_U \theta(x) \bullet \big\langle D^2 Q_a(a), C_F(DQ_a(a)) \big\rangle \,\mathrm{d}\mathcal{L}^n x$$

for $\theta \in \mathcal{D}(U, \mathbb{R}^m)$ where C_F is defined as in 1.2.

Proof. Let $\varepsilon = \min\{\varepsilon_{3.1}(n,m,p), 1/2, \varepsilon_{1.11}(n,m,p,2), \varepsilon_{1.6}(n,m,2,2)\}$. Suppose F and u satisfy the hypotheses with ε. Clearly A and B are Borels sets. C is a Borel set by B.2. Abbreviate $\Lambda = \operatorname{Lip} D^2 F$ and $T = L_F(u)$.

To prove part (1), the criterion 1.11 will be verified with $q = 2$, $j = 0$. For this purpose let $a \in A \cap C$, $0 < r < \infty$ such that $\bar{B}_r(a) \subset U$ and $u_{a,r} = u | B_r(a) \in W^{1,2}(B_r(a), \mathbb{R}^m)$. Using the direct method of the calculus of variation, one constructs $v_{a,r} \in W^{1,2}(B_r(a), \mathbb{R}^m)$ such that

$$\begin{gathered} v_{a,r} - u_{a,r} \in W^{1,2}_0(B_r(a), \mathbb{R}^m), \\ L_F(v_{a,r}) = 0. \end{gathered}$$

Define $P_a : \mathbb{R}^n \to \mathbb{R}^m$ by $P_a(x) = \langle x, Du(a) \rangle$ for $x \in \mathbb{R}^n$. By 3.4 one estimates

$$\begin{gathered} r^{-1-n/p} |v_{a,r} - u_{a,r}|_{p;a,r} \\ \leq \Gamma_1 r^{-n} \big(|T|_{-1,1;a,r} + \Lambda(|D(u_{a,r} - P_a)|_{2;a,r} + |D(v_{a,r} - P_a)|_{2;a,r})^2 \big). \end{gathered}$$

with $\Gamma_1 = \Gamma_{3.4}(n,m,p)$. By 3.3 with $c = 1/2$, $M = 2$ one infers

$$|D(v_{a,r} - u_{a,r})|_{2;a,r} \leq 4 |D(u_{a,r} - P_a)|_{2;a,r},$$

hence

$$r^{-1-n/p} |v_{a,r} - u_{a,r}|_{p;a,r} \leq \Gamma_1 r^{-n} \big(|T|_{-1,1;a,r} + \Lambda(6 |D(u_{a,r} - P_a)|_{2;a,r})^2 \big).$$

Since $a \in A \cap C$, this implies

$$\limsup_{r \downarrow 0} r^{-2-n/p} |v_{a,r} - u_{a,r}|_{p;a,r} < \infty.$$

Therefore part (1) follows from 1.11.

To prove part (2), assume now additionally that the assumptions of (2) are valid for a, i.e. $a \in B \cap C$, and Q_a, T_a satisfy the conclusions of part (1) and B.2 respectively. Choose $y \in \mathbb{R}^m$ such that

$$T_a(\theta) = \textstyle\int_U \theta(x) \bullet y \,\mathrm{d}\mathcal{L}^n x \quad \text{for } \theta \in \mathcal{D}(U, \mathbb{R}^m).$$

Using the direct method of the calculus of variation, one constructs functions $w_{a,r} \in W^{1,2}(B_r(a), \mathbb{R}^m)$ such that

$$\begin{gathered} w_{a,r} - u_{a,r} \in W^{1,2}_0(B_r(a), \mathbb{R}^m), \\ L_F(w_{a,r})(\theta) = \textstyle\int_{B_r(a)} \theta(x) \bullet y \,\mathrm{d}\mathcal{L}^n x \quad \text{whenever } \theta \in \mathcal{D}(B_r(a), \mathbb{R}^m). \end{gathered}$$

By 3.4 one estimates

$$\begin{gathered} r^{-1-n/p} |w_{a,r} - u_{a,r}|_{p;a,r} \\ \leq \Gamma_1 r^{-n} \big(|T - T_a|_{-1,1;a,r} + \Lambda(|D(u_{a,r} - P_a)|_{2;a,r} + |D(w_{a,r} - P_a)|_{2;a,r})^2 \big). \end{gathered}$$

Since, by Poincaré's inequality,

$$\big| \textstyle\int_{B_r(a)} \theta(x) \bullet y \,\mathrm{d}\mathcal{L}^n x \big| \leq |y| \Gamma_2 r^{1+n/2} |D\theta|_{2;a,r}$$

where Γ_2 is a positive, finite number depending only on m and n, one infers from 3.3

$$|D(w_{a,r}-u_{a,r})|_{2;a,r} \le 4|D(u_{a,r}-P_a)|_{2;a,r} + 2\Gamma_2|y|r^{1+n/2},$$

hence

$$\begin{aligned} & r^{-1-n/p}|w_{a,r}-u_{a,r}|_{p;a,r} \\ & \le \Gamma_1 r^{-n}\big(|T-T_a|_{-1,1;a,r} + \Lambda(6|D(u_{a,r}-P_a)|_{2;a,r} + 2\Gamma_2|y|r^{1+n/2})^2\big). \end{aligned}$$

Since $a \in B \cap C$, this implies

$$\lim_{r\downarrow 0} r^{-2-n/p}|w_{a,r}-u_{a,r}|_{p;a,r} = 0.$$

Therefore by the assumption on Q_a

$$\lim_{r\downarrow 0} r^{-2-n/p}|w_{a,r}-Q_a|_{p;a,r} = 0.$$

Define $P : \mathbb{R}^n \to \mathbb{R}^m$ by $P(x) = Q_a(a) + \langle x-a, DQ_a(a)\rangle$ for $x \in \mathbb{R}^n$, $R = Q_a - P$, $S : \mathbb{R}^n \to \mathbb{R}^m$ by $S(x) = \frac{1}{2}\left\langle (x,x), D^2Q_a(a)\right\rangle$ for $x \in \mathbb{R}^n$ and note $r^{-2}R \circ \eta_{a,r}^{-1} = S$

$$r^{-2}(w_{a,r}-P)\circ\eta_{a,r}^{-1}|B_1^n(0) \to S|B_1^n(0) \quad \text{in } L^p(B_1^n(0),\mathbb{R}^m)$$

as $r \downarrow 0$. By 1.6

$$r^{-n/2}|D^2(w_{a,r}-P)|_{2;a,r/2} \le \Gamma_3(r^{-2-n/p}|w_{a,r}-P|_{p;a,r} + |y|)$$

where $\Gamma_3 = \max\{\omega_n^{1/p}, \omega_n^{1/2}\}\Gamma_{1.6}(n,m,2)$, hence

$$\limsup_{r\downarrow 0} r^{-n/2}|D^2(w_{a,r}-P)|_{2;a,r/2} < \infty.$$

This implies by use of an interpolation inequality and weak compactness properties of Sobolev spaces

$$r^{-2}(w_{a,r}-P)\circ\eta_{a,r}^{-1}|B_{1/2}^n(0) \to S|B_{1/2}^n(0)$$

weakly in $W^{2,2}(B_{1/2}^n(0),\mathbb{R}^m)$ as $r \downarrow 0$. By Rellich's embedding theorem

$$r^{-2}(w_{a,r}-Q_a)\circ\eta_{a,r}^{-1}|B_{1/2}^n(0) \to 0 \quad \text{in } W^{1,2}(B_{1/2}^n(0),\mathbb{R}^m)$$

as $r \downarrow 0$. Using this convergence, one computes for $\theta \in \mathcal{D}(B_{1/2}^n(0),\mathbb{R}^m)$

$$\begin{aligned} & \textstyle\int_{B_{1/2}^n(0)}\theta(x)\bullet y\,\mathrm{d}\mathcal{L}^n x = r^{-n}\int_{B_{r/2}(a)}(\theta\circ\eta_{a,r})(x)\bullet y\,\mathrm{d}\mathcal{L}^n x \\ & \textstyle= -r^{-n-1}\int_{B_{r/2}(a)}\big\langle (D\theta)\circ\eta_{a,r}, DF(Dw_{a,r}(x))\big\rangle\,\mathrm{d}\mathcal{L}^n x, \\ & \textstyle\big|r^{-n-1}\int_{B_{r/2}(a)}\big\langle (D\theta)\circ\eta_{a,r}, DF(Dw_{a,r}(x)) - DF(DQ_a(x))\big\rangle\,\mathrm{d}\mathcal{L}^n x\big| \\ & \le r^{-n-1}(\operatorname{Lip} DF)r^{n/2}|D\theta|_{2;0,1}|D(w_{a,r}-Q_a)|_{2;a,r} \to 0 \quad \text{as } r\downarrow 0, \\ & \textstyle -r^{-n-1}\int_{B_{r/2}(a)}\big\langle (D\theta)\circ\eta_{a,r}, DF(DQ_a(x))\big\rangle\,\mathrm{d}\mathcal{L}^n x \\ & \textstyle= r^{-n}\int_{B_{r/2}(a)}(\theta\circ\eta_{a,r})(x)\bullet\left\langle D^2Q_a(x), C_F(DQ_a(x))\right\rangle\,\mathrm{d}\mathcal{L}^n x \\ & \textstyle\to \int_{B_{1/2}^n(0)}\theta(x)\bullet\left\langle D^2Q_a(a), C_F(DQ_a(a))\right\rangle\,\mathrm{d}\mathcal{L}^n x \quad \text{as } r\downarrow 0, \end{aligned}$$

hence

$$y = \left\langle D^2Q_a(a), C_F(DQ_a(a))\right\rangle,$$

as claimed. □

3.6. *Remark.* Clearly, by [Reš68] for $\mathcal{L}^n$ almost all $a \in A \cap C$

$$Q_a(a) = u(a), \quad DQ_a(a) = Du(a).$$

Also by [CZ61, Theorem 9] (see also [Zie89, 3.6–8]), there exists a sequence of functions $u_i : \mathbb{R}^n \to \mathbb{R}^m$ of class $\mathcal{C}^2$ such that

$$\mathcal{L}^n \left(A \cap C \sim \bigcup_{i=1}^{\infty} \{ a : D^k u_i(a) = D^k Q_a(a) \text{ for } k \in \{0,1,2\} \} \right) = 0.$$

3.7. **Theorem.** *Suppose $n, m \in \mathbb{N}$, U is an open subset of $\mathbb{R}^{n+m}$, and μ is an integral n varifold in U of locally bounded first variation.*

Then μ is countably rectifiable of class $\mathcal{C}^2$ and for every n dimensional submanifold M of U of class $\mathcal{C}^2$ there holds

$$\vec{\mathbf{H}}_\mu(x) = \vec{\mathbf{H}}_M(x) \quad \textit{for } \mu \textit{ almost all } x \in M$$

where $-\vec{\mathbf{H}}_\mu$ corresponds to the absolutely continuous part of $\delta\mu$ with respect to μ and $\vec{\mathbf{H}}_M$ denotes the mean curvature of M.

Proof. It is enough to prove the existence of a countable collection of n dimensional submanifolds of U of class $\mathcal{C}^2$ such that for each member M

$$\vec{\mathbf{H}}_\mu(x) = \vec{\mathbf{H}}_M(x) \quad \text{for } \mu \text{ almost all } x \in M.$$

For this purpose define $p = 1$, $\varepsilon = \varepsilon_{3.5}(m, n, p)$, $\Gamma = \Gamma_{1.8}(m \cdot n, 2)$, $s = \varepsilon / \Gamma$. Denote by $F : \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m) \to \mathbb{R}$ the nonparametric area integrand, and choose $0 < \delta < \infty$ such that

$$\| D^2 F(\sigma) - D^2 F(0) \| \leq s \quad \text{whenever } \sigma \in \bar{B}_\delta(0) \cap \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m).$$

Using 1.8, there exists $G : \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m) \to \mathbb{R}$ of class $\mathcal{C}^3$ such that

$$D^i G(\sigma) = D^i F(\sigma) \quad \text{for } i = \{0,1,2\},\ \sigma \in \bar{B}_{\delta/2}(0) \cap \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m),$$
$$\| D^2 G(\sigma) - D^2 F(0) \| \leq \Gamma s = \varepsilon \quad \text{whenever } \sigma \in \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^m),$$
$$D^3 G \text{ has compact support},$$

hence $\operatorname{Lip} D^2 G < \infty$. Now, the conclusion is obtained by combining 2.12, 2.13 with $L = m^{-1/2}\delta/2$ and 3.5, 3.6 with F replaced by G. □

Appendix A. Almgren's notation for Q valued functions

In this appendix the part of Almgren's notation for Q valued functions used in the body of the text is summarised for the convenience of the reader.

A.1 (cf. [Alm00, 1.1 (1) (3)]). Suppose $Q \in \mathbb{N}$ and V is a finite dimensional Euclidian vector space.

$Q_Q(V)$ is defined to be the set of all 0 dimensional integral currents R such that $R = \sum_{i=1}^Q [\![x_i]\!]$ for some $x_1, \ldots, x_Q \in V$. A metric $\mathcal{G}$ on $Q_Q(V)$ is defined such that

$$\mathcal{G}\big(\textstyle\sum_{i=1}^Q [\![x_i]\!], \sum_{i=1}^Q [\![y_i]\!]\big) = \min \left\{ \big(\sum_{i=1}^Q |x_i - y_{\pi(i)}|^2 \big)^{1/2} : \pi \in \mathcal{S}(Q) \right\}$$

whenever $x_1, \ldots, x_Q, y_1, \ldots, y_Q \in V$ where $\mathcal{S}(Q)$ denotes the set of permutations of $\{1, \ldots, Q\}$. The function $\boldsymbol{\eta}_Q : Q_Q(V) \to V$ is defined by

$$\boldsymbol{\eta}_Q(R) = Q^{-1} \textstyle\int x \, \mathrm{d}\|R\|(x) \quad \text{whenever } R \in Q_Q(V).$$

If $R = \sum_{i=1}^Q [\![x_i]\!]$ for some $x_1, \ldots, x_Q \in V$, then $\boldsymbol{\eta}_Q(R) = \frac{1}{Q} \sum_{i=1}^Q x_i$. $\operatorname{Lip} \boldsymbol{\eta}_Q = Q^{-1/2}$.

Whenever $f : X \to Q_Q(V)$ the Q *valued graph of* f is defined by

$$\operatorname{graph}_Q f = \{ (x, v) \in X \times V : v \in \operatorname{spt} f(x) \}.$$

A.2 (cf. [Alm00, 1.1 (9) (10)]). Suppose $n, m, Q \in \mathbb{N}$.

A function $f : \mathbb{R}^n \to Q_Q(\mathbb{R}^m)$ is called *affine* if and only if there exist affine functions $f_i : \mathbb{R}^n \to \mathbb{R}^m$, $i = 1, \dots, Q$ such that

$$f(x) = \textstyle\sum_{i=1}^Q [\![f_i(x)]\!] \quad \text{whenever } x \in \mathbb{R}^n.$$

$f_1, \dots, f_Q$ are uniquely determined up to order.

Let $a \in A \subset \mathbb{R}^n$, $f : A \to Q_Q(\mathbb{R}^m)$. f is called *approximately affinely approximable at* a if and only if there exists an affine function $g : \mathbb{R}^n \to Q_Q(\mathbb{R}^m)$ such that (see [Fed69, 3.1.2])

$$\operatorname{ap} \lim_{x \to a} \mathcal{G}(f(x), g(x))/|x - a| = 0.$$

The function g is unique and denoted by $\operatorname{ap} A f(a)$. f is called *approximately strongly affinely approximable at* a if and only if $\operatorname{ap} A f(a)$ has the following property: If $\operatorname{ap} A f(a)(x) = \sum_{i=1}^Q [\![g_i(x)]\!]$ for some affine functions $g_i : \mathbb{R}^n \to \mathbb{R}^m$ and $g_i(a) = g_j(a)$ for some i and j, then $D g_i(a) = D g_j(a)$.

A.3. **Definition** (cf. [Alm00, T.1 (23)]). Whenever $f : X \to Y$, $g : X \to Z$ the *join* $f \bowtie g$ *of* f *and* g is defined by

$$(f \bowtie g)(x) = (f(x), g(x)) \quad \text{whenever } x \in X.$$

A.4. The following proposition in [Men08c, 1.11] or [Men08a, D.11] will be used for calculations involving Lipschitzian Q valued functions.

If $n, m, Q \in \mathbb{N}$, A is $\mathcal{L}^n$ measurable, $f : A \to Q_Q(\mathbb{R}^m)$ is Lipschitzian, I is countable, and to each $i \in I$ there corresponds a function $f_i \subset \operatorname{graph}_Q f$ with $\mathcal{L}^n$ measurable domain and $\operatorname{Lip} f_i \leq \operatorname{Lip} f$ *such that*

$$\#\{ i : f_i(x) = y \} = \theta^0(\|f(x)\|, y) \quad \textit{whenever } (x, y) \in A \times \mathbb{R}^m,$$

then f is approximately strongly affinely approximable with

$$\operatorname{ap} A f(a)(v) = \textstyle\sum_{i \in I(a)} [\![f_i(x) + \langle v, \operatorname{ap} D f_i(x) \rangle]\!] \quad \textit{whenever } v \in \mathbb{R}^n$$

at $\mathcal{L}^n$ almost all $a \in A$ where $I(a) = \{ i \in I : a \in \operatorname{dmn} \operatorname{ap} D f_i \}$. Moreover, such functions f_i do exist whenever n, m, Q, A, and f are as above.

A.5. Suppose U is an open subset of $\mathbb{R}^n$, Y is a Banach space and $T \in \mathcal{D}'(U, Y)$. Then T has a unique extension S to

$$\{ \theta \in \mathcal{E}(U, Y) : \operatorname{spt} \theta \cap \operatorname{spt} T \text{ is compact} \}$$

characterised by the requirement

$$S(\theta) = S(\eta) \quad \text{whenever } \operatorname{spt} T \subset \operatorname{Int}\{ x : \theta(x) = \eta(x) \}.$$

The extension will usually be denoted by the same symbol T.

A.6. Suppose $n, m, Q \in \mathbb{N}$, U is an open subset of $\mathbb{R}^n$, A is an $\mathcal{L}^n$ measurable subset of U, $\mathcal{L}^n(A) < \infty$, $f : A \to Q_Q(\mathbb{R}^m)$ is Lipschitzian, f_i for $i \in I$ are as in A.4, and $\pi \in O^*(n+m, n)$, $\sigma \in O^*(n+m, m)$ such that $\sigma \circ \pi^* = 0$ (see 2.8).

Defining an integral n varifold μ in $\pi^{-1}(U)$ by the requirement

$$\mu(X) = \textstyle\int_{X \cap \pi^{-1}(A)} \theta^0(\|f(\pi(x))\|, \sigma(x)) \, d\mathcal{H}^n x$$

for every Borel subset X of $\pi^{-1}(U)$, a simple calculation shows

$$(\delta \mu)(\sigma^* \circ \theta \circ \pi) = \textstyle\sum_{i \in I} \int_{\operatorname{dmn} f_i} \langle D\theta(x), DF(\operatorname{ap} D f_i(x)) \rangle \, d\mathcal{L}^n x$$

whenever $\theta \in \mathcal{D}(U, \mathbb{R}^m)$; here F denotes the nonparametric area integrand and the convention A.5 is used.

Appendix B. Lebesgue points for distributions

In general, for a distribution $T \in \mathcal{D}'(\mathbb{R}^n, \mathbb{R}^m)$ one cannot determine a value $y \in \mathbb{R}^m$ at a given point $a \in \mathbb{R}^n$. However, in case the rescaled distributions $r^{-n}(\eta_{a,r})_\# T$ whose values at $\theta \in \mathcal{D}(\mathbb{R}^n, \mathbb{R}^m)$ equal $r^{-n} T(\theta \circ \eta_{a,r})$ converge in $\mathcal{D}'(\mathbb{R}^n, \mathbb{R}^m)$ to a constant distribution T_a, this distribution T_a can be called the value of T at a. The main theorem of this appendix, B.2, asserts that if the distributions $r^{-n}(\eta_{a,r})_\# T$ are locally bounded with respect to the norm $|\cdot|_{-1,1;\cdot,\cdot}$ defined in 1.1 as $r \downarrow 0$ for all a in a set A then they actually converge for $\mathcal{L}^n$ almost all $a \in A$ to a constant distribution T_a with respect to this norm. As the author could not find this result in the literature, it is included here. Its proof uses techniques from [Fed69, 2.9.18] or [Men08b, 3.1].

B.1. **Lemma.** *Suppose $n, m \in \mathbb{N}$, A is a closed subset of $\mathbb{R}^n$, $R \in \mathcal{D}'(\mathbb{R}^n, \mathbb{R}^m)$, $\operatorname{dist}(\operatorname{spt} R, A) > 0$, $0 \leq \gamma < \infty$, and $0 < r < \infty$ such that*

$$|R|_{-1,1;x,\varrho} \leq \gamma\, \varrho^{n+1} \quad \textit{whenever } 0 < \varrho < 5r,\ x \in A.$$

Then

$$|R|_{-1,1;a,r} \leq \Gamma\, \gamma\, r\, \mathcal{L}^n(\bar{B}_{4r}(a) \sim A) \quad \textit{for } a \in A$$

where Γ is a positive, finite number depending only on n.

Proof. Assume $r \leq \frac{2}{9}$, let $a \in A$, $\theta \in \mathcal{D}(\mathbb{R}^n, \mathbb{R}^m)$ with $\operatorname{spt}\theta \subset B_r(a)$, choose $0 < \varepsilon \leq \min\{r, \operatorname{dist}(\operatorname{spt} R, A)\}$, define

$$B = \{x \in \mathbb{R}^n : \operatorname{dist}(x, \operatorname{spt}(R \llcorner \theta)) \leq \varepsilon/2\}$$

where $R \llcorner \theta \in \mathcal{E}'(\mathbb{R}^n, \mathbb{R})$ is defined by $(R \llcorner \theta)(v) = R(v\theta)$ for $v \in \mathcal{E}(\mathbb{R}^n, \mathbb{R})$, and apply [Fed69, 3.1.13] to obtain S, v_s, and h with $\Phi = \{\mathbb{R}^n \sim A, \mathbb{R}^n \sim B\}$; in particular S is a countable subset of $\bigcup \Phi$,

$$h(x) = \tfrac{1}{20} \max\{\min\{1, \operatorname{dist}(x, A)\}, \min\{1, \operatorname{dist}(x, B)\}\} \quad \text{for } x \in \textstyle\bigcup \Phi$$

and v_s for $s \in S$ form a partition of unity on $\bigcup \Phi$ with $\operatorname{spt} v_s \subset \bar{B}_{10h(s)}(s)$ for $s \in S$. Noting $\bigcup \Phi = \mathbb{R}^n$ one defines $T = \{s \in S : B \cap \operatorname{spt} v_s \neq \emptyset\}$ and infers

$$\textstyle\sum_{s \in S \sim T} v_s(x) = 0 \quad \text{for } x \in \mathbb{R}^n \text{ with } \operatorname{dist}(x, \operatorname{spt}(R \llcorner \theta)) < \varepsilon/2,$$

hence $(R \llcorner \theta)(\sum_{s \in S \sim T} v_s) = 0$ and

$$R(\theta) = R\big((\textstyle\sum_{s \in T} v_s)\theta\big) = \textstyle\sum_{s \in T} R(v_s \theta).$$

Choose $\xi(s) \in A$ for each $s \in T$ such that $|s - \xi(s)| = \operatorname{dist}(s, A)$. If $s \in T$ then there exists $y \in B \cap \operatorname{spt} v_s \subset \bar{B}_{r+\varepsilon/2}(a)$ and one observes

$$\begin{gathered}
\operatorname{dist}(y, A) \leq |y - a| \leq r + \varepsilon/2 \leq (3/2)r \leq \tfrac{1}{3} < 1, \quad h(y) = \tfrac{1}{20} \operatorname{dist}(y, A), \\
|s - y| \leq 10h(s) \leq 10h(y) + \tfrac{1}{2}|s - y|, \quad |s - y| \leq 20h(y) = \operatorname{dist}(y, A) \leq |y - a|, \\
\operatorname{dist}(s, A) \leq |s - y| + \operatorname{dist}(y, A) \leq 2\operatorname{dist}(y, A) \leq 3r \leq \tfrac{2}{3} < 1, \\
B \cap \bar{B}_{10h(s)}(s) \neq \emptyset, \quad \tfrac{1}{20}\operatorname{dist}(s, B) \leq \tfrac{1}{2}h(s), \quad 0 < h(s) = \tfrac{1}{20}\operatorname{dist}(s, A), \\
|s - \xi(s)| \leq |s - a| \leq |s - y| + |y - a| \leq 2r + \varepsilon \leq 3r \leq \tfrac{2}{3}, \\
\bar{B}_{h(s)}(s) \subset \bar{B}_{4r}(a) \sim A.
\end{gathered}$$

Moreover, for any $x \in \bar{B}_{10h(s)}(s)$, $s \in T$

$$|x - \xi(s)| \leq |x - s| + |s - \xi(s)| \leq (3/2)|s - \xi(s)| < 5r,$$
$$\operatorname{spt} v_s \subset \bar{B}_{(3/2)|s-\xi(s)|}(\xi(s)),$$
$$\operatorname{dist}(s, A) \leq \operatorname{dist}(x, A) + |x - s| \leq \operatorname{dist}(x, A) + \tfrac{1}{2} \operatorname{dist}(s, A),$$
$$|s - \xi(s)| = \operatorname{dist}(s, A) \leq 2 \operatorname{dist}(x, A),$$
$$\operatorname{dist}(x, A) \leq \operatorname{dist}(s, A) + |x - s| \leq \tfrac{3}{2} \operatorname{dist}(s, A) \leq 1,$$
$$h(x) \geq \tfrac{1}{20} \operatorname{dist}(x, A) \geq \tfrac{1}{40} |s - \xi(s)|.$$

Using the estimates of the preceding paragraph and the estimates of $|Dv_s|$ given in [Fed69, 3.1.13], one infers for $s \in T$, since θ has compact support in $B_r(a)$,

$$|(Dv_s)\theta|_{\infty;a,r} \leq V_1 40 |s - \xi(s)|^{-1} r |D\theta|_{\infty;a,r},$$
$$|D(v_s\theta)|_{\infty;a,r} \leq V_1 40 (|s - \xi(s)|^{-1} r + 1) |D\theta|_{\infty;a,r}$$

where V_1 is a positive, finite number depending only on n with $V_1 40 \geq 1$, hence

$$\begin{aligned} |R(v_s\theta)| &\leq \gamma (3/2)^{n+1} |s - \xi(s)|^{n+1} V_1 40 (|s - \xi(s)|^{-1} r + 1) |D\theta|_{\infty;a,r} \\ &= \gamma (3/2)^{n+1} V_1 40 |s - \xi(s)|^n (r + |s - \xi(s)|) |D\theta|_{\infty;a,r} \\ &\leq \gamma V_1 160 (3/2)^{n+1} \omega_n^{-1} (20)^n r\, \mathcal{L}^n(\bar{B}_{h(s)}(s))\, |D\theta|_{\infty;a,r}. \end{aligned}$$

Recalling from [Fed69, 3.1.13] that the family $\{\bar{B}_{h(s)}(s) : s \in S\}$ is disjointed, one concludes

$$|R(\theta)| \leq \Gamma\, \gamma\, r\, \mathcal{L}^n(\bar{B}_{4r}(a) \sim A) |D\theta|_{\infty;a,r}$$

where $\Gamma = 8(30)^{n+1} V_1 \omega_n^{-1}$. ☐

B.2. **Theorem.** *Suppose $n, m \in \mathbb{N}$, U is an open subset of $\mathbb{R}^n$, $T \in \mathcal{D}'(U, \mathbb{R}^m)$, and A denotes the set of all $a \in U$ such that*

$$\limsup_{r \downarrow 0} r^{-1-n} |T|_{-1,1;a,r} < \infty.$$

Then A is a Borel set and for $\mathcal{L}^n$ almost all $a \in A$ there exists a unique $T_a \in \mathcal{D}'(U, \mathbb{R}^m)$ with $D_i T_a = 0$ for $i \in \{1, \ldots, n\}$ such that

$$\lim_{r \downarrow 0} r^{-1-n} |T - T_a|_{-1,1;a,r} = 0.$$

Moreover, T_a depends $\mathcal{L}^n \llcorner A$ measurably on a.

Proof. The conclusion is local, hence one may assume $\operatorname{spt} T$ to be compact and $U = \mathbb{R}^n$. Since $|T|_{-1,1;a,r}$ depends lower semicontinuously on (a, r), the sets

$$A_i = \left\{ a \in \mathbb{R}^n : |T|_{-1,1;a,r} \leq i\, r^{n+1} \text{ for } 0 < r < (10)/i \right\}$$

defined for $i \in \mathbb{N}$ are closed. Observing $A = \bigcup \{A_i : i \in \mathbb{N}\}$, the conclusion will be shown to hold for $\mathcal{L}^n$ almost all $a \in A_i$.

Let $0 < \varepsilon < 5/i$, choose $\Phi \in \mathcal{D}(\mathbb{R}^n, \mathbb{R})$ with $\int \Phi \, d\mathcal{L}^n = 1$, $\operatorname{spt} \Phi \subset B^n_1(0)$ and define $\Phi_\varepsilon(x) = \varepsilon^{-n} \Phi(\varepsilon^{-1} x)$ for $x \in \mathbb{R}^n$,

$$T_\varepsilon(\theta) = T(\Phi_\varepsilon * \theta) = \textstyle\int f_\varepsilon \bullet \theta \, d\mathcal{L}^n \quad \text{for } \theta \in \mathcal{D}(\mathbb{R}^n, \mathbb{R}^m)$$

with $f_\varepsilon \in \mathcal{E}(\mathbb{R}^n, \mathbb{R}^m)$ given by

$$z \bullet f_\varepsilon(x) = T_y(\Phi_\varepsilon(y - x) z) \quad \text{whenever } x \in \mathbb{R}^n \text{ and } z \in \mathbb{R}^m,$$

see [Fed69, 4.1.2]. Clearly $T_\varepsilon \to T$ as $\varepsilon \downarrow 0$ and

$$|f_\varepsilon(x)| \leq i 2^{n+1} |D\Phi|_{\infty;0,1} \quad \text{for } x \in \mathbb{R}^n,\ a \in A_i \text{ with } |x - a| \leq \varepsilon.$$

One defines a_ε to be the characteristic function of $\{x \in \mathbb{R}^n : \operatorname{dist}(x, A_i) \le \varepsilon\}$ and $S_\varepsilon, R_\varepsilon \in \mathcal{D}'(\mathbb{R}^n, \mathbb{R}^m)$ by

$$S_\varepsilon(\theta) = \textstyle\int a_\varepsilon f_\varepsilon \bullet \theta \,\mathrm{d}\mathcal{L}^n \quad \text{for } \theta \in \mathcal{D}(\mathbb{R}^n, \mathbb{R}^m), \quad R_\varepsilon = T_\varepsilon - S_\varepsilon.$$

Estimating for $a \in A_i$, $0 < \varrho < 5r < 5/i$, $\theta \in \mathcal{D}(\mathbb{R}^n, \mathbb{R}^m)$ with $\operatorname{spt} \theta \subset B_\varrho(a)$ and $|D\theta|_{\infty;a,\varrho} \le 1$

$$\begin{gathered} \operatorname{spt}(\Phi_\varepsilon * \theta) \subset B_{\varepsilon+\varrho}(a), \quad |T_\varepsilon(\theta)| \le i(\varepsilon + \varrho)^{n+1} \le i 2^{n+1} \varrho^{n+1} \quad \text{if } \varepsilon \le \varrho, \\ \{x \in \operatorname{spt} R_\varepsilon : \operatorname{dist}(x, A_i) < \varepsilon\} = \emptyset, \quad R_\varepsilon(\theta) = 0 \quad \text{if } \varepsilon > \varrho, \\ |S_\varepsilon(\theta)| \le |a_\varepsilon f_\varepsilon|_{\infty;a,\varrho} \, |\theta|_{1;a,\varrho} \le i 2^{n+1} |D\Phi|_{\infty;0,1} \omega_n \varrho^{n+1} \\ |R_\varepsilon|_{-1,1;a,\varrho} \le \gamma \, \varrho^{n+1} \quad \text{with } \gamma = 2^{n+1} i \big(1 + |D\Phi|_{\infty;0,1} \, \omega_n\big), \end{gathered}$$

B.1 may be applied with to obtain

$$|R_\varepsilon|_{-1,1;a,r} \le \Gamma \, \gamma \, r \, \mathcal{L}^n(\bar{B}_{4r}(a) \sim A_i) \quad \text{for } 0 < r < 1/i.$$

Since $L^1(\mathcal{L}^n, \mathbb{R}^m)$ is separable, one can use [DS88, V.4.2, V.5.1, IV.8.3] to infer the existence of $S \in \mathcal{D}'(\mathbb{R}^n, \mathbb{R}^m)$, $f \in L^\infty(\mathcal{L}^n, \mathbb{R}^m)$ and a sequence $\varepsilon_j \downarrow 0$ as $j \to \infty$ such that

$$S(\theta) = \textstyle\int f \bullet \theta \,\mathrm{d}\mathcal{L}^n \quad \text{for } \theta \in \mathcal{D}(\mathbb{R}^n, \mathbb{R}^m), \quad S_{\varepsilon_j} \to S \quad \text{as } j \to \infty.$$

Defining $R = T - S$ and noting $R_{\varepsilon_j} \to R$ as $j \to \infty$,

$$|R|_{-1,1;a,r} \le \Gamma \, \gamma \, r \, \mathcal{L}^n(\bar{B}_{4r}(a) \sim A_i) \quad \text{for } 0 < r < 1/i$$

and [Fed69, 2.9.11] implies

$$\lim_{r \downarrow 0} r^{-1-n} |R|_{-1,1;a,r} = 0 \quad \text{for } \mathcal{L}^n \text{ almost all } a \in A_i.$$

Moreover,

$$\big| \textstyle\int (f(x) - f(a)) \bullet \theta(x) \,\mathrm{d}\mathcal{L}^n x \big| \le \big(\textstyle\int_{B_r(a)} |f(x) - f(a)| \,\mathrm{d}\mathcal{L}^n x \big) \, r \, |D\theta|_{\infty;a,r}$$

whenever $a \in A$, $0 < r < \infty$, $\theta \in \mathcal{D}(\mathbb{R}^n, \mathbb{R}^m)$ with $\operatorname{spt} \theta \subset B_r(a)$ and [Fed69, 2.9.9] implies that one can take T_a defined by $T_a(\theta) = \int \theta(x) \bullet f(a) \,\mathrm{d}\mathcal{L}^n x$ for $\theta \in \mathcal{D}(\mathbb{R}^n, \mathbb{R}^m)$ for $\mathcal{L}^n$ almost all $a \in A_i$ in the existence part of the conclusion.

The uniqueness follows, since every T_a admissible in the conclusion satisfies

$$r^{-n} (\eta_{a,r})_\# T_a = T_a, \quad r^{-n} (\eta_{a,r})_\# T \to T_a \quad \text{as } r \downarrow 0. \qquad \square$$

References

[ADN59] S. Agmon, A. Douglis, and L. Nirenberg. Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions. I. *Comm. Pure Appl. Math.*, 12:623–727, 1959.

[ADN64] S. Agmon, A. Douglis, and L. Nirenberg. Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions. II. *Comm. Pure Appl. Math.*, 17:35–92, 1964.

[All72] William K. Allard. On the first variation of a varifold. *Ann. of Math. (2)*, 95:417–491, 1972.

[Alm00] Frederick J. Almgren, Jr. *Almgren's big regularity paper*, volume 1 of *World Scientific Monograph Series in Mathematics*. World Scientific Publishing Co. Inc., River Edge, NJ, 2000. Q-valued functions minimizing Dirichlet's integral and the regularity of area-minimizing rectifiable currents up to codimension 2, With a preface by Jean E. Taylor and Vladimir Scheffer.

[AS94] Gabriele Anzellotti and Raul Serapioni. $\mathcal{C}^k$-rectifiable sets. *J. Reine Angew. Math.*, 453:1–20, 1994.

[Bra78] Kenneth A. Brakke. *The motion of a surface by its mean curvature*, volume 20 of *Mathematical Notes*. Princeton University Press, Princeton, N.J., 1978.

[CV77] C. Castaing and M. Valadier. *Convex analysis and measurable multifunctions.* Springer-Verlag, Berlin, 1977. Lecture Notes in Mathematics, Vol. 580.

[CZ61] A.-P. Calderón and A. Zygmund. Local properties of solutions of elliptic partial differential equations. *Studia Math.*, 20:171–225, 1961.

[DS88] Nelson Dunford and Jacob T. Schwartz. *Linear operators. Part I.* Wiley Classics Library. John Wiley & Sons Inc., New York, 1988. General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication.

[Fed69] Herbert Federer. *Geometric measure theory.* Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.

[GT01] David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order.* Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[LM08] G. P. Leonardi and S. Masnou. Locality of the mean curvature of rectifiable varifolds. URL: http://cvgmt.sns.it/papers/leomas08/ , 2008.

[Men08a] Ulrich Menne. *$\mathcal{C}^2$ rectifiability and Q valued functions.* PhD thesis, Universität Tübingen, 2008. URL: http://tobias-lib.ub.uni-tuebingen.de/volltexte/2008/3518 .

[Men08b] Ulrich Menne. Some Applications of the Isoperimetric Inequality for Integral Varifolds, 2008. `arXiv:0808.3652v1 [math.DG]`.

[Men08c] Ulrich Menne. A Sobolev Poincaré type inequality for integral varifolds, 2008. `arXiv:0808.3660v1 [math.DG]`.

[Mor66] Charles B. Morrey, Jr. *Multiple integrals in the calculus of variations.* Die Grundlehren der mathematischen Wissenschaften, Band 130. Springer-Verlag New York, Inc., New York, 1966.

[Reš68] Yurii G. Rešetnyak. Generalized derivatives and differentiability almost everywhere. *Math. USSR, Sb.*, 4:293–302, 1968.

[Sch01] Reiner Schätzle. Hypersurfaces with mean curvature given by an ambient Sobolev function. *J. Differential Geom.*, 58(3):371–420, 2001.

[Sch04a] Reiner Schätzle. Quadratic tilt-excess decay and strong maximum principle for varifolds. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 3(1):171–231, 2004.

[Sch04b] Reiner Schätzle. Lower Semicontinuity of the Willmore Functional for Currents. Preprint No. 167, Sonderforschungsbereich 611, Bonn, 2004.

[Sim83] Leon Simon. *Lectures on geometric measure theory*, volume 3 of *Proceedings of the Centre for Mathematical Analysis, Australian National University.* Australian National University Centre for Mathematical Analysis, Canberra, 1983.

[Zie89] William P. Ziemer. *Weakly differentiable functions*, volume 120 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1989. Sobolev spaces and functions of bounded variation.

Max-Planck-Institut für Gravitationsphysik (Albert-Einstein-Institut), Am Mühlenberg 1, D-14476 Golm, Deutschland

E-mail address: `Ulrich.Menne@aei.mpg.de`