CHAPTER 5

THE ALLARD REGULARITY THEOREM

Here we discuss Allard's ([AW1]) regularity theorem, which says roughly that if the generalized mean curvature of a rectifiable n-varifold $V = \underline{\underline{v}}(M,\theta)$ is in $L^p_{loc}(\mu_V)$ in U, $p > n$, if $\theta \geq 1$ μ_V a.e. in U, if $\xi \in \operatorname{spt} V \cap U$, and if $\omega_n^{-1} \rho^{-n} \mu_V(B_\rho(\xi))$ is sufficiently close to 1 for *some* sufficiently small* ρ, then V is *regular* near V in the sense that $\operatorname{spt} V$ is a $C^{1,1-n/p}$ n-dimensional submanifold near ξ.

A key idea of the proof is to show that V is well-approximated by the graph of a harmonic function near ξ. The background results needed for this are given in §20 (where it is shown that it is possible to approximate $\operatorname{spt} V$ by the graph of a Lipschitz function) and in §21 (which gives the relevant results about approximation by harmonic functions). The actual harmonic approximation is made as a key step in proving the central "tilt-excess decay" theorem in §22.

The idea of approximating by harmonic functions (in roughly the sense used here) goes back to De Giorgi [DG] who proved a special case of the above theorem (when $k = 1$ and when V corresponds to the reduced boundary of a set of least perimeter - see the previous discussion in §14 and the discussion in §37 below. Almgren used analogous approximations in his work [A1] for arbitrary $k \geq 1$. Reifenberg [R1, R2] used approximation by harmonic functions in a rather different way in his work or regularity of minimal surfaces.

---

* Depending on $\|\underline{\underline{H}}\|_{L^p(\mu_V)}$

§20 LIPSCHITZ APPROXIMATION

In this section $V = \underline{\underline{v}}(M,\theta)$ is a rectifiable n-varifold with generalized mean curvature $\underline{\underline{H}}$ in U (see 16.5), and we assume $p > n$, and

20.1 $$\begin{cases} 0 \in \operatorname{spt}\mu_V \ , \quad \bar{B}_R(0) \subset U \\ \left(\omega_n^{-1}\int_{B_R(0)} |\underline{\underline{H}}|^p d\mu_V\right)^{1/p} \le (1-n/p)\Gamma \ , \quad \Gamma R^{1-n/p} \le 1/2 \\ \theta \ge 1 \ , \quad \omega_n^{-1} R^{-n} \mu_V(B_R(0)) \le 2(1-\alpha) \ , \end{cases}$$

where $\alpha \in (0,1)$. We also subsequently write μ for μ_V, and

$$E = R^{-n}\int_{B_R(0)} \|p_x - p\|^2 d\mu + \left(\Gamma R^{1-n/p}\right)^2 ,$$

where $p = p_{\mathbb{R}^n}$, $p_x = p_{T_xV}$ $(= p_{T_xM}$ μ-a.e. $x)$. Notice that then the first term in the definition of E measures the "mean-square deviation" of T_xV away from $\mathbb{R}^n$ over $B_R(0)$. (This is called the "tilt-excess" of V over $B_R(0)$ - see §22).

20.2 THEOREM *Assuming* 20.1, *there is a constant* $\gamma = \gamma(n,\alpha,k,p) \in (0,1/2)$ *such that if* $\ell \in (0,1]$ *then there is a Lipschitz function* $f = (f^1,\ldots,f^k) : B^n_{\gamma R}(0) \to \mathbb{R}^k$ *with*

$$\operatorname{Lip} f \le \ell \ , \quad \sup|f| \le c\, E^{\frac{1}{2n+2}} R$$

and

$$\mathcal{H}^n(((\operatorname{graph} f \sim \operatorname{spt} V) \cup (\operatorname{spt} V \sim \operatorname{graph} f)) \cap B_{\gamma R}(0)) \le c\ell^{-2n-2} E ,$$

where $c = c(n,\alpha,k,p)$.

20.3 REMARK Notice that this is trivial (by setting $f \equiv 0$ and taking suitable c) unless $\ell^{-2n-2}E$ is small. In particular we may assume $E \le \delta\ell^{2n+2}$, which δ is as small as we please, so long as our eventual choice of δ depends only on n, k, α, p.

Proof of 20.2 By virtue of the above remark we can assume

$$E \le \delta_0^2 , \tag{1}$$

δ_0 to be chosen depending only on n, k, α, p. Set

$$\ell_0 = (\delta_0^{-2}E)^{\frac{1}{2n+2}} < 1 ,$$

and take any two points $x,y \in B_{\beta R}(0) \cap \operatorname{spt} V$ with $|q(x-y)| \ge \ell_0|x-y|$, $|x-y| \ge \beta R/4$, where $\beta \in (0,1/2)$ is for the moment arbitrary. By Lemma 19.5 we have

$$\Theta^n(\mu,x) + \Theta^n(\mu,y) \le (1+c(\ell_0\beta)^{-n}\Gamma R^{1-n/p})(1-\beta)^{-n}\omega_n^{-1}R^{-n}\mu(B_R(0))$$

$$+ c(\ell_0\beta)^{-n-1}R^{-n}\int_{B_R(0)} \|p_x-p\|\,d\mu .$$

Using Cauchy inequality $ab \le \frac{\alpha}{4}a^2 + \frac{1}{\alpha}b^2$ in the last term, together with the assumption (in 20.1) that $\omega_n^{-1}R^{-n}\mu(B_R(0)) \le 2(1-\alpha)$, this gives

$$\Theta^n(\mu,x) + \Theta^n(\mu,y) \le 2(1+c(\ell_0\beta)^{-n}\sqrt{E})(1-\beta)^{-n}(1-\alpha)$$

$$+ \frac{\alpha}{2} + \frac{c}{\alpha}(\ell_0\beta)^{-2n-2}E .$$

Since $\ell_0^{2n+2} = \delta_0^{-2}E$ and $\Theta^n(\mu,\xi) \ge 1 \ \forall\, \xi \in \operatorname{spt} V \cap U$ (by 17.8 and the assumption that $\theta \ge 1$ μ-a.e.) this gives

$$2 \le 2(1+c\delta_0)(1-\beta)^{-n}(1-\alpha)$$

$$+ \frac{\alpha}{2} + c\alpha^{-1}\delta_0\beta^{-2n-2}$$

which is clearly impossible if we take $\beta = \beta(n,k,p,\alpha)$ and $\delta_0 = \delta_0(n,k,\beta,p,\alpha)$ small enough. Thus for such a choice of β, δ_0 we have

$$|q(x-y)| \le cE^{\frac{1}{2n+2}} R \ , \ x,y \in \operatorname{spt}\mu \cap B_{\beta R}(0) \ , \ |x-y| \ge \beta R/4 \ , \tag{2}$$

where $c = c(n,k,p,\alpha)$, $\beta = \beta(n,k,p,\alpha)$. (Formally we derived this subject to assumption (1), but if (1) fails then (2) is trivial with $c = \delta_0^{-1}$.) Noting the arbitrariness of x,y in (2) and noting also that $0 \in \operatorname{spt}\mu$ and that $\operatorname{spt}\mu \cap \partial B_{\beta R/2}(0) \ne \emptyset$ (which follows for example by selecting suitable ϕ in 17.2), we conclude (after replacing β by $\beta/4$)

$$|q(x)| \le c\, E^{\frac{1}{2n+2}} R \ , \quad x \in B_{\beta R}(0) \cap \operatorname{spt} V \ , \ \beta = \beta(n,k,p,\alpha) \in (0,1) \ . \tag{3}$$

Next let $\delta, \ell \in (0,1]$ be arbitrary and assume

$$\left(\Gamma R^{1-n/p}\right)^2 \le \ell^{2n+2}\delta \tag{4}$$

(which we can do by Remark 20.3, provided we eventually choose $\delta = \delta(n,k,\alpha,p)$). Set $E_0(\sigma,\xi) = \sigma^{-n}\int_{B_\sigma(\xi)} \|p_x - p\|^2 d\mu(x)$ for any $\xi \in \operatorname{spt} V$, $B_\sigma(\xi) \subset B_R(0)$, and define

$$G = \{\xi \in \operatorname{spt} V \cap B_{\beta R/2}(0) : E_0(\sigma,\xi) \le \delta\ell^{2n+2} \quad \forall\, \sigma \in (0,R/2)\} \ .$$

Notice that if $\xi \in \operatorname{spt} V \cap B_{\beta R}(0)$ then by (4) and the monotonicity formula 17.6(1) (see Remark 17.9(2) to justify the application of 17.6(1))

$$\begin{aligned} \tfrac{1}{2} \le \omega_n^{-1}\sigma^{-n}\mu(B_\sigma(\xi)) &\le (1+c\delta)\omega_n^{-1}((1-\beta)R)^{-n}\mu(B_{(1-\beta)R}(\xi)) \\ &\le (1+c\delta)(1-\beta)^{-n}\omega_n^{-1}R^{-n}\mu(B_R(0)) \\ &\le 2(1+c\delta)(1-\beta)^{-n}(1-\alpha) \\ &\le 2(1-\alpha/2) \ , \end{aligned} \tag{5}$$

for δ, β small enough (depending on n,k,p,α).

Now let $x,y \in G$. In view of (4), (5) we may now apply the previous argument with $\alpha/2$, $\beta^{-1}|x-y|/2$, x in place of α , R , 0 in order to deduce from (3) that

$$|q(x-y)| \le c\ell\delta^{1/(2n+2)}|x-y| , \; x,y \in G , \; c = c(n,k,p,\alpha) \tag{6}$$

(because $E_0(\sigma,x) + \left(\Gamma\sigma^{1-n/p}\right)^2 \le 2\delta\ell^{2n+2}$, $\sigma = \beta^{-1}|x-y|/2$, by virtue of (4) and the fact that $x \in G$) .

Choosing δ so that $2\delta^{1/(2n+2)}(1+c)(n+k) < 1$ (c as in (6)), we thus deduce

$$|q(x-y)| \le \frac{\ell}{2(n+k)} |x-y| , \; x,y \in G , \quad c = c(n,k,p,\alpha) .$$

Since $|x-y| \le |q(x-y)| + |p(x-y)|$, this implies

$$|q(x) - q(y)| \le \frac{\ell}{(n+k)} |p(x) - p(y)| \tag{7}$$

and so (by the extension theorem 5.1)

$$G \subset \text{graph } f ,$$

where f is a Lipschitz function $B_{\beta R/2}(0) \to \mathbb{R}^k$ with $\text{Lip } f \le \ell$. By virtue of (3) we can assume (by truncating f if necessary) that $\sup|f| \le cE^{\frac{1}{2n+2}}R$.

Next we note that (by definition of G) for each $\xi \in (B_{\beta R/2}(0) \sim G) \cap \text{spt } V$ we have $\sigma(\xi) \in (0,R/10)$ such that

$$\ell^{2n+2}\delta\sigma(\xi)^n \le \int_{B_{\sigma(\xi)}(\xi)} \|p_x - p\|^2 d\mu(x)$$

and by (5) we therefore have

$$\mu(\bar{B}_{5\sigma(\xi)}(\xi)) \le c\ell^{-2n-2}\delta^{-1}\int_{B_{\sigma(\xi)}(\xi)} \|p_x - p\|^2 d\mu(x) .$$

By definition the collection of balls $\{B_{\sigma(\xi)}(\xi)\}_{\xi\in B_{\beta R/2}(0)\sim G}$ is a cover for $B_{\beta R/2}(0) \sim G$, and hence by the covering Theorem 3.3 we can select points $\xi_1,\xi_2,\ldots \in B_{\beta R/2}(0) \sim G$ such that $\{B_{\sigma_j}(\xi_j)\}$ is a disjoint collection $(\sigma_j = \sigma(\xi_j))$ and $\{\bar{B}_{5\sigma_j}(\xi_j)\}$ still covers $B_{\beta R/2}(0) \sim G$. Then setting $\xi = \xi_j$ and summing over j , we conclude

$$(8) \qquad \mu(B_{\beta R/2}(0)\sim G) \le c\,\ell^{-2n-2}\delta^{-1}\int_{B_R(0)} \|p_x - p\|^2 d\mu(x) .$$

Since $\Theta^n(\mu,\xi) \ge 1$ for $\xi \in \operatorname{spt} V \cap U$ we have

$$\mathcal{H}^n((\operatorname{spt}\mu \sim \operatorname{graph} f) \cap B_{\beta R/2}(0)) \le \mu(B_{\beta R/2}(0) \sim \operatorname{graph} f)$$

(by Theorem 3.2(1)) and it thus remains only to prove

$$(9) \qquad \mathcal{H}^n((\operatorname{graph} f \sim \operatorname{spt}\mu) \cap B_{\beta R/2}(0)) \le c\,\ell^{-2n-2} E_R{}^n .$$

(Then the theorem will be established with $\gamma = \beta/2$.)

To check this, take any $\eta \in (\operatorname{graph} f \sim \operatorname{spt}\mu) \cap B_{\beta R/4}(0)$ and let $\sigma \in (0,\beta R/2)$ be such that $B_{\sigma/2}(\eta) \cap \operatorname{spt}\mu = \emptyset$ and $B_{3\sigma/4}(\eta) \cap \operatorname{spt}\mu \ne \emptyset$. (Such σ exists because $0 \in \operatorname{spt}\mu$.) Then the monotonicity formula 17.6(2) (See Remark 17.9(2)) implies

$$\mu(B_\sigma(\eta)) \le c\,\sigma^n \int_{B_\sigma(\eta)\sim B_{\sigma/2}(\eta)} |x-\eta|^{-n} \left| p_{(T_xM)^\perp}\left(\frac{x-\eta}{|x-\eta|}\right)\right|^2 d\mu$$

$$\le c \int_{B_\sigma(\eta)} \left| p_{(T_xM)^\perp}\left(\frac{x-\eta}{\sigma}\right)\right|^2 d\mu$$

$$\le c\left(\int_{B_\sigma(\eta)} \left| p_{(\mathbb{R}^n)^\perp}\left(\frac{x-\eta}{\sigma}\right)\right|^2 d\mu + \int_{B_\sigma(\eta)} \|p_{T_xM} - p_{\mathbb{R}^n}\|^2 d\mu\right)$$

$$\le c\left(\int_{B_\sigma(\eta)\cap F}\left|p_{(\mathbb{R}^n)^\perp}\left(\frac{x-\eta}{\sigma}\right)\right|^2 d\mu + \mu(B_\sigma(\eta)\sim F)\right.$$

$$\left. + \int_{B_\sigma(\eta)}\|p_{T_xM} - p_{\mathbb{R}^n}\|^2 d\mu\right) ,$$

where $F = \text{graph } f$, and where we used $p_{T^\perp}(x) = x - p_T(x)$ for any subspace $T \subset \mathbb{R}^{n+k}$. Since $\left|p_{(\mathbb{R}^n)^\perp}\left(\frac{x-y}{\sigma}\right)\right| \le c\ell$ for $x,y \in F \cap B_\sigma(\eta)$ (because $\text{Lip } f \le \ell$) , this implies

$$\mu(B_\sigma(\eta)) \le c\left(\ell\mu(B_\sigma(\eta)) + \mu(B_\sigma(\eta)\sim F) + \int_{B_\sigma(\eta)}\|p_{T_xM} - p_{\mathbb{R}^n}\|^2 d\mu\right) .$$

Since we can take $c\,\ell \le 1/2$ (notice again the validity of the theorem in this case automatically implies its validity for larger values of $\ell \in (0,1]$) , we thus get

(10) $$\mu(B_\sigma(\eta)) \le c\left(\mu(B_\sigma(\eta)\sim F) + \int_{B_\sigma(\eta)}\|p_{T_xM}-p_{\mathbb{R}^n}\|^2\, d\mu\right) ,$$

where $F = \text{graph } f$. Now since $\text{spt }\mu \cap B_{3\sigma/4}(\eta) \ne \emptyset$, the monotonicity (5) implies $\mu(B_\sigma(\eta)) \ge \frac{1}{2}\sigma^n$, and hence (10) gives

(11) $$\sigma^n \le c\,T ,$$

where T is the expression on the right of (10). Thus, writing $\eta' = p_{\mathbb{R}^n}(\eta)$, we get

$$\begin{aligned} L^n(B^n_{5\sigma}(\eta')) &\le c\,T \\ &\le c\left(\mu((B^n_\sigma(\eta')\times\mathbb{R}^k)\cap B_{\beta R/2}(0)\sim F)\right. \\ &\qquad \left. + \int_{(B^n_\sigma(\eta')\times\mathbb{R}^k)\cap B_{\beta R/2}(0)}\|p_{T_xM}-p_{\mathbb{R}^n}\|^2\, d\mu\right) . \end{aligned}$$

Since we have this for each $\eta \in (\text{graph } f \sim \text{spt }\mu) \cap B_{\beta R/4}(0)$, it follows from the Covering Theorem 3.3 in the usual way that

$$\mathcal{L}^n(p_{\mathbb{R}^n}((\text{graph } f \sim \text{spt } \mu) \cap B_{\beta R/4}(0)) \le c\, \mu(B_{\beta R/2}(0) \sim F)$$

$$+ c \int_{B_{\beta R/2}(0)} \| p_{T_x M} - p_{\mathbb{R}^n} \|^2 \, d\mu$$

$$\le c\, \ell^{-2n-2} E_R^{\,n} \quad \text{by} \quad (8) .$$

Since $\text{Lip } f \le 1$, this gives (9) with $\beta/4$ in place of β. Thus the theorem is established for suitable γ depending only on n, k, α, p.

§21. APPROXIMATION BY HARMONIC FUNCTIONS

The main result we shall need is given in the following lemma, which is an almost trivial consequence of Rellich's theorem:

21.1 LEMMA *Given any* $\varepsilon > 0$ *there is a constant* $\delta = \delta(n,\varepsilon) > 0$ *such that if* $f \in W^{1,2}(B)$, $B \equiv B_1(0)$ = *open unit ball in* $\mathbb{R}^n$, *satisfies*

$$\int_B |\text{grad } f|^2 \le 1$$

$$\left| \int_B \text{grad } f \cdot \text{grad } \zeta \, d\mathcal{L}^n \right| \le \delta \sup|\text{grad } \zeta|$$

for any $\zeta \in C_c^\infty(B)$, *then there is a harmonic function* u *on* B *such that*

$$\int_B |\text{grad } u|^2 \le 1$$

and

$$\int_B (u-f)^2 \le \varepsilon .$$

Proof Suppose the lemma is false. Then we can find $\varepsilon > 0$ and a sequence $\{f_k\} \in W^{1,2}(B)$ such that

(1) $$\left|\int_B \operatorname{grad} f_k \cdot \operatorname{grad} \zeta \, dL^n\right| \le k^{-1} \sup|\operatorname{grad} \zeta|$$

for each $\zeta \in C_c^\infty(B)$, and

(2) $$\int_B |\operatorname{grad} f_k|^2 \le 1 ,$$

but so that

(3) $$\int_B |f_k - u|^2 > \varepsilon$$

whenever u is a harmonic function on B with $\int_B |\operatorname{grad} u|^2 \le 1$.

Let $\lambda_k = \omega_n^{-1} \int_B f_k \, dL^n$. Then by the Poincaré inequality (see e.g. [GT]) we have

(4) $$\int_B |f_k - \lambda_k|^2 \le c \int_B |\operatorname{grad} f_k|^2 \le c ,$$

and hence, by Rellich's theorem (see [GT]), we have a subsequence $\{k'\} \subset \{k\}$ such that $f_{k'} - \lambda_{k'} \to w$ in $L^2(B)$, where $w \in W^{1,2}(B)$ with $\int_B |\operatorname{grad} w|^2 \le 1$. Also by (1) we evidently have

$$\int \operatorname{grad} w \cdot \operatorname{grad} \zeta \, dL^n = 0$$

for each $\zeta \in C_c^\infty(B)$. Thus w is harmonic in B and $\int_B |f_{k'} - w - \lambda_{k'}|^2 \to 0$. Since $w + \lambda_{k'}$ is harmonic, this contradicts (3).

We also recall the following standard estimates for harmonic functions (which follow directly from the mean-value property - see e.g. [GT]): If u is harmonic on $B \equiv B_1(0)$, then

$$\sup_{B_{\frac{1}{2}}(0)} |D^q(u)| \le c\|u\|_{L^1(B)}$$

for each integer $q \ge 1$, where $c = c(q,n)$. Indeed applying this with Du in place of u we get

21.2 $$\sup_{B_{\frac{1}{2}}(0)} |D^q u| \le c\|Du\|_{L^1(B)} \quad (\le c'\|Du\|_{L^2(B)})$$

for $q \ge 2$. Using an order 2 Taylor series expansion for u , we see that this implies

21.3 $$\sup_{B_\eta(0)} |u-\ell| \le c\eta^2 \|Du\|_{L^2(B)}$$

for each $\eta \in (0,1/2]$, where $c = c(n)$ is independent of η and where ℓ is the affine function given by $\ell(x) = u(0) + x \cdot \text{grad}\, u(0)$.

§22. THE TILT-EXCESS DECAY LEMMA

We define tilt-excess $E(\xi,\rho,T)$ (relative to the rectifiable n-varifold $V = \underline{\underline{v}}(M,\theta))$ by

$$E(\xi,\rho,T) = \frac{1}{2}\rho^{-n} \int_{B_\rho(\xi)} |p_{T_xM} - p_T|^2 d\mu_V \ , \ ^*$$

whenever $\rho > 0$, $\xi \in \mathbb{R}^{n+k}$ and T is an n-dimensional subspace of $\mathbb{R}^{n+k}$. Thus E measures the mean-square deviation of the approximate tangent space T_xM away from the given subspace T . Notice that if we have $T = \mathbb{R}^n$ then $|p_{T_xM} - p_{\mathbb{R}^n}|^2$ is just $2\sum_{j=1}^{k} |\nabla^M x^{n+j}|^2$, so that in this case

22.1 $$E(\xi,\rho,T) = \rho^{-n} \int_{B_\rho(\xi)} \sum_{j=1}^{k} |\nabla^M x^{n+j}|^2 \, d\mu_V$$

(∇^M = gradient operator on M as defined in §12.)

---

* $|p_{T_xM} - p|^2$ denotes the inner product norm trace $(p_{T_x} - p)^2$; this differs from $\|p_{T_xM} - p\|^2$ by at most a constant factor depending on $n+k$.

In this section we continue to assume V has generalized mean curvature $\underline{\underline{H}} \in L^1_{loc}(\mu_V)$ in U, and we write μ for μ_V.

We shall need the following simple lemma relating tilt-excess and height; note that we do not need $\theta \geq 1$ for this.

22.2 LEMMA *Suppose* $B_\rho(\xi) \subset U$. *Then for any* n*-dimensional subspace* $T \subset \mathbb{R}^{n+k}$ *we have*

$$E(\xi,\rho/2,T) \leq c\left[\rho^{-n}\int_{B_\rho(\xi)}\left(\frac{\operatorname{dist}(x-\xi,T)}{\rho}\right)^2 d\mu + \rho^{2-n}\int_{B_\rho(\xi)}|\underline{\underline{H}}|^2\, d\mu\right].$$

22.3 REMARK Note that in case $\rho^{-n}\mu(B_\rho(\xi)) \leq c$, we can use the Hölder inequality to estimate the term $\int_{B_\rho(\xi)}|\underline{\underline{H}}|^2\, d\mu$, giving

$\rho^{2-n}\int_{B_\rho(\xi)}|\underline{\underline{H}}|^2 d\mu \leq c\left[\left(\int_{B_\rho(\xi)}|\underline{\underline{H}}|^p d\mu\right)^{1/p}\rho^{1-n/p}\right]^2$, $p > 2$. Thus 22.2 gives

$$E(\xi,\rho/2,T) \leq c\left[\rho^{-n}\int_{B_\rho(\xi)}\left(\frac{\operatorname{dist}(x-\xi,T)}{\rho}\right)^2 d\mu + \left(\left(\int_{B_\rho(\xi)}|\underline{\underline{H}}|^p d\mu\right)^{1/p}\rho^{1-n/p}\right)^2\right],$$

$p \geq 2$.

Proof of 22.2 It evidently suffices to prove the result with $\xi = 0$ and $T = \mathbb{R}^n$. The proof simply involves making a suitable choice of X in the formula of 16.5. In fact we take

$$X_x = \zeta^2(x)x', \quad x' = (0, x^{n+1}, \ldots, x^{n+k})$$

for $x = (x^1, \ldots, x^{n+k}) \in U$, where $\zeta \in C_0^\infty(U)$ will be chosen.

By the definition of div_M (see §12) we have

$$\operatorname{div}_M x' = \sum_{i=n+1}^{n+k} e^{ii}, \quad \mu\text{-a.e. } x \in M,$$

where (e^{ij}) is the matrix of the projection p_{T_xM} (relative to the usual orthonormal basis for $\mathbb{R}^{n+k}$). Thus by the definition 16.5 of $\underline{\underline{H}}$ we have

(1) $$\int \sigma\zeta^2\, d\mu = \int \left(-2\zeta \sum_{i=n+1}^{n+k} \sum_{j=1}^{n+k} x^i e^{ij} D_j\zeta + \zeta^2 x'\cdot\underline{\underline{H}}\right) d\mu ,$$

with

(2) $$\sigma \equiv \sum_{i=n+1}^{n+k} e^{ii} = \frac{1}{2}\sum_{i,j=1}^{n+k} (e^{ij}-\varepsilon^{ij})^2 = \frac{1}{2}\,|p_{T_xM}-p_{\mathbb{R}^n}|^2 ,$$

where (ε^{ij}) = matrix of $p_{\mathbb{R}^n}$ and where we used $(e^{ij})^2 = (e^{ij})$ and $\text{trace}(e^{ij}) = n$. We thus have for $\zeta \geq 0$

$$\int \sigma\zeta^2\, d\mu \leq \int (2\sqrt{\sigma}|x'|\,|\text{grad}\,\zeta|\zeta + |x'|\,|\underline{\underline{H}}|\zeta^2)\, d\mu ,$$

and hence (using $ab \leq \frac{1}{2}a^2 + \frac{1}{2}b^2$)

$$\int \sigma\zeta^2\, d\mu \leq 4\int (|x'|^2|\text{grad}\,\zeta|^2 + |x'|\,|\underline{\underline{H}}|\zeta^2)\, d\mu ,$$

The lemma now follows by choosing $\zeta \equiv 1$ in $B_{\rho/2}(0)$, $\zeta \equiv 0$ outside $B_\rho(0)$, and $|\text{grad}\,\zeta| \leq 3/\rho$, and then noting that $|x'|\,|\underline{\underline{H}}| = (\rho^{-1}|x'|)\,(|\underline{\underline{H}}|\rho) \leq \frac{1}{2}\rho^{-2}|x'|^2 + \frac{1}{2}(|\underline{\underline{H}}|\rho)^2$.

We are now ready to discuss the following *tilt-excess decay theorem*, which is the main result concerning tilt-excess needed for the regularity theorem of the next section. (The Lipschitz approximation result of the previous section will play an important rôle in the proof.)

In order to state this result in a convenient manner, we let $\varepsilon,\alpha \in (0,1)$, $\rho > 0$, $p > n$, and T, an n-dimensional subspace of $\mathbb{R}^{n+k}$, be fixed, and we shall consider the hypotheses

$$\text{22.4}\qquad \begin{cases} 1 \le \theta \le 1+\varepsilon \qquad \mu\text{-a.e. in } U \\ \xi \in \operatorname{spt}\mu\,,\ B_\rho(\xi) \subset U\,,\ \dfrac{\mu(B_\rho(\xi))}{\omega_n\rho^n} \le 2(1-\alpha)\,, \\ E_*(\xi,\rho,T) \le \varepsilon\,, \end{cases}$$

where $E_*(\xi,\rho,T) = \max\left\{E(\xi,\rho,T)\ ,\ \varepsilon^{-1}\left(\int_{B_\rho(\xi)} |\underline{\underline{H}}|^p d\mu\right)^{2/p} \rho^{2(1-n/p)}\right\}$.

22.5 THEOREM *For any* $\alpha \in (0,1)$, $p > n$ *there are constants* $\eta,\varepsilon \in (0,1/2)$, *depending only on* n, k, α, p , *such that if hypotheses* 22.4 *hold, then*

$$E_*(\xi,\eta\rho,S) \le \eta^{2(1-n/p)} E_*(\xi,\rho,T)$$

for some n-*dimensional subspace* $S \subset \mathbb{R}^{n+k}$

22.6 REMARK Notice that any such S automatically satisfies

$$(*)\qquad |p_S - p_T|^2 \le c(\eta) E_*(\xi,\rho,T)\ .$$

Indeed we trivially have

$$(\eta\rho)^{-n} \int_{B_{\eta\rho}(\xi)} |p_{T_xM} - p_T|^2 d\mu \le \eta^{-n} E(\xi,\rho,T)\ ,$$

while by 22.5 we have

$$(\eta\rho)^{-n} \int_{B_{\eta\rho}(\xi)} |p_{T_xM} - p_S|^2 d\mu \le E_*(\xi,\rho,T)\ ,$$

and hence by adding these inequalities and using the fact that $\mu(B_{\eta\rho}(\xi)) \ge c\rho^n$ (see 19.6) we get (*) as required.

Proof of Theorem 22.5 Throughout the proof, $c = c(n,k,\alpha,p)$. We can suppose $\xi = 0$, $T = \mathbb{R}^n$. By the Lipschitz approximation theorem 20.2 there is a $\beta = \beta(n,k,\alpha,p) > 0$ and a Lipschitz function $f : B^n_{\beta\rho}(0) \to \mathbb{R}^k$ with

(1) $$\operatorname{Lip} f \le 1 \ , \ \sup|f| \le c\, E_*^{\frac{1}{2n+2}} \rho \le c\, \varepsilon^{\frac{1}{2n+2}} \rho$$

and

(2) $$\mathcal{H}^n(((\operatorname{spt}\mu \sim \operatorname{graph} f) \cup (\operatorname{graph} f \sim \operatorname{spt}\mu)) \cap B_{\beta\rho}(0)) \le c\, E_* \rho^n \ ,$$

where $E_* = E_*(0,\rho,\mathbb{R}^n) \left(\equiv \max\left\{ \rho^{-n} \int_{B_\rho(0)} |p_{T_xM} - p_{\mathbb{R}^n}|^2 d\mu \ , \right.\right.$ $\left.\left. \varepsilon^{-1} \left(\int_{B_\rho(0)} |\underline{\underline{H}}|^p d\mu \right)^{2/p} \rho^{2(1-n/p)} \right\}\right)$. Furthermore by the height estimate (3) in the proof of 20.2 we have

(3) $$\sup_{B_{\beta\rho}(0)\cap \operatorname{spt}\mu} |x^j| \le c\, E_*^{\frac{1}{2n+2}} \rho \le c\, \varepsilon^{\frac{1}{2n+2}} \rho \ ,$$

$j = n+1,\ldots,n+k$. Let us agree that

(4) $$c\, \varepsilon^{\frac{1}{2n+2}} \le \beta/4 \qquad (c \text{ as in } (3)) \ .$$

Then (3) implies

(5) $$\sup_{B_{\beta\rho}(0)\cap \operatorname{spt}\mu} |x^j| \le \beta\rho/4 \ ,$$

so that

(6) $$\mathbb{R}^k \times B^n_{\beta\rho/2}(0) \cap \operatorname{spt}\mu \cap \partial B_{\beta\rho}(0) = \emptyset \ .$$

Our aim now is to prove that each component of the approximating function f is well-approximated by a harmonic function. Preparatory to this, note that the defining identity for $\underline{\underline{H}}$ (see 16.5), with $X = \zeta\, e_{n+j}$, implies

$$\int_M \nabla^M_{n+j} \zeta \, d\mu = - \int e_{n+j} \cdot \underline{\underline{H}}\, \zeta \, d\mu \ , \quad \zeta \in C^1_0(U) \ ,$$

$j = 1,\ldots,k$, where $\nabla^M_{n+j} = e_{n+j} \cdot \nabla^M = p_{T_xM}(e_{n+j}) \cdot \nabla^M = (\nabla^M x^{n+j}) \cdot \nabla^M$ (∇^M = gradient operator for M as in §12) . Thus we can write

(7) $$\int_M (\nabla^M x^{n+j}) \cdot \nabla^M \zeta \, d\mu = - \int_M e_{n+j} \cdot \underline{\underline{H}}\, \zeta \, d\mu \ .$$

Since $x^{n+j} \equiv \tilde{f}^j(x)$ on $M_1 = M \cap \text{graph } f$ (where $\tilde{f}^j$ is defined on $\mathbb{R}^{n+k}$ by $\tilde{f}^j(x^1,\ldots,x^{n+k}) = f^j(x^1,\ldots,x^n)$ for $x = (x^1,\ldots,x^{n+k}) \in \mathbb{R}^{n+k}$), we have by the definition of ∇^M (see §12) that

$$(8) \qquad \nabla^M x^{n+j} = \nabla^M \tilde{f}^j(x) \quad \mu \text{ a.e. } x \in M_1 = M \cap \text{graph } f .$$

Hence (7) can be written

$$\int_{M_1} (\nabla^M \tilde{f}^j) \cdot \nabla^M \zeta \, d\mu = - \int_{M \sim M_1} (\nabla^M x^{n+j}) \cdot \nabla^M \zeta \, d\mu - \int_M e_{n+j} \cdot \underline{\underline{H}} \zeta \, d\mu ,$$

and hence by (2), together with the fact that (by 22.4)

$$(9) \qquad \int_{B_\rho(\xi)} |\underline{\underline{H}}| \, d\mu \le \left(\int_{B_\rho(\xi)} |\underline{\underline{H}}|^p d\mu \right)^{1/p} (\mu(B_\rho(\xi)))^{1-1/p} \le c\varepsilon^{\frac{1}{2}} E_*^{\frac{1}{2}} \rho^{n-1} ,$$

we obtain

$$(10) \qquad \rho^{-n} \int_{M_1} (\nabla^M \tilde{f}^j) \cdot \nabla^M \zeta \, d\mu \le c(\rho^{-1} \sup|\zeta| + \sup|\text{grad}\zeta|) \varepsilon^{\frac{1}{2}} E_*^{\frac{1}{2}}$$

$$\le c \sup|\text{grad}\zeta| \varepsilon^{\frac{1}{2}} E_*^{\frac{1}{2}} ,$$

for any smooth ζ with $\text{spt } \zeta \subset B_{\beta\rho}(0)$.

Furthermore by (8), 22.1, we evidently have

$$(11) \qquad \rho^{-n} \int_{M_1 \cap B_{\beta\rho}(0)} |\nabla^M \tilde{f}^j|^2 d\mu \le E_* .$$

Now suppose that ζ_1 is an arbitrary $C_c^1(B^n_{\beta\rho/2}(0))$ function, and note that (by (6)) there is a function $\zeta \in C_c^1(B_{\beta\rho}(0))$ such that $\zeta \equiv \tilde{\zeta}_1$ in some neighbourhood of $B^n_{\beta\rho/2}(0) \times \mathbb{R}^k \cap \text{spt } \mu \cap B_{\beta\rho}(0)$ where $\tilde{\zeta}_1(x^1,\ldots,x^{n+k}) \equiv \zeta_1(x^1,\ldots x^n)$. Hence (10) holds with $\tilde{\zeta}_1$ in place of ζ . Also, since $p_{\mathbb{R}^n} \text{grad } \tilde{\zeta}_1 = \text{grad } \tilde{\zeta}_1$ and $p_{\mathbb{R}^n} \text{grad } \tilde{f}^j = \text{grad } \tilde{f}^j$, we have

(12) $$|\nabla^M \tilde{f}^j \cdot \nabla^M \tilde{\zeta}_1 - \text{grad } f^j \cdot \text{grad } \zeta_1|$$

$$\equiv |p_{(T_xM)^\perp}(\text{grad } \tilde{f}^j) \cdot p_{(T_xM)^\perp}(\text{grad } \tilde{\zeta}_1)|$$

$$\equiv \left|\left(p_{(T_xM)^\perp} \circ p_{\mathbb{R}^n}(\text{grad } \tilde{f}^j)\right) \cdot \left(p_{(T_xM)^\perp} \circ p_{\mathbb{R}^n}(\text{grad } \tilde{\zeta}_1)\right)\right|$$

$$\leq |p_{(T_xM)^\perp} \circ p_{\mathbb{R}^n}|^2 |\text{grad } \tilde{f}^j| \; |\text{grad } \tilde{\zeta}_1|$$

$$\leq |p_{T_xM} - p_{\mathbb{R}^n}|^2 |\text{grad } \tilde{f}^j| \; |\text{grad } \tilde{\zeta}_1|$$

μ-a.e. on $\text{spt } \mu \cap B_{\beta\rho}(0)$, and hence (10) implies

(13) $$|\rho^{-n} \int_{M_1} \text{grad } \tilde{f}^j \cdot \text{grad } \tilde{\zeta}_1 \, d\mu \; | \leq c \, \varepsilon^{\frac{1}{2}} E_*^{\frac{1}{2}} \sup|\text{grad } \zeta_1| \; .$$

Also since (12) is valid with $\zeta_1 = \tilde{f}^j$, we conclude from (11) that

(14) $$\rho^{-n} \int_{M_1 \cap B_{\beta\rho}} |\text{grad } \tilde{f}^j|^2 \, d\mu \leq c \, E_* \; .$$

From (13), (14) and the area formula 8.5 we then have (using also (1),(2))

(15) $$|\rho^{-n} \int_{B^n_{\beta\rho}(0)} \text{grad } f^j \cdot \text{grad } \zeta_1 \quad \theta \circ F \quad J(F) \; dL^n|$$

$$\leq c \, \varepsilon^{\frac{1}{2}} E_*^{\frac{1}{2}} \sup|\text{grad } \zeta_1|$$

and

(16) $$\rho^{-n} \int_{B^n_{\beta\rho}(0)} |\text{grad } f^j|^2 \; \theta \circ F \quad J(F) \; dL^n \leq c \, E_* \; ,$$

where $F : \mathbb{R}^n \to \mathbb{R}^{n+k}$ is defined by $F(x) = (x, f(x))$, $x \in B^n_{\beta\rho}(0)$, and where $J(F)$ is the Jacobian $(\det((dF_x)^* \circ dF_x))^{1/2}$ as in §8 . Since $1 \leq J(F) \leq 1 + c|\text{grad } f|^2$ on $B^n_{\beta\rho}(0)$, as one checks by directly computing the matrix of dF_x (relative to the usual orthonormal bases for $\mathbb{R}^n$, $\mathbb{R}^{n+k}$) in terms of the partial derivatives of f , and since $1 \leq \theta \leq 1+\varepsilon$, we conclude

$$(17)\qquad \left|\rho^{-n}\int_{B^n_{\beta\rho}(0)} \operatorname{grad} f^j \cdot \operatorname{grad} \zeta_1 \, dL^n\right|$$

$$\le c\left(\varepsilon^{\frac{1}{2}}E_*^{\frac{1}{2}} + \varepsilon\rho^{-n}\int_{B^n_{\beta\rho/2}(0)} |\operatorname{grad} f^j|\, dL^n\right)\sup|\operatorname{grad} \zeta_1|$$

$$\le c\ \varepsilon^{\frac{1}{2}}E_*^{\frac{1}{2}}\ \sup|\operatorname{grad} \zeta_1|$$

by (16), because by (16) (and the fact that $\theta \ge 1$, $J(F) \ge 1$) we have

$$(18)\qquad \rho^{-n}\int_{B^n_{\beta\rho}(0)} |\operatorname{grad} f^j|^2\, dL^n \le c\ E_* \ .$$

Now (17), (18) and the harmonic approximation lemma 21.1 (with $E_*^{-\frac{1}{2}}f^j$ in place of f) we know that for any given $\delta > 0$ there is $\varepsilon_0 = \varepsilon_0(n,k,\delta)$ such that, if the hypotheses 22.1 hold with $\varepsilon \le \varepsilon_0$, there are harmonic functions $u^1,\ldots,u^k$ on $B_{\beta\rho/2}(0)$ such that

$$(19)\qquad \sigma^{-n}\int_{B^n_\sigma(0)} |\operatorname{grad} u^j|^2\, dL^n \le c\ E_*$$

and

$$(20)\qquad \sigma^{-n-2}\int_{B^n_\sigma(0)} |f^j-u^j|^2\, dL^n \le \delta\ E_* \ ,$$

where $\sigma = \beta\rho/2$.

Using 21.3 we then conclude that

$$(21)\qquad (\eta\sigma)^{-n-2}\int_{B^n_{\eta\sigma}(0)} |f^j-\ell^j|^2\, dL^n \le 2\eta^{-n-2}\ \delta\ E_* +$$

$$+\ c\eta^2\sigma^{-n}\int_{B^n_\sigma(0)} |\operatorname{grad} u^j|^2\, dL^n$$

$$\le\ 2\eta^{-n-2}\ \delta\ E_* + c\eta^2\ E_* \quad \text{(by (19))} \ ,$$

where $\ell^j(x) = u^j(0) + x\cdot\text{grad } u^j(0)$. Notice that, since $\sup|f| \le c\, E_*^{\frac{1}{2n+2}}\rho$, (19), (20) in particular imply (using 21.3 again)

(22) $$\sum_{j=1}^{k} |\ell^j(0)| \le c\, E_*^{\frac{1}{2n+2}}\rho \le c\, \varepsilon^{\frac{1}{2n+2}}\rho .$$

Now let $\ell = (\ell^1,\ldots,\ell^k) : \mathbb{R}^n \to \mathbb{R}^k$ and let S be the n-dimensional subspace graph $(\ell-\ell(0))$. In view of (1), (2), (3) and (22) it is clear that (21) implies

(23) $$(\eta\sigma)^{-n-2}\int_{B_{\eta\sigma/2}(\tau)} \text{dist}(x-\tau,S)^2 d\mu \le c\eta^{-n-2}\delta E_* + c\eta^2 E_* ,$$

where $\tau = (0,\ell(0))$, provided $c\varepsilon^{\frac{1}{2n+2}} < \eta/2$. Then by 22.3 we get

$$E(\tau,\eta\sigma/2,S) \le c\eta^{-n-2}\delta E_* + c\eta^2 E_* .$$

If we in fact require

(24) $$(1+c)\varepsilon^{\frac{1}{2n+2}} < \eta \qquad (c \text{ as in } (22)) .$$

then $B_{\eta\sigma/4}(0) \subset B_{\eta\sigma/2}(\tau)$ (by (22)) and this gives

(25) $$E(0,\eta\sigma/4,S) \le c\eta^{-n-2}\delta E_* + c\eta^2 E_* .$$

The proof of the theorem is now completed as follows:

First select η so that $c\eta^2 \le \frac{1}{2}(\eta\beta/8)^{2(1-n/p)}$ (with c as in (25)), then select δ so that $c\eta^{-n-2}\delta < \frac{1}{2}(\eta\beta/8)^{2(1-n/p)}$ (c again as in (25)). Then, provided the hypotheses 22.4 hold with ε satisfying the conditions required during the above argument (in particular (4), (24) must hold, and $\varepsilon \le \varepsilon_0(n,k,\delta)$, $\varepsilon_0(n,k,\delta)$ as in the discussion leading to (19)) we get

$$E(0,\tilde{\eta}\rho,S) \le \tilde{\eta}^{2(1-n/p)} E_* ,$$

where $\tilde{\eta} = \eta\beta/8$. Since we trivially have

$$\left(\int_{B_{\tilde{\eta}\rho}(0)} |\underline{\underline{H}}|^p d\mu\right)^{1/p} (\tilde{\eta}\rho)^{1-n/p} \le \tilde{\eta}^{1-n/p}\left(\int_{B_\rho(0)} |\underline{\underline{H}}|^p d\mu\right)^{1/p} \rho^{1-n/p}$$

we thus conclude that

$$E_*(0,\tilde{\eta}\rho,S) \le \tilde{\eta}^{2(1-n/p)} E_*(0,\rho,T)$$

as required.

This completes the proof of 22.5 (with $\tilde{\eta}$ in place of η).

§23. MAIN REGULARITY THEOREM: FIRST VERSION

We here show that one useful form of Allard's theorem follows very directly from the tilt-excess decay theorem 22.5 of the previous section.

23.1 THEOREM *Suppose* $\alpha \in (0,1)$ *and* $p > n$ *are given. There are constants* $\varepsilon = \varepsilon(n,k,\alpha,p)$, $\gamma = \gamma(n,k,\alpha,p) \in (0,1)$ *such that if hypotheses* 22.4 *hold with* $T = \mathbb{R}^n$ *and* $\xi = 0$, *then there is a* $C^{1,1-n/p}$ *function* $u = (u^1,\ldots,u^k) : \bar{B}^n_{\gamma\rho}(0) \to \mathbb{R}^k$ *such that* $u(0) = 0$,

(1) $$\operatorname{spt} V \cap B_{\gamma\rho}(0) = \operatorname{graph} u \cap B_{\gamma\rho}(0) ,$$

and

(2) $$\rho^{-1}\sup|u| + \sup|Du| + \rho^{1-n/p}\sup_{\substack{x,y\in B^n_{\gamma\rho}(0)\\ x\neq y}} |x-y|^{-(1-n/p)}|Du(x)-Du(y)|$$

$$\le c\left\{E^{\frac{1}{2}}(0,\rho,\mathbb{R}^n) + \left(\int_{B_\rho(0)} |\underline{\underline{H}}|^p d\mu\right)^{1/p} \rho^{1-n/p}\right\} .$$

Before giving the proof we make a couple of important remarks concerning removability of the hypothesis $\theta \le 1+\varepsilon$ in 22.4: *

23.2 REMARKS

(1) The monotonicity formula in 17.6(1), together with Remark 17.9(1), evidently implies that if $\left(\omega_n^{-1}\int_{B_\rho(\xi)}|\underline{\underline{H}}|^p d\mu\right)^{1/p}\rho^{1-n/p} \le \varepsilon < \frac{1}{2}$, then, for $0<\sigma<\tau<(1-\beta)\rho$

$$(*)\qquad \omega_n^{-1}\sigma^{-n}\mu(B_\sigma(\zeta)) \le (1+c\varepsilon)\omega_n^{-1}\tau^{-n}\mu(B_\tau(\zeta))$$

$$\le (1+c\varepsilon)^2\,\omega_n^{-1}((1-\beta)\rho)^{-n}\,\mu(B_{(1-\beta)\rho}(\zeta))$$

$$\le (1+c\varepsilon)^2\,(1-\beta)^{-n}\,\omega_n^{-1}\,\rho^{-n}\,\mu(B_\rho(\xi))$$

provided $\zeta \in \operatorname{spt} V \cap B_{\beta\rho}(\xi)$. Then the hypothesis $\omega_n^{-1}\rho^{-n}\mu(B_\rho(\xi)) \le 2(1-\alpha)$ (in 22.4) gives

$$(**)\qquad \omega_n^{-1}\sigma^{-n}\mu(B_\sigma(\zeta)) \le 2(1-\alpha/2)\ ,\quad 0<\sigma<\rho/2\ ,\ \zeta\in \operatorname{spt} V\cap B_{\beta\rho}(\xi)\ ,$$

provided $\beta = \beta(n,k,\alpha,p)$ is sufficiently small. Thus letting $\sigma \downarrow 0$ we have

$$\theta(\zeta) \le 2(1-\alpha/2)\qquad \mu\text{-a.e. } \zeta \in B_{\beta\rho}(\xi)\ .$$

If θ is integer-valued (i.e. if V is an integer multiplicity n-varifold) then this evidently implies $\theta = 1$ μ-a.e. in $B_{\beta\rho}(\xi)$. Thus, with $\beta\rho$ in place of ρ , the hypothesis $\theta \le 1+\varepsilon$ in 22.4 is automatically satisfied, hence the conclusion of Theorem 23.1 holds with $\beta\rho$ in place of ρ , even without the hypothesis $\theta \le 1+\varepsilon$.

---

* J. Duggan in his Ph.D. thesis [DJ] has shown that in any case the hypothesis $\theta \le 1+\varepsilon$ can be dropped entirely.

(2) Quite generally, even if θ is not necessarily integer-valued, we note that if we make the stronger hypothesis $\omega_n^{-1}\rho^{-n}\mu(B_\rho(\xi)) < 1+\varepsilon$ (instead of $\omega_n^{-1}\rho^{-n}\mu(B_\rho(\xi)) \le 2(1-\alpha))$, then (*) above gives (taking $\beta=\varepsilon$)

$$\omega_n^{-1}\sigma^{-n}\mu(B_\sigma(\zeta)) \le 1+c\varepsilon \ , \ 0<\sigma<\rho/2 \ , \ \zeta\in B_{\varepsilon\rho}(\xi)\cap \operatorname{spt} V \ .$$

Thus again we can drop the restriction $\theta \le 1+\varepsilon$ in 22.4, provided we make the assumption $\omega_n^{-1}\sigma^{-n}\mu(B_\rho(\xi)) < 1+\varepsilon$; then Theorem 23.1 holds with $\varepsilon\rho$ in place of ρ.

Proof of 23.1 Throughout the proof $c = c(n,k,\alpha,p) > 0$. We are assuming

(1) $$1\le\theta\le 1+\varepsilon \quad \mu\text{-a.e. in } B_\rho(0)\cap\operatorname{spt} V$$

(ε to be chosen) and by Remark 23.2(1) (**) we can select $\varepsilon=\varepsilon(n,k,\alpha,p)$ and $\beta=\beta(n,k,\alpha,p)$ such that

(2) $$\omega_n^{-1}\sigma^{-n}\mu(B_\sigma(\zeta)) \le 2(1-\alpha/2) \ , \ 0<\sigma\le\rho/2 \ , \ \zeta\in B_{\beta\rho}(0)\cap\operatorname{spt} V \ .$$

By (1), (2) and the tilt excess decay theorem 22.5 (with σ in place of ρ, $\alpha/2$ in place of α, ζ in place of ξ) we then know that there are $\varepsilon=\varepsilon(n,k,\alpha,p)$, $\eta=\eta(n,k,\alpha,p)$ so that, for $\sigma<\rho/2$, $\zeta\in \operatorname{spt} B_{\beta\rho}(0)\cap\operatorname{spt} V$,

(3) $$E_*(\zeta,\sigma,S_0) < \varepsilon \ \Rightarrow \ E_*(\zeta,\eta\sigma,S_1) < \eta^{2(1-n/p)} E_*(\zeta,\sigma,S_0)$$

for suitable S_1. Notice that this can be repeated; by induction we prove that if $\zeta\in\operatorname{spt} V\cap B_{\beta\rho}(0)$, $\sigma<\rho/2$, and $E_*(\zeta,\sigma,S_0)<\varepsilon$, then there is a sequence $S_1,S_2\ldots$ of n-dimensional subspaces such that

(4) $$E_*(\zeta,\eta^j\sigma,S_j) \le \eta^{2(1-n/p)}E_*(\zeta,\eta^{j-1}\sigma,S_{j-1}) \le \eta^{2(1-n/p)j}E_*(\zeta,\sigma,S_0)$$

for each $j\ge 1$, and (by Remark 22.6)

(5) $$|p_{S_j}-p_{S_{j-1}}|^2 \le cE_*(\zeta,\eta^{j-1}\sigma,S_{j-1}) \le c\eta^{2(1-n/p)j}E_*(\zeta,\sigma,S_0) \ .$$

Next we note that $E_*(\zeta,\rho/2,\mathbb{R}^n) \le 2^n E_*(0,\rho,\mathbb{R}^n)$ for $\zeta \in B_{\rho/2}(0)$, and hence the above discussion shows that if 22.4 holds with $\xi = 0$, $T = \mathbb{R}^n$ and $2^{-n}\varepsilon$ in place of ε (ε as above) then (4), (5) hold with $S_0 = \mathbb{R}^n$ and $\sigma = \rho/2$. Thus

(6) $$E(\zeta,\eta^j\rho/2,S_j) \le \eta^{2(1-n/p)j} E_*(\zeta,\rho/2,\mathbb{R}^n) \le c\eta^{2(1-n/p)j} E_*(0,\rho,\mathbb{R}^n)$$

and

(7) $$|p_{S_j}-p_{S_{j-1}}|^2 \le c\eta^{2(1-n/p)j} E_*(0,\rho,\mathbb{R}^n)$$

for each $j \ge 1$ (with $S_0 = \mathbb{R}^n$) . Notice that (7) gives

(8) $$|p_{S_\ell}-p_{S_j}|^2 \le c\eta^{2(1-n/p)j} E_*(0,\rho,\mathbb{R}^n)$$

for $\ell \ge j \ge 0$. It evidently follows from (8) that there is $S(\zeta)$ such that

(9) $$|p_{S(\zeta)}-p_{S_j}|^2 \le c\eta^{2(1-n/p)j} E_*(0,\rho,\mathbb{R}^n) \ .$$

In particular (setting $j=0$)

(10) $$|p_{S(\zeta)}-p_{\mathbb{R}^n}|^2 \le c\, E_*(0,\rho,\mathbb{R}^n) \ .$$

Now if $r \in (0,\rho)$ is arbitrary we can choose $j \ge 0$ such that $\eta^{j+1}\rho < r \le \eta^j\rho$. Then (6) and (9) evidently imply

(11) $$E_*(\zeta,r,S(\zeta)) \le c(r/\rho)^{2(1-n/p)} E_*(0,\rho,\mathbb{R}^n)$$

for each $\zeta \in B_{\beta\rho}(0) \cap \operatorname{spt} V$ and each $0<r\le\rho$. Notice also that (10), (11) and (2), with $\sigma = r$, imply

(12) $$E_*(\zeta,r,\mathbb{R}^n) \le c\, E_*(0,\rho,\mathbb{R}^n) (\le c\varepsilon) \ .$$

Hence for sufficiently small ε we have from (12) that if G is as in the proof of Theorem 20.2 (with $\ell = \varepsilon^{\frac{1}{2n+3}}$) then $\mu(B_{\beta\rho} \sim G) = 0$ $(\beta = \beta(n,k,\alpha,p)$, $\varepsilon = \varepsilon(n,k,\alpha,p)$ sufficiently small). That is

$$(13) \qquad \operatorname{spt} V \cap B_{\beta\rho}(0) \subset \operatorname{graph} f$$

for $\varepsilon = \varepsilon(n,k,\alpha,p)$ and $\beta = \beta(n,k,\alpha,p)$ sufficiently small, where f is a Lipschitz function $B^n_{\beta\rho}(0) \to \mathbb{R}^k$ with

$$(14) \qquad \operatorname{Lip} f \le \varepsilon^{\frac{1}{2n+3}}, \quad \sup|f| \le c\, \varepsilon^{\frac{1}{2n+2}} \rho .$$

Now we claim that in fact

$$(15) \qquad \operatorname{spt} V \cap B_{\beta\rho}(0) = \operatorname{graph} f \cap B_{\beta\rho}(0) .$$

Indeed otherwise by (13) we could choose $\zeta \in B^n_{\beta\rho/2}(0)$ and $0 < \sigma < \beta\rho/2$ such that

$$(16) \qquad \begin{cases} (B^n_\sigma(\zeta) \times \mathbb{R}^k) \cap B_{\beta\rho}(0) \cap \operatorname{spt} V = \emptyset \\ (\bar{B}^n_\sigma(\zeta) \times \mathbb{R}^k) \cap B_{\beta\rho}(0) \cap \operatorname{spt} V \ne \emptyset . \end{cases}$$

Then taking $\zeta_* \in (\bar{B}^n_\sigma(\zeta) \times \mathbb{R}^k) \cap B_{\beta\rho}(0) \cap \operatorname{spt} V$ and using (1), (13), (14), (16) we would evidently have $\Theta^n(\mu,\zeta_*) < 1$ (if ε is sufficiently small). This contradicts the fact that $\Theta^n(\mu,\zeta) \ge 1 \quad \forall\, \zeta \in \operatorname{spt} V \cap B_\rho(0)$.

Having established (15) we can now easily check (using the area formulae) that for any linear subspace $S = \operatorname{graph} \ell$, where $\ell = (\ell^1, \dots, \ell^k) : \mathbb{R}^n \to \mathbb{R}^k$ is linear and $|\operatorname{grad} \ell^j| \le 1$ for each j , we have

$$(17) \qquad \sigma^{-n} \int_{B^n_{\sigma/2}(p_{\mathbb{R}^n}(\zeta))} \sum_{j=1}^k |\operatorname{grad} f^j(x) - \operatorname{grad} \ell^j|^2 \, dL^n(x) \le c\, E(\zeta,\sigma,S)$$

for $\sigma \in (0,\beta\rho/2)$ (again provided ε in (14) is sufficiently small). Using

(17) and (11) we conclude that for each $\zeta \in B_{\beta\rho/2}(0) \cap \operatorname{spt} V$ there is a linear function $\ell_\zeta = (\ell_\zeta^1, \ldots, \ell_\zeta^k) : \mathbb{R}^n \to \mathbb{R}^k$ such that

$$(18) \qquad r^{-n} \int_{B_r^n(p_{\mathbb{R}^n}(\zeta))} \sum_{j=1}^{k} |\operatorname{grad} f^j(x) - \operatorname{grad} \ell_\zeta^j|^2 dL^n(x) \le c(r/\rho)^{2(1-n/p)} E_*(0,\rho,\mathbb{R}^n)$$

for $0 < r < \beta\rho/4$. It evidently follows from this, by letting $r \downarrow 0$ in (18), that $\operatorname{grad} f_j(p_{\mathbb{R}^n}(\zeta)) = \operatorname{grad} \ell_\zeta^j$ for μ-a.e. $\zeta \in \operatorname{spt} V \cap B_{\beta\rho/4}(0)$. Hence using (18) again, we easily conclude $|\operatorname{grad} f^j(x_1) - \operatorname{grad} f^j(x_2)| \le c(r/\rho)^{1-n/p} E_*(0,\rho,\mathbb{R}^n)^{\frac{1}{2}}$ for $x_1, x_2 \in B_r^n(0)$, and so

$$(19) \qquad |\operatorname{grad} f^j(x_1) - \operatorname{grad} f^j(x_2)| \le c\left(\frac{|x_1 - x_2|}{\rho}\right)^{1-n/p} E_*(0,\rho,\mathbb{R}^n)^{\frac{1}{2}}$$

for L^n-a.e. $x_1, x_2 \in B^n_{\beta\rho/8}(0)$. Since f is Lipschitz it follows from this that $f \in C^{1,1-n/p}$ with (19) holding for *every* $x_1, x_2 \in B^n_{\beta\rho/8}(0)$. The theorem now follows with $f = u$ and $\gamma = \beta/8$.

§24. MAIN REGULARITY THEOREM: SECOND VERSION

In this section we continue to assume $V = \underline{v}(M,\theta)$ is a rectifiable n-varifold with generalized mean curvature $\underline{H}$ in U. With $\delta \in (0,1/2)$ a constant to be specified below, we consider the hypotheses:

$$24.1 \qquad \begin{cases} 1 \le \theta \quad \mu\text{-a.e.}, \quad 0 \in \operatorname{spt} V, \quad B_\rho(0) \subset U \\ \omega_n^{-1}\rho^{-n}\mu(B_\rho(0)) \le 1+\delta, \quad \left(\int_{B_\rho(0)} |\underline{H}|^p d\mu\right)^{1/p} \rho^{1-n/p} \le \delta. \end{cases}$$

24.2 THEOREM *If* $p > n$ *is arbitrary, then there are* $\delta = \delta(n,k,p)$, $\gamma = \gamma(n,k,p) \in (0,1)$ *such that the hypotheses* 24.1 *imply the existence of a linear isometry* q *of* $\mathbb{R}^{n+k}$ *and a* $u = (u^1,\dots,u^k) \in C^{1,1-n/p}(\bar{B}^n_{\gamma\rho}(0);\mathbb{R}^k)$ *with* $u(0) = 0$, $\operatorname{spt} V \cap B_{\gamma\rho}(0) = q(\operatorname{graph} u) \cap B_{\gamma\rho}(0)$, *and*

$$\rho^{-1}\sup|u| + \sup|Du| + \rho^{1-n/p}\sup_{\substack{x,y\in B^n_{\gamma\rho}(0)\\ x\neq y}} |x-y|^{-(1-n/p)}|Du(x)-Du(y)| \le c\delta^{1/4n} ,$$

$c = c(n,k,p)$.

Before giving the proof of 24.2, we shall need the following lemma.

24.3 LEMMA *Suppose* $\delta \in (0,1/2)$ *and that* 24.1 *holds. Then there is* $\beta = \beta(n,k,p,\delta) \in (0,1/2)$ *such that*

$$(1+c\delta)^{-1} < \omega_n^{-1}\sigma^{-n}\mu(B_\sigma(\zeta)) < 1+c\delta \ , \quad 0<\sigma\le\beta\rho \ , \quad \zeta\in \operatorname{spt} V\cap B_{\beta\rho}(0)$$

and such that, for any $\zeta \in \operatorname{spt} V \cap B_{\beta\rho}(0)$, $\sigma \in (0,\beta\rho)$ *there is an* n*-dimensional subspace* $T = T(\zeta,\sigma)$ *with*

$$\sigma^{-1}\sup\{\operatorname{dist}(x,T) \,:\, x\in \operatorname{spt} V\cap B_\sigma(\zeta)\} \le c\delta^{1/4n} .$$

Proof First note that by the monotonicity formulae of §17 (see in particular 23.2(1)(*)) we have, subject to 24.1, that

$$(1) \qquad (1+c\delta)^{-1} \le \omega_n^{-1}\sigma^{-n}\mu(B_\sigma(\zeta)) \le 1+c\delta \ , \quad 0<\sigma<\rho/2 \ ,$$

$\zeta \in \operatorname{spt} V \cap B_{\beta\rho}(0)$, $\beta = \beta(n,k,p,\delta) \in (0,1/4)$, so the first part of the lemma is proved.

Now take any fixed $\sigma \in (0,\beta\rho)$ and suppose for convenience of notation (by changing scale and translating the origin) that $\sigma = 1/2$ and $\zeta = 0$. Then by (1) and 17.6(1) (see in particular Remark 17.9(1)) we have

$$(2)\qquad \int_{B_{1/2}(0)} |p_{T_x^\perp}(x-\zeta)|^2 \le \int_{B_1(\zeta)} |p_{T_x^\perp}(x-\zeta)|^2\, |x-\zeta|^{-n-2} d\mu \le c\delta$$

for $\zeta \in \operatorname{spt} V \cap B_{1/2}(0)$, where $T_x^\perp = (T_x M)^\perp$. Next note that we can select N points $\zeta_1,\ldots,\zeta_N \in \operatorname{spt} V \cap B_{1/2}(0) \sim B_{\delta^{1/4n}}(0)$, $N \le c\delta^{-1/4}$, such that

$$(3)\qquad \operatorname{spt} V \cap B_{1/2}(0) \sim B_{\delta^{1/4n}}(0) \subset \bigcup_{j=1}^{N} B_{\delta^{1/4n}}(\zeta_j) .$$

(To achieve this, just take a *maximal* disjoint collection of balls of radius $\delta^{1/4n}/4$ centred in $\operatorname{spt} V \cap B_{1/2}(0) \sim B_{\delta^{1/4n}}(0)$.) Then by using (2) with $\zeta = \zeta_j$ we have

$$\int_{B_{1/2}(0)} \sum_{j=1}^{N} |p_{T_x^\perp}(x-\zeta_j)|^2 d\mu \le c\,\delta\, N \le c'\delta^{1/2} ,$$

so that for any given $R \ge 1$ we have

$$(4)\qquad \sum_{j=1}^{N} |p_{T_x^\perp}(x-\zeta_j)|^2 \le R\, \delta^{1/4}$$

except possibly for a set of $x \in B_{1/2}(0) \cap \operatorname{spt} V$ of μ-measure $\le c\, R^{-1}\delta^{1/4}$. Taking $R = R(n,k)$ sufficiently large and noting $\mu(B_{\delta^{1/4n}}(0)) \ge c\delta^{1/4}$ (by (1)), we can therefore find $x_0 \in \operatorname{spt} V \cap B_{\delta^{1/4n}}(0)$ such that

$$|p_{T_{x_0}^\perp}(x_0-\zeta_j)| \le c\delta^{1/8} , \quad j = 1,\ldots,N .$$

Since $|x_0| < \delta^{1/4n}$, we then have

$$(5)\qquad |p_{T_{x_0}^\perp}\, \zeta_j| \le c\delta^{1/4n} , \quad j = 1,\ldots,N .$$

That is, all points $\zeta_1,\ldots,\zeta_N$ are in the $c\delta^{1/4n}$ neighbourhood of the subspace T_{x_0} , and the required result now follows from (3).

Proof of Theorem 24.2 Theorem 24.2 in fact now follows directly from Theorem 23.1, because by combining Lemma 22.2 and the above lemma we see that for any $\varepsilon > 0$ there is $\delta = c\,\varepsilon^{2n}$ $(c = c(n,k,p))$ such that the hypotheses 24.1 imply 22.4 with $\xi = 0$, ρ replaced by $\beta\rho$ and with suitable T.