Analysis 101:

Functions of Several Variables

ABSTRACT

Apart from an account of classical preliminaries, these notes contain a systematic introduction to Sobolev spaces and functions of bounded variation, along with selected applications.

FUNCTIONS OF SEVERAL VARIABLES

SECTION 1: MEASURE THEORY

§1.1. FACTS

Let X be a nonempty set and let $\mathcal{E} \subset \mathcal{P}(X)$ be a collection of subsets of X.

1.1.1. DEFINITION The pair $(X,\mathcal{E})$ is called a <u>measurable space</u> if $\mathcal{E}$ is a σ-algebra.

1.1.2 EXAMPLE If (X,τ) is a topological space, then $(X,\mathcal{B}(X))$ is a measurable space, $\mathcal{B}(X)$ the σ-algebra of Borel subsets of X, i.e., the σ-algebra generated by the open subsets of X.

1.1.3. DEFINITION Let $(X,\mathcal{E})$ be a measurable space.

- A function $\mu:\mathcal{E} \to [0, +\infty]$ is a <u>positive measure</u> provided $\mu(\emptyset) = 0$ and μ is σ-additive on $\mathcal{E}$.

1.1.4. LEMMA Let $(X,\mathcal{E})$ be a measurable space and suppose that $\mu:\mathcal{E} \to [0, +\infty]$ $(\mu(\emptyset) = 0)$ is σ-subadditive and additive -- then μ is σ-additive, hence μ is a positive measure.

PROOF Let $E_1,E_2,\ldots$ be a sequence of pairwise disjoint elements of $\mathcal{E}$ -- then

$$\begin{aligned}
\mu\Big(\bigcup_{n=1}^{\infty} E_n\Big) &\leq \sum_{n=1}^{\infty} \mu(E_n) \\
&\leq \lim_{N\to\infty} \sum_{n=1}^{N} \mu(E_n) \\
&= \lim_{N\to\infty} \mu\Big(\bigcup_{n=1}^{N} E_n\Big) \\
&\leq \mu\Big(\bigcup_{n=1}^{\infty} E_n\Big).
\end{aligned}$$

1.1.5. DEFINITION Let $(X,\mathcal{E})$ be a measurable space.

- A function $\mu:\mathcal{E} \to R^m$ $(m \geq 1)$ is a <u>vector measure</u> provided $\mu(\emptyset) = 0$ and μ is σ-additive.

[Note: If μ is a vector measure and if $m = 1$, then μ is a <u>real measure</u>. Since $+\infty$ is admitted, a positive measure is not necessarily a real measure.]

1.1.6. REMARK If μ is a vector measure and if $E_1,E_2,\ldots$ is a sequence of pairwise disjoint elements of $\mathcal{E}$, then

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n).$$

Here the series on the right is absolutely convergent since its sum does not depend on the order of its terms (this being the case of the union on the left).

1.1.7. DEFINITION Suppose that $\mu:\mathcal{E} \to R^m$ $(m \geq 1)$ is a vector measure -- then its <u>total variation</u> $||\mu||$ is the arrow $\mathcal{E} \to [0, +\infty]$ defined by the prescription

$$||\mu||(E) = \sup\left[\sum_{n=1}^{\infty} ||\mu(E_n)||:\{E_n\} \text{ pairwise disjoint, } E = \bigcup_{n=1}^{\infty} E_n\right].$$

1.1.8. THEOREM $||\mu||$ is a positive finite measure (hence $||\mu||(X) < +\infty$).

1.1.9. REMARK Denote by $M(X;R^m)$ the set of R^m-valued vector measures $\mu:X \to R^m$ $(m \geq 1)$ -- then $M(X;R^m)$ is a real vector space and the total variation is a norm on $M(X;R^m)$ under which it is a Banach space.

1.1.10. NOTATION Given a real measure μ, let

$$\begin{cases} \mu^+ = \dfrac{|\mu| + \mu}{2} & \text{(positive part)} \\[2ex] \mu^- = \dfrac{|\mu| - \mu}{2} & \text{(negative part).} \end{cases}$$

1.1.11. <u>N.B.</u> Therefore μ^+ and μ^- are positive finite measures and

$$\mu = \mu^+ - \mu^-,$$

the <u>Jordan decomposition</u> of μ.

1.1.12. SCHOLIUM If μ is an R^m-valued vector measure, say

$$\mu = (\mu_1, \ldots, \mu_m),$$

put

$$\int_X f \, d\mu = (\int_X f \, d\mu_1, \ldots, \int_X f \, d\mu_m).$$

Then

$$||\int_X f \, d\mu|| \leq \int_X |f| \, d||\mu||.$$

1.1.13. NOTATION Let μ be a positive measure on $(X, \mathcal{E})$. Given an $f \in L^1(X,\mu)^m$, say

$$f = (f_1, \ldots, f_m),$$

and an $E \in \mathcal{E}$, put

$$\int_E f \, d\mu = (\int_E f_1 \, d\mu, \ldots, \int_E f_m \, d\mu).$$

1.1.14. SUBLEMMA The assignment

$$E \to \int_E f \, d\mu$$

is an R^m-valued vector measure, call it $f\mu$.

1.1.15. LEMMA The total variation $||f\mu||$ of $f\mu$ is the assignment

$$E \to \int_E ||f|| \, d\mu.$$

PROOF First

$$||f\mu|| \leq ||f||\mu.$$

This said, fix a countable dense set $\{u_k\} \subset S^{m-1}$ $(\subset R^m)$ and let $E \in \mathcal{E}$. Given $\varepsilon > 0$, put

$$\sigma(x) = \min\{k \in N: \langle f(x), u_k\rangle \geq (1-\varepsilon)||f(x)||\}$$

and write

$$E_k = \sigma^{-1}(\{k\}) \cap E \in \mathcal{E}.$$

Then

$$(1-\varepsilon)||f||\mu(E) = (1-\varepsilon) \int_E ||f||d\mu = \sum_{k=1}^{\infty} (1-\varepsilon) \int_{E_k} ||f||d\mu$$

$$\leq \sum_{k=1}^{\infty} \langle f\mu(E_k), u_k\rangle$$

$$\leq \sum_{k=1}^{\infty} ||f\mu(E_k)|| < ||f\mu||(E).$$

1.1.16. DEFINITION Let $(X,\mathcal{E})$ be a measurable space.

- Let ν be a positive measure and let μ be a vector measure -- then μ is <u>absolutely continuous</u> w.r.t. ν, denoted $\mu << \nu$, if for every $E \in \mathcal{E}$, the implication

$$\nu(E) = 0 \Rightarrow ||\mu||(E) = 0$$

obtains.

1.1.17. EXAMPLE If $f \in L^1(X,\nu)^m$, then $f\nu <<< \nu$.

1.1.18. CRITERION μ is absolutely continuous w.r.t. ν iff for every $\varepsilon > 0$ there exists $\delta > 0$ such that for every $E \in \mathcal{E}$,

$$\nu(E) < \delta \Rightarrow ||\mu||(E) < \varepsilon.$$

1.1.19. DEFINITION Let $(X,\mathcal{E})$ be a measurable space.

- Let μ_1,μ_2 be positive measures -- then μ_1,μ_2 are <u>mutually singular</u>, denoted $\mu_1 \perp \mu_2$, if there exists $E \in \mathcal{E}$ such that

$$\mu_1(E) = 0 \text{ and } \mu_2(X - E) = 0.$$

1.1.20. N.B. Vector measures μ_1, μ_2 are mutually singular provided this is the case of $||\mu_1||$, $||\mu_2||$.

[Note: If ν, μ are as above, write $\nu \perp \mu$ when $\nu \perp ||\mu||$.]

1.1.21. RADON-NIKODYM Let $(X,\mathcal{E})$ be a measurable space.

- Let ν be a positive measure and let μ be a vector measure, say $\mu:\mathcal{E} \to R^m$ $(m \geq 1)$. Assume: ν is σ-finite -- then there is a unique pair μ^a, μ^s of R^m-valued vector measures such that

$$\mu^a << \nu, \quad \mu^s \perp \nu$$

and

$$\mu = \mu^a + \mu^s.$$

1.1.22. N.B. In addition, there is a unique $f \in L^1(X,\nu)^m$ such that $\mu^a = f\nu$, the so-called density of μ w.r.t. ν, denoted $\frac{d\mu}{d\nu}$.

[Note: Uniqueness is taken to mean in the sense of equivalence classes of functions which agree ν - a.e.]

1.1.23 LEMMA Let μ be an R^m-valued vector measure --- then there is a unique S^{m-1}-valued function $f \in L^1(X,||\mu||)^m$ such that $\mu = f\ ||\mu||$.

PROOF Trivially, $\mu << ||\mu||$, so $\mu = f\ ||\mu||$ $(f \in L^1(X,||\mu||)^m)$. Therefore

$$\begin{aligned} ||\mu|| &= ||f||\mu||\ || \\ &= ||f||\ ||\mu|| \quad \text{(cf. 1.1.15.)} \end{aligned}$$

thus $||f|| = 1$ ($||\mu||$ - a.e.).

1.1.24. DEFINITION Let $(X,\mathcal{E})$ be a measurable space.

● A function $\mu:\mathcal{E} \to [-\infty, +\infty]$ is a <u>signed measure</u> provided $\mu(\emptyset) = 0$, μ takes at most one of the two values $+\infty$ and $-\infty$, i.e., either $\mu:\mathcal{E} \to]-\infty, +\infty]$ or $\mu:\mathcal{E} \to [-\infty, +\infty[$, and μ is σ-additive on $\mathcal{E}$.

1.1.25. <u>N.B.</u> A positive measure is a signed measure.

1.1.26. <u>N.B.</u> A real measure is a signed measure.

1.1.27. DEFINITION Suppose that μ is a signed measure on $(X,\mathcal{E})$.

● A set $E \in \mathcal{E}$ is a <u>positive set</u> for μ if $\mu(E_0) \geq 0$ for every $E_0 \in \mathcal{E}$ such that $E_0 \subset E$.

● A set $E \in \mathcal{E}$ is a <u>negative set</u> for μ if $\mu(E_0) \leq 0$ for every $E_0 \in \mathcal{E}$ such that $E_0 \subset E$.

1.1.28. DEFINITION Suppose that μ is a signed measure on $(X,\mathcal{E})$ — then a set $E \in \mathcal{E}$ is a <u>null set</u> for μ if $\mu(E_0) = 0$ for every $E_0 \in \mathcal{E}$ such that $E_0 \subset E$.

1.1.29. DEFINITION Let $(X,\mathcal{E})$ be a measurable space.

● Given a signed measure μ, sets E_+, E_- are said to constitute a <u>Hahn decomposition</u> for μ provided $E_+ \cap E_- = \emptyset$, $E_+ \cup E_- = X$, and

$$\left[\begin{array}{l} E_+ \text{ is a positive set for } \mu \\ E_- \text{ is a negative set for } \mu. \end{array}\right.$$

1.1.30. THEOREM Hahn decompositions exist. Moreover, if (E_+,E_-) and

(E'_+, E'_-) are two such, then $E_+ \Delta E'_+$ and $E_- \Delta E'_-$ are null sets for μ.

1.1.31. LEMMA If μ,ν are signed measures on $(X,\mathcal{E})$, at least one of which is finite, then the set function $\mu - \nu$ is a signed measure on $(X,\mathcal{E})$.

1.1.32. LEMMA Suppose that μ is a signed measure on $(X,\mathcal{E})$ and let $E_1, E_2 \in \mathcal{E}$ with $E_1 \subset E_2$ -- then $\mu(E_2) \in R \Rightarrow \mu(E_1) \in R$. And

$$\left[\begin{array}{l} \mu(E_1) = +\infty \Rightarrow \mu(E_2) = +\infty \\ \\ \mu(E_1) = -\infty \Rightarrow \mu(E_2) = -\infty. \end{array}\right.$$

1.1.33. LEMMA Suppose that μ is a signed measure -- then there exist unique positive measures μ^+,μ^- such that $\mu = \mu^+ - \mu^-$ and $\mu^+ \perp \mu^-$.

[Let $X = E_+ \cup E_-$ be a Hahn decomposition for μ and put

$$\left[\begin{array}{l} \mu^+(E) = \mu(E \cap E_+) \\ \\ \mu^-(E) = -\mu(E \cap E_-) \end{array}\right. \qquad (E \in \mathcal{E}).]$$

1.1.34. REMARK If μ omits the value $+\infty$ $(-\infty)$, then $\mu^+(\mu^-)$ is a finite positive measure. So if the range of μ is contained in R, then μ is bounded, i.e., μ is a real measure.

1.1.35. DEFINITION Let $(X,\mathcal{E})$ be a measurable space.

- A signed measure μ is <u>finite</u> if $|\mu|$ is a finite positive measure.

1.1.36. LEMMA μ is finite iff $\mu(X) \in R$.

1.1.37. RESTRICTION Let $(X,\mathcal{E})$ be a measurable space and suppose that

μ is a positive, real, signed, or vector measure in $(X,\mathcal{E})$. Given $E \in \mathcal{E}$, put

$$(\mu \llcorner E)(S) = \mu(E \cap S) \quad (S \in \mathcal{E}).$$

Then

$$\mu \llcorner E = \chi_E \mu.$$

In fact,

$$(\chi_E \mu)(S) = \int_S \chi_E d\mu = \mu(E \cap S).$$

1.1.38. EXAMPLE Per Radon-Nikodym, consider ν and μ -- then there exists a set $E \in \mathcal{E}$ such that $\nu(E) = 0$ and $\mu^S = \mu \llcorner E$.

§1.2. BOREL MEASURES

Let X be a locally compact Hausdorff space (LCH-space).

1.2.1. NOTATION

- $\mathcal{O}(X)$ is the collection of open subsets of X.
- $\mathcal{K}(X)$ is the collection of compact subsets of X.
- $\mathcal{B}(X)$ is the collection of Borel subsets of X.

1.2.2. DEFINITION A positive measure on $(X,\mathcal{B}(X))$ is referred to simply as a Borel measure on X.

1.2.3. DEFINITION Let μ be a Borel measure on X and let $E \in \mathcal{B}(X)$.

- μ is outer regular on E if

$$\mu(E) = \inf\{\mu(U):U \supset E,\ U \in \mathcal{O}(X)\}.$$

- μ is inner regular on E if

$$\mu(E) = \sup\{\mu(K):K \subset E,\ K \in \mathcal{K}(X)\}.$$

1.2.4. DEFINITION Let μ be a Borel measure on X and let $E \in \mathcal{B}(X)$ -- then μ is regular on E if μ is both inner and outer regular on E.

1.2.5. DEFINITION Let μ be a Borel measure on X and let $\mathcal{C}$ be a subset of $\mathcal{B}(X)$ -- then μ is outer regular for $\mathcal{C}$, inner regular for $\mathcal{C}$, or regular for $\mathcal{C}$ according to whether μ is outer regular, inner regular, or regular for every $E \in \mathcal{C}$.

1.2.6. TERMINOLOGY A Borel measure μ on X is outer regular, inner regular, or regular if μ is outer regular, inner regular, or regular for $\mathcal{B}(X)$.

1.2.7. LEMMA If every open subset of X is σ-compact, then every Borel measure on X is inner regular for $\mathcal{O}(X)$ ($\subset \mathcal{B}(X)$).

1.2.8. EXAMPLE Every open subset of R^n is σ-compact.

1.2.9. REMARK If X is σ-compact, then the σ-algebra generated by $\mathcal{K}(X)$ is $\mathcal{B}(X)$.

1.2.10. LEMMA Let μ_1 and μ_2 be two Borel measures on X.

- If μ_1 and μ_2 are outer regular for $\mathcal{B}(X)$ and $\mu_1 = \mu_2$ on $\mathcal{O}(X)$, then $\mu_1 = \mu_2$ on $\mathcal{B}(X)$.
- If μ_1 and μ_2 are inner regular for $\mathcal{B}(X)$ and $\mu_1 = \mu_2$ on $\mathcal{K}(X)$, then $\mu_1 = \mu_2$ on $\mathcal{B}(X)$.

[To establish the first point, let $E \in \mathcal{B}(X)$ and write

$$\begin{aligned}\mu_1(E) &= \inf\{\mu_1(U):U \supset E,\ U \in \mathcal{O}(X)\}\\ &= \inf\{\mu_2(U):U \supset E,\ U \in \mathcal{O}(X)\}\\ &= \mu_2(E).]\end{aligned}$$

* * * * * * * * * *

APPENDIX

A locally compact Hausdorff space X is <u>σ-compact</u> if X can be expressed as the union of at most countably many compact subspaces.

[Note: $Q = \bigcup\limits_{q \in Q} \{q\}$ and $\forall\ q$, $\{q\}$ is compact but Q is not locally compact.]

LEMMA Every open subset of a second countable LCH space X is σ-compact.

E.g.: This is the case of R^n.

RAPPEL A second countable topological space is separable (but, in general, not conversely).

RAPPEL Every separable metric space is second countable.

So, if X is a metrizable separable LCH space, then every open subset of X is σ-compact.

N.B. Every compact metric space is separable.

§1.3. RADON MEASURES

Let X be a locally compact Hausdorff space (LCH space).

1.3.1. DEFINITION A Borel measure μ on X is said to be locally finite if $\forall K \in \mathcal{K}(X)$, $\mu(K) < +\infty$.

1.3.2. DEFINITION A locally finite Borel measure μ on X is a Radon measure provided

- μ is outer regular for $\mathcal{B}(X)$
- μ is inner regular for $\mathcal{O}(X)$.

1.3.3. N.B. A Radon measure is a positive measure.

1.3.4. REMARK A finite Borel measure μ on a compact Hausdorff space X is locally finite but it need not be Radon.

1.3.5. EXAMPLE Take $X = R^n$ -- then the restriction of Lebesgue measure to $\mathcal{B}(X)$ is a Radon measure.

[Note: Counting measure on R^n is not locally finite, hence is not Radon.]

1.3.6. LEMMA Every σ-finite Radon measure is inner regular for $\mathcal{B}(X)$, hence is regular for $\mathcal{B}(X)$.

In particular: Every finite Radon measure is regular for $\mathcal{B}(X)$, thus every Radon measure on a compact Hausdorff space is regular for $\mathcal{B}(X)$.

1.3.7. LEMMA Suppose that X is σ-compact -- then every Radon measure on X is inner regular for $\mathcal{B}(X)$, hence is regular for $\mathcal{B}(X)$.

[A Radon measure is locally finite, so here

$$\sigma\text{-compact} \Rightarrow \sigma\text{-finite.}]$$

1.3.8. LEMMA Suppose that X is σ-compact, let μ be a Radon measure on X, and let ν be a locally finite Borel measure on X. Assume: $\nu = \mu$ on $\mathcal{O}(X)$ — then ν is regular for $\mathcal{B}(X)$.

1.3.9. RIESZ REPRESENTATION THEOREM (RRT) If I is a positive linear functional on $C_c(X)$, then there exists a unique Radon measure μ on X such that

$$I(f) = \int_X f \, d\mu$$

for all $f \in C_c(X)$.

1.3.10. SUBLEMMA Let U be an open σ-compact subset of X — then there is an increasing sequence $f_1, f_2, \ldots$ on $C_c(X)$ such that $\lim_{n \to \infty} f_n = \chi_U$.

[Note: An open subset of a compact Hausdorff space need not be σ-compact.]

1.3.11. THEOREM If every open subset of X is σ-compact, then every locally finite Borel measure ν on X is a regular Radon measure.

PROOF The issue is outer and inner regularity for $\mathcal{B}(X)$ per ν. Define a positive linear functional I in $C_c(X)$ by the prescription

$$I(f) = \int_X f \, d\nu.$$

Then by the RRT, there exists a unique Radon measure μ on X such that $\forall\, f \in C_c(X)$,

$$I(f) = \int_X f \, d\mu.$$

The claim now is that $\mu = \nu$ on $\mathcal{O}(X)$. So let $U \in \mathcal{O}(X)$ and choose as above $\{f_n\}$, $\lim_{n \to \infty} f_n = \chi_U$ -- then by monotone convergence,

$$\nu(U) = \int_X \chi_U \, d\nu = \lim_{n \to \infty} \int_X f_n \, d\nu$$

$$= \lim_{n \to \infty} \int_X f_n \, d\mu$$

$$= \int_X \chi_U \, d\mu = \mu(U).$$

Therefore $\nu = \mu$ on $\mathcal{O}(X)$, thus ν is regular for $\mathcal{B}(X)$.

[Note: Consequently, $\nu = \mu$ (cf. 1.2.10).]

1.3.12. RAPPEL Let $(X,\mathcal{E})$ be a measurable space -- then a simple function is a finite linear combination with real coefficients of characteristic functions of sets in $\mathcal{E}$.

1.3.13. LEMMA For any positive measure $\mu:\mathcal{E} \to [0, +\infty]$, the simple functions are dense in $L^p(X,\mu)$ $(1 \leq p < +\infty)$.

1.3.14. THEOREM If μ is a Radon measure on X, then $C_c(X)$ is dense in $L^p(X,\mu)$ $(1 \leq p < +\infty)$.

PROOF It is enough to show that for any Borel set E with $\mu(E) < +\infty$, χ_E can be approximated in the L^p-norm by elements of $C_c(X)$. Given $\varepsilon > 0$, choose a compact $K \subset E$ and an open $U \supset E$ such that $\mu(U-K) < \varepsilon$ and using Urysohn, choose an $f \in C_c(X)$ such that $\chi_K \leq f \leq \chi_U$ -- then

$$||\chi_E - f||_p \leq \mu(U-K)^{1/p} < \varepsilon^{1/p}.$$

§1.4. OUTER MEASURES

Let X be a nonempty set — then the pair $(X, \mathcal{P}(X))$ is a measurable space.

1.4.1. DEFINITION A monotone function $\mu^*: \mathcal{P}(X) \to [0, +\infty]$ is an <u>outer measure</u> provided $\mu^*(\emptyset) = 0$ and μ^* is σ-subadditive on $\mathcal{P}(X)$.

1.4.2. DEFINITION Let μ^* be an outer measure — then a set $E \in \mathcal{P}(X)$ is <u>μ^*-measurable</u> if for every $A \in \mathcal{P}(X)$,

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^C).$$

1.4.3. NOTATION $M(\mu^*)$ is the collection of all μ^*-measurable sets $E \in \mathcal{P}(X)$.

1.4.4. THEOREM $M(\mu^*)$ is a σ-algebra.

1.4.5. NOTATION Let μ be the restriction of μ^* to $M(\mu^*)$.

1.4.6. THEOREM μ is a positive measure.

1.4.7. THEOREM μ is a complete measure.

1.4.8. DEFINITION An outer measure μ^* is said to be <u>regular</u> if every $E \in \mathcal{P}(X)$ is contained in a μ^*-measurable set F of equal outer measure.

[In symbols: $\forall\ E \in \mathcal{P}(X)\ \exists\ F \in M(\mu^*): F \supset E\ \&\ \mu^*(F) = \mu^*(E)$.]

1.4.9. DEFINITION An outer measure μ^* on a topological space (X,τ) is <u>Borel</u> if $\mathcal{B}(X) \subset M(\mu^*)$ and is <u>Borel regular</u> if in addition for every $E \in \mathcal{P}(X)$ there exists an $F \in \mathcal{B}(X)$ such that $F \supset E$ and $\mu^*(F) = \mu^*(E)$.

1.4.10. DEFINITION An outer measure μ^* on a metric space (X,d) is a <u>metric outer measure</u> if

$$\mu^*(E \cup F) = \mu^*(E) + \mu^*(F)$$

for all sets $E,F \in \mathcal{P}(X)$ such that $\mathrm{dist}(E,F) > 0$.

1.4.11. THEOREM An outer measure on a metric space (X,d) is Borel iff μ^* is a metric outer measure.

§1.5. LEBESGUE MEASURE

1.5.1. NOTATION L^{n*} is outer Lebesgue measure on R^n.

1.5.2. DEFINITION M_L^n (= $M(L^{n*})$) is the σ-algebra comprised of the L^{n*}-measurable subsets of R^n, the members of M_L^n being referred to as the <u>Lebesgue measurable subsets</u> of R^n.

1.5.3. THEOREM L^{n*} is a metric outer measure, hence

$$\mathcal{B}(R^n) \subset M_L^n.$$

1.5.4. NOTATION L^n is the restriction of L^{n*} to M_L^n, the <u>Lebesgue measure</u> on R^n.

1.5.5. THEOREM L^n is a complete measure and is the completion of the restriction of L^n to $\mathcal{B}(R^n)$.

The restriction of L^n to $\mathcal{B}(R^n)$ is locally finite and Borel regular, thus is Radon.

1.5.6. NOTATION Put

$$\omega_n = \frac{\pi^{n/2}}{\Gamma(1 + n/2)},$$

the Lebesgue measure of the unit ball in R^n.

1.5.7. ISODIAMETRIC INEQUALITY For every bounded Borel set $E \subset R^n$,

$$L^n(E) \leq \omega_n \left(\frac{\mathrm{diam}(E)}{2}\right)^n.$$

If

$$B(x,r) = \{y \in R^n : ||y - x|| \leq r\},$$

then

$$L^n(B(x,r)) = \omega_n r^n.$$

Thus the interpretation of the isodiametric inequality is that the Lebesgue measure of E cannot exceed the Lebesgue measure of a ball with the same diameter as that of E, i.e., among all E with a given diameter d, a ball B(x,r) with diameter d has Lebesgue measure $\omega_n(\frac{d}{2})^n$ and

$$L^n(E) \leq \omega_n(\frac{d}{2})^n.$$

1.5.8. NOTATION Given a nonsingular linear transformation $T:R^n \to R^n$, let M_T be the matrix of T per the standard basis of R^n.

1.5.9. LEMMA

$$E \in \mathcal{B}(R^n) \Rightarrow T(E) \in \mathcal{B}(R^n)$$

and

$$L^n(T(E)) = |\det(M_T)| L^n(E).$$

1.5.10 LEMMA

$$E \in M_L^n \Rightarrow T(E) \in M_L^n$$

and

$$L^n(T(E)) = |\det(M_T)| L^n(E).$$

1.5.11. LEMMA $\forall E \subset R^n$,

$$L^{n*}(T(E)) = |\det(M_T)| L^{n*}(E).$$

On general grounds, $C_c(R^n)$ is dense in $L^p(R^n)$ $(1 \leq p < +\infty)$ (cf. 1.3.14). But more is true:

1.5.12. THEOREM $C_c^\infty(R^n)$ is dense in $L^p(R^n)$ $(1 \leq p < +\infty)$.

§1.6. HAUSDORFF MEASURES

In what follows, take $X = R^n$.

1.6.1. NOTATION Given $s \in [0, +\infty[$, put

$$\omega_s = \frac{\pi^{s/2}}{\Gamma(1 + s/2)}.$$

1.6.2. NOTATION Given $0 < \delta \leq +\infty$ and a subset $E \subset X$, put

$$H_\delta^s(E) = \frac{\omega_s}{2^s} \inf\{\sum_{k=1}^\infty (\mathrm{diam}(E_k))^s : E \subset \bigcup_{k=1}^\infty E_k,\ \mathrm{diam}(E_k) \leq \delta\}.$$

1.6.3. SUBLEMMA

$$\delta_1 \leq \delta_2 \Rightarrow H_{\delta_1}^s(E) \geq H_{\delta_2}^s(E).$$

1.6.4. LEMMA $\forall\, E \subset X$,

$$H^s(E) \equiv \lim_{\delta \downarrow 0} H_\delta^s(E)$$

$$= \sup_{\delta > 0} H_\delta^s(E)$$

exists.

1.6.5. THEOREM

$$H^s : P(X) \to [0, +\infty]$$

is a metric outer measure, the <u>s-dimensional Hausdorff outer measure</u> on X, hence H^s is Borel, hence

$$B(X) \subset M(H^s).$$

1.6.6. LEMMA $\mathcal{H}^s$ is Borel regular.

[In fact, if $E \in \mathcal{P}(X)$, then there exists a G_δ set $G \supset E$ such that $\mathcal{H}^s(G) = \mathcal{H}^s(E)$ and, of course, $G \in \mathcal{B}(X)$.]

1.6.7. N.B. The restriction of $\mathcal{H}^s$ to $M(\mathcal{H}^s)$ is a complete measure.

1.6.8. LEMMA $\forall\, x \in X$,

$$\mathcal{H}^s(x + E) = \mathcal{H}^s(E)$$

and $\forall\, t > 0$,

$$\mathcal{H}^s(tE) = t^s \mathcal{H}^s(E).$$

1.6.9. LEMMA

$$\mathcal{H}^s \equiv 0$$

if $s > n$.

Therefore matters reduce to the consideration of $\mathcal{H}^s$ in the range $0 \leq s \leq n$.

1.6.10. LEMMA $\mathcal{H}^0$ is counting measure. Moreover, $M(\mathcal{H}^0) = \mathcal{P}(X)$.

Therefore matters reduce to the consideration of $\mathcal{H}^s$ in the range $0 < s \leq n$.

Recall that L^{n*} is outer Lebesgue measure on R^n.

1.6.11. THEOREM

$$L^{n*} = \mathcal{H}^n.$$

Therefore matters reduce to the consideration of $\mathcal{H}^s$ in the range $0 < s < n$.

1.6.12. LEMMA Let $E \subset R^n$ and let $0 \leq s < t < +\infty$.

- If $\mathcal{H}^s(E) < +\infty$, then $\mathcal{H}^t(E) = 0$.
- If $\mathcal{H}^t(E) > 0$, then $\mathcal{H}^s(E) = +\infty$.

PROOF The second point is implied by the first point. To arrive at the latter, choose sets E_k such that $\operatorname{diam}(E_k) \leq \delta$, $E \subset \bigcup_{k=1}^{\infty} E_k$, and

$$\frac{\omega_s}{2^s} \sum_{k=1}^{\infty} (\operatorname{diam}(E_k))^s \leq \mathcal{H}_\delta^s(E) + 1 \leq \mathcal{H}^s(E) + 1.$$

Then

$$\begin{aligned}\mathcal{H}_\delta^t(E) &\leq \frac{\omega_t}{2^t} \sum_{k=1}^{\infty} (\operatorname{diam}(E_k))^t \\ &= \frac{\omega_t}{\omega_s} 2^{s-t} \frac{\omega_s}{2^s} \sum_{k=1}^{\infty} (\operatorname{diam}(E_k))^s (\operatorname{diam}(E_k))^{t-s} \\ &\leq \frac{\omega_t}{\omega_s} 2^{s-t} (\mathcal{H}^s(E) + 1)\delta^{t-s}.\end{aligned}$$

Noting that $t - s > 0$, send $\delta \downarrow 0$ to conclude that $\mathcal{H}^t(E) = 0$.

1.6.13. LEMMA Let $E \subset R^n$ — then there exists at most one point $s^* \in [0, +\infty[$ such that $\mathcal{H}^{s^*}(E) \in]0, +\infty[$.

PROOF Take two distinct points $s, t \in [0, +\infty[$ with $s < t$. If $\mathcal{H}^s(E) \in]0, +\infty[$, then $\mathcal{H}^t(E) = 0$ while if $\mathcal{H}^t(E) \in]0, +\infty[$, then $\mathcal{H}^s(E) = +\infty$.

1.6.14. NOTATION Given $E \in \mathcal{P}(X)$, denote by $\mathcal{H}^\bullet(E)$ the function

$$\begin{cases} [0, +\infty[\rightarrow [0, +\infty] \\ \quad s \rightarrow \mathcal{H}^s(E). \end{cases}$$

1.6.15. LEMMA $\mathcal{H}^\bullet(E)$ is a decreasing function on $[0, +\infty[$ which vanishes on $]n, +\infty[$.

1.6.16. THEOREM There are three possibilities for the range of $H^{\bullet}(E)$.

(i) $H^{\bullet}(E)$ assumes one value, viz. 0.

(ii) $H^{\bullet}(E)$ assumes two values, viz. $+\infty$ and 0.

(iii) $H^{\bullet}(E)$ assumes three values, viz. $+\infty$, 0 and one finite positive value s^*.

1.6.17. EXAMPLE

$$\left[\begin{array}{ll} H^{s}(R^{n}) = +\infty & (s \in [0,n]) \\ H^{s}(R^{n}) = 0 & (s \in]n, +\infty[). \end{array}\right.$$

1.6.18. LEMMA If $H^{\bullet}(E)$ assumes a finite positive value at some point $s^* \in [0, +\infty[$, then

$$\left[\begin{array}{ll} H^{s}(E) = +\infty & (s \in [0, s^*[) \\ H^{s}(E) = 0 & (s \in]s^*, +\infty[). \end{array}\right.$$

1.6.19. <u>N.B.</u> $H^{\bullet}(E)$ is identically zero on $[0, +\infty[$ iff $E = \emptyset$.

1.6.20. <u>N.B.</u> If $E \neq \emptyset$, then $H^{\bullet}(E)$ has exactly one point of discontinuity in $[0, +\infty[$ and it belongs to $[0,n]$.

1.6.21. DEFINITION The <u>Hausdorff dimension</u> of a nonempty subset E of R^n, denoted $\dim_H(E)$, is the unique number $s^* \in [0, +\infty[$ at which $H^{\bullet}$ is discontinuous.

1.6.22. <u>N.B.</u>

$$\dim_H(\emptyset) = 0.$$

1.6.23. LEMMA

$$\dim_H(E) = \sup\{s \in [0, +\infty[: \mathcal{H}^s(E) > 0\}$$

$$= \sup\{s \in [0, +\infty[: \mathcal{H}^s(E) = +\infty\}$$

and

$$\dim_H(E) = \inf\{s \in [0, +\infty[: \mathcal{H}^s(E) = 0\}$$

$$= \inf\{s \in [0, +\infty[: \mathcal{H}^s(E) < +\infty\}.$$

1.6.24. LEMMA If $\mathcal{H}^s(E) \in]0, +\infty[$, then $s = \dim_H(E)$.

1.6.25. EXAMPLE

$$\dim_H(R^n) = n \text{ but } \mathcal{H}^n(R^n) = +\infty.$$

1.6.26. LEMMA If $E \in \mathcal{P}(X)$ is countable, then $\dim_H(E) = 0$.

1.6.27. LEMMA If $E \in \mathcal{P}(X)$ has a nonempty interior, then $\dim_H(E) = n$.

In particular: If $U \in \mathcal{P}(X)$ is open and nonempty, then $\dim_H(U) = n$, so $\mathcal{H}^s(U) = +\infty\ (0 < s < n)$.

1.6.28. THEOREM For every $s \in]0,n[$, there exists a compact $K \subset R^n$ such that $\dim_H(K) = s$.

1.6.29. EXAMPLE Take $n = 1$ and let $C \subset R^1$ be the Cantor set -- then

$$\dim_H(C) = \frac{\log 2}{\log 3}.$$

1.6.30. DEFINITION A metric outer measure μ^* on X is <u>locally finite</u> if $\mu^*(K) < +\infty$ for every $K \in \mathcal{K}(X)$.

1.6.31. THEOREM Suppose that μ^* is locally finite -- then for every Borel set $E \in \mathcal{B}(X)$,

$$\mu^*(E) = \inf\{\mu^*(U) : U \supset E,\ U \in \mathcal{O}(X)\}$$

and

$$\mu^*(E) = \sup\{\mu^*(K) : K \subset E,\ K \in \mathcal{K}(X)\}.$$

If $U \subset X$ is open and nonempty, then

$$\mathcal{H}^s(U) = +\infty \quad (0 \leq s < \dim_H(U) = n).$$

1.6.32. SCHOLIUM $\mathcal{H}^s$ is not locally finite if $0 < s < n$.

[Pretend it was -- then for a generic $K \in \mathcal{K}(X)$,

$$\begin{aligned} \mathcal{H}^s(K) &= \inf\{\mathcal{H}^s(U) : U \supset K,\ U \in \mathcal{O}(X)\} \\ &= +\infty \ \ldots\ .] \end{aligned}$$

1.6.33. <u>N.B.</u> Bearing in mind that $\mathcal{B}(X) \subset M(\mathcal{H}^s)$, it follows that the restriction of $\mathcal{H}^s$ to $\mathcal{B}(X)$ is not Radon.

1.6.34. THEOREM Let $\Phi : R^n \to R^n$ be an isometry (a distance preserving bijection) and suppose that $E \in \mathcal{P}(X)$ -- then

$$\Phi(E) \in M(\mathcal{H}^s) \iff E \in M(\mathcal{H}^s)\ (s \in [0, +\infty[).$$

[Note: The assumption that Φ is an isometry implies that

$$\mathcal{H}^s(\Phi(E)) = \mathcal{H}^s(E).]$$

SECTION 2: DIFFERENTIATION THEORY

§2.1. SCALAR FUNCTIONS

Let Ω be a nonempty open subset of R^n and let $f:\Omega \to R$ be a function.

2.1.1. DEFINITION f is <u>differentiable</u> at a point $x_0 \in \Omega$ if there exists a linear function $T:R^n \to R$ (depending on x_0) such that

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - T(h)}{||h||} = 0.$$

2.1.2. <u>N.B.</u> Consider the situation when $n = 1$, $\Omega = R$ and suppose that $f:R \to R$ is differentiable at x_0 in the traditional sense, i.e.,

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Then f is differentiable at x_0. Thus view the number $f'(x_0)$ as the linear map $R \to R$ that sends h to $f'(x_0)(h)$, hence

$$\begin{aligned}
&\frac{f(x_0 + h) - f(x_0) - f'(x_0)(h)}{h} \\
&= \frac{f(x_0 + h) - f(x_0)}{h} - \frac{f'(x_0)(h)}{h} \\
&= \frac{f(x_0 + h) - f(x_0)}{h} - f'(x_0)(1) \\
&\to f'(x_0) - f'(x_0) \quad (h \to 0) \\
&= 0.
\end{aligned}$$

T is called the <u>differential</u> of f at x_0 and is denoted by $df(x_0)$.

[Note: The differential is unique, if it exists at all. Proof: Per the definition, suppose that $T = T_1$ and $T = T_2$ -- then $\forall\ h \neq 0$,

$$|(T_1 - T_2)(h)|$$

$$< |f(x_0 + h) - f(x_0) - T_1(h)| + |f(x_0 + h) - f(x_0) - T_2(h)|$$

=>

$$\frac{|(T_1 - T_2)(h)|}{||h||} \to 0 \text{ as } h \to 0$$

=>

$$\frac{|(T_1 - T_2)(th)|}{||th||} \to 0 \text{ as } t \to 0$$

=>

$$(T_1 - T_2)(h) = 0 \Rightarrow T_1 = T_2.]$$

2.1.3. N.B. f is differentiable in Ω if f is differentiable at every point of Ω.

2.1.4. EXAMPLE Take $\Omega = R^n$ -- then polynomials in several variables are everywhere differentiable.

2.1.5. EXAMPLE Take $\Omega = R^n$ and let $T:R^n \to R$ be linear -- then $dT(x_0) = T$.

2.1.6. LEMMA If f is differentiable at $x_0 \in \Omega$, then f is continuous at x_0.

[Given $h \neq 0$, write

$$|f(x_0 + h) - f(x_0) - T(h)|$$

$$\leq ||h|| \ \left|\frac{f(x_0 + h) - f(x_0) - T(h)}{h}\right|$$

to conclude that

$$f(x_0 + h) - f(x_0) \to 0 \quad (h \to 0).]$$

[Note: Since T is linear, $\lim_{h \to 0} T(h) = 0$.]

Given $x_0 \in \Omega$, suppose that $B(x_0, r_0)$ is contained in Ω -- then for each nonzero $v \in \underset{\sim}{R}^n$, $x_0 + tv \in \Omega$ for $|t| \leq r_0/\|v\|$.

2.1.7. DEFINITION The <u>directional derivative</u> of f at x_0 in the direction v is

$$\lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0)}{t},$$

denoted

$$\frac{\partial f}{\partial v}(x_0).$$

[Note: The underlying assumption is that the limit exists and is finite.]

2.1.8. <u>N.B.</u> $\forall \lambda \neq 0$,

$$\frac{\partial f}{\partial(\lambda v)}(x_0) = \lambda \frac{\partial f}{\partial v}(x_0).$$

2.1.9. LEMMA If f is differentiable at x_0, then $\frac{\partial f}{\partial v}(x_0)$ exists for all $v \neq 0$ and

$$\frac{\partial f}{\partial v}(x_0) = df(x_0)(v).$$

[Observe that

$$\left|\frac{f(x_0 + tv) - f(x_0)}{t} - T(v)\right|$$

$$= \left|\frac{f(x_0 + tv) - f(x_0) - T(tv)}{t}\right|.]$$

2.1.10. EXAMPLE The function

$$f(x,y) = \begin{cases} \dfrac{x^2 y}{x^2 + y^2} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

is continuous at (0,0) and all its directional derivatives exist at (0,0). Still, the differential df(0,0) does not exist.

[To see the last point, suppose instead that df(0,0) does exist, thus being linear,

$$df(0,0)(1,0) + df(0,0)(0,1) = df(0,0)(1,1).$$

On the other hand,

$$\begin{cases} df(0,0)(1,0) = \dfrac{\partial f}{\partial (1,0)} (0,0) = 0 \\ df(0,0)(0,1) = \dfrac{\partial f}{\partial (0,1)} (0,0) = 0. \end{cases}$$

Meanwhile

$$\frac{\partial f}{\partial (1,1)} (0,0) = \frac{1}{2}$$

$$\Rightarrow$$

$$df(0,0)(1,1) = \frac{1}{2}.$$

I.e.:

$$0 + 0 = \frac{1}{2}.$$

Contradiction.]

2.1.11 REMARK If Ω is convex and if $f:\Omega \to R$ is convex, then f is differentiable at x_0 iff f has ordinary partial derivatives at x_0.

Suppose that $\langle\ ,\ \rangle$ is the standard inner product in R^n. Since the differential of f at x_0 is a linear function from R^n to R, there is a unique vector $\nabla f(x_0) \in R^n$ such that for all $h \in R^n$,

$$df(x_0)(h) = \langle h, \nabla f(x_0)\rangle.$$

2.1.12. DEFINITION $\nabla f(x_0)$ is called the gradient of f at x_0.

2.1.13. NOTATION Let $(e_1, e_2, \ldots, e_n)$ be the standard basis for R^n and let $(x_1, x_2, \ldots, x_n)$ be the associated system of coordinates.

2.1.14. DEFINITION The derivative of f at x_0 in the direction e_i is called the partial derivative of f w.r.t. x_i, denoted

$$\frac{\partial f}{\partial x_i}(x_0).$$

2.1.15. LEMMA

$$\nabla f(x_0) = \left(\frac{\partial f}{\partial x_1}(x_0), \frac{\partial f}{\partial x_2}(x_0), \ldots, \frac{\partial f}{\partial x_n}(x_0)\right).$$

2.1.16. DEFINITION The Jacobian matrix of f at x_0 is the $1 \times n$ matrix

$$Df(x_0) = \left[\frac{\partial f}{\partial x_1}(x_0), \frac{\partial f}{\partial x_2}(x_0), \ldots, \frac{\partial f}{\partial x_n}(x_0)\right].$$

2.1.17. LEMMA For all $h \in R^n$,

$$df(x_0)(h) = \sum_{i=1}^{n} df(x_0)(e_i)h_i$$

$$= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} (x_0) h_i$$

$$= Df(x_0) \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix} = \langle h, \nabla f(x_0) \rangle.$$

Consider two points x_0, $x_0 + h$ -- then the line segment ℓ joining x_0 and $x_0 + h$ is the curve $x_0 + th$ $(0 \leq t \leq 1)$.

2.1.18. MEAN VALUE THEOREM Suppose that f is continuous at the points of ℓ and differentiable at the points of ℓ except perhaps the endpoints -- then there exists an $s \in]0,1[$ such that

$$f(x_0 + h) - f(x_0) = df(x_0 + sh)(h).$$

PROOF Introduce

$$\phi(t) = f(x_0 + th) \quad (0 \leq t \leq 1).$$

Then ϕ is continuous in $[0,1]$ and

$$\phi(0) = f(x_0), \ \phi(1) = f(x_0 + h),$$

$$\phi'(t) = df(x_0 + th)(h) \quad (0 < t < 1).$$

By the mean value theorem for functions of one variable, there exists an $s \in]0,1[$ such that

$$\phi(1) - \phi(0) = \phi'(s).$$

2.1.19. APPLICATION Suppose that Ω is not only open but is connected as well. Assume: f is differentiable in Ω and that $df(x) = 0$ for every $x \in \Omega$ -- then f is a constant function.

* * * * * * * * * * *

APPENDIX

What has been said in 2.1.10 can be substantially generalized. Indeed, there are continuous functions of 2 variables which have partial derivatives almost everywhere but for which the differential fails to exist anywhere.

§2.2. VECTOR FUNCTIONS

Let Ω be a nonempty open subset of R^n and let $f:\Omega \to R^m$ be a function.

2.2.1. DEFINITION f is <u>differentiable</u> at a point $x_0 \in \Omega$ if there exists a linear function $T:R^n \to R^m$ (depending on x_0) such that

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - T(h)}{||h||} = 0.$$

T is called the <u>differential</u> of f at x_0 and is denoted by $df(x_0)$.

[Note: As in the scalar case, the differential is unique, if it exists at all.]

2.2.2. <u>N.B.</u> f is differentiable in Ω if f is differentiable at every point of Ω.

Let $f^1(x), f^2(x), \ldots, f^m(x)$ be the components of f and let $T^1, T^2, \ldots, T^m$ be the components of T — then the defining relation for the differential of f at x_0 amounts to the relations

$$\lim_{h \to 0} \frac{f^1(x_0 + h) - f^1(x_0) - T^1(h)}{||h||} = 0$$

$$\lim_{h \to 0} \frac{f^2(x_0 + h) - f^2(x_0) - T^2(h)}{||h||} = 0$$

$$\vdots$$

$$\lim_{h \to 0} \frac{f^m(x_0 + h) - f^m(x_0) - T^m(h)}{||h||} = 0.$$

Therefore f is differentiable at x_0 iff all the components of f are differentiable at x_0 and when this is so, f is continuous at x_0.

2.2.3. SCHOLIUM For all $h = (h_1, h_2, \ldots, h_n)$,

$$T^1(h) = df^1(x_0)(h) = \frac{\partial f^1}{\partial h}(x_0) = \sum_{i=1}^{n} \frac{\partial f^1}{\partial x_i}(x_0)h_i$$

$$T^2(h) = df^2(x_0)(h) = \frac{\partial f^2}{\partial h}(x_0) = \sum_{i=1}^{n} \frac{\partial f^2}{\partial x_i}(x_0)h_i$$

$$\vdots$$

$$T^m(h) = df^m(x_0)(h) = \frac{\partial f^m}{\partial h}(x_0) = \sum_{i=1}^{n} \frac{\partial f^m}{\partial x_i}(x_0)h_i,$$

2.2.4. DEFINITION The <u>Jacobian matrix</u> of f at x_0 is the $m \times n$ matrix

$$\begin{bmatrix} \frac{\partial f^1}{\partial x_1}(x_0) & \frac{\partial f^1}{\partial x_2}(x_0) & \cdots & \frac{\partial f^1}{\partial x_n}(x_0) \\ \frac{\partial f^2}{\partial x_1}(x_0) & \frac{\partial f^2}{\partial x_2}(x_0) & \cdots & \frac{\partial f^2}{\partial x_n}(x_0) \\ \vdots & & & \\ \frac{\partial f^m}{\partial x_1}(x_0) & \frac{\partial f^m}{\partial x_2}(x_0) & \cdots & \frac{\partial f^m}{\partial x_n}(x_0) \end{bmatrix},$$

denoted by $Df(x_0)$.

[Note: The partial derivatives of f are the partial derivatives of its components, i.e., the

$$\frac{\partial f^i}{\partial x_j} \equiv D_j f^i.]$$

2.2.5. DEFINITION Suppose that $n = m$ -- then the determinant of the Jacobian matrix $Df(x_0)$ is called the Jacobian of f at x_0, denoted by

$$J_f(x_0) \text{ or } \frac{\partial(f^1,f^2,\ldots,f^n)}{\partial(x_1,x_2,\ldots,x_n)} .$$

2.2.6. OPEN MAPPING THEOREM Suppose that $n = m$ and suppose that $J_f(x) \neq 0$ for all $x \in \Omega$ -- then the image $f(U)$ of any open set $U \subset \Omega$ is open.

2.2.7. CHAIN RULE Let $U \subset R^n$ and $V \subset R^m$ be nonempty open sets and let $f:U \to R^m$ and $g:V \to R^p$ subject to $f(U) \subset V$. Assume: f is differentiable at $x_0 \in U$ and g is differentiable at $f(x_0)$ -- then $g \circ f$ is differentiable at x_0 and

$$d(g \circ f)(x_0) = dg(f(x_0)) \circ df(x_0)$$

or, in terms of the Jacobian matrices,

$$D(g \circ f)(x_0) = Dg(f(x_0))Df(x_0).$$

2.2.8. RAPPEL The set $\mathrm{Hom}(R^n,R^m)$ of linear transformations from R^n to R^m is a vector space of dimension nm. Moreover, it is a Banach space under the norm

$$||A|| = \max\{||Ax||:||x|| \leq 1\}.$$

And $\forall$ x,

$$||Ax|| \leq ||A|| \; ||x||.$$

2.2.9. EXAMPLE Given $f:\Omega \to R^m$,

$$df(x_0) \in \mathrm{Hom}(R^n,R^m).$$

[Note: If $f = A \in \mathrm{Hom}(R^n,R^m)$, then $df(x_0) = A$.]

2.2.10. DEFINITION A differentiable function $f:\Omega \to R^m$ is <u>continuously differentiable</u> if

$$df:\Omega \to \mathrm{Hom}(R^n,R^m)$$

is continuous.

[Spelled out, given $x_0 \in \Omega$ and $\varepsilon > 0$, there is a $\delta > 0$ such that

$$||df(x) - df(x_0)|| < \varepsilon$$

if $||x - x_0|| < \delta$.]

2.2.11. NOTATION $C^1(\Omega;R^m)$ is the set of continuously differentiable functions from Ω to R^m, often referred to as the C'-functions (a vector space over R).

2.2.12. THEOREM $f:\Omega \to R^m$ is C' iff the partial derivatives of f exist and are continuous through out Ω.

PROOF That the differentiability of f implies the continuity of the partials can be seen by noting that

$$|(D_jf^i)(y) - (D_jf^i)(x)| \le ||df(y) - df(x)||.$$

In the other direction, take m = 1, fix $x_0 \in \Omega$, let $\varepsilon > 0$, and choose $r_0 > 0:B(x_0,r_0)^o \subset \Omega$ and

$$|(D_jf)(x) - (D_jf)(x_0)| < \frac{\varepsilon}{n} \quad (x \in B(x_0,r_0)^o, \ 1 \le j \le n).$$

Write

$$h = \sum_{j=1}^{n} h_je_j, \ ||h|| < r_0,$$

and put

$$v_0 = 0, \ v_k = h_1e_1 + \cdots + h_ke_k \quad (1 \le k \le n).$$

Then

$$f(x_0 + h) - f(x_0)$$

$$= \sum_{j=1}^{n} [f(x_0 + v_j) - f(x_0 + v_{j-1})].$$

Since $||v_k|| < r_0$, the line segments with endpoints $x_0 + v_{j-1}$ and $x_0 + v_j$ lie in $B(x_0,r_0)^{\circ}$. Taking into account that

$$v_j = v_{j-1} + h_j e_j,$$

the MVT implies that

$$f(x_0 + v_j) - f(x_0 + v_{j-1})$$

$$= h_j (D_j f)(x_0 + v_{j-1} + \theta_j h_j e_j)$$

for some $\theta_j \in]0,1[$. Next

$$|h_j (D_j f)(x_0 + v_{j-1} + \theta_j h_j e_j) - h_j (D_j f)(x_0)| < \frac{|h_j|\varepsilon}{n}.$$

Consequently

$$|f(x_0 + h) - f(x_0) - \sum_{j=1}^{n} h_j (D_j f)(x_0)|$$

$$= |\sum_{j=1}^{n} [f(x_0 + v_j) - f(x_0 + v_{j-1})] - \sum_{j=1}^{n} h_j (D_j f)(x_0)|$$

$$= |\sum_{j=1}^{n} h_j (D_j f)(x_0 + v_{j-1} + \theta_j h_j e_j) - \sum_{j=1}^{n} h_j (D_j f)(x_0)|$$

$$\leq \sum_{j=1}^{n} |h_j (D_j f)(x_0 + v_{j-1} + \theta_j h_j e_j) - h_j (D_j f)(x_0)|$$

$$\leq \sum_{j=1}^{n} \frac{|h_j|\varepsilon}{n} \leq \frac{1}{n} (\sum_{j=1}^{n} |h_j|)\varepsilon$$

$$\leq \frac{1}{n} (\sqrt{n}\ ||h||)\varepsilon$$

$$\leq ||h||\varepsilon.$$

Therefore f is differentiable at x_0:

$$df(x_0)(h) = df(x_0)(\sum_{j=1}^{n} h_j e_j)$$

$$= \sum_{j=1}^{n} h_j (D_j f)(x_0).$$

Since $m = 1$, the Jacobian matrix is a row:

$$Df(x_0) = [(D_1 f)(x_0),\ (D_2 f)(x_0), \ldots, (D_n f)(x_0)\ ]$$

or still,

$$Df(x_0) = [\frac{\partial f}{\partial x_1}(x_0), \frac{\partial f}{\partial x_2}(x_0), \ldots, \frac{\partial f}{\partial x_n}(x_0)].$$

Its entries are continuous functions of x_0, thus f is a C'-function.

2.2.13. DEFINITION Take $m = n$ and suppose that $f:\Omega \to R^n$ is a C'-function -- then a point $x_0 \in \Omega$ is a <u>critical point</u> for f if the rank of $Df(x_0)$ is not maximal, i.e., if the rank of $Df(x_0)$ is $<$ n or still, if $J_f(x_0) = 0$.

2.2.14. NOTATION Write Z_f for the set of critical points of f.

2.2.15. SARD $f(Z_f)$ is a set of Lebesgue measure 0.

There are numerous variants on this theme which need not be considered at this juncture. However:

2.2.16. LEMMA Under the above assumptions, for any Lebesgue measurable set $E \subset \Omega$, the set $f(E)$ is Lebesgue measurable and

$$L^n(f(E)) \leq \int_E |J_f| dL^n \quad \text{(cf. 12.3.3)}.$$

2.2.17. N.B. SARD is an immediate conseqence of this result.

The mean value theorem does not hold in general for a vector valued function $f:\Omega \to R^m$ ($m > 1$) (but it does hold if the number of auxiliary points is increased (details omitted)). However:

2.2.18. THEOREM Suppose that $f:[a,b] \to R^k$ is continuous and that its restriction to $]a,b[$ is differentiable — then there exists an $x \in]a,b[$ such that

$$||f(b) - f(a)|| \leq (b - a)||f'(x)||.$$

PROOF Let

$$\varphi(t) = \langle f(b) - f(a), f(t)\rangle \quad (a \leq t \leq b).$$

Then φ satisfies the assumptions of the MVT, hence

$$\begin{aligned}\varphi(b) - \varphi(a) &= (b - a)\varphi'(x) \\ &= (b - a) \langle f(b) - f(a), f'(x)\rangle\end{aligned}$$

for some $x \in]a,b[$. On the other hand,

$$\begin{aligned}\varphi(b) - \varphi(a) &= \langle f(b) - f(a), f(b)\rangle - \langle f(b) - f(a), f(a)\rangle \\ &= \langle f(b) - f(a), f(b) - f(a)\rangle \\ &= ||f(b) - f(a)||^2.\end{aligned}$$

Then

$$\begin{aligned}||f(b) - f(a)||^2 &= (b - a) \langle f(b) - f(a), f'(x)\rangle \\ &\leq (b - a) ||f(b) - f(a)|| \, ||f'(x)||,\end{aligned}$$

hence

$$||f(b) - f(a)|| \leq (b - a) \, ||f'(x)||.$$

§2.3. LIPSCHITZ FUNCTIONS

Let E be a nonempty subset of R^n.

2.3.1. DEFINITION A function $f:E \to R^m$ is said to be <u>L-Lipschitz</u> ($L \geq 0$) if for all $x,y \in E$,

$$||f(x) - f(y)|| \leq L||x - y||.$$

2.3.2. EXAMPLE A constant function $x \to C(\in R^m)$ is 0-Lipschitz.

2.3.3. EXAMPLE $||\cdot||:R^n \to R$ is 1-Lipschitz.

[In fact,

$$\left| \, ||x|| - ||y|| \, \right| \leq ||x - y||.]$$

2.3.4. LEMMA Let $\{f_i:i \in I\}$ be a collection of L-Lipschitz functions $f_i:E \to R$ -- then the functions

$$\begin{bmatrix} x \to \sup\limits_{i\in I} f_i(x) \equiv F(x) \\ \\ x \to \inf\limits_{i\in I} f_i(x) \equiv f(x) \end{bmatrix} \quad (x \in E)$$

are L-Lipschitz if finite at one point.

PROOF To establish the first assertion, note that for all $x,y \in E$,

$$f_i(y) \leq f_i(x) + L||x - y||.$$

Take now the supremum on the RHS and then on the LHS to get:

$$F(y) \leq F(x) + L||x - y||.$$

If $F(x) < +\infty$, then $F(y) < +\infty$ for all $y \in E$, hence $F(y) - F(x) < L||x - y||$, hence $F(x) - F(y) \leq L||x - y||$, hence $|F(x) - F(y)| \leq L||x - y||$.

2.3.5. APPLICATION The function from R^n to R defined by the rule

$$y \to \operatorname{dist}(y,E) = \inf\{||x - y||: x \in E\}$$

is 1-Lipschitz.

2.3.6. THEOREM If $f:E \to R$ is L-Lipschitz, then there is an L-Lipschitz function $F:R^n \to R$ such that $F \mid E = f$.

[Consider

$$F(y) = \inf_{x \in E} (f(x) + L||x - y||).]$$

2.3.7. NOTATION Given $f:E \to R^m$, put

$$\operatorname{Lip}(f;E) = \sup_{\substack{x,y \in E \\ x \neq y}} \frac{||f(x) - f(y)||}{||x - y||} \quad (\equiv \inf\{L\}).$$

[Note: Omit the "E" if it is R^n: $\operatorname{Lip}(f)$.]

2.3.8. DEFINITION A function $f:E \to R^m$ is <u>Lipschitz</u> if it is L-Lipschitz for some $L \geq 0$.

2.3.9. <u>N.B.</u> If $f:E \to R$ is Lipschitz, then f is uniformly continuous.

[Conversely, it can be shown that if $f:E \to R$ is bounded and uniformly continuous, then f is the uniform limit of a sequence of Lipschitz functions.]

[Note: The function $f(x) = \sqrt{x}$ $(0 \leq x \leq 1)$ is not Lipschitz but it is uniformly continuous.]

2.3.10. THEOREM Let $f:R^n \to R^m$ be Lipschitz -- then for any nonempty $E \subset R^n$,

$$\mathcal{H}^s(f(E)) \leq (\operatorname{Lip}(f))^s\, \mathcal{H}^s(E) \quad (s \in [0, +\infty[).$$

PROOF Fix $\delta > 0$ and choose sets $\{E_k\} \subset R^n$ such that

$$E \subset \bigcup_{k=1}^{\infty} E_k, \; \operatorname{diam}(E_k) \leq \delta.$$

Then

$$\operatorname{diam}(f(E_k)) \leq \operatorname{Lip}(f)\operatorname{diam}(E_k).$$

Since

$$f(E) \subset \bigcup_{k=1} f(E_k), \; \operatorname{diam}(f(E_k)) < \operatorname{Lip}(f)\delta,$$

it therefore follows that

$$\mathcal{H}^s_{\operatorname{Lip}(f)\delta}(f(E)) \leq \frac{\omega_s}{2^s} \sum_{k=1}^{\infty} (\operatorname{diam}(f(E_k)))^s$$

$$\leq \frac{\omega_s}{2^s} \operatorname{Lip}(f)^s \sum_{k=1}^{\infty} (\operatorname{diam}(E_k))^s.$$

Now take the infimum over this data to arrive at

$$\mathcal{H}^s_{\operatorname{Lip}(f)\delta}(f(E)) \leq \operatorname{Lip}(f)^s \mathcal{H}^s_\delta(E),$$

from which the assertion upon sending $\delta \downarrow 0$.

2.3.11. EXAMPLE If $n > m$ and if $P: R^n \to R^m$ is the usual projection, then for all $E \subset R^n$,

$$\mathcal{H}^s(P(E)) \leq \mathcal{H}^s(E).$$

[In fact, $\operatorname{Lip}(P) = 1$.]

2.3.12. SUBLEMMA Let

$$\begin{cases} \{x_1,\ldots,x_k\} \subset R^n \\ \{y_1,\ldots,y_k\} \subset R^m \end{cases}$$

subject to the condition that

$$||y_i - y_j|| \le ||x_i - x_j||$$

for all $i,j \in \{1,\ldots,k\}$. Suppose that $r_1,\ldots,r_k$ are positive numbers such that

$$\bigcap_{i=1}^{k} B(x_i,r_i) \neq \emptyset.$$

Then

$$\bigcap_{j=1}^{k} B(y_j,r_j) \neq \emptyset.$$

2.3.13. LEMMA Let $E \subset R^n$ be a nonempty finite set and let $f:E \to R^m$ be a 1-Lipschitz function -- then for any $x \in R^n$, there is an extension of f to a 1-Lipschitz function on $E \cup \{x\}$.

PROOF Let $E = \{x_1,\ldots,x_k\}$ and assume that $\forall\ i,\ x \neq x_i$. Put $r_i = ||x - x_i|| > 0$ and let $y_i = f(x_i)$ -- then there exists a point $y \in R^m$ such that

$$||y - f(x_i)|| \le ||x - x_i||$$

for each i, so it remains only to let $f(x) = y$.

2.3.14 EXTENSION PRINCIPLE Let $f:E \to R^m$ be an L-Lipschitz function -- then there exists an L-Lipschitz function $f:R^n \to R^m$ such that $F \mid E = f$.

PROOF Upon dividing f by L, it can be assumed that f is 1-Lipschitz and it will be enough to deal explicitly with the situation when E and $R^n \backslash E$ are both infinite. Accordingly, choose a countable dense set $\{x_1,x_2,\ldots\}$ in E and a countable dense set $\{y_1,y_2,\ldots\}$ in $R^n \backslash E$. This done, for each $k = 1,2,\ldots$, use the previous lemma repeatedly to obtain a 1-Lipschitz function

$$f_k:\{x_1,\ldots,x_k,\ y_1,\ldots,y_k\} \to R^m$$

such that $f_k(x_i) = f(x_i)$ $(i = 1,\ldots,k)$. Claim: The sequence $\{f_k(y_1)\} \subset R^m$ is bounded. Proof:

$$f_k(y_1) = f_k(y_1) - f_k(x_1) + f_k(x_1)$$

$$\Rightarrow$$

$$\begin{aligned} ||f_k(y_1)|| &\leq ||f_k(y_1) - f_k(x_1)|| + ||f_k(x_1)|| \\ &\leq ||y_1 - x_1|| + ||f(x_1)|| \\ &< +\infty \end{aligned}$$

independently of k. Proceeding, extract a convergent subsequence, say $\{f_{k_j^1}(y_1)\}$, and then extract from it yet another convergent subsequence $\{f_{k_j^2}(y_2)\}$. ETC. Pass to the diagonal sequence $\{g_j\}: g_j = f_{k_j^j}$, hence for every

$$c \in C = \{x_1, x_2, \ldots\} \cup \{y_1, y_2, \ldots\},$$

there follows

$$g(c) = \lim_{j \to \infty} g_j(c) \in R^m.$$

In addition, $g: C \to R^m$ is 1-Lipschitz and $g(x_i) = f(x_i)$ $(i = 1,2,\ldots)$. And finally, in view of the density of C in R^m and the density of $\{x_1, x_2, \ldots\}$ in E, g extends to a 1-Lipschitz function $F: R^n \to R^m$ such that $F \mid E = f$.

2.3.15. THEOREM Suppose that $f: \Omega \to R^m$ is differentiable (thus, by definition, Ω is open). Assume: Ω is convex and that there is an $L \geq 0$ such that

$$||df(x)|| \leq L$$

for all $x \in \Omega$ -- then f is L-Lipschitz.

PROOF Given $x,y \in \Omega$, the convexity of Ω implies that

$$tx + (1 - t)y \in \Omega \quad (0 \leq t \leq 1).$$

Let

$$g(t) = f(tx + (1 - t)y) \quad (0 \leq t \leq 1).$$

Then

$$\frac{d}{dt} g(t) = df(tx + (1 - t)y)\ (x - y)$$

=>

$$\left|\left| \frac{d}{dt} g(t) \right|\right| \leq ||df(tx + (1 - t)y)||\ ||x - y||$$

$$\leq \quad L||x - y||.$$

Take now in 2.2.13., $[a,b] = [0,1]$ and apply it to g, thus for some $t_0 \in]0,1[$,

$$||g(1) - g(0)|| \leq (1 - 0)||g'(t_0)||$$

$$= ||g'(t_0)||$$

$$\leq L||x - y||.$$

But

$$g(1) = f(x),\ g(0) = f(y)$$

=>

$$||f(x) - f(y)|| \leq L||x - y||.$$

2.3.16. EXAMPLE The sine function is 1-Lipschitz (since its derivative is the cosine which is bounded by 1).

2.3.17. DEFINITION A function $f:\Omega \to R^m$ is said to be <u>locally Lipschitz</u> if for each compact set $K \subset \Omega$, there exists a constant $C_K \geq 0$ such that for all $x,y \in K$,

$$||f(x) - f(y)|| \leq C_K||x - y||.$$

[Note: If there exists L such that $C_K = L$ for all K, then f is said to be locally L-Lipschitz.]

2.3.18. EXAMPLE In R^2, let

$$E = \{(r,\theta):0 < r < +\infty, -\pi < \theta < \pi\}.$$

Then the function $E \to R^2$ given by

$$(r,\theta) \to (r,\theta/2)$$

is locally 1-Lipschitz but not Lipschitz.

2.3.19. THEOREM If $\Omega \subset R^n$ is convex and if $f:\Omega \to R$ is convex, then f is locally Lipschitz.

2.3.20. RAPPEL Suppose that $E \subset R^n$ is Lebesgue measurable -- then there exists an increasing sequence $\{F_k\}$ of closed sets F_k contained in E and a set N of Lebesgue measure 0 such that

$$E = (\bigcup_k F_k) \cup N.$$

2.3.21. N.B. A closed set is the union of a countable family of compact sets and a continuous function $f:R^n \to R^n$ sends a countable union of compact sets to a countable union of compact sets.

2.3.22. DEFINITION Let $\Omega \subset R^n$ be nonempty and open -- then a continuous function $f:\Omega \to R^n$ is said to have property (N) if f sends sets of Lebesgue measure 0 to sets of Lebesgue measure 0.

If $f:R^n \to R^n$ is continuous and if $E \subset R^n$ is closed, then f(E) is Lebesgue measurable. Consequently, in the presence of property (N), it follows that f sends

Lebesgue measurable sets to Lebesgue measurable sets.

2.3.23. THEOREM If $f:R^n \to R^n$ is locally Lipschitz, then f has property (N).

PROOF The claim is that

$$L^n(N) = 0 \Rightarrow L^{n*}(f(N)) = 0.$$

To this end, fix a closed cube K in R^n, write

$$||f(x) - f(y)|| \leq C_K ||x - y|| \quad (x,y \in K),$$

and note that a cube I of side r in K has diameter of length $r\sqrt{n}$. Since f is Lipschitz, f(I) has diameter at most $r\sqrt{n}\, C_K$, thus is contained in a cube of side $r\sqrt{n}\, C_K$, and so

$$L^{n*}(f(I)) \leq n^{n/2} C_K^n r^n = n^{n/2} C_K^n L^{n*}(I),$$

or still,

$$L^{n*}(f(N \cap K)) \leq n^{n/2} C_K^n L^{n*}(N \cap K)$$

$$\Rightarrow L^{n*}(f(N \cap K)) = 0.$$

Choose now an increasing sequence $\{K_j\}$ of closed cubes K_j such that $R^n = \bigcup_i K_j$, hence

$$f(N) = \bigcup_j f(N \cap K_j),$$

and therefore

$$L^n(N) \leq \sum_j L^{n*}(f(N \cap K_j)) = 0.$$

2.3.24. LEMMA Suppose that f is a C'-function, i.e., $f \in C^1(\Omega;R^m)$ -- then f is locally Lipschitz.

§2.4. RADEMACHER

If Ω is a nonempty open subset of R and if $f:\Omega \to R$ is L-Lipschitz, then f is absolutely continuous, hence is differentiable almost everywhere (per L^1).

2.4.1. THEOREM If Ω is a nonempty open subset of R^n and if $f:\Omega \to R^m$ is L-Lipschitz, then f is differentiable at L^n almost all points in Ω.

The proof will be given in the lines below.

First Step: It can be assumed that $m = 1$,

[For f is Lipschitz (or differentiable) iff every component of f is Lipschitz (or differentiable).]

Second Step: It can be assumed that $\Omega = R^n$, so $f:R^n \to R$.

[Invoke the Extension Principle.]

2.4.2. RAPPEL A Lebesgue measurable function $f:R^n \to R$ is locally integrable if

$$\int_K |f| \, dL^n < +\infty$$

for every compact $K \subset R^n$.

Denote the space of such by

$$L^1_{\ell oc}(R^n).$$

2.4.3. LEMMA If $f \in L^1_{\ell oc}(R^n)$ and if

$$\int_{R^n} f\varphi \, dL^n = 0$$

for all $\varphi \in C_c^\infty(R^n)$, then $f = 0$ almost everywhere.

<u>Third Step:</u> Given $x \in R^n$, $v \in S^{n-1}$, form

$$f_{x,v}(t) = f(x + tv) \quad (t \in R).$$

Then $f_{x,v}$ as a function of t is Lipschitz, hence is differentiable almost everywhere (per L^1).

<u>Fourth Step:</u> Recall that by definition,

$$\frac{\partial f}{\partial v}(x) = \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t}$$

whenever the limit exists.

<u>Fifth Step:</u> Let

$$E_v = \{x \in R^n : \frac{\partial f}{\partial v}(x) \text{ exists}\}.$$

Then E_v is Borel and the function

$$\left[\begin{array}{l} E_v \to R \\ \\ x \to \frac{\partial f}{\partial v}(x) \end{array}\right.$$

is Lebesgue measurable.

<u>Sixth Step:</u> Write

$$R^n = Rv \oplus v^{\perp}.$$

Then

$$\begin{aligned} L^n(R^n \backslash E_v) &= \int_{R^n} \chi_{R^n \backslash E_v} \, dL^n \\ &= \int_{v^{\perp}} \int_{Rv} \chi_{R^n \backslash E_v}(tv + w) \, dt dw \end{aligned}$$

$$= \int_{v^\perp} L^1(S_w)\ dw$$

$$= 0,$$

where

$$S_w = \{t \in R : tv + w \in R^n \setminus E_v\}.$$

Seventh Step: Therefore E_v is of full measure in that $\frac{\partial f}{\partial v}(x)$ exists for almost every $x \in R^n$ per L^n.

Eighth Step: In particular, the partial derivatives $\frac{\partial f}{\partial x_i}(x)$ $(i = 1,2,\ldots,n)$ exist for almost all x per L^n, hence the same is so of the formal gradient

$$\nabla f(x) = (\frac{\partial f}{\partial x_1}(x), \frac{\partial f}{\partial x_2}(x), \ldots, \frac{\partial f}{\partial x_n}(x)).$$

2.4.4. LEMMA For each $v \in S^{n-1}$,

$$\frac{\partial f}{\partial v}(x) = \langle v, \nabla f(x) \rangle$$

almost everywhere (per L^n).

PROOF Both functions are in $L^1_{\ell oc}(R^n)$. E.g.:

$$\frac{|f(x + tv) - f(x)|}{|t|} \leq L \frac{||x + tv - x||}{|t|}$$

$$= L \frac{||tv||}{|t|} = L||v||.$$

Bearing in mind 2.4.3, it suffices to show that

$$\int_{R^n} \frac{\partial f}{\partial v}(x) \varphi(x)\ dL^n(x)$$

$$= \int_{R^n} \langle v, \nabla f(x) \rangle \varphi(x)\ dL^n(x)$$

for all $\varphi \in C_c^\infty(R^n)$. Start with the left hand side and proceed:

$$\int_{R^n} \frac{\partial f}{\partial v}(x)\varphi(x)\ dL^n(x)$$

$$= \int_{R^n} \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t}\varphi(x)\ dL^n(x)$$

$$= \lim_{t \to 0} \int_{R^n} \frac{f(x + tv) - f(x)}{t}\varphi(x)\ dL^n(x)$$

$$= \lim_{t \to 0} \int_{R^n} - f(x)\frac{\varphi(x) - \varphi(x - tv)}{t}\ dL^n(x)$$

$$= - \int_{R^n} f(x) \lim_{t \to 0} \frac{\varphi(x) - \varphi(x - tv)}{t}\ dL^n(x)$$

$$= - \int_{R^n} f(x) \frac{\partial \phi}{\partial v}(x)\ dL^n(x)$$

$$= - \sum_{i=1}^n v_i \int_{R^n} f(x)\frac{\partial\varphi}{\partial x_i}(x)\ dL^n(x)$$

$$= \sum_{i=1}^n v_i \int_{R^n} \frac{\partial f}{\partial x_i}(x)\varphi(x)\ dL^n(x)$$

$$= \int_{R^n} \langle v,\nabla f(x)\rangle\varphi(x)\ dL^n(x).$$

[Note: The justification of the formalities is left to the reader.]

<u>Ninth Step:</u> Let $D \subset S^{n-1}$ be a countable dense set -- then there is a Lebesgue measurable set $E \subset R^n$ such that $L^n(R^n\backslash E) = 0$ and $\forall\ v \in D$,

$$\frac{\partial f}{\partial v}(x) = \langle v,\nabla f(x)\rangle \quad (x \in E).$$

<u>Tenth Step:</u> Fix $x_0 \in E$ -- then the claim is that f is differentiable at x_0:

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - \langle h, \nabla f(x_0)\rangle}{||h||} = 0,$$

the ambient linear function $T: R^n \to R$ being the arrow

$$h \to \langle h, \nabla f(x_0)\rangle.$$

2.4.5. LEMMA f is differentiable at x_0 if $\forall\ \varepsilon > 0$, $\exists\ \delta > 0$ such that

$$|f(x_0 + tv) - f(x_0) - t\langle v, \nabla f(x_0)\rangle| \leq \varepsilon|t|$$

provided $|t| \leq \delta$ and $v \in S^{n-1}$.

To verify that this condition is satisfied, fix $\varepsilon > 0$ and choose a finite set $D_0 \subset D$ with the property that for every $v \in S^{n-1}$ there is a $v_0 \in D_0$ such that $||v - v_0|| \leq \varepsilon$. Since the directional derivatives indexed by the $v_0 \in D_0$ are finite in number, there is a $\delta > 0$ such that $\forall\ v_0 \in D_0$,

$$|f(x_0 + tv_0) - f(x_0) - t\langle v_0, \nabla f(x_0)\rangle| \leq \varepsilon|t|$$

if $|t| \leq \delta$. Given now $v \in S^{n-1}$, determine $v_0 \in D_0$ for which $||v - v_0|| \leq \varepsilon$ -- then

$$\begin{aligned} &|f(x_0 + tv) - f(x_0) - t\langle v, \nabla f(x_0)\rangle| \\ &\leq \varepsilon|t| + |f(x_0 + tv) - f(x_0 + tv_0)| \\ &\qquad + |t|\ |\langle v - v_0, \nabla f(x_0)\rangle| \\ &\leq (1 + \text{Lip}(f) + ||\nabla f(x_0)||)\varepsilon|t| \end{aligned}$$

for all $|t| \leq \delta$.

§2.5. STEPANOFF

Let Ω be a nonempty open subset of R^n and let $f:\Omega \to R^m$ be a Lebesgue measurable function.

2.5.1. DEFINITION The pointwise Lipschitz constant of f is

$$\mathrm{Lip}f(x) = \limsup_{y \to x, y \in \Omega} \frac{||f(x) - f(y)||}{||x - y||} \quad (x \in \Omega).$$

2.5.2. THEOREM f is differentiable almost everywhere in the set

$$L_f = \{x \in \Omega : \mathrm{Lip}f(x) < +\infty\}.$$

2.5.3. REMARK

$$L_f = \bigcup_{k,\ell} E_{k,\ell},$$

where

$$E_{k,\ell} = \{x \in L_f : ||f(x)|| \leq k \text{ and } \frac{||f(x) - f(y)||}{||x - y||} \leq k \text{ if } ||x - y|| \leq \frac{1}{\ell}\}.$$

Moreover $f|E_{k,\ell}$ is Lipschitz:

- $||x - y|| < \frac{1}{\ell} \Rightarrow ||f(x) - f(y)|| \leq k||x - y||$
- $||x - y|| \geq \frac{1}{\ell} \Rightarrow ||f(x) - f(y)|| \leq 2k \leq 2k\ell\ ||x - y||,$

and it turns out that f is differentiable almost everywhere in each $E_{k,\ell}$ (details omitted).

[Note: The $E_{k,\ell}$ are Lebesgue measurable, hence the same is true of L_f.]

2.5.4. SUBLEMMA Let g,f,h be functions from Ω to R. Suppose that $g \leq f \leq h$, $g(x_0) = f(x_0) = h(x_0)$, and g,h are differentiable at x_0 -- then f is differentiable at x_0.

PROOF Since $h - g \geq 0$ and $(h - g)(x_0) = 0$, it follows that $d(h - g)(x_0) = 0$, hence $dh(x_0) = dg(x_0)$, call it T -- then

$$\frac{g(x) - g(x_0) - dg(x_0)(x - x_0)}{||x - x_0||}$$

$$\leq \frac{f(x) - f(x_0) - T(x - x_0)}{||x - x_0||}$$

$$\leq \frac{h(x) - h(x_0) - dh(x_0)(x - x_0)}{||x - x_0||}.$$

The first and third terms converge to 0 when $x \to x_0$, thus so does the second term.

Passing to the proof of the theorem, take $m = 1$ and assume that L_f is nonempty. Consider the countable collection $\{B_1, B_2, \ldots\}$ of all open balls $B(x,r)^\circ$ contained in Ω with $x \in Q^n$ and $r \in Q \cap]0, +\infty[$ such that $f|B(x,r)^\circ$ is bounded -- then $L_f \subset \bigcup_{n=1}^{\infty} B_n$. Given $x \in B_n$, introduce

$$\begin{cases} u_n(x) = \sup\{u(x) : u \leq f \text{ on } B_n,\ \text{Lip}(u;B_n) \leq n\} \\ v_n(x) = \inf\{v(x) : v \geq f \text{ on } B_n,\ \text{Lip}(v;B_n) \leq n\}. \end{cases}$$

Here the "sup" ("inf") is over all the $u(v)$ with the stated properties, thus

$$u_n \leq f|B_n \leq v_n$$

and

$$\begin{cases} \text{Lip}(u_n;B_n) \leq n \\ \text{Lip}(v_n;B_n) \leq n. \end{cases}$$

Let E_n stand for the set of $x \in B_n$ at which both u_n and v_n are differentiable at x, hence by Rademacher, the set

$$Z = \bigcup_{n=1}^{\infty} B_n \backslash E_n$$

has Lebesgue measure 0. The claim now is that f is differentiable at all points of $L_f \backslash Z$. So let $x_0 \in L_f \backslash Z$ -- then it need only be shown that there is an index n such that $x_0 \in E_n$ and $u_n(x_0) = v_n(x_0)$. This said, choose an $r_0 > 0$ and an $M > 0$ such that

$$|f(x) - f(x_0)| \leq M||x - x_0|| \quad (x \in B(x_0,r_0)^{\circ}).$$

Next, choose $n \geq M$:

$$x_0 \in B_n \subset B(x_0,r_0)^{\circ}.$$

Then $x_0 \notin Z \Rightarrow x_0 \in E_n$ (for $x_0 \notin E_n \Rightarrow x_0 \in B_n \backslash E_n \subset Z$). Proceeding, $\forall\, x \in B_n$

$$f(x) \leq f(x_0) + M||x - x_0|| \leq f(x_0) + n||x - x_0||$$

$\Rightarrow$

$$f(x) \leq v_n(x) \leq f(x_0) + n||x - x_0||$$

$\Rightarrow$

$$f(x_0) \leq v_n(x_0) \leq f(x_0)$$

$\Rightarrow$

$$f(x_0) = v_n(x_0).$$

Therefore $u_n(x_0) = v_n(x_0)$, completing the proof.

2.5.5. APPLICATION Suppose that $\mathrm{Lip}f(x) < +\infty$ almost everywhere -- then f is differentiable almost everywhere.

2.5.6. EXAMPLE Quasiconformal maps are differentiable almost everywhere.

2.5.7. REMARK It can be shown that the subset $E \subset \Omega$ consisting of those x at which f is differentiable is Lebesgue measurable, as are the partial derivatives

$$\frac{\partial f}{\partial x_i} : E \to R \quad (i = 1,\ldots,n).$$

2.5.8. N.B. The set of points where a given first order partial derivative f exists need not be Lebesgue measurable.

2.5.9. EXAMPLE Let $S \subset R$ be a non Lebesgue measurable set and let

$$f(x,y) = \chi_Q(x)\chi_S(y) \quad ((x,y) \in R^2).$$

Then f is Lebesgue measurable but the set of points (x,y) at which $\frac{\partial f}{\partial x}$ exists is not Lebesgue measurable.

2.5.10. REMARK It can be shown that if $f:\Omega \to R$ is continuous and if E_i is the set of all $x \in \Omega$ such that $\frac{\partial f}{\partial x_i}$ exists, then E_i is a Borel set and $\frac{\partial f}{\partial x_i}$ is a Borel function in E_i.

[Note: If instead $f:\Omega \to R$ is merely Borel measurable, then the $\frac{\partial f}{\partial x_i}$ are Lebesgue measurable.]

§2.6. LUSIN

Convention: Be it a set or a function, measurable means Lebesgue measurable.

2.6.1. THEOREM Suppose given a function $f:R^n \to R$ — then the following conditions are equivalent: f is measurable or

- For every $\varepsilon > 0$ and any compact $K \subset R^n$, there is an open set $G \subset R^n$ such that $L^n(G) < \varepsilon$ and $f|K\backslash G$ is continuous.

- For every $\varepsilon > 0$ and any compact $K \subset R^n$, there exists a continuous function $\varphi:R^n \to R$ such that

$$L^n(\{x \in K:f(x) \neq \varphi(x)\}) < \varepsilon.$$

- For every compact $K \subset R^n$, there exists a sequence $\{\varphi_n\}$ of continuous functions $\varphi_n:R^n \to R$ such that $\varphi_n \to f$ almost everywhere on K.

2.6.2. THEOREM Suppose given a function $f:R^n \to R$ — then the following conditions are equivalent: f is measurable or

- For every $\varepsilon > 0$, there exists an open set $G \subset R^n$ such that $L^n(G) < \varepsilon$ and $f|R^n\backslash G$ is continuous.

- For every $\varepsilon > 0$, there exists a continuous function $\varphi:R^n \to R$ and an open set $G \subset R^n$ such that $L^n(G) < \varepsilon$ and $\varphi = f$ on $R^n\backslash G$.

- There exists a sequence $\{\varphi_n\}$ of continuous functions $\varphi_n:R^n \to R$ such that $\varphi_n \to f$ almost everywhere on R^n.

2.6.3. CHARACTERIZATION Let $f:E \to R$ be a function defined on a measurable set $E \subset R^n$ — then f is measurable iff for every $\varepsilon > 0$ there exists a closed set

$F \subset E$ such that $L^n(E \backslash F) < \varepsilon$ and the restriction of f to F is continuous.

2.6.4. SCHOLIUM Suppose that $f:R^n \to R$ is measurable -- then for any measurable set E and any $\varepsilon > 0$, there exists a continuous function $g:R^n \to R$ such that

$$L^n(\{x \in E: f(x) \neq g(x)\}) < \varepsilon$$

and

$$||g||_{\infty,R^n} \leq ||f||_{\infty,E}.$$

In particular: Take $E = R^n$ -- then the conclusion is that a measurable function coincides with a continuous function outside a set of arbitrarily small measure.

There is also a C' version of this result, the proof of which depends on an extension theorem due to Whitney.

2.6.5. THEOREM Let $K \subset R^n$ be a compact set and let $f:K \to R$, $T:K \to R^n$ be continuous functions. Assume: For every $\varepsilon > 0$, there is a $\delta > 0$ such that

$$\frac{|f(y) - f(x) - T(x)(y - x)|}{||y - x||} \leq \varepsilon$$

whenever $x,y \in K$, $x \neq y$, and $||y - x|| \leq \delta$ -- then there exists a C' function $g:R^n \to R$ such that

$$g|K = f \text{ and } \nabla g|K = T.$$

2.6.6. NOTATION As usual, Ω is a nonempty open subset of R^n.

2.6.7. APPLICATION Suppose that $f:\Omega \to R$ is measurable and differentiable almost everywhere -- then for any $\varepsilon > 0$, there is a function $g \in C^1(\Omega;R)$ such that

$$L^n(\{x \in \Omega: f(x) \neq g(x)\}) < \varepsilon.$$

2.6.8. N.B. Thanks to Rademacher, this applies in the special case when f is Lipschitz.

SECTION 3: DENSITY THEORY

§3.1. LEBESGUE POINTS

Let $f \in L^1_{\ell oc}(R^n)$.

3.1.1. DEFINITION A point $x \in R^n$ such that

$$\lim_{r \to 0} \frac{1}{\omega_n r^n} \int_{B(x,r)} |f - f(x)| \, dL^n = 0$$

is called a <u>Lebesgue point</u> of f.

[Note: Recall that

$$L^n(B(x,r)) = \omega_n r^n.$$

In particular, if $n = 1$, then

$$\omega_1 = \frac{\pi^{1/2}}{\Gamma(1 + 1/2)} = \frac{\pi^{1/2}}{\Gamma(3/2)} = \frac{\pi^{1/2}}{\frac{1}{2}\Gamma(1/2)} = \frac{\pi^{1/2}}{\frac{1}{2}\pi^{1/2}} = 2.]$$

3.1.2. DEFINITION The <u>Lebesgue set</u> of f is the set of its Lebesgue points, denoted $\Lambda(f)$.

3.1.3. THEOREM

$$\Lambda(f) \in M^n_L$$

and

$$L^n(R^n \backslash \Lambda(f)) = 0.$$

3.1.4. <u>N.B.</u> Every continuity point of f is a Lebesgue point of f.

[Supposing that f is continuous at x, given $\varepsilon > 0$, there exists $\delta > 0$ such that $|f(y) - f(x)| < \varepsilon$ if $y \in B(x,\delta)$, so

$$r \in \,]0,\delta[\; \Rightarrow B(x,r) \subset B(x,\delta)$$

$\Rightarrow$

$$\frac{1}{\omega_n r^n} \int_{B(x,r)} |f - f(x)| \, dL^n$$

$$\leq \frac{1}{\omega_n r^n} \int_{B(x,r)} \varepsilon \, dL^n = \varepsilon.]$$

3.1.5. DEFINITION If $E \subset R^n$ is Lebesgue measurable, then the <u>density</u> of E at a point $x \in R^n$ (not necessarily in E) is

$$D_E(x) = \lim_{r \to 0} \frac{L^n(E \cap B(x,r))}{\omega_n r^n} \quad (\in [0,1])$$

provided the limit exists.

3.1.6. EXAMPLE Work in R and let

$$E = \bigcup_{k=0}^{\infty} I_k,$$

where

$$I_k = \left[\frac{1}{2^{2k+1}}, \frac{1}{2^{2k}}\right],$$

hence

$$L^1(I_k) = \frac{1}{2^{2k+1}},$$

but $D_E(0)$ does not exist. In fact,

$$\frac{L^1(E \cap B(0,2^{-2k}))}{L^1(B(0,2^{-2k}))} = \frac{1}{3}$$

and

$$\frac{L^1(E \cap B(0,2^{-2k-1}))}{L^1(B(0,2^{-2k-1}))} = \frac{1}{6}.$$

Therefore

$$\frac{L^1(E \cap B(0,r))}{L^1(B(0,r))}$$

assumes the value $\frac{1}{3}$ for $r = 2^{-2k}$ and the value $\frac{1}{6}$ for $r = 2^{-2k-1}$, so $D_E(0)$ does not exist.

3.1.7. N.B.

$$E \in M_L^n \Rightarrow \chi_E \in L^1_{\ell oc}(R^n).$$

3.1.8. LEMMA Let $E \in M_L^n$ -- then

$$E^\circ \cup (R^n \backslash E)^\circ \subset \Lambda(\chi_E).$$

PROOF If $x \in E^\circ$ (or if $x \in (R^n \backslash E)^\circ$), then χ_E is continuous at x, thus $x \in \Lambda(\chi_E)$.

3.1.9. EXAMPLE It can happen that $D_E(x)$ exists for some $x \notin \Lambda(\chi_E)$.

[Work in R and let $E = [0, +\infty[$ -- then

$$R\backslash\{0\} = E^\circ \cup (R\backslash E)^\circ \subset \Lambda(\chi_E).$$

On the other hand,

$$\lim_{r \to 0} \frac{1}{L^1(B(0,r))} \int_{B(0,r)} |\chi_E - \chi_E(0)| \, dL^1$$

$$= \lim_{r \to 0} \frac{1}{2r} \int_{-r}^{r} (1 - \chi_E(y)) \, dL^1$$

$$= \lim_{r \to 0} \frac{1}{2r} (2r - \int_0^r dL^1)$$

$$= \frac{1}{2} \neq 0 \Rightarrow 0 \notin \Lambda(\chi_E).$$

Nevertheless

$$D_E(0) = \lim_{r \to 0} \frac{L^1(E \cap B(0,r))}{L^1(B(0,r))}$$

$$= \lim_{r \to 0} \frac{r}{2r} = \frac{1}{2} .]$$

3.1.10. LEMMA If $E \subset R^n$ is a set of Lebesgue measure 0, then $E \subset R^n \backslash \Lambda(\chi_E)$.

PROOF The assertion is trivial if $E = \emptyset$, so take an $x \in E$ -- then $\chi_E(x) = 1$, while $\chi_E = 0$ almost everywhere in R^n, hence

$$\lim_{r \to 0} \frac{1}{\omega_n r^n} \int_{B(x,r)} |\chi_E - \chi_E(x)| \, dL^n$$

$$= \lim_{r \to 0} \frac{1}{\omega_n r^n} \int_{B(x,r)} 1 \, dL^n$$

$$= 1 \neq 0 \Rightarrow x \in R^n \backslash \Lambda(\chi_E).$$

3.1.11. EXAMPLE Take for E the Cantor set in R -- then $\Lambda(\chi_E) = R \backslash E$. Moreover, $\forall \ x \in R$, $D_E(x)$ exists and is equal to zero.

[Recall that E is a closed subset of R, thus $R \backslash E$ is an open subset of R, thus $R \backslash E = (R \backslash E)^\circ \subset \Lambda(\chi_E)$. But E is also a set of Lebesgue measure 0, hence

$$E \subset R \backslash \Lambda(E)$$

$$\Rightarrow$$

$$R \backslash E \supset \Lambda(E) \Rightarrow \Lambda(E) \supset R \backslash E \supset \Lambda(E)$$

$$\Rightarrow \Lambda(\chi_E) = R \backslash E.$$

As for the other contention, simply note that

$$L^1(E) = 0 \Rightarrow L^1(E \cap B(x,r)) = 0.]$$

3.1.12. LEMMA Let $E \in M_L^n$ -- then

$$D_E(x) = \begin{cases} 1 \text{ for } x \in E \cap \Lambda(\chi_E) \\ \\ 0 \text{ for } x \in (R^n \backslash E) \cap \Lambda(\chi_E). \end{cases}$$

3.1.13. SCHOLIUM

$$D_E(x) = \begin{cases} 1 \text{ for almost all } x \in E \\ \\ 0 \text{ for almost all } x \in R^n \backslash E. \end{cases}$$

[It is a question of establishing that

- $L^n(E\backslash(E \cap \Lambda(\chi_E))) = 0$

and

- $L^n((R^n\backslash E)\backslash((R^n\backslash E) \cap \Lambda(\chi_E))) = 0.$

E.g.:

$$\begin{aligned} & E\backslash(E \cap \Lambda(\chi_E)) \\ & = R^n\backslash E \cap (E \cap \Lambda(\chi_E)) \\ & = E \cap ((R^n\backslash E) \cup R^n\backslash\Lambda(\chi_E)) \\ & = E \cap (R^n\backslash\Lambda(\chi_E)). \end{aligned}$$

But

$$L^n(R^n\backslash\Lambda(\chi_E)) = 0.]$$

3.1.14. NOTATION Given $E \in M_L^n$, let

$$\mathcal{D}_E = \{x \in R^n : D_E(x) \text{ exists}\}.$$

3.1.15. <u>N.B.</u>

$$\mathcal{D}_E \supset \Lambda(\chi_E).$$

3.1.16. LEMMA

$$\mathcal{D}_E \in M_L^n.$$

3.1.17. LEMMA The function

$$x \to D_E(x) \quad (x \in \mathcal{D}_E)$$

is Lebesgue measurable.

3.1.18. THEOREM

$$\int_{R^n} D_E \, dL^n = L^n(E).$$

PROOF Write

$$R^n = (E \cap \Lambda(\chi_E)) \cup (R^n \backslash E) \cap \Lambda(\chi_E) \cup R^n \backslash \Lambda(\chi_E).$$

Then this is a disjoint union of Lebesgue measurable sets, the third of which, viz. $R^n \backslash \Lambda(\chi_E)$, being of Lebesgue measure 0. Therefore

$$\int_{R^n} D_E \, dL^n$$

$$= \int_{E \cap \Lambda(\chi_E)} D_E \, dL^n + \int_{(R^n \backslash E) \cap \Lambda(\chi_E)} D_E \, dL^n$$

$$= \int_{E \cap \Lambda(\chi_E)} 1 \, dL^n + \int_{(R^n \backslash E) \cap \Lambda(\chi_E)} 0 \, dL^n$$

$$= L^n(E \cap \Lambda(\chi_E)).$$

Write

$$E \cap \Lambda(\chi_E) = E \backslash (R^n \backslash \Lambda(\chi_E))$$

$$= E\backslash(E \cap (R^n\backslash\Lambda(\chi_E)),$$

from which

$$L^n(E \cap \Lambda(\chi_E)) = L^n(E) - L^n(E \cap (R^n\backslash\Lambda(\chi_E))$$
$$= L^n(E),$$

thereby completing the proof.

3.1.19. DEFINITION If $E \subset R^n$ is Lebesgue measurable, then a point $x \in R^n$ (not necessarily in E) is a point of density 1 for E, denoted $x \in E^1$, if $D_E(x) = 1$ and a point of density 0 for E, denoted $x \in E^0$, if $D_E(x) = 0$.

3.1.20. DEFINITION

- E^1 is the measure theoretic interior of E.
- E^0 is the measure theoretic exterior of E.

3.1.21. DEFINITION The measure theoretic boundary of E, denoted $\partial_M E$, is the set of points where the density is neither 0 nor 1.

3.1.22. DEFINITION A Lebesgue measurable set $E \subset R^n$ is d-open if each point of E is a point of density 1, i.e., if $\forall\ x \in E$, $D_E(x) = 1$.

3.1.23. EXAMPLE Take n = 1 -- then the set of irrational numbers is d-open.

3.1.24. LEMMA Every open subset of R^n is d-open.

3.1.25. THEOREM The collection of all d-open sets forms a topology, the density topology.

3.1.26. N.B. The density topology is strictly finer than the euclidean topology.

Let $E \subset R^n$ be Lebesgue measurable -- then

$$\frac{L^n(E \cap B(x,r))}{\omega_n r^n} + \frac{L^n((R^n \backslash E) \cap B(x,r))}{\omega_n r^n}$$

$$= \frac{L^n(R^n \cap B(x,r))}{\omega_n r^n}$$

$$= \frac{L^n(B(x,r))}{\omega_n r^n} = \frac{\omega_n r^n}{\omega_n r^n} = 1.$$

Thus it follows that

$$\left[\begin{array}{l} D_E(x) = 1 \text{ iff } D_{(R^n \backslash E)}(x) = 0 \\ \\ D_E(x) = 0 \text{ iff } D_{(R^n \backslash E)}(x) = 1. \end{array}\right.$$

§3.2. *APPROXIMATE LIMITS*

Let $f: R^n \to R$ be a Lebesgue measurable function.

3.2.1. DEFINITION An element $\ell \in R$ is the <u>approximate limit</u> of f as $y \to x$, denoted

$$\underset{y \to x}{\text{ap} \lim} f(y) = \ell,$$

if for every $\varepsilon > 0$, the set

$$\{y : |f(y) - \ell| \geq \varepsilon\}$$

has density 0 at x, i.e.,

$$\lim_{r \to 0} \frac{L^n(\{|f - \ell| \geq \varepsilon\} \cap B(x,r))}{\omega_n r^n} = 0.$$

3.2.2. LEMMA Approximate limits are unique (thereby justifying the use of "the" in the definition).

PROOF Let ℓ_1 and ℓ_2 be two candidates for the approximate limit per the definition. Assume that $\ell_1 \neq \ell_2$ and take $\varepsilon = |\ell_1 - \ell_2|/3$ -- then for each $y \in R^n$,

$$3\varepsilon = |\ell_1 - \ell_2| \leq |f(y) - \ell_1| + |f(y) - \ell_2|$$

$\Rightarrow$

$$B(x,r) \subset \{|f - \ell_1| \geq \varepsilon\} \cup \{|f - \ell_2| \geq \varepsilon\}.$$

Proof: If there were a $y \in B(x,r)$ which was not in the union, then

$$\left[\begin{array}{l} |f(y) - \ell_1| < \varepsilon \\ |f(y) - \ell_2| < \varepsilon \end{array}\right.$$

=>

$$|f(y) - \ell_1| + |f(y) - \ell_2| < 2\varepsilon$$

$$\Rightarrow \quad 3\varepsilon < 2\varepsilon \Rightarrow 3 < 2 \ldots .$$

Therefore

$$\omega_n r^n = L^n(B(x,r))$$

$$\leq L^n(\{|f - \ell_1| \geq \varepsilon\} \cap B(x,r)) + L^n(\{|f - \ell_2| \geq \varepsilon\} \cap B(x,r)).$$

Now divide through by $\omega_n r^n$ and send r to zero to get $1 \leq 0$.

3.2.3. THEOREM

$$\underset{y \to x}{\text{ap lim}}\ f(y) = \ell$$

iff there exists a Lebesgue measurable set $E \subset R^n$ with $D_E(x) = 1$ such that

$$\lim_{\substack{y \to x \\ y \in E}} f(y) = \ell.$$

[The discussion infra supplies the proof.]

3.2.4. N.B. In view of established principles, x may or may not belong to E. As for the symbol

$$\lim_{\substack{y \to x \\ y \in E}} f(y) = \ell,$$

it means: $\forall\ \varepsilon > 0,\ \exists\ r > 0$ such that

$$|f(y) - \ell| < \varepsilon$$

if $y \in E \cap (B(x,r)\backslash\{x\})$.

Start matters by assuming that the limit above is in force -- then the claim is that for every $\varepsilon > 0$, the set

$$\{y: |f(y) - \ell| \geq \varepsilon\}$$

has density 0 at x or, equivalently, that the set

$$\{y: |f(y) - \ell| < \varepsilon\}$$

has density 1 at x. This set, however, contains $E \cap (B(x,r)\backslash\{x\})$ for small r. Therefore

$$\frac{L^n(\{|f - \ell| < \varepsilon\} \cap B(x,r))}{\omega_n r^n}$$

$$\geq \frac{L^n(E \cap (B(x,r)\backslash\{x\}))}{\omega_n r^n} .$$

But

$$L^n(E \cap B(x,r)) = L^n(E \cap (B(x,r)\backslash\{x\})) + L^n(E \cap \{x\})$$

$$= L^n(E \cap (B(x,r)\backslash\{x\}))$$

=>

$$\frac{L^n(\{|f - \ell| < \varepsilon\} \cap B(x,r))}{\omega_n r^n}$$

$$\geq \frac{L^n(E \cap B(x,r))}{\omega_n r^n} \rightarrow 1 \quad (r \rightarrow 0).$$

In the other direction, assume that

$$\underset{y \rightarrow x}{\text{ap lim}} \; f(x) = \ell,$$

the objective being to construct an $E \in M_L^n$ with the stated property. To this end, choose a strictly decreasing sequence $\{r_k\}$ such that

$$L^n(\{|f - \ell| \geq \frac{1}{k}\} \cap B(x,r_k)) \leq \frac{L^n(B(x,r_k))}{2^k}$$

and put

$$E = R^n \backslash \bigcup_{k=1}^{\infty} (B(x,r_k)\backslash B(x,r_{k+1})) \cap \{|f - \ell| \geq \frac{1}{k}\}.$$

Then E is Lebesgue measurable and

$$\lim_{\substack{y \to x \\ y \in E}} f(y) = \ell.$$

There remains the contention that $D_E(x) = 1$ or still, that

$$D_{(R^n\backslash E)}(x) = 0.$$

By definition,

$$R^n\backslash E = \bigcup_{k=1}^{\infty} (B(x,r_k)\backslash B(x,r_{k+1})) \cap \{y:|f(y) - \ell| \geq \frac{1}{k}\}.$$

Given $r > 0$, denote by K the integer for which $r_{K+1} < r < r_K$ -- then

$$\begin{aligned}
&L^n((R^n\backslash E) \cap B(x,r)) \\
&\leq \sum_{k=K}^{\infty} L^n((B(x,r_k)\backslash B(x,r_{k+1})) \cap \{|f - \ell| \geq \frac{1}{k}\}) \\
&\leq \sum_{k=K}^{\infty} L^n(B(x,r_k) \cap \{|f - \ell| \geq \frac{1}{k}\}) \\
&\leq \sum_{k=K}^{\infty} \frac{L^n(B(x,r_k))}{2^k} \\
&\leq L^n(B(x,r)) \sum_{k=K}^{\infty} \frac{1}{2^k} \to 0 \ (r \to 0).
\end{aligned}$$

3.2.5. DEFINITION A Lebesgue measurable function $f:R^n \to R$ is <u>approximately</u>

<u>continuous</u> at $x \in R^n$ if f is defined at x and

$$\operatorname{ap}\lim_{y \to x} f(y) = f(x).$$

3.2.6. SCHOLIUM f is approximately continuous at x iff there exists a Lebesgue measurable set $E \subset R^n$ with $D_E(x) = 1$ such that $f|E$ is continuous at x.

3.2.7. REMARK In terms of the density topology, f is approximately continuous at x iff f is d-continuous at x.

3.2.8. THEOREM A Lebesgue measurable function $f:R^n \to R$ is approximately continuous L^n almost everywhere.

PROOF Given $\varepsilon > 0$, there is a continuous function $g:R^n \to R$ and an open set $G \subset R^n$ such that $L^n(G) < \varepsilon$ and $f = g$ in $R^n \backslash G$. On general grounds, almost every point of $R^n \backslash G$ is a point of density 1, thus f is approximately continuous at almost all points of $R^n \backslash G$. The arbitrariness of the data then implies that f is approximately continuous at almost all points of R^n.

3.2.9. LEMMA Let $f \in L^1_{\ell oc}(R^n)$ and let $\ell \in R^n$. Assume:

$$\lim_{r \to 0} \frac{1}{\omega_n r^n} \int_{B(x,r)} |f - \ell| \, dL^n = 0.$$

Then

$$\operatorname{ap}\lim_{y \to x} f(y) = \ell.$$

PROOF Thanks to Chebyshev, $\forall\ \varepsilon > 0$,

$$\varepsilon \frac{L^n(\{|f - \ell| \geq \varepsilon\} \cap B(x,r))}{\omega_n r^n} \leq \frac{1}{\omega_n r^n} \int_{B(x,r)} |f - \ell| \, dL^n.$$

3.2.10. APPLICATION The approximate limit exists at each Lebesgue point x of f and coincides with the value f(x).

The literal converse to 3.2.9. is false in general, i.e., it can happen that $\forall\ \varepsilon > 0$,

$$\lim_{r \to 0} \frac{L^n(\{|f - \ell| \geq \varepsilon\} \cap B(x,r))}{\omega_n r^n} = 0,$$

yet the relation

$$\lim_{r \to 0} \frac{1}{\omega_n r^n} \int_{B(x,r)} |f - \ell|\, dL^n = 0$$

fails or, what amounts to the same, it can happen that at some x, there is no ℓ such that

$$\lim_{r \to 0} \frac{1}{\omega_n r^n} \int_{B(x,r)} |f - \ell|\, dL^n = 0$$

but for some ℓ and $\forall\ \varepsilon > 0$,

$$\lim_{r \to 0} \frac{L^n(\{|f - \ell| \geq \varepsilon\} \cap B(x,r))}{\omega_n r^n} = 0.$$

3.2.11. EXAMPLE In R^2, take $\alpha > 0$ and consider

$$f(x,y) = \begin{cases} 0 & \text{if } y \leq 0 \text{ or } y \geq x^2 \\ |y|^{-\alpha}, & \text{otherwise.} \end{cases} \qquad (x,y) \in R^2$$

Then

$$\frac{1}{\omega_2 r^2} \int_{B((0,0),r)} |f|\, dxdy$$

tends to $+\infty$ as $r \to 0$ if $1/2 < \alpha < 1$ while choosing $\ell = 0$, the sets $\{|f - 0| \geq \varepsilon\}$

have density 0 at (0,0).

3.2.12. LEMMA Suppose that the sets

$$E_\varepsilon = \{|f - \ell| \geq \varepsilon\}$$

have density 0 at x and f is bounded in a neighborhood of x, say $|f| \leq M$ -- then

$$\lim_{r \to 0} \frac{1}{\omega_n r^n} \int_{B(x,r)} |f - \ell| \, dL^n = 0.$$

PROOF Write

$$\frac{1}{\omega_n r^n} \int_{B(x,r)} |f - \ell| \, dL^n$$

$$\leq (M + |\ell|) \frac{L^n(E_\varepsilon \cap B(x,r))}{\omega_n r^n} + \varepsilon \frac{L^n(B(x,r) \backslash E_\varepsilon)}{\omega_n r^n}$$

from which

$$\limsup_{r \to 0} \frac{1}{\omega_n r^n} \int_{B(x,r)} |f - \ell| \, dL^n \leq \varepsilon$$

Now let $\varepsilon \to 0$.

3.2.13. DEFINITION f has an <u>AFP approximate limit ℓ</u> at x if

$$\lim_{r \to 0} \frac{1}{\omega_n r^n} \int_{B(x,r)} |f - \ell| \, dL^n = 0.$$

3.2.14. NOTATION S_f is the set of points x which do not possess an AFP approximate limit.

3.2.15. <u>N.B.</u> If f has an AFP approximate limit ℓ at x, then

$$\operatorname{ap}\lim_{y \to x} f(y) = \ell,$$

the converse being false in general (cf. supra).

3.2.16. LEMMA AFP approximate limits are unique.

3.2.17. NOTATION Write $\bar{f}(x)$ in place of ℓ.

3.2.18. OBSERVATION If $\bar{f}(x) = f(x)$, then x is a Lebesgue point of f.

3.2.19. LEMMA The set of points where the AFP approximate limit exists does not depend on the representative in the equivalence class of f, i.e., if $f = g$ L^n almost everywhere in Ω, then $x \notin S_f$ iff $x \notin S_g$ and $\bar{f}(x) = \bar{g}(x)$.

3.2.20. LEMMA S_f is a Borel set of Lebesgue measure 0.

PROOF The complement of the Lebesgue set $\Lambda(f)$ of f is a set of Lebesgue measure 0, hence $L^n(S_f) = 0$. As for S_f being Borel, write

$$R^n \backslash S_f = \bigcap_{k=1}^{\infty} \bigcup_{q \in Q} \{x: \limsup_{r \to 0} \frac{1}{\omega_n r^n} \int_{B(x,r)} |f - q| \, dL^n < \frac{1}{k}\}.$$

[The inclusion $\subset$ is trivial. On the other hand, if x belongs to the set on the RHS, then for any integer $k \geq 1$, there is a $q_k \in Q$ such that

$$\limsup_{r \to 0} \frac{1}{\omega_n r^n} \int_{B(x,r)} |f - q_k| \, dL^n < \frac{1}{k} .$$

The sequence $\{q_k\}$ obtained in this way is Cauchy and its limit ℓ has the property that

$$\lim_{r \to 0} \int_{B(x,r)} |f - \ell| \, dL^n = 0,$$

so $x \notin S_f$, i.e., $x \in R^n \backslash S_f$.]

3.2.21. LEMMA $\bar{f}: R^n \backslash S_f \to R$ is a Borel function which coincides L^n almost everywhere with $f | R^n \backslash S_f$.

PROOF In fact, for any $x \in R^n \backslash S_f$,

$$\lim_{r \to 0} \frac{1}{\omega_n r^n} \int_{B(x,r)} f \, dL^n = \ell = \bar{f}(x),$$

thus $\bar{f}$ is the pointwise limit as $r \to 0$ of the continuous function

$$x \to \frac{1}{\omega_n r^n} \int_{B(x,r)} f \, dL^n.$$

3.2.22. EXAMPLE Suppose that $f = \chi_E$ is a characteristic function (E a Lebesgue measurable set) -- then S_f is the measure theoretic boundary $\partial_M E$ of E.

On occasion, it will be necessary to consider a generalization of "ap lim".

3.2.23. DEFINITION Let $f:R^n \to R$ be a Lebesgue measurable function.

- An element $\ell \in R$ is the <u>approximate lim sup</u> of f as $y \to x$, denoted

$$\operatorname*{ap\,lim\,sup}_{y \to x} f(y),$$

if ℓ is the infimum of the real numbers t such that

$$\lim_{r \to 0} \frac{L^n(\{f > t\} \cap B(x,r))}{\omega_n r^n} = 0.$$

- An element $\ell \in R$ is the <u>approximate lim inf</u> of f as $y \to x$, denoted

$$\operatorname*{ap\,lim\,inf}_{y \to x} f(y),$$

if ℓ is the supremum of the real numbers t such that

$$\lim_{r \to 0} \frac{L^n(\{f < t\} \cap B(x,r))}{\omega_n r^n} = 0.$$

Obviously

$$\underset{y \to x}{\text{ap lim inf}}\, f(y) \le \underset{y \to x}{\text{ap lim sup}}\, f(y)$$

and if

$$\underset{y \to x}{\text{ap lim inf}}\, f(y) = \underset{y \to x}{\text{ap lim sup}}\, f(y)$$

and if their common value is ℓ, then the approximate limit exists and

$$\underset{y \to x}{\text{ap lim}}\, f(y) = \ell.$$

* * * * * * * * * * *

APPENDIX

The preceding considerations have been formulated under the assumption that $f:R^n \to R$ is Lebesgue measurable. Matters can be generalized. Thus let $S \subset R^n$ be Lebesgue measurable and suppose that $f:S \to R$ is Lebesgue measurable. Fix a point $x \in R^n$ such that $D_S(x) = 1$.

DEFINITION An element $\ell \in R$ is the <u>approximate limit</u> of f as $y \to x$ in S, denoted

$$\underset{\substack{y \to x \\ y \in S}}{\text{ap lim}}\, f(y) = \ell,$$

if for every $\varepsilon > 0$, the set

$$\{y \in S: |f(y) - \ell| \ge \varepsilon\}$$

has density 0 at x.

<u>N.B.</u> If $S = R^n$, then the demand that $D_{R^n}(x) = 1$ is automatic. Proof:

$$D_{R^n}(x) = \frac{L^n(R^n \cap B(x,r))}{\omega_n r^n}$$

$$= \frac{L^n(B(x,r))}{\omega_n r^n} = \frac{\omega_n r^n}{\omega_n r^n} = 1.$$

The earlier developments carry over modulo minor changes here and there. In particular: Approximate limits are unique and the notion of approximate continuity is clear.

THEOREM

$$\operatorname*{ap\,lim}_{\substack{y \to x \\ y \in S}} f(y) = \ell$$

iff there exists a Lebesgue measurable set $E \subset R^n$ with $D_E(x) = 1$ such that

$$\lim_{\substack{y \to x \\ y \in E}} f(y) = \ell.$$

§3.3. *APPROXIMATE DERIVATIVES*

Let $f: R^n \to R$ be a Lebesgue measurable function.

3.3.1. DEFINITION f is approximately differentiable at a point $x \in R^n$ if there exists a linear function $T: R^n \to R$ (depending on x) such that

$$\operatorname*{ap\,lim}_{y \to x} \frac{|f(y) - f(x) - T(y - x)|}{||y - x||} = 0.$$

T is called the approximate differential of f at x and is denoted by

$$\text{ap } df(x).$$

[Note: If f is differentiable at x in the ordinary sense, then f is approximately differentiable at x and

$$df(x) = \text{ap } df(x).]$$

3.3.2. N.B. Existence is implied by demanding that

$$\operatorname*{ap\,lim\,sup}_{y \to x} \frac{|f(y) - f(x) - T(y - x)|}{||y - x||} = 0.$$

3.3.3. LEMMA An approximate differential is unique (if it exists at all).

PROOF Let T_1 and T_2 be two candidates for the approximate differential -- then $\forall\, \varepsilon > 0$,

$$\lim_{r \to 0} \frac{L^n(\{y: \dfrac{|f(y) - f(x) - T_1(y - x)|}{||y - x||} \geq \varepsilon\} \cap B(x,r))}{\omega_n r^n} = 0$$

and

$$\lim_{r \to 0} \frac{L^n(\{y: \frac{|f(y) - f(x) - T_2(y - x)|}{||y - x||} \geq \varepsilon\} \cap B(x,r))}{\omega_n r^n} = 0.$$

To get a contradiction, suppose that $T_1 \neq T_2$ and take $\varepsilon = ||T_1 - T_2||/6$. Let

$$S = \{y: |(T_1 - T_2)(y - x)| \geq \frac{||T_1 - T_2|| \; ||y - x||}{2}\}.$$

Then

$$\frac{L^n(S \cap B(x,r))}{\omega_n r^n} \equiv C > 0$$

for all $r > 0$. On the other hand,

$$y \in S \Rightarrow 3\varepsilon||y - x|| = \frac{||T_1 - T_2|| \; ||y - x||}{2}$$

$$\leq |(T_1 - T_2)(y - x)|$$

$$\leq |f(y) - f(x) - T_1(y - x)| + |f(y) - f(x) - T_2(y - x)|$$

$\Rightarrow$

$$S \subset \{y: \frac{|f(y) - f(x) - T_1(y - x)|}{||y - x||} \geq \varepsilon\} \cup \{y: \frac{|f(y) - f(x) - T_2(y - x)|}{||y - x||} \geq \varepsilon\}$$

$\Rightarrow$

$$\lim_{r \to 0} \frac{L^n(S \cap B(x,r))}{\omega_n r^n} = 0,$$

a contradiction... .

[Note: Here is a different proof. Suppose that $T_1 \neq T_2$ and put $T = T_1 - T_2$, hence

$$\operatorname{ap} \lim_{y \to x} \frac{|T(y - x)|}{||y - x||} = 0$$

or still,

$$\operatorname*{ap\,lim}_{v \to 0} \frac{|T(v)|}{||v||} = 0.$$

So, if $0 < \varepsilon < 1$, then there exists $r > 0$ such that

$$\frac{L^n(B(0,r) \cap \{v: |T(v)| \geq \varepsilon||v||\})}{\omega_n r^n} < \varepsilon$$

and for every $u \in R^n$ with $||u|| = r - r\varepsilon$, there exists

$$v \in B(u,\varepsilon r) \subset B(0,r)$$

with

$$|T(v)| \leq \varepsilon||v||$$

which implies that

$$\begin{aligned} |T(u)| = |T(u - v + v)| &\leq |T(u - v)| + |T(v)| \\ &\leq ||T||\ ||u - v|| + \varepsilon||v|| \\ &\leq ||T||\ \varepsilon r + \varepsilon||v|| \\ &\leq ||T||\ \varepsilon r + \varepsilon r = (||T|| + 1)\varepsilon r. \end{aligned}$$

And

$$r = \frac{r(1 - \varepsilon)}{1 - \varepsilon} = \frac{r - r\varepsilon}{1 - \varepsilon} = \frac{||u||}{1 - \varepsilon}$$

=>

$$\varepsilon r = \frac{\varepsilon}{1 - \varepsilon}\ ||u||$$

=>

$$\begin{aligned} |T(u)| &\leq (||T|| + 1)\varepsilon r \\ &\leq (||T|| + 1)\ \frac{\varepsilon}{1 - \varepsilon}\ ||u|| \end{aligned}$$

=>

$$||T|| \le (||T|| + 1) \frac{\varepsilon}{1 - \varepsilon}$$

=>

$$||T|| = 0 \Rightarrow T_1 = T_2.]$$

3.3.4. REMARK If f is approximately differentiable at x, then f is approximately continuous at x.

3.3.5. THEOREM Let $f,g:R^n \to R$ be Lebesgue measurable functions. Assume: f is approximately differentiable almost everywhere and f = g almost everywhere -- then g is approximately differentiable almost everywhere and

$$\text{ap } df = \text{ap } dg$$

almost everywhere in R^n.

Therefore the notion of approximate differentiability does not depend on the particular choice of the representative in the equivalence class.

3.3.6. THEOREM f is approximately differentiable at x iff there exists a Lebesgue measurable set $E \subset R^n$ and a linear function $T:R^n \to R$ with $D_E(x) = 1$ such that

$$\lim_{\substack{y \to x \\ y \in E}} \frac{|f(y) - f(x) - T(y - x)|}{||y - x||} = 0.$$

3.3.7. DEFINITION For $i = 1,\ldots,n$, the <u>approximate partial derivative</u> $\text{ap } D_i f(x)$ of f at a point $x \in R^n$ is defined by the condition

$$\operatorname*{ap\,lim}_{t \to 0} \frac{|f(x + te_i) - f(x) - \text{ap } D_i f(x)t|}{|t|} = 0.$$

3.3.8. THEOREM The following conditions are equivalent.

(a) The function f has approximate partial derivatives almost everywhere in R^n.

(b) The function f is approximately differentiable almost everywhere in R^n.

[Note: Work in R^n ($n > 1$) -- then it can happen that the partial derivatives of f exist almost everywhere in R^n, yet f might be nowhere differentiable (but, of course, f will be approximately differentiable almost everywhere in R^n).]

3.3.9. N.B. The equivalent conditions (a) and (b) are also equivalent to

(c) For every $\varepsilon > 0$ there is a locally Lipschitz function $g:R^n \to R$ such that

$$L^n(\{x:f(x) \neq g(x)\}) < \varepsilon$$

or even

(d) For every $\varepsilon > 0$ there is a C'-function $g:R^n \to R$ such that

$$L^n(\{x:f(x) \neq g(x)\}) < \varepsilon.$$

3.3.10. NOTATION $A_D(f)$ is the domain of existence of ap df.

3.3.11. LEMMA If f is approximately differentiable at L^n almost all points in R^n, i.e., if $L^n(R^n \backslash A_D(f)) = 0$, then there exist Lebesgue measurable sets E_0, E_k ($k = 1,2,\ldots$) such that

$$A_D(f) = E_0 \cup \bigcup_{k=1}^{\infty} E_k,$$

where $L^n(E_0) = 0$ and for every k, the restriction $f|E_k$ is Lipschitz.

3.3.12. N.B.

$$\text{ap } df:A_D(f) \to R$$

is Lebesgue measurable.

Owing to 2.5.1., f is differentiable at almost all points where

$$\limsup_{y \to x} \frac{|f(x) - f(y)|}{||x - y||} < +\infty.$$

3.3.13. LEMMA f is approximately differentiable at almost all points where

$$\operatorname{ap}\limsup_{y \to x} \frac{|f(x) - f(y)|}{||x - y||} < +\infty.$$

* * * * * * * * * *

APPENDIX

Suppose that f has ordinary partial derivatives almost everywhere -- then f is approximately differentiable almost everywhere, thus if f is approximately differentiable at x, there exists a Lebesgue measurable set $E \subset R^n$ with $D_E(x) = 1$ such that $f|E$ is differentiable at x in the ordinary sense. Moreover,

$$d(f|E)(x) = \operatorname{ap} df(x).$$

Assume now that n = 2 -- then in this special case f admits a <u>regular</u> approximate differential at x. Here "regular" means that the ubiquitous set E is comprised of the boundaries of oriented squares centered at x.

SUMMARY If $f:R^2 \to R$ has ordinary partial derivatives almost everywhere, then it has a regular approximate differential almost everywhere.

SECTION 4: WEAK PARTIAL DERIVATES

Let Ω be a nonempty open subset of R^n.

4.1. DEFINITION A Lebesgue measurable function $f:\Omega \to R$ is <u>locally integrable</u> if

$$\int_K |f| \, dL^n < +\infty$$

for every compact $K \subset \Omega$.

Denote the space of such by

$$L^1_{\ell oc}(\Omega).$$

4.2. EXAMPLE Take $\Omega = R$ — then $\ell n|x| \in L^1_{\ell oc}(R)$ but $x^{-1} \notin L^1_{\ell oc}(R)$.

4.3. DEFINITION Let $1 \leq p < +\infty$ — then a Lebesgue measurable function $f:\Omega \to R$ is <u>locally L^p</u> if

$$\int_K |f|^p \, dL^n < +\infty$$

for every compact $K \subset \Omega$.

4.4. LEMMA Every locally L^p function f is locally L^1 (i.e., is locally integrable).

PROOF Given a compact $K \subset \Omega$, by Hölder's inequality

$$\int_K |f| \, dL^n \leq ||\chi_K f||_{L^p} ||\chi_K||_{L^{p'}},$$

where $p' = +\infty$ if $p = 1$ and $p' = p/(p-1)$ if $1 < p < +\infty$.

4.5. <u>N.B.</u> The product of two functions in $L^1_{\ell oc}(\Omega)$ need not be locally integrable.

Every $f \in L^1_{\ell oc}(\Omega)$ determines a distribution via the arrow

$$\varphi \to \int_\Omega \varphi \, f dL^n.$$

Moreover two locally integrable functions define the same distribution iff they are equal almost everywhere.

[Note: A distribution T "is a function" if there exists an element $f \in L^1_{\ell oc}(\Omega)$ such that $T = fdL^n$.]

4.6. NOTATION Let $T:C_c^\infty(\Omega) \to R$ be a distribution -- then

$$\frac{\partial T}{\partial x_i} \quad (i = 1,\ldots,n)$$

is the distributional derivative of $T:\forall\ \varphi \in C_c^\infty(\Omega)$,

$$< \varphi, \frac{\partial T}{\partial x_i} > = - < \frac{\partial \varphi}{\partial x_i}, T >.$$

4.7. DEFINITION Given an $f \in L^1_{\ell oc}(\Omega)$, denote by

$$\frac{\partial f}{\partial x_i} \quad (i = 1,\ldots,n)$$

its distributional derivative (per $T = fdL^n$) -- then $\frac{\partial f}{\partial x_i}$ is said to be a <u>weak partial derivative</u> of f if $\frac{\partial f}{\partial x_i} \in L^1_{\ell oc}(\Omega)$, thus $\forall\ \varphi \in C_c^\infty(\Omega)$,

$$< \varphi, \frac{\partial f}{\partial x_i} > = - < \frac{\partial \varphi}{\partial x_i}, fdL^n >$$

or still,

$$\int_\Omega \varphi \frac{\partial f}{\partial x_i} dL^n = - \int_\Omega \frac{\partial \varphi}{\partial x_i} fdL^n.$$

4.8. EXAMPLE Take $\Omega = R$ and consider the function

$$h(x) = \begin{cases} 0 \text{ if } x \le 0 \\ x \text{ if } x > 0. \end{cases}$$

Then $h \in L^1_{\ell oc}(R)$ and its distributional derivative $\frac{dh}{dx}$ is the Heaviside function

$$H(x) = \begin{cases} 0 \text{ if } x \le 0 \\ 1 \text{ if } x > 0, \end{cases}$$

which is therefore the weak derivative of h. Since $H \in L^1_{\ell oc}(R)$, one can form its distributional derivative $\frac{dH}{dx}$, so

$$\begin{aligned} < \varphi, \frac{dH}{dx} > &= - < \frac{d\varphi}{dx}, HdL^1 > \\ &= - \int_R \frac{d\varphi}{dx} H(x) dL^1 \\ &= - \int_0^\infty \frac{d\varphi}{dx} dL^1 \\ &= - [\varphi(\infty) - \varphi(0)] = \varphi(0). \end{aligned}$$

Consequently $\frac{dH}{dx} = \delta$, the Dirac measure concentrated at the origin. However there is no $f \in L^1_{\ell oc}(R)$ such that

$$\int_R \varphi \, f dL^1 = \varphi(0)$$

for all $\varphi \in C_c^\infty(R)$, hence H does not have a weak derivative.

4.9. DEFINITION Suppose that $f \in L^1_{\ell oc}(\Omega)$ admits weak partial derivatives $\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}$ -- then the <u>distributional gradient</u> attached to f is the n-tuple

$$\nabla f = (\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}).$$

4.10. EXAMPLE Working in R^n, take $\Omega = B(0,1)^\circ \backslash \{0\}$ and define $f \in C^\infty(B(0,1)^\circ \backslash \{0\})$ by the rule

$$f(x) = ||x||^{-\alpha} \quad (\alpha > 0) \quad (||x|| = (x_1^2 + \cdots + x_n^2)^{1/2}).$$

Then f is unbounded in every neighborhood of the origin $(0 < ||x|| < 1)$ and

$$\frac{\partial f}{\partial x_i}(x) = -\alpha \frac{x_i}{||x||^{\alpha + 2}} \quad (i = 1, \ldots, n).$$

Therefore

$$\nabla f(x) = -\alpha \frac{x}{||x||^{\alpha + 2}}$$

$\Rightarrow$

$$||\nabla f(x)|| = \frac{|\alpha|}{||x||^{\alpha + 1}},$$

where

$$||\nabla f(x)|| = (\sum_{i=1}^{n} \left| \frac{\partial f}{\partial x_i}(x) \right|^2)^{1/2}.$$

4.11. RAPPEL Let S^{n-1} $(= \partial B(0,1))$ be the unit sphere and let σ^{n-1} be its surface measure, thus

$$\begin{aligned} \sigma^{n-1}(S^{n-1}) = nL^n(B(0,1)) &= n \frac{\pi^{n/2}}{\Gamma(1 + n/2)} \\ &= n\omega_n. \end{aligned}$$

4.12. APPLICATION Given $\alpha > 0$ subject to $n > \alpha$, put $f(x) = ||x||^{-\alpha}$ and write

$$\int_{B(0,1)} |f| \, dL^n$$

$$= \int_{B(0,1)} ||x||^{-\alpha} \, dL^n$$

$$= \int_0^1 \int_{S^{n-1}} ||rx||^{-\alpha} r^{n-1} \, d\sigma^{n-1}(x) dr$$

$$= \sigma^{n-1}(S^{n-1}) \int_0^1 r^{-\alpha + n-1} \, dr$$

$$= \sigma^{n-1}(S^{n-1}) \left. \frac{r^{-\alpha + n}}{-\alpha + n} \right|_0^1 < + \infty.$$

Therefore $f \in L^1(B(0,1))$.

4.13. EXAMPLE Consider again $f(x) = ||x||^{-\alpha}$ $(\alpha > 0)$ but replace $\Omega = B(0,1)^\circ \backslash \{0\}$ by $\Omega = B(0,1)^\circ$ -- then

$$n > \alpha \Rightarrow f \in L^1(B(0,1))$$

$$\Rightarrow f \in L^1(\Omega) \Rightarrow f \in L^1_{\ell oc}(\Omega).$$

Next

$$\int_\Omega \left|\frac{\partial f}{\partial x_i}\right| \, dL^n$$

$$\leq \int_\Omega ||\nabla f|| \, dL^n$$

$$= |\alpha| \int_\Omega \frac{1}{||x||^{\alpha + 1}} \, dL^n$$

$$= |\alpha| \sigma^{n-1}(S^{n-1}) \left. \frac{r^{n - \alpha - 1}}{n - \alpha - 1} \right|_0^1 < + \infty$$

if $n > \alpha + 1$, so

$$\frac{\partial f}{\partial x_i} \in L^1(\Omega) \Rightarrow \frac{\partial f}{\partial x_i} \in L^1_{\ell oc}(\Omega).$$

Let T be the distribution corresponding to f, hence $\forall\ \varphi \in C_c^\infty(\Omega)$,

$$< \varphi, \frac{\partial T}{\partial x_i} > = - < \frac{\partial \varphi}{\partial x_i}, T >$$

$$= - < \frac{\partial \varphi}{\partial x_i}, f\, dL^n >$$

$$= < \varphi, \frac{\partial f}{\partial x_i}\, dL^n > \quad \text{(dominated convergence)}.$$

Accordingly, as distributions,

$$\frac{\partial T}{\partial x_i} = \frac{\partial f}{\partial x_i}\, dL^n.$$

Therefore f admits weak partial derivatives in $B(0,1)^\circ$ (and not just in $B(0,1)\backslash\{0\}$).

4.14. LEMMA If $f \in C^1(\Omega)$, then the ordinary partial derivatives $\frac{\partial f}{\partial x_i}$ of f are also the corresponding weak partial derivatives of f.

SECTION 5: MOLLIFIERS

Let Ω be a nonempty open subset of R^n.

5.1. NOTATION Given $\varepsilon > 0$ and a nonnegative even bounded function $\varphi \in L^1(R^n)$ with

$$\operatorname{spt} \varphi \subset B(0,1), \ \int_{R^n} \varphi \, dL^n = 1,$$

put

$$\varphi_\varepsilon(x) = \frac{1}{\varepsilon^n} \varphi\left(\frac{x}{\varepsilon}\right) \quad (x \in R^n).$$

5.2. DEFINITION The φ_ε are called <u>mollifiers</u>.

5.3. <u>N.B.</u> Mollifiers exist... .

[The <u>standard</u> choice for φ is

$$\varphi(x) = \begin{cases} C(n) \exp\left(\dfrac{1}{||x||^2 - 1}\right) \text{ if } ||x|| < 1 \\ \\ 0 \text{ if } ||x|| \geq 1, \end{cases}$$

where $C(n) > 0$ is so chosen that

$$\int_{R^n} \varphi \, dL^n = 1.$$

Here $\varphi \in C_c^\infty(R^n)$. Another possibility is

$$\varphi(x) = \frac{1}{\omega_n} \chi_{B(0,1)^\circ}(x).]$$

5.4. NOTATION Put

$$\Omega_\varepsilon = \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > \varepsilon\}.$$

[Note: If $\Omega = R^n$, then $\Omega_\varepsilon = R^n$.]

5.5. DEFINITION Given a function $f \in L^1_{\ell oc}(\Omega)$, write

$$f_\varepsilon(x) = (f * \varphi_\varepsilon)(x) = \int_\Omega \varphi_\varepsilon(x - y)f(y)\, dL^n,$$

where $x \in \Omega_\varepsilon$.

[Note:

$$\int_\Omega \varphi_\varepsilon(x - y)f(y)\, dL^n = \int_{B(x,\varepsilon)} \varphi_\varepsilon(x - y)f(y)\, dL^n.]$$

The function $f_\varepsilon:\Omega_\varepsilon \to R$ is said to be a <u>mollification</u> of f, the <u>standard mollification</u> of f being the f_ε per the standard choice for φ per supra.

[Note: Given $x \in \Omega$, $f_\varepsilon(x)$ is well defined for all $0 < \varepsilon < \mathrm{dist}(x,\partial\Omega)$, thus it makes sense to consider

$$\lim_{\varepsilon > 0} f_\varepsilon(x).]$$

5.6. THEOREM If f_ε is the standard mollification of f, then $f_\varepsilon \in C^\infty(\Omega_\varepsilon)$ $(0 < \varepsilon < 1)$ and for every multi index α and for every $x \in \Omega_\varepsilon$,

$$\partial^\alpha f_\varepsilon(x) = (f * \partial^\alpha \varphi_\varepsilon)(x) = \int_\Omega \frac{\partial^{|\alpha|}\varphi_\varepsilon}{\partial x^\alpha}(x - y)f(y)\, dL^n .$$

5.7. LEMMA If the standard choice for φ is used and if $f \in L^1_{\ell oc}(\Omega)$ admits a weak partial derivative $\frac{\partial f}{\partial x_i}$ (hence, by definition, $\frac{\partial f}{\partial x_i} \in L^1_{\ell oc}(\Omega)$), then the derivative of the mollification coincides with the mollification of the weak partial derivative, i.e.,

$$\frac{\partial f_\varepsilon}{\partial x_i} = \frac{\partial f}{\partial x_i} * \varphi_\varepsilon.$$

In fact,

$$\frac{\partial f_\varepsilon}{\partial x_i}(x) = \frac{\partial}{\partial x_i}\left(\int_\Omega \varphi_\varepsilon(x - y)f(y)\ dL^n\right)$$

$$= \int_\Omega \frac{\partial}{\partial x_i}\varphi_\varepsilon(x - y)f(y)\ dL^n$$

$$= (-1)\int_\Omega \frac{\partial}{\partial y_i}\varphi_\varepsilon(x - y)f(y)\ dL^n$$

$$= (-1)(-1)\int_\Omega \varphi_\varepsilon(x - y)\frac{\partial f}{\partial y_i}(y)\ dL^n$$

$$= \left(\frac{\partial f}{\partial x_i} * \varphi_\varepsilon\right)(x).$$

5.8. APPLICATION Work with the standard choice for φ, suppose that Ω is connected, let $f \in L^1_{\ell oc}(\Omega)$, assume that the weak partial derivatives $\frac{\partial f}{\partial x_i}$ $(i = 1,\ldots,n)$ exist and are equal to 0 almost everywhere -- then f coincides almost everywhere in Ω with a constant function.

[To begin with,

$$\frac{\partial f_\varepsilon}{\partial x_i} = \frac{\partial f}{\partial x_i} * \varphi_\varepsilon = 0 * \varphi_\varepsilon = 0,$$

thus f_ε, being smooth, must be constant in each connected component of Ω_ε (in general, Ω_ε is not connected). Consider now a pair of points $x,y \in \Omega$ -- then there exists a polygonal path γ in Ω joining x and y and for small enough ε, γ is in Ω_ε, so $f_\varepsilon(x) = f_\varepsilon(y)$. But $f_\varepsilon \rightarrow f$ $(\varepsilon \downarrow 0)$ almost everywhere (see below). Therefore $\lim f_\varepsilon$ $(\varepsilon \downarrow 0)$ is a constant function in Ω.]

5.9. LEMMA Suppose given $f \in C(\Omega)$ $(\subset L^1_{\ell oc}(\Omega))$ -- then for every choice of φ,

$$f_\varepsilon \to f \quad (\varepsilon \downarrow 0)$$

uniformly on compact subsets of Ω.

5.10. LEMMA Suppose given $f \in L^1_{\ell oc}(\Omega)$ -- then for every choice of φ and for every Lebesgue point $x \in \Omega$,

$$f_\varepsilon(x) \to f(x) \quad (\varepsilon \downarrow 0),$$

hence $f_\varepsilon \to f$ almost everywhere.

PROOF Write

$$|f_\varepsilon(x) - f(x)| = |\textstyle\int_\Omega \varphi_\varepsilon(x - y) f(y) \, dL^n - f(x)|$$

$$= |\textstyle\int_{B(x,\varepsilon)} \varphi_\varepsilon(x - y) f(y) \, dL^n - f(x) \int_{B(x,\varepsilon)} \varphi_\varepsilon(x - y) \, dL^n|$$

$$= |\textstyle\int_{B(x,\varepsilon)} \varphi_\varepsilon(x - y)(f(y) - f(x)) \, dL^n|$$

$$\leq \frac{1}{\varepsilon^n} \int_{B(x,\varepsilon)} \varphi_\varepsilon\left(\frac{x - y}{\varepsilon}\right) |f(y) - f(x)| \, dL^n$$

$$\leq ||\varphi||_\infty \frac{\omega_n}{\omega_n} \frac{1}{\varepsilon^n} \int_{B(x,\varepsilon)} |f - f(x)| \, dL^n$$

$$= ||\varphi||_\infty \, \omega_n \frac{1}{\omega_n \varepsilon^n} \int_{B(x,\varepsilon)} |f - f(x)| \, dL^n$$

$$\to 0 \quad (\varepsilon \downarrow 0).$$

5.11. LEMMA Suppose given $f \in L^p_{\ell oc}(\Omega)$ $(1 \leq p < +\infty)$ $(\Rightarrow f \in L^1_{\ell oc}(\Omega))$ -- then for every choice of φ,

$$\lim_{\varepsilon \downarrow 0} ||f_\varepsilon - f||_{L^p(\Omega)} = 0.$$

SECTION 6: $W^{1,\infty}(R^n)$

Let $(X,\mathcal{E})$ be a measurable space and let μ be a measure on $(X,\mathcal{E})$.

6.1. NOTATION Given a measurable function $f:X \to R$, put

$$||f||_\infty = \inf\{t \geq 0:\mu(\{x:|f(x)| > t\}) = 0\},$$

with the convention that $\inf \emptyset = \infty$.

6.2. DEFINITION $||f||_\infty$ is the <u>essential supremum</u> of f and is written

$$||f||_\infty = \operatorname*{ess\,sup}_{x\in X} |f(x)|.$$

6.3. NOTATION

$$L^\infty(X) \equiv L^\infty(X,\mathcal{E},\mu)$$

is the set of measurable functions defined on X for which $||f||_\infty < \infty$.

[Note: Such functions are said to be <u>essentially bounded</u> and if f is one such, then

$$|f(x)| \leq ||f||_\infty$$

almost everywhere.]

6.4. LEMMA $f \in L^\infty(X)$ iff there is a bounded measurable function g such that f = g almost everywhere.

6.5. LEMMA $L^\infty(X)$ is a Banach space.

Henceforth the pair $(X,\mathcal{E})$ will be the pair (R^n,M_L^n), μ being L^n.

6.6. <u>N.B.</u> The set of bounded continuous functions $f:R^n \to R$ carries the uniform norm

$$||f||_u = \sup_{x \in X} |f(x)|$$

and

$$||f||_u = ||f||_\infty.$$

6.7. NOTATION $W^{1,\infty}(R^n)$ is the space consisting of all essentially bounded functions $f:R^n \to R$ whose distributional derivatives

$$\frac{\partial f}{\partial x_i} \quad (i = 1,\ldots,n)$$

are also essentially bounded functions in R^n as well.

In $W^{1,\infty}(R^n)$ introduce the norm

$$||f||_{W^{1,\infty}} \equiv ||f||_\infty + \sum_{i=1}^{n} ||\frac{\partial f}{\partial x_i}||_\infty.$$

6.8. THEOREM $W^{1,\infty}(R^n)$ is a Banach space.

6.9. NOTATION Given $f \in L^1_{\ell oc}(R^n)$, the $\underline{i^{th}\text{ difference quotient}}$ is

$$D_i^h f(x) = \frac{f(x + he_i) - f(x)}{h} \quad (i = 1,\ldots,n).$$

6.10. LEMMA $\forall\ \varphi \in C_c^\infty(R^n)$,

$$\int_{R^n} \frac{\varphi(x + he_i) - \varphi(x)}{h} f(x)\ dL^n$$

$$= - \int_{R^n} \varphi(x) \frac{f(x - he_i) - f(x)}{-h}\ dL^n.$$

Consequently

$$\int_{R^n} (D_i^h \varphi) f \, dL^n = - \int_{R^n} \varphi (D_i^{-h} f) \, dL^n \quad (i = 1,\ldots,n).$$

6.11. THEOREM Suppose that f is a bounded Lipschitz continuous function, say

$$|f(x) - f(y)| \leq L||x - y|| \quad (x,y \in R^n).$$

Then

$$f \in W^{1,\infty}(R^n).$$

Since by hypothesis $f \in L^\infty(R^n)$, the problem is to show that its distributional derivatives

$$\frac{\partial f}{\partial x_i} \quad (i = 1,\ldots,n)$$

are (essentially) bounded functions as well.

To begin with, $\forall \; x \in R^n$,

$$\left|\frac{f(x - he_i) - f(x)}{-h}\right| \leq L \quad (i = 1,\ldots,n)$$

$$\Rightarrow$$

$$||D_i^{-h} f||_{L^\infty(R^n)} \leq L \quad (i = 1,\ldots,n),$$

so if Ω is open and bounded,

$$||D_i^{-h} f||_{L^2(\Omega)} \leq ||D_i^{-h} f||_{L^\infty(R^n)} (L^n(\Omega))^{1/2} \leq L(L^n(\Omega))^{1/2}.$$

Let $h_k = 1/k$ $(k = 1,2,\ldots)$ — then

$$\{D_i^{-h_k} f\}$$

is a bounded sequence in $L^2(\Omega)$, thus there is a subsequence

$$\{D_i^{-h_{k_\ell}} f\}$$

that converges weakly in $L^2(\Omega)$ as $\ell \to \infty$ to $g_i \in L^2(\Omega)$. To simplify, put $h_j = h_{k_\ell}$, hence $h_j \to 0$ $(j \to \infty)$.

Accordingly, $\forall\ \varphi \in C_c^\infty(\Omega)$,

$$- < \varphi, \frac{\partial f}{\partial x_i} > = \int_\Omega \frac{\partial \varphi}{\partial x_i} \, f dL^n$$

$$= \int_\Omega \left(\lim_{j \to \infty} D_i^{h_j} \varphi \right) f dL^n$$

$$= \lim_{j \to \infty} \int_\Omega (D_i^{h_j} \varphi) \, f dL^n$$

$$= - \lim_{j \to \infty} \int_\Omega \varphi (D_i^{-h_j} f) \, dL^n$$

$$= - \int_\Omega \varphi g_i \, dL^n,$$

the last equality following from weak convergence. Therefore the weak partial derivative $\frac{\partial f}{\partial x_i}$ exists and is represented by g_i.

Because

$$f_j \equiv D_i^{-h_j} f \quad (i = 1,2,\ldots)$$

converges weakly in $L^2(\Omega)$ to $\frac{\partial f}{\partial x_i}$, there exists a subsequence $\{f_{j_n}\}$ such that the convex combinations

$$\sum_{n=1}^{N} a_n f_{j_n} \to \frac{\partial f}{\partial x_i} \text{ in } L^2(\Omega)$$

as $N \to \infty$. Here

$$\left|\left|\sum_{n=1}^{N} a_n f_{j_n}\right|\right|_{L^\infty(\Omega)} \leq \sum_{n=1}^{N} a_n \left|\left|D_i^{-h_j} f\right|\right|_{L^\infty(\Omega)}$$

$$\leq \left(\sum_{n=1}^{N} a_n\right) L = L.$$

Summary:

$$\left|\frac{\partial f}{\partial x_i}\right| \leq L \quad (i = 1,\ldots,n)$$

for almost every $x \in \Omega$, hence

$$\frac{\partial f}{\partial x_i} \in L^\infty(R^n).$$

And then

$$f \in W^{1,\infty}(R^n).$$

6.12. REMARK Let $f \in W^{1,\infty}(R^n)$ — then on the basis of "embedding theory", it can be shown that f has a bounded continuous representative $\bar{f}$. Moreover, $\bar{f}$ can be taken Lipschitz continuous.

[Working with the standard mollification $\bar{f}_\varepsilon$ and assuming that the support of $\bar{f}$ is compact, note that

$$||\nabla \bar{f}_\varepsilon||_{L^\infty(R^n)} \leq ||\nabla \bar{f}||_{L^\infty(R^n)} \quad (\varepsilon > 0)$$

=>

$$|\bar{f}_\varepsilon(x) - \bar{f}_\varepsilon(y)| = \left|\int_0^1 < \nabla \bar{f}_\varepsilon(tx + (1 - t)y, x - y > dt\right|$$

$$\leq ||\nabla \bar{f}_\varepsilon||_{L^\infty(R^n)} ||x - y||$$

$$\leq ||\nabla \bar{f}||_{L^\infty(R^n)} ||x - y||$$

$\Rightarrow$ $(\varepsilon \downarrow 0)$

$$|\bar{f}(x) - \bar{f}(y)| \leq ||\nabla\bar{f}||_{L^\infty(R^n)} ||x - y||$$

for all $x, y \in R^n$.]

SECTION 7: SOBOLOV SPACES

§7.1. FORMALITIES

Let Ω be a nonempty open subset of R^n.

7.1.1. DEFINITION Let $1 \leq p < +\infty$ -- then the Sobolov space

$$W^{1,p}(\Omega)$$

consists of those $f \in L^1_{\ell oc}(\Omega)$ such that f belongs to $L^p(\Omega)$ and such that the distributional derivatives $\frac{\partial f}{\partial x_i}$ are weak partial derivatives and also belong to $L^p(\Omega)$ $(i = 1,\ldots,n)$.

[Note: There is a local version of this definition, namely call

$$W^{1,p}_{\ell oc}(\Omega)$$

the set comprised of all $f \in L^1_{\ell oc}(\Omega)$ with the property that the restriction $f|\Omega' \in W^{1,p}(\Omega')$ for every nonempty open set $\Omega' \subset \Omega$ whose closure is a compact subset of Ω.]

7.1.2. N.B. Spelled out, $W^{1,p}(\Omega)$ consists of those $f \in L^p(\Omega)$ for which there exist functions $\frac{\partial f}{\partial x_1},\ldots,\frac{\partial f}{\partial x_n}$ in $L^p(\Omega)$ such that $\forall\ \varphi \in C_c^\infty(\Omega)$,

$$\int_\Omega \varphi \frac{\partial f}{\partial x_i} dL^n = - \int_\Omega \frac{\partial \varphi}{\partial x_i} f dL^n \quad (i = 1,\ldots,n).$$

[Note: Another point is this: $W^{1,p}(\Omega)$ is closed under taking absolute values, i.e.,

$$f \in W^{1,p}(\Omega) \Rightarrow |f| \in W^{1,p}(\Omega).]$$

Depending on the parameters, it can happen that there exists a function in $W^{1,p}(\Omega)$ which is nowhere continuous.

7.1.3. EXAMPLE Take $\Omega = B(0,1)^{\circ}$, let $\{q_k\}$ be a countable dense subset of Ω, and consider

$$f(x) = \sum_{k=1}^{\infty} \frac{1}{2^k} ||x - q_k||^{-\alpha} \quad (\alpha > 0) \; (x \in \Omega \backslash \bigcup_{k=1}^{\infty} \{q_k\}).$$

Then

$$f \in W^{1,p}(\Omega) \text{ if } \alpha < \frac{n-p}{p}$$

but f is unbounded in every nonempty open subset of Ω.

7.1.4. N.B. It will be seen later on that each function in $W^{1,p}(\Omega)$ $(p > n)$ coincides with a continuous function almost everywhere.

7.1.5. LEMMA Let $f \in W^{1,p}(\Omega)$ -- then there is a partition

$$\Omega = (\bigcup_{k=1}^{\infty} E_k) \cup Z,$$

where the E_k are Lebesgue measurable sets such that $f|E_k$ is Lipschitz and Z has Lebesgue measure 0.

7.1.6. THEOREM Let $f \in W^{1,p}(\Omega)$ -- then f is approximately differentiable almost everywhere.

[Extend $f|E_k$ to all of R^n and use Rademacher.]

7.1.7. THEOREM The prescription

$$||f||_{W^{1,p}} \equiv ||f||_{L^p} + \sum_{i=1}^{n} ||\frac{\partial f}{\partial x_i}||_{L^p}$$

endows $W^{1,p}(\Omega)$ with the structure of a Banach space.

[Note: An equivalent norm is the prescription

$$f \to ||f||_{L^p} + ||\nabla f||_{L^p},$$

where ∇f is the weak gradient attached to f.]

7.1.8. LEMMA $W^{1,p}(\Omega)$ is separable.

7.1.9. THEOREM Let $f \in W^{1,p}(\Omega)$ -- then there exists a sequence $\{f_k\} \subset W^{1,p}(\Omega) \cap C^\infty(\Omega)$ such that

$$f_k \to f \text{ in } W^{1,p}(\Omega).$$

7.1.10. REMARK It can be shown that $C_c^\infty(R^n)$ is dense in $W^{1,p}(R^n)$.

7.1.11. PRODUCT RULE Let $f,g \in W^{1,p}(\Omega) \cap L^\infty(\Omega)$ -- then $fg \in W^{1,p}(\Omega) \cap L^\infty(\Omega)$ and

$$\frac{\partial(fg)}{\partial x_i} = \frac{\partial f}{\partial x_i} g + f \frac{\partial g}{\partial x_i} \quad (i = 1,\ldots,n)$$

L^n almost everywhere in Ω.

7.1.12. CHAIN RULE Let $f \in W^{1,p}(\Omega)$ and let $g \in C^1(R)$ subject to $g' \in L^\infty(R)$, $g(0) = 0$ -- then $g \circ f \in W^{1,p}(\Omega)$ and

$$\frac{\partial(g \circ f)}{\partial x_i} = (g' \circ f) \frac{\partial f}{\partial x_i} \quad (i = 1,\ldots,n)$$

L^n almost everywhere in Ω.

[Note: The assumption that $g(0) = 0$ is not needed if $L^n(\Omega) < +\infty$.]

To formulate the next result, let Ω' be another nonempty open subset of R^n.

7.1.13. CHANGE OF VARIABLE Suppose that $\Psi:\Omega' \to \Omega$ is invertible, where Ψ and Ψ' are Lipschitz continuous functions, and let $f \in W^{1,p}(\Omega)$ -- then $f \circ \Psi \in W^{1,p}(\Omega')$ and

$$\frac{\partial(f \circ \Psi)}{\partial x_i'}(x') = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\Psi(x')) \frac{\partial \Psi_k}{\partial x_i'}(x') \qquad (i = 1,\ldots,n)$$

L^n almost everywhere in Ω'.

7.1.14. TERMINOLOGY Given normed spaces $(X,||\cdot||_X)$ and $(Y,||\cdot||_Y)$, one says that X is <u>embedded</u> in Y, denoted $X \hookrightarrow Y$, if X is a subspace of Y and there exists a constant $C > 0$ such that for all $x \in X$,

$$||x||_Y \leq C||x||_X.$$

7.1.15. EXAMPLE Suppose that $L^n(\Omega) < +\infty$ and $1 \leq p < q < +\infty$ — then $\forall\, f \in L^q(\Omega)$,

$$||f||_{L^p(\Omega)} \leq (L^n(\Omega))^{(1/p) - (1/q)} ||f||_{L^q(\Omega)}$$

$\Rightarrow$

$$L^q(\Omega) \hookrightarrow L^p(\Omega).$$

[Note: Here $X = L^q(\Omega)$, $Y = L^p(\Omega)$, and

$$C = (L^n(\Omega))^{(1/p) - (1/q)}.]$$

Looking ahead:

$$W^{1,p}(R^n) \hookrightarrow L^{p^*}(R^n) \text{ if } p < n \left(\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}\right)$$

$$W^{1,p}(R^n) \hookrightarrow L^{q}(R^n) \text{ if } p = n \ (q \in [p, +\infty[)$$

$$W^{1,p}(R^n) \hookrightarrow L^{\infty}(R^n) \text{ if } p > n.$$

§7.2. EMBEDDINGS: GNS

7.2.1. DEFINITION Let $1 \leq p < n$ -- then the <u>conjugate exponent</u> of p is

$$p^* = \frac{np}{n-p} = p + \frac{p^2}{n-p}.$$

[Note: $p^* > p$ and

$$\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}.]$$

7.2.2. THEOREM Let $1 \leq p < n$ -- then there exists a constant $C(n,p) > 0$ such that for all $f \in W^{1,p}(R^n)$,

$$\left(\int_{R^n} |f|^{p^*} dL^n\right)^{1/p^*} \leq C(n,p) \left(\int_{R^n} ||\nabla f||^p dL^n\right)^{1/p} < +\infty.$$

7.2.3. SCHOLIUM When $1 \leq p < n$,

$$W^{1,p}(R^n) \hookrightarrow L^{p^*}(R^n).$$

7.2.4. RAPPEL If $1 \leq p_1, \ldots, p_k < +\infty$ with

$$\frac{1}{p_1} + \cdots + \frac{1}{p_k} = 1$$

and if $f_j \in L^{p_j}(R^n)$ $(j = 1, \ldots, k)$, then

$$\int_{R^n} |f_1 \cdots f_k| \, dL^n \leq \prod_{j=1}^{k} ||f_j||_{L^{p_j}}.$$

The proof of the theorem can be divided into three parts.

<u>Step 1:</u> $p = 1$, $f \in C_c^1(R^n)$.

[For each $i \in \{1,\dots,n\}$ and each point $x = (x_1,\dots,x_i,\dots,x_n) \in R^n$, write

$$f(x_1,\dots,x_i,\dots,x_n) = \int_{-\infty}^{x_i} \frac{\partial f}{\partial x_i}(x_1,\dots,t_i,\dots,x_n)\, dt_i,$$

hence

$$|f(x)| \leq \int_R \left|\frac{\partial f}{\partial x_i}(x_1,\dots,t_i,\dots,x_n)\right| dt_i \qquad (1 \leqq i \leq n)$$

$\Rightarrow$

$$|f(x)|^n \leq \prod_{i=1}^{n} \int_R \left|\frac{\partial f}{\partial x_i}(x_1,\dots,t_i,\dots,x_n)\right| dt_i$$

$\Rightarrow$

$$|f(x)|^{n/(n-1)} \leq \prod_{i=1}^{n} \left(\int_R \left|\frac{\partial f}{\partial x_i}(x_1,\dots,t_i,\dots,x_n)\right| dt_i\right)^{1/(n-1)}$$

$\Rightarrow$

$$\int_R |f|^{n/(n-1)} dx_1 \leq \left(\int_R \left|\frac{\partial f}{\partial x_1}\right| dt_1\right)^{1/(n-1)} \int_R \prod_{i=2}^{n} \left(\int_R \left|\frac{\partial f}{\partial x_i}\right| dt_i\right)^{1/(n-1)} dx_1$$

$$\leq \left(\int_R \left|\frac{\partial f}{\partial x_1}\right| dt_1\right)^{1/(n-1)} \prod_{i=2}^{n} \left(\int_R \int_R \left|\frac{\partial f}{\partial x_i}\right| dt_i dx_1\right)^{1/(n-1)}$$

$\Rightarrow \dots$

$$\int_R \cdots \int_R |f|^{n/(n-1)}\, dx_1 \dots dx_n$$

$$\leq \prod_{i=1}^{n} \left(\int_R \cdots \int_R \left|\frac{\partial f}{\partial x_i}\right| dx_1 \dots dx_n\right)^{1/(n-1)}$$

$\Rightarrow$

$$\int_{R^n} |f|^{n/(n-1)}\, dL^n$$

$$\leq \prod_{i=1}^{n} \left(\int_{R^n} \left|\frac{\partial f}{\partial x_i}\right| \, dL^n\right)^{1/(n-1)}$$

$$\leq \left(\int_{R^n} ||\nabla f|| \, dL^n\right)^{n/(n-1)}$$

=>

$$\left(\int_{R^n} |f|^{n/(n-1)} \, dL^n\right)^{(n-1)/n} \leq \int_{R^n} ||\nabla f|| \, dL^n.]$$

[Note:

$$1^* = \frac{n}{n-1} \Rightarrow 1/p^* = \frac{n-1}{n}$$

and

$$C(n,1) = 1.]$$

<u>Step 2:</u> $1 < p < n$, $f \in C_c^1(R^n)$.

[Put

$$\gamma = \frac{p(n-1)}{n-p} .$$

Then $\gamma > 1$ and

$$\frac{\gamma n}{n-1} = p^* = \frac{np}{n-p} = \frac{(\gamma - 1)p}{p-1} .$$

Now apply Step 1 to $|f^\gamma| = |f|^\gamma$:

$$\left(\int_{R^n} |f|^{p^*} \, dL^n\right)^{(n-1)/n}$$

$$= \left(\int_{R^n} |f^\gamma|^{n/(n-1)} \, dL^n\right)^{(n-1)/n}$$

$$\leq \int_{R^n} ||\nabla |f^\gamma| \,|| \, dL^n$$

$$= \gamma \int_{R^n} |f|^{\gamma-1} ||\nabla f|| \, dL^n$$

$$\leq \gamma\left(\int_{R^n} |f|^{(\gamma-1)p/(p-1)}\, dL^n\right)^{(p-1)/p} \left(\int_{R^n} ||\nabla f||^p\, dL^n\right)^{1/p}$$

$$= \gamma\left(\int_{R^n} |f|^{p^*}\, dL^n\right)^{(p-1)/p} \left(\int_{R^n} ||\nabla f||^p\, dL^n\right)^{1/p}.$$

And

$$\frac{n-1}{n} - \frac{p-1}{p} = \frac{p(n-1) - n(p-1)}{p}$$

$$= \frac{n-p}{np} = \frac{1}{p^*}$$

=>

$$\left(\int_{R^n} |f|^{p^*}\, dL^n\right)^{1/p^*} \leq \gamma\left(\int_{R^n} ||\nabla f||^p\, dL^n\right)^{1/p},$$

where $C(n,p) = \gamma$.]

<u>Step 3:</u> $1 \leq p < n$, $f \in W^{1,p}(R^n)$.

[Given an $f \in W^{1,p}(R^n)$, there exists a sequence $\{f_k\} \subset C_c^1(R^n)$ such that

$$||f_k - f||_{W^{1,p}} \to 0 \quad (k \to \infty).$$

So

$$||f_k - f||_{L^p} \to 0 \quad (k \to \infty)$$

and, upon passing to a subsequence if necessary, it can be assumed that $f_k \to f$ almost everywhere in R^n. This said, we then claim that $\{f_k\}$ is a Cauchy sequence in $L^{p^*}(R^n)$. For $f_k - f_\ell \in C_c^1(R^n)$, thus it follows that

$$||f_k - f_\ell||_{L^{p^*}} \leq C(n,p)||\nabla f_k - \nabla f_\ell||_{L^p}$$

$$\leq C(n,p)(||\nabla f_k - \nabla f||_{L^p} + ||\nabla f - \nabla f_\ell||_{L^p})$$

$$\to 0 \ (k,\ell \to \infty).$$

Consequently there exists a $g \in L^{p^*}(R^n)$ such that $f_k \to g$ in $L^{p^*}(R^n)$. Therefore $f = g$ almost everywhere in R^n, which implies that $f_k \to f$ in $L^{p^*}(R^n)$. Finally

$$\begin{aligned}||f||_{L^{p^*}} &\leq ||f - f_k||_{L^{p^*}} + ||f_k||_{L^{p^*}} \\ &\leq ||f - f_k||_{L^{p^*}} + C(n,p)||\nabla f_k||_{L^p} \\ &\leq ||f - f_k||_{L^{p^*}} + C(n,p)(||\nabla f_k - \nabla f||_{L^p} + ||\nabla f||_{L^p}) \\ &\to 0 + C(n,p)(0 + ||\nabla f||_{L^p}).\end{aligned}$$

I.e.:

$$||f||_{L^{p^*}} \leq C(n,p)(||\nabla f||_{L^p}).]$$

7.2.5. APPLICATION If $f \in W^{1,p}(R^n)$ with $1 \leq p < n$ and if $\nabla f = 0$ almost everywhere in R^n, then $f = 0$ almost everywhere in R^n.

7.2.6. RAPPEL If $1 \leq p < q < r < +\infty$, then

$$L^p \cap L^r \subset L^q$$

and

$$||f||_q \leq (||f||_p)^{\lambda}(||f||_r)^{1-\lambda},$$

where

$$\frac{1}{q} = \frac{\lambda}{p} + \frac{1-\lambda}{r} \quad (0 < \lambda < 1).$$

7.2.7. RAPPEL If $a \geq 0$, $b \geq 0$ and if $0 \leq \lambda \leq 1$, then

$$a^{\lambda}b^{1-\lambda} \leq \lambda a + (1-\lambda)b.$$

Therefore

$$||f||_q \leq \lambda||f||_p + (1-\lambda)||f||_r$$

$$< \quad ||f||_p + ||f||_r.$$

Specialize now and take $r = p^*$ (recall that $p < p^*$) and let $p < q < p^*$ -- then it follows that

$$W^{1,p}(R^n) \hookrightarrow L^q(R^n).$$

§7.3. EMBEDDINGS: BMO

Having dealt with the case when $1 \leq p < n$, the next item on the agenda is the case when $p = n$, which necessitates some preparation.

7.3.1. DEFINITION The set

$$Q = [a_1,b_1] \times \ldots \times [a_n,b_n],\ b_1 - a_1 = \cdots = b_n - a_n$$

is a <u>cube</u> in R^n (if $n = 1$, a cube is a bounded closed interval in R, if $n = 2$, a cube is a square in R^2 etc.). The <u>side length</u> $\ell(Q)$ of Q is the common value

$$b_i - a_i \ (i = 1,\ldots,n).$$

7.3.2. NOTATION

$$Q(x,\ell) = \{y \in R^n: |y_i - x_i| \leq \frac{\ell}{2}\ (i = 1,\ldots,n)\}$$

is a cube with center x and side length ℓ. Here

$$L^n(Q(x,\ell)) = \ell^n \text{ and } \operatorname{diam} Q(x,\ell) = \sqrt{n}\ \ell.$$

7.3.3. DEFINITION Given $f \in L^1_{\ell oc}(R^n)$, its <u>integral average</u> over the cube $Q(x,\ell)$ is the entity

$$f_{Q(x,\ell)} = \frac{1}{\ell^n} \int_{Q(x,\ell)} f\, dL^n.$$

7.3.4. LEMMA Let $1 \leq p < +\infty$ -- then there exists a constant $C(n,p) > 0$ such that for all $f \in W^{1,p}(R^n)$,

$$\left(\int_{Q(x,\ell)} |f - f_{Q(x,\ell)}|^p\, dL^n\right)^{1/p} \leq C(n,p)\ell \left(\int_{Q(x,\ell)} ||\nabla f||^p\, dL^n\right)^{1/p} < +\infty .$$

PROOF Take $f \in C_c^1(R^n)$, put $Q = Q(x,\ell)$, let $z,y \in Q$, and write

$$|f(z) - f(y)|$$

$$\leq |f(z) - f(z_1,\ldots,z_{n-1},y_n)|$$

$$+ \cdots + |f(z_1,y_2,\ldots,y_n) - f(y)|$$

$$\leq \sum_{i=1}^{n} \int_{a_i}^{b_i} ||\nabla f(z_1,\ldots,z_{i-1},t,y_{i+1},\ldots,y_n)|| \, dt$$

=>

$$|f(z) - f(y)|^p$$

$$\leq \left(\sum_{i=1}^{n} \int_{a_i}^{b_i} ||\nabla f(z_1,\ldots,z_{i-1},t,y_{i+1},\ldots,y_n)|| \, dt\right)^p$$

$$\leq \left(\sum_{i=1}^{n} \left(\int_{a_i}^{b_i} ||\nabla f(z_1,\ldots,z_{i-1},t,y_{i+1},\ldots,y_n)||^p \, dt\right)^{1/p} (b_i - a_i)^{1-1/p}\right)^p$$

$$\leq n^p \ell^{p-1} \sum_{i=1}^{n} \int_{a_i}^{b_i} ||\nabla f(z_1,\ldots,z_{i-1},t,y_{i+1},\ldots,y_n)||^p \, dt$$

=>

$$\int_Q |f - f_Q|^p \, dL^n$$

$$= \int_Q |f(z) - f_Q|^p \, dz$$

$$= \int_Q \left|\frac{1}{\ell^n} \int_Q (f(z) - f(y)) \, dy\right|^p dz$$

$$\leq \int_Q \left(\frac{1}{\ell^n} \int_Q |f(z) - f(y)| \, dy\right)^p dz$$

$$\leq \int_Q \frac{1}{\ell^n} \int_Q |f(z) - f(y)|^p \, dzdy$$

$$\leq \frac{n^p \ell^{p-1}}{\ell^n} \sum_{i=1}^{n} \int_Q \int_Q \int_{a_i}^{b_i} ||\nabla f(z_1,\ldots,z_{i-1},t,y_{i+1},\ldots,y_n)||^p \, dtdydz$$

$$\leq \frac{n^p \ell^{p-1}}{\ell^n} \sum_{i=1}^{n} (b_i - a_i) \int_Q \int_Q ||\nabla f(z)||^p \, dzdy$$

$$\leq \frac{n^p \ell^{p-1}}{\ell^n} \cdot n\ell^{n+1} \int_Q ||\nabla f||^p \, dL^n$$

$$= n^{p+1} \ell^p \int_Q ||\nabla f||^p \, dL^n$$

=>

$$(\int_Q |f - f_Q|^p \, dL^n)^{1/p} \leq C(n,p)\ell(\int_Q ||\nabla f||^p \, dL^n)^{1/n}$$

if

$$C(n,p) = (n^{p+1})^{1/p}.$$

7.3.5. SCHOLIUM If $f \in W^{1,n}(R^n)$, then for every cube Q,

$$\frac{1}{L^n(Q)} \int_Q |f - f_Q| \, dL^n$$

$$\leq (\frac{1}{L^n(Q)} \int_Q |f - f_Q|^n \, dL^n)^{1/n}$$

$$\leq C(n)\ell(Q) (\frac{1}{L^n(Q)} \int_Q ||\nabla f||^n \, dL^n)^{1/n}$$

$$= C(n)\ell(Q) (\frac{1}{\ell(Q)^n} \int_Q ||\nabla f||^n \, dL^n)^{1/n}$$

$$= C(n) (\int_Q ||\nabla f||^n \, dL^n)^{1/n}$$

$$\leq C(n) ||\nabla f||_{L^n(R^n)} < +\infty.$$

7.3.6. DEFINITION A function $f \in L^1_{\ell oc}(R^n)$ is of <u>bounded mean oscillation</u> provided

$$||f||_{BMO} \equiv \sup_Q \frac{1}{L^n(Q)} \int_Q |f - f_Q| \, dL^n < +\infty,$$

where the supremum is taken over all cubes Q in R^n.

7.3.7. NOTATION $BMO(R^n)$ is the set of functions of bounded mean oscillation.

7.3.8. <u>N.B.</u> $||\cdot||_{BMO}$ is a seminorm, not a norm (constant functions have vanishing bounded mean oscillation).

7.3.9. LEMMA $BMO(R^n)$ is a vector space over R.

[If $f,g \in BMO(R^n)$, then

$$||f + g||_{BMO} \leq ||f||_{BMO} + ||g||_{BMO}$$

and $||\cdot||_{BMO}$ is scale invariant, i.e., $\forall\, r \in R$,

$$||f(r\cdot)||_{BMO} = ||f||_{BMO}.]$$

7.3.10. THEOREM

$$BMO(R^n)/R$$

is a Banach space.

7.3.11. LEMMA

$$f \in BMO(R^n) \Rightarrow |f| \in BMO(R^n).$$

7.3.12. THEOREM $L^\infty(R^n)$ is contained in $BMO(R^n)$.

[If $f \in L^\infty(R^n)$, then

$$||f||_{BMO} \leq 2\ ||f||_{L^\infty}.]$$

[Note: Therefore

$$L^\infty(R^n) \hookrightarrow BMO(R^n).]$$

7.3.13. N.B. The containment is strict.

[The unbounded function

$$\log ||x||$$

belongs to $BMO(R^n)$.]

7.3.14. EXAMPLE Take $n = 1$ -- then the function

$$f(x) = \begin{cases} \log x & (x > 0) \\ 0 & (x \leq 0) \end{cases}$$

is not of bounded mean oscillation.

Let $f \in W^{1,n}(R^n)$ -- then, as has been seen above, for every cube Q,

$$\frac{1}{L^n(Q)} \int_Q |f - f_Q|\ dL^n \leq C(n)||\nabla f||_{L^n(R^n)},$$

so upon taking the supremum over Q, it follows that $f \in BMO(R^n)$, where

$$||f||_{BMO} \leq C(n)||\nabla f||_{L^n(R^n)}.$$

7.3.15. SCHOLIUM

$$W^{1,n}(R^n) \hookrightarrow BMO(R^n).$$

7.3.16. APPLICATION If $f \in W^{1,n}(R^n)$ and if $\nabla f = 0$ almost everywhere in R^n, then f = some constant almost everywhere in R^n.

§7.4. EMBEDDINGS: MOR

It remains to consider the situation when $p > n$.

7.4.1. RAPPEL Let E be a nonempty subset of R^n -- then a function $f:E \to R$ is <u>Hölder continuous</u> with exponent $0 < \alpha \leq 1$ if there is a constant $C > 0$ such that

$$|f(x) - f(y)| \leq C||x - y||^{\alpha}$$

for all $x,y \in E$.

[Note: Of course if $\alpha = 1$, then it is a question of Lipschitz continuous.]

A Hölder continuous function is continuous but it can be nowhere differentiable.

7.4.2. NOTATION $C^{0,\alpha}(E)$ is the set of all bounded functions that are Hölder continuous with exponent α and norm

$$||f||_{C^{0,\alpha}} = \sup_{x \in E} |f(x)| + \sup_{\substack{x,y \in E \\ x \neq y}} \frac{|f(x) - f(y)|}{||x - y||^{\alpha}} .$$

[Note: When so equipped, $C^{0,\alpha}(E)$ is a Banach space.]

7.4.3. LEMMA Let $p > n$ -- then there is a constant $C(n,p) > 0$ such that for all $f \in W^{1,p}(R^n) \cap C^{\infty}(R^n)$ and for all $z,y \in R^n$,

$$|f(z) - f(y)| \leq C(n,p) \ ||z - y||^{1-n/p}||\nabla f||_{L^p(R^n)} .$$

PROOF To begin with,

$$f(z) - f(y)$$

$$= \int_0^1 \frac{\partial}{\partial t} (f(tz + (1-t)y))dt$$

$$= \int_0^1 \langle \nabla f(tz + (1-t)y),\ z - y\rangle\, dt.$$

Assume now that $z, y \in Q(x,\ell)$ -- then

$$|f_{Q(x,\ell)} - f(y)|$$

$$= \left| \frac{1}{\ell^n} \int_{Q(x,\ell)} (f(z) - f(y))\, dz \right|$$

$$= \left| \frac{1}{\ell^n} \int_{Q(x,\ell)} \int_0^1 \langle \nabla f(tz + (1-t)y),\ z - y\rangle\, dt dz \right|$$

$$\leq \sum_{i=1}^n \frac{1}{\ell^n} \int_{Q(x,\ell)} \int_0^1 \left| \frac{\partial f}{\partial x_i} (tz + (1-t)y) \right| |z_i - y_i|\, dt dz$$

$$\leq \sum_{i=1}^n \frac{1}{\ell^{n-1}} \int_0^1 \int_{Q(x,\ell)} \left| \frac{\partial f}{\partial x_i} (tz + (1-t)y) \right| dz dt$$

$$= \sum_{i=1}^n \frac{1}{\ell^{n-1}} \int_0^1 \frac{1}{t^n} \int_{Q(tx + (1-t)y, t\ell)} \left| \frac{\partial f}{\partial x_i} (w) \right| dw dt$$

$$\leq \sum_{i=1}^n \frac{1}{\ell^{n-1}} \int_0^1 \frac{1}{t^n} \left(\int_{Q(tx + (1-t)y, t\ell)} \left| \frac{\partial f}{\partial x_i} (w) \right|^p dw\right)^{1/p}$$

$$\times\ (L^n(Q(tx + (1-t)y, t\ell)))^{1-1/p}\, dt$$

$$\leq n||\nabla f||_{L^p(Q(x,\ell))} \frac{\ell^{n(1-1/p)}}{\ell^{n-1}} \int_0^1 \frac{t^{n(1-1/p)}}{t^n}\, dt$$

$$= \frac{np}{p - n} \ell^{1-n/p} ||\nabla f||_{L^p(Q(x,\ell))}.$$

Since the same estimate obtains if the roles of y and z are interchanged, write

$$|f(z) - f(y)| \leq |f(z) - t_{Q(x,\ell)}| + |f_{Q(x,\ell)} - f(y)|$$

$$\leq 2\,\frac{np}{p-n}\,\ell^{1-n/p}||\nabla f||_{L^p(Q(x,\ell))},$$

where $z,y \in Q(x,\ell)$. Proceed finally to when $z,y \in R^n$ are arbitrary — then there exists a cube $Q(x,\ell)$ such that $z,y \in Q(x,\ell)$ and $\ell = ||z - y||$ (e.g., take $x = \frac{z+y}{2}$), hence

$$|f(z) - f(y)| \leq C(n,p)||z - y||^{1-n/p}||\nabla f||_{L^p(Q(x,\ell))}$$

$$\leq C(n,p)||z - y||^{1-n/p}||\nabla f||_{L^p(R^n)},$$

where

$$C(n,p) = 2\,\frac{np}{p-n}.$$

In the foregoing, replace z,y by x,y.

7.4.4. THEOREM Let $f \in W^{1,p}(R^n)$ $(p > n)$ and let $x,y \in \Lambda(f)$ — then

$$|f(x) - f(y)| \leq C(n,p)||x - y||^{1-n/p}||\nabla f||_{L^p(R^n)}.$$

[Utilize the standard mollification f_ε of f and apply it to Lebesgue points x,y of f.]

The restriction $f|\Lambda(f)$ can be extended uniquely to R^n as a Hölder continuous function $\bar{f}$ of exponent $1 - n/p$ in such a way that

$$|\bar{f}(x) - \bar{f}(y)| \leq C(n,p)||x - y||^{1-n/p}||\nabla\bar{f}||_{L^p}$$

for all $x,y \in R^n$.

[Bearing in mind that $\Lambda(f)$ is dense, given an $x \in R^n$, choose a sequence $\{f_k\} \subset \Lambda(f)$ such that $x_k \to x$ $(k \to \infty)$. From what has been said above, $\{f(x_k)\}$ is a Cauchy sequence, thus the prescription

$$\bar{f}(x) = \lim_{k \to \infty} f(x_k)$$

makes sense and has the desired property.]

7.4.5. THEOREM Let $f \in W^{1,p}(R^n)$ $(p > n)$ -- then in the equivalence class of f there is a unique function $\bar{f}$ which is Hölder continuous with exponent $1 - n/p$.

In particular: Every element of $W^{1,p}(R^n)$ $(p > n)$ coincides with a continuous function almost everywhere.

7.4.6. LEMMA If $f \in W^{1,p}(R^n)$ with $p > n$, then f is essentially bounded.

PROOF Let $y \in Q(x,1)$, take $f \in W^{1,p}(R^n) \cap C^\infty(R^n)$, and write

$$|f(z)| \leq |f(z) - f_{Q(x,1)}| + |f_{Q(x,1)}|$$

$$\leq \frac{np}{p-n} 1^{1-n/p} ||\nabla f||_{L^p(Q(x,1))} + \int_{Q(x,1)} |f(y)| \, dy$$

$$\leq \frac{np}{p-n} ||\nabla f||_{L^p(R^n)} + (\int_{Q(x,1)} |f(y)|^p \, dy)^{1/p}$$

$$\leq \frac{np}{p-n} ||\nabla f||_{L^p(R^n)} + ||f||_{L^p(R^n)}.$$

But one choice for the norm of f in $W^{1,p}(R^n)$ is

$$||f||_{W^{1,p}(R^n)} = ||f||_{L^p(R^n)} + ||\nabla f||_{L^p(R^n)}.$$

This said, there are then two possibilities.

- $\frac{np}{p-n} \geq 1$

=>

$$|f(z)| \le \frac{np}{p-n} ||\nabla f||_{L^p(R^n)} + ||f||_{L^p(R^n)}$$

$$\le \frac{np}{p-n} ||\nabla f||_{L^p(R^n)} + \frac{np}{p-n} ||f||_{L^p(R^n)}$$

$$= \frac{np}{p-n} (||f||_{W^{1,p}(R^n)}).$$

- $\frac{np}{p-n} < 1$

=>

$$|f(z)| \le ||\nabla f||_{L^p(R^n)} + ||f||_{L^p(R^n)}$$

$$= ||f||_{W^{1,p}(R^n)}.$$

Therefore the L^∞-norm of f is bounded by a constant

$$\bar{C}(n,p) = \begin{cases} \frac{np}{p-n} & (\text{if} \ge 1) \\ 1 & (\text{if} < 1) \end{cases}$$

depending on n and p times the $W^{1,p}$-norm of f.

7.4.7. SCHOLIUM When $p > n$,

$$W^{1,p}(R^n) \hookrightarrow L^\infty(R^n).$$

7.4.8. THEOREM

$$\bar{f} \in C^{0,1-n/p}(R^n) \quad (p > n).$$

PROOF

$$\|\bar{f}\|_{C^{0,1-n/p}}$$

$$= \sup_{x\in R^n} |\bar{f}(x)| + \sup_{\substack{x,y\in R^n \\ x \neq y}} \frac{|\bar{f}(x) - \bar{f}(y)|}{\|x - y\|^{1-n/p}}$$

$$\leq \bar{C}(n,p)\|f\|_{W^{1,p}} + C(n,p)\|f\|_{W^{1,p}}.$$

7.4.9. N.B. If $p > n$, then this state of affairs is symbolized by writing

$$W^{1,p}(R^n) \hookrightarrow C^{0,1-n/p}(R^n)$$

since

$$\|\bar{f}\|_{C^{0,1-n/p}} \leq C\|f\|_{W^{1,p}}.$$

7.4.10. THEOREM ($p > n$)

$$\bar{f}(x) \to 0 \text{ as } \|x\| \to +\infty.$$

PROOF Given $f \in W^{1,p}(R^n)$, choose a sequence $\{f_k\}$ in $C_c^\infty(R^n)$ that converges to f in $W^{1,p}(R^n)$ -- then

$$\|f - f_k\|_{L^\infty} \leq C\|f - f_k\|_{W^{1,p}} \to 0 \quad (k \to \infty).$$

Fix $\varepsilon > 0$ and choose $\bar{k}$ such that

$$\|f - f_k\|_{L^\infty} \leq \varepsilon \quad (k \geq \bar{k}).$$

Next choose $R_{\bar{k}} > 0$ such that $f_{\bar{k}}(x) = 0$ for all $x:\|x\| \geq R_{\bar{k}}$. So, for L^n almost every $x \in R^n$ with $\|x\| \geq R_{\bar{k}}$, it follows that

$$|\bar{f}(x)| = |\bar{f}(x) - f_{\bar{k}}(x)| \leq \|f - f_k\|_{L^\infty} \leq \varepsilon$$

and since $\bar{f}$ is continuous, this inequality holds for all $x: ||x|| \geq R_{\bar{k}}$.

7.4.11. LEMMA $(p > n)$ $\forall\, x,y \in R^n$,

$$|\bar{f}(x) - \bar{f}(y)| \leq C(n,p)\,||x - y||^{1-n/p}\Big(\int_{B(x,||y-x||)} ||\nabla\bar{f}||^p \, dL^n\Big)^{1/p}.$$

7.4.12. THEOREM $(p > n)$ $\bar{f}$ is differentiable almost everywhere.

PROOF It suffices to show that $\bar{f}$ is differentiable at every L^p-Lebesgue point x_0 of ∇f, where by definition,

$$\lim_{r \to 0} \frac{1}{\omega_n r^n} \int_{B(x,r)} ||\nabla f - \nabla f(x_0)||^p \, dL^n = 0.$$

To this end, note that

$$|\bar{f}(x) - \bar{f}(x_0) - \langle \nabla f(x_0),\, x - x_0 \rangle|$$

$$\leq C(n,p)\,||x - x_0||^{1-n/p}\Big(\int_{B(x,||x-x_0||)} ||\nabla f - \nabla f(x_0)||^p \, dL^n\Big)^{1/p}.$$

And

$$C(n,p)\,||x - x_0||^{1-n/p} = C(n,p)\,||x - x_0|| \Big(\frac{1}{||x-x_0||^n}\Big)^{1/p}$$

$$= C(n,p)\,(\omega_n)^{1/p}\,||x - x_0|| \Big(\frac{1}{\omega_n ||x-x_0||^n}\Big)^{1/p}$$

$\Rightarrow$

$$\frac{|\bar{f}(x) - \bar{f}(x_0) - \langle \nabla f(x_0),\, x - x_0 \rangle|}{||x - x_0||}$$

$$\leq C(n,p)\,(\omega_n)^{1/p}\Big(\frac{1}{\omega_n ||x-x_0||^n} \int_{B(x,||x-x_0||)} ||\nabla f - \nabla f(x_0)||^p \, dL^n\Big)^{1/p}$$

$$\longrightarrow 0$$

as $x \to x_0$.

7.4.13. N.B. The weak derivatives of f coincide with the ordinary partial derivatives of $\bar{f}$ almost everywhere in R^n.

SECTION 8: ACL

Working in R^n, let

$$I_k = [a_k, b_k] \quad (k = 1, \ldots, n)$$

be closed intervals and put

$$Q = [a_1, b_1] \times \ldots \times [a_n, b_n],$$

a rectangular box.

8.1. DEFINITION A function $f: Q \to R$ is said to be ACL (absolutely continuous on lines) if for each $k = 1, \ldots, n$, and almost every point

$$(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n)$$
$$\in I_1 \times \ldots \times I_{k-1} \times I_{k+1} \times \ldots \times I_n \in R^{n-1}$$

with respect to L^{n-1} measure, the function

$$x_k \to f(x_1, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_n) \quad (a_k \leq x_k \leq b_k)$$

is absolutely continuous.

8.2. N.B. In the literature, the foregoing situation is sometimes referred to as saying that f is absolutely continuous on almost every line segment in Q parallel to the coordinate axes.

Let Ω be a nonempty open subset of R^n.

8.3. DEFINITION A function $f: \Omega \to R$ is ACL if the restriction $f|Q$ is ACL for every $Q \subset \Omega$.

8.4. NOTATION $ACL(\Omega)$ is the set of ACL functions in Ω.

8.5. EXAMPLE A quasiconformal map belongs to $ACL(\Omega)$.

8.6. NOTATION Let $1 \leq p < +\infty$ -- then $L^{1,p}(\Omega)$ consists of those $f \in L^1_{\ell oc}(\Omega)$ such that the distributional derivatives $\frac{\partial f}{\partial x_i}$ are weak partial derivatives and also belong to $L^p(\Omega)$ $(i = 1,\ldots,n)$.

[Note: Evidently

$$L^{1,p}(\Omega) \subset L^p_{\ell oc}(\Omega) \subset L^1_{\ell oc}(\Omega).]$$

8.7. <u>N.B.</u> Obviously

$$W^{1,p}(\Omega) \subset L^{1,p}(\Omega).$$

8.8. THEOREM Let $1 \leq p < +\infty$ -- then a function $f \in L^{1,p}(\Omega)$ admits a representative $\bar{f}:\Omega \to R$ in $ACL(\Omega)$.

[Note: The ordinary partial derivatives of $\bar{f}$ exist almost everywhere.]

The proof in general is notationally involved so to simplify the bookkeeping, take $n = 2$, assume that f is continuous, suppose that

$$Q = [0,1] \times [0,1] \subset \Omega =]-\varepsilon, 1 + \varepsilon[\times]-\varepsilon, 1 + \varepsilon[,$$

let $(x_1,x_2) = (x,y)$, thus the distributional derivatives are $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$, and the claim is that

$$\left[\begin{array}{l} x \to f(x,y) \text{ is absolutely continuous for almost every } y \in [0,1] \\ y \to f(x,y) \text{ is absolutely continuous for almost every } x \in [0,1]. \end{array}\right.$$

The discussion in either case is conceptually the same, hence it will suffice to deal with the second of these.

For use below:

8.9. CRITERION If $f \in L^1_{\ell oc}]a,b[$ and if $\forall\ \varphi \in C_c^\infty]a,b[$,

$$\int_{]a,b[} \varphi f \, dL^1 = 0,$$

then $f = 0$ almost everywhere in $]a,b[$.

8.10. APPLICATION If $f \in L^1]a,b[$ and if $\forall\ \varphi \in C_c^\infty]0,1[$,

$$\int_0^1 f\varphi' \, dL^1 = 0,$$

then there exists a constant C such that $f = C$ almost everywhere in $]0,1[$.

[Let Φ,ψ be functions in $C_c^\infty]0,1[$ with $\int_0^1\psi = 1$. Put

$$\Psi(x) = \int_0^x\Phi - (\int_0^1\Phi)\int_0^x\psi.$$

Then $\Psi(0) = 0$ and

$$\Psi(1) = \int_0^1\Phi - (\int_0^1\Phi)\int_0^1\psi$$

$$= 0.$$

Therefore $\Psi \in C_c^\infty]0,1[$ and

$$\Psi' = \Phi - (\int_0^1\Phi)\psi$$

$\Rightarrow$

$$0 = \int_0^1 f\Psi' \quad \text{(by assumption)}$$

$$= \int_0^1 (f - \int_0^1 f\psi)\Phi$$

$\Rightarrow$

$$f - \int_0^1 f\psi = 0 \text{ almost everywhere in }]0,1[$$

$\Rightarrow$

$$f = \int_0^1 f\psi \text{ almost everywhere in }]0,1[.$$

By hypothesis,

$$f \in L^{1,p}(\Omega) \Rightarrow \frac{\partial f}{\partial y} |Q \in L^p(Q),$$

from which it follows that

$$\int_0^1 |\frac{\partial f}{\partial y} (x,y)| \; dy < + \infty$$

for almost all $x \in [0,1]$ (Fubini). Therefore the function

$$y \to g(x,y) = \int_0^y \frac{\partial f}{\partial t} (x,t) \; dt$$

is absolutely continuous on the segment $[0,1]$ for almost all $x \in [0,1]$ and its ordinary derivative $g' = \frac{\partial g}{\partial y}$ coincides with the distributional derivative $\frac{\partial f}{\partial g}$ for almost all $y \in]0,1[$. Consider now a test function ϕ of the form

$$\phi = \xi\eta \;\; (\xi \in C_c^\infty]0,1[, \; \eta \in C_c^\infty]0,1[),$$

i.e.,

$$\phi(x,y) = \xi(x)\eta(y).$$

Then

$$\int_Q f(x,y)\xi(x)\eta'(y) \; dxdy$$

$$= - \int_Q \xi(x)\eta(y) \frac{\partial f}{\partial y} \; dxdy$$

$$= - \int_Q \phi(x,y) \frac{\partial g}{\partial y} \; dxdy.$$

On the other hand,

$$\int_0^1 g(x,y)\eta'(y) \; dy = - \int_0^1 \eta(y) \frac{\partial g}{\partial y} (x,y) \; dy$$

$\Rightarrow$

$$\xi(x) \int_0^1 g(x,y)\eta'(y) \; dy = \xi(x) \;\; (- \int_0^1 \eta(y) \frac{\partial g}{\partial y} (x,y) \; dy)$$

=>

$$\int_Q g(x,y)\xi(x)\eta'(y)\ dxdy$$

$$= -\int_Q \xi(x)\eta(y) \frac{\partial g}{\partial y}\ dxdy$$

$$= -\int_Q \phi(x,y) \frac{\partial g}{\partial y}\ dxdy$$

=>

$$\int_Q f(x,y)\xi(x)\eta'(y)\ dxdy = \int_Q g(x,y)\xi(x)\eta'(y)\ dxdy$$

=>

$$\int_Q \xi(x)f(x,y)\eta'(y)\ dxdy = \int_Q \xi(x)g(x,y)\eta'(y)\ dxdy$$

=>

$$\int_0^1 [f(x,y) - g(x,y)]\eta'(y)\ dy = 0$$

for almost all $x \in [0,1]$. Denote by $E \subset [0,1]$ the set of x for which equality obtains -- then $\forall\ x \in E$,

$$y \to f(x,y)$$

is absolutely continuous. In fact, for any such x,

$$\int_0^1 [f(x,y) - g(x,y)]\eta'(y)\ dy = 0$$

=>

$$f(x,y) - g(x,y) = C_x (\in \mathbb{R})$$

for almost all $y \in [0,1]$

=>

$$f(x,y) = g(x,y) + C_x$$

$$= \int_0^y \frac{\partial f}{\partial t}(x,t)\ dt + C_x$$

for almost all $y \in [0,1]$. The right hand side is an absolutely continuous function of y and the left hand side is a continuous function of y. Since equality holds for a subset of $[0,1]$ of full measure and such a set is dense in $[0,1]$, the conclusion is that

$$f(x,y) = \int_0^y \frac{\partial f}{\partial t}(x,t)\, dt + C_x$$

for all $y \in [0,1]$.

Summary:

$y \to f(x,y)$ is absolutely continuous for almost every $x \in [0,1]$ (viz., $\forall\, x \in E$), thereby completing the proof.

The preceding result also admits an easy converse (where, as above, $1 \le p < +\infty$).

8.11. THEOREM If $f:\Omega \to R$ has an ACL representative $\bar{f}$ whose ordinary partial derivatives belong to $L^p(\Omega)$, then these derivatives coincide almost everywhere with the corresponding distributional derivatives of f, hence $f \in L^{1,p}(\Omega)$.

As for Sobolev spaces, there is a characterization.

8.12. THEOREM Let $1 \le p < +\infty$ -- then a function $f \in L^p(\Omega)$ belongs to $W^{1,p}(\Omega)$ iff it has a representative $\bar{f}$ that is ACL and whose ordinary partial derivatives belong to $L^p(\Omega)$.

8.13. CRITERION Suppose that $f:\Omega \to R$ is continuous and ACL -- then the ordinary partial derivatives of f exist almost everywhere in Ω and they are Borel functions.

PROOF Fix $i \in \{1,\ldots,n\}$, put

$$R_i^{n-1} = \{x \in R^n : x_i = 0\},$$

and let P_i be the orthogonal projection of R^n onto R_i^{n-1} (so $P_i(x) = x - x_i e_i$).
Let E_i be the set of all $x \in \Omega$ at which $\frac{\partial f}{\partial x_i}$ does not exist, the claim being that $L^n(E \cap Q) = 0$ for all $Q \subset \Omega$. Since f is continuous, E is Borel and by Fubini,

$$L^n(E_i \cap Q) = \int_{P_i\Omega} L^1(P_i^{-1}(x) \cap E_i \cap Q)\, dL^{n-1}(x).$$

If f is absolutely continuous on the segment $P_i^{-1}(x) \cap Q$, then $\frac{\partial f}{\partial x_i}$ exists almost everywhere on this segment, hence $L^1(P_i^{-1}(x) \cap E \cap Q) = 0$, implying thereby that $L^n(E \cap Q) = 0$, f being ACL.

8.14. REMARK Recall that without some assumption, the set of points in Ω where $\frac{\partial f}{\partial x_i}$ exists need not be Lebesgue measurable (let alone Borel).

8.15. CRITERION Let

$$R =]a_1,b_1[\times \ldots \times]a_n,b_n[\subset R^n$$

be an open rectangle. Fix $i \in \{1,\ldots,n\}$ and let $f:R \to R$ be a Lebesgue measurable function that is monotone on almost every line in R parallel to the x_i axis -- then the ordinary partial derivative $\frac{\partial f}{\partial x_i}$ exists almost everywhere in R (and is Lebesgue measurable).

8.16. REMARK The assumption "monotone on almost every line in R" cannot be replaced by "bounded variation on almost every line in R".

[Note: But if f is of bounded variation on almost every line in R parallel to the x_i axis, then there is an equivalent function $\bar{f}$ which does have an ordinary partial derivative $\frac{\partial \bar{f}}{\partial x_i}$ almost everywhere in R (and is Lebesgue measurable).]

SECTION 9: BV SPACES

§9.1. PROPERTIES

Let Ω be a nonempty open subset of R^n.

9.1.1. DEFINITION Let $f \in L^1(\Omega)$ -- then f is said to be a <u>function of bounded variation</u> if its partial derivatives in the sense of distributions are finite signed Radon measures in Ω of finite total variation.

9.1.2. NOTATION $BV(\Omega)$ is the set of functions of bounded variation in Ω.

[Note: There is a local version of this definition, namely call

$$BV_{\ell oc}(\Omega)$$

the set comprised of all $f \in L^1_{\ell oc}(\Omega)$ with the property that $f|\Omega' \in BV(\Omega')$ for every nonempty open set $\Omega' \subset \Omega$ whose closure is a compact subset of Ω.]

9.1.3. <u>N.B.</u> Let $f \in BV(\Omega)$ -- then there are finite signed Radon measures

$$D_i f \quad (i = 1,\ldots,n)$$

of finite total variation such that $\forall\ \varphi \in C_c^\infty(\Omega)$,

$$\int_\Omega \varphi \, dD_i f = - \int_\Omega \frac{\partial\varphi}{\partial x_i} f d\mathcal{L}^n \quad (i = 1,\ldots,n).$$

[Note:

$$Df \equiv (D_1 f,\ldots,D_n f)$$

is an R^n-valued vector measure and on general grounds, $||Df||$ is a positive finite measure (hence $||Df||(\Omega) < +\infty$).]

9.1.4. LEMMA Let $f \in BV(\Omega)$ — then $||Df||$ is absolutely continuous w.r.t. Lebesgue measure iff each of the $D_i f$ is, in which case the distributional partial derivatives can be represented by L^1 functions.

9.1.5. LEMMA Let $f \in BV(\Omega)$ — then $f \in W^{1,1}(\Omega)$ iff $||Df||$ is absolutely continuous w.r.t. Lebesgue measure, in which case

$$||Df||(\Omega) = \int_\Omega ||\nabla f|| \, dL^n.$$

9.1.6. <u>N.B.</u> The containment

$$W^{1,1}(\Omega) \subset BV(\Omega)$$

is strict and every $f \in C^\infty(\Omega) \cap BV(\Omega)$ belongs to $W^{1,1}(\Omega)$.

9.1.7. NOTATION Given $\Phi \in C_c^\infty(\Omega;R^n)$, put

$$\text{div } \Phi = \sum_{i=1}^{n} \frac{\partial \Phi_i}{\partial x_i},$$

the <u>divergence</u> of Φ.

9.1.8. DEFINITION Let $f \in L^1(\Omega)$ -- then the <u>variation</u> of f in Ω is the entity

$$V(f;\Omega) = \sup\{\int_\Omega f \text{ div } \Phi \, dL^n : \Phi \in C_c^\infty(\Omega;R^n), ||\Phi||_\infty \leq 1\}.$$

9.1.9. THEOREM Let $f \in L^1(\Omega)$ -- then

$$V(f;\Omega) < +\infty$$

iff $f \in BV(\Omega)$. And when this is so,

$$V(f;\Omega) = ||Df||(\Omega).$$

9.1.10. LSC PRINCIPLE Suppose that $\{f_k\}$ is a sequence in $BV(\Omega)$ which converges in $L^1(\Omega)$ to a function f -- then

$$||Df||(\Omega) \leq \liminf_{k \to \infty} ||Df_k||(\Omega).$$

PROOF Choose a $\Phi \in C_c^\infty(\Omega;R^n)$ with $||\Phi||_\infty \leq 1$, thus

$$\int_\Omega f \operatorname{div} \Phi \, dL^n = \lim_{k \to \infty} \int_\Omega f_k \operatorname{div} \Phi \, dL^n$$

$$\leq \liminf_{k \to \infty} V(f_k;\Omega)$$

$$= \liminf_{k \to \infty} ||Df_k||(\Omega).$$

Now take the supremum over Φ.

9.1.11. REMARK To conclude that $f \in BV(\Omega)$, it suffices to assume that the f_k have equibounded total variations, say $\forall$ k,

$$||Df_k||(\Omega) \leq M.$$

For then

$$V(f;\Omega) = ||Df||(\Omega)$$

$$\leq \liminf_{k \to \infty} ||D_k||(\Omega) \leq M < +\infty.$$

9.1.12. NOTATION Given $f \in BV(\Omega)$, put

$$||f||_{BV} = ||f||_{L^1} + ||Df||(\Omega).$$

9.1.13. THEOREM Under the norm $||\cdot||_{BV}$, $BV(\Omega)$ is a Banach space.

PROOF Completeness is the issue so suppose that $\{f_k\}$ is a Cauchy sequence in $BV(\Omega)$ -- then by the definition of $||\cdot||_{BV}$, it must also be a Cauchy sequence in $L^1(\Omega)$, hence by the completeness of $L^1(\Omega)$, there exists a function $f \in L^1(\Omega)$ such that $f_k \to f$ in $L^1(\Omega)$. On the other hand, since $\{f_k\}$ is a Cauchy sequence in $BV(\Omega)$, $\{||f_k||_{BV}\}$ is bounded: $\exists$ M > 0 such that $\forall$ k:

$$||f_k||_{BV} = ||f_k||_{L^1} + ||Df_k||(\Omega)$$

$$\leq M$$

$$\Rightarrow$$

$$||Df_k||(\Omega) \leq M$$

$$\Rightarrow$$

$$||Df||(\Omega) \leq M$$

$$\Rightarrow$$

$$f \in BV(\Omega).$$

The claim now is that $f_k \to f$ in $BV(\Omega)$. Because we already have convergence in $L^1(\Omega)$, matters reduce to showing that

$$||D(f_k - f)||(\Omega) \to 0 \quad (k \to \infty).$$

To this end, let $\varepsilon > 0$ -- then there exists N:

$$k,j \geq N \Rightarrow ||f_k - f_j||_{BV} < \varepsilon$$

$$\Rightarrow ||D(f_k - f_j)||(\Omega) < \varepsilon.$$

By construction,

$$f_j \to f \text{ in } L^1(\Omega),$$

so

$$f_k - f_j \to f_k - f \text{ in } L^1(\Omega),$$

thus

$$||D(f_k - f)||(\Omega) \leq \liminf_{j \to \infty} ||D(f_k - f_j)||(\Omega) \leq \varepsilon$$

from which the conclusion, $\varepsilon > 0$ being arbitrary.

9.1.14. REMARK $BV(\Omega)$ is not separable.

[To illustrate, work in R and consider the family F of characteristic functions χ_α of the interval $]\alpha,1[$ $(0 < \alpha < 1)$ -- then $F \subset BV]0,1[$ and for $\alpha \neq \beta$,

$$||\chi_\alpha - \chi_\beta||_{BV} = 2 + |\alpha - \beta|.]$$

9.1.15. N.B. The closure of $BV(\Omega) \cap C^\infty(\Omega)$ in $BV(\Omega)$ is $W^{1,1}(\Omega)$, hence is not dense in $BV(\Omega)$.

[Note: By way of comparison, recall that $W^{1,1}(\Omega) \cap C^\infty(\Omega)$ is dense in $W^{1,1}(\Omega)$.]

9.1.16. THEOREM Let $f \in BV(\Omega)$ -- then there exists a sequence $\{f_k\} \subset BV(\Omega) \cap C^\infty(\Omega)$ such that

$$f_k \to f \text{ in } L^1(\Omega) \quad (k \to \infty)$$

and

$$\lim_{k \to \infty} ||Df_k||(\Omega) = ||Df||(\Omega).$$

[Note: It is not claimed nor is it true in general that

$$||D(f_k - f)||(\Omega) \to 0 \quad (k \to \infty).]$$

9.1.17. APPLICATION Take $\Omega = R^n$ and in GNS, take $p = 1$, hence

$$1^* = \frac{n}{n-1} \quad (n \geq 2).$$

Choose the f_k above in $C_c^\infty(R^n)$ -- then

$$\begin{aligned} ||f_k||_{L^{n/(n-1)}} &\leq C(n)||\nabla f_k||_{L^1} \\ &= C(n) \int_{R^n} ||\nabla f_k|| \, dL^n \\ &= C(n)||Df_k||(R^n), \end{aligned}$$

so, upon passing to the limit, it follows that $\forall\ f \in BV(\Omega)$,

$$||f||_{L^{n/(n-1)}} \leq C(n)\,||Df||\,(R^n),$$

=>

$$BV(R^n) \hookrightarrow L^{n/(n-1)}(R^n).$$

9.1.18. HEURISTICS Let $(X,\mathcal{E})$ be a measurable space, let μ be a σ-finite positive measure on $(X,\mathcal{E})$, and let $f:X \to [0,+\infty]$ be a μ-measurable function -- then (Cavalieri)

$$\int_X f d\mu = \int_0^\infty \mu(\{x \in X: f(x) > t\})\, dL^1.$$

[Let

$$E_t = \{x \in X: f(x) > t\}.$$

Then $\chi_{E_t}(x) = 1$ iff $x \in E_t$ iff $f(x) > t$

=>

$$\int_0^\infty \chi_{E_t}(x)\, dt = \int_0^{f(x)} 1\, dt = f(x)$$

=>

$$\begin{aligned}\int_X f\, d\mu &= \int_X \left(\int_0^\infty \chi_{E_t}(x)\, dt\right) d\mu \\ &= \int_0^\infty \left(\int_X \chi_{E_t}(x)\, d\mu\right) dt \quad \text{(Fubini)} \\ &= \int_0^\infty \mu(E_t)\, dt.]\end{aligned}$$

9.1.19. LEMMA Let $E \subset R^n$ be Lebesgue measurable -- then $\chi_E|\Omega \in L^1(\Omega)$ iff $L^n(E \cap \Omega) < +\infty$.

PROOF

$$\int_\Omega \chi_E|\Omega \, dL^n = \int_{R^n} \chi_E \chi_\Omega \, dL^n \quad \text{(by definition)}$$

$$= \int_{R^n} \chi_{E\cap\Omega} \, dL^n = L^n(E \cap \Omega).$$

[Note: For the record,

$$L^n(E \cap \Omega) = (L^n \llcorner E)(\Omega) = (\chi_E L^n)(\Omega).]$$

9.1.20. DEFINITION Let $E \subset R^n$ be a Lebesgue measurable set and suppose that $\chi_E|\Omega \in L^1(\Omega)$ -- then the <u>perimeter</u> of E in Ω, denoted $P(E;\Omega)$, is the variation of $\chi_E|\Omega$ in Ω, i.e.,

$$P(E;\Omega) = V(\chi_E|\Omega;\Omega).$$

[Note: The set E is said to have <u>finite perimeter</u> in Ω if $P(E,\Omega) < +\infty$.]

9.1.21. NOTATION Given $f \in BV(\Omega)$, put

$$\Omega_t(f) = \{x \in \Omega : f(x) > t\}.$$

9.1.22. SUBLEMMA The function

$$(x,t) \rightarrow \chi_{\Omega_t(f)}(x)$$

is $(L^n \times L^1)$-measurable, thus for each $\Phi \in C_c^\infty(\Omega;R^n)$, the function

$$t \rightarrow \int_{\Omega_t} \operatorname{div} \Phi \, dL^n = \int_\Omega \chi_{\Omega_t(f)} \operatorname{div} \Phi \, dL^n$$

is L^1-measurable.

9.1.23. LEMMA The function

$$t \rightarrow ||D\chi_{\Omega_t(f)}||(\Omega)$$

is Lebesgue measurable.

9.1.24. THEOREM (COAREA) Let $f \in BV(\Omega)$ -- then the set $\Omega_t(f)$ has finite perimeter in Ω for almost all t and

$$||Df||(\Omega) = \int_R ||D\chi_{\Omega_t(f)}||(\Omega)\, dL^1.$$

The proof proceeds in two steps.

Step 1: Consider

$$\int_\Omega f \operatorname{div} \Phi \, dL^n,$$

where

$$\Phi \in C_c^\infty(\Omega; R^n), \quad ||\Phi||_\infty \leq 1,$$

and recall that

$$\int_\Omega \operatorname{div} \Phi \, dL^n = 0.$$

- $f \geq 0 \Rightarrow$

$$f(x) = \int_0^\infty \chi_{\Omega_t(f)}(x)\, dt$$

$\Rightarrow$

$$\int_\Omega f \operatorname{div} \Phi \, dL^n$$

$$= \int_\Omega \left(\int_0^\infty \chi_{\Omega_t(f)}(x)\, dt\right) \operatorname{div} \Phi \, dL^n$$

$$= \int_0^\infty \left(\int_\Omega \chi_{\Omega_t(f)}(x) \operatorname{div} \Phi \, dL^n\right) dt$$

$$= \int_0^\infty \left(\int_{\Omega_t(f)} \operatorname{div} \Phi \, dL^n\right) dt.$$

- $f \le 0 \Rightarrow$

$$f(x) = \int_{-\infty}^{0} (\chi_{\Omega_t(f)}(x) - 1)\, dt$$

$\Rightarrow$

$$\int_\Omega f \operatorname{div} \Phi \, dL^n$$

$$= \int_\Omega \left(\int_{-\infty}^{0} (\chi_{\Omega_t(f)}(x) - 1)\, dt\right) \operatorname{div} \Phi \, dL^n$$

$$= \int_{-\infty}^{0} \left(\int_\Omega (\chi_{\Omega_f(t)}(x) - 1) \operatorname{div} \Phi \, dL^n\right) dt$$

$$= \int_{-\infty}^{0} \left(\int_{\Omega_t(f)} \operatorname{div} \Phi \, dL^n\right) dt.$$

So, upon writing $f = f^+ + (- f^-)$, it follows that

$$\int_\Omega f \operatorname{div} \Phi \, dL^n$$

$$= \int_R \left(\int_{\Omega_t(f)} \operatorname{div} \Phi \, dL^n\right) dL^1$$

or still, by the definition of the variation of the perimeter of $\Omega_t(f)$ in Ω,

$$\int_\Omega f \operatorname{div} \Phi \, dL^n \le \int_R P(\Omega_t(f);\Omega)\, dL^1$$

or still, upon taking the supremum over Φ,

$$||Df||(\Omega) \le \int_R P(\Omega_t(f);\Omega)\, dt$$

$$= \int_R ||D\chi_{\Omega_t(f)}||(\Omega)\, dL^1.$$

It remains to reverse this inequality and for that, as an intermediary, one first shows that for $f \in BV(\Omega) \cap C^\infty(\Omega)$,

$$||Df||(\Omega) \geq \int_R ||D\chi_{\Omega_t(f)}||(\Omega)\; dL^1,$$

a point of detail that will be admitted without proof.

<u>Step 2:</u> Choose a sequence $\{f_k\} \subset BV(\Omega) \cap C^\infty(\Omega)$ such that

$$f_k \to f \text{ in } L^1(\Omega) \quad (k \to \infty)$$

and

$$\lim_{k \to \infty} ||Df_k||(\Omega) = ||Df||(\Omega).$$

Then $\forall\, k$,

$$||Df_k||(\Omega) = \int_R ||D\chi_{\Omega_t(f_k)}||(\Omega)\; dL^1.$$

Next

$$f_k(x) - f(x) = \int_R (\chi_{\Omega_t(f_k)}(x) - \chi_{\Omega_t(f)}(x))\; dL^1$$

and moreover

$$|f_k(x) - f(x)| = \int_R |\chi_{\Omega_t(f_k)}(x) - \chi_{\Omega_t(f)}(x)|\; dL^1$$

since

$$\text{sign}(f_k(x) - f(x)) = \text{sign}(\chi_{\Omega_t(f_k)}(x) - \chi_{\Omega_t(f)}(x))$$

for all t. Therefore

$$\int_\Omega |f_k(x) - f(x)|\; dL^1$$

$$= \int_R \left(\int_\Omega |\chi_{\Omega_t(f_k)}(x) - \chi_{\Omega_t(f)}(x)|\; dL^n\right) dL^1.$$

Bearing in mind that $f_k \to f$ in $L^1(\Omega)$, there exists a subsequence, not relabeled, with the property that

$$\chi_{\Omega_t(f_k)} \to \chi_{\Omega_t}(f) \text{ in } L^1(\Omega) \quad (k \to \infty)$$

for almost every t. Finally

$$\int_R ||D\chi_{\Omega_t(f)}||(\Omega)\ dL^1$$

$$\leq \int_R \liminf_{k \to \infty} ||D\chi_{\Omega_t(f_k)}||(\Omega)\ dL^1 \quad \text{(LSC)}$$

$$\leq \liminf_{k \to \infty} \int_R ||D\chi_{\Omega_t(f_k)}||(\Omega)\ dL^1 \quad \text{(Fatou)}$$

$$= \liminf_{k \to \infty} ||Df_k||(\Omega) \quad \text{(cf. supra)}$$

$$= \lim_{k \to \infty} ||Df_k||(\Omega) = ||Df||(\Omega).$$

9.1.15. EXAMPLE Given $f \in BV(\Omega)$, let

$$f_r(x) = \begin{cases} r \text{ if } f(x) > r \\ f(x) \text{ if } -r \leq f(x) \leq r \\ -r \text{ if } f(x) < -r \end{cases}$$

and put

$$H_r(x) = f(x) - f_r(x).$$

Then $H_r \in BV(\Omega)$ and

$$||DH_r||(\Omega) = \int_R ||D\chi_{\Omega_t(H_r)}||(\Omega)\ dL^1$$

$$= \int_{|t|>r} ||D\chi_{\Omega_t(f)}||(\Omega)\ dL^1.$$

9.1.26. PRODUCT RULE Let $f,g \in BV(\Omega) \cap L^\infty(\Omega)$ -- then $fg \in BV(\Omega) \cap L^\infty(\Omega)$ and

$$||D(fg)||(\Omega) \leq ||f||_\infty ||Dg||(\Omega) + ||g||_\infty ||Df||(\Omega).$$

9.1.27. REMARK $BV(\Omega) \cap L^{\infty}(\Omega)$ is dense in $BV(\Omega)$.

9.1.28. PRODUCT RULE If $f \in BV(\Omega)$ and if $\phi \in C_c^{\infty}(\Omega)$, then $\phi f \in BV(\Omega)$ and

$$||D(\phi f)||(\Omega) = \ldots .$$

§9.2. *DECOMPOSITION THEORY*

We shall first review matters in R, with $\Omega =]a,b[$. So fix an $f \in BV(\Omega)$.

- $Df = D^a f + D^s f$ is the decomposition of Df into its absolutely continuous part w.r.t. Lebesgue measure L^1 and its singular part $D^s f$.

Recall next that AT_f is the set of atoms of the theory, i.e., the $x \in]a,b[$ such that $D(\{x\}) \neq 0$.

- $D^s f = D^j f + D^c f$,

where

$$\begin{cases} D^j f = D^s f \llcorner AT_f \\ \\ D^c f = D^s f \llcorner (\Omega \backslash AT_f). \end{cases}$$

9.2.1. <u>N.B.</u> The measures $D^a f$, $D^j f$, $D^c f$ are mutually singular and

$$||Df||(\Omega) = ||D^a f||(\Omega) + ||D^j f||(\Omega) + ||D^c f||(\Omega).$$

9.2.2. DEFINITION f is a <u>jump function</u> if $Df = D^j f$.

9.2.3. DEFINITION f is a <u>Cantor function</u> if $Df = D^c f$.

9.2.4. THEOREM Each $f \in BV(\Omega)$ can be represented as a sum

$$f^a + f^j + f^s,$$

where f^a belongs to $W^{1,1}(]a,b[)$, f^j is a jump function, and f^s is a Cantor function.

9.2.5. <u>N.B.</u> These functions are uniquely determined up to additive constants and if $\bar{f}$ is an admissible representative of f, then

$$||D^a f||(\Omega) = \int_a^b |\bar{f}'| \, dL^1$$

and

$$||D^j f||(\Omega) = \sum_{x \in AT_f} |\bar{f}(x+) - \bar{f}(x-)|.$$

9.2.6. EXAMPLE Work in $\Omega =]0,1[$ and let $\{r_n\} \subset]0,1[$ be a sequence. Define $f \in BV(\Omega)$ by the prescription

$$f(x) = \sum_{\{n : r_n < x\}} 2^{-n},$$

Then f is a jump function and its distributional derivative Df is

$$\sum_n 2^{-n} \delta_{r_n}.$$

9.2.7. EXAMPLE Work in $\Omega =]0,1[$ and take for f the Cantor function -- then its distributional derivative has no absolutely continuous part and no jump part.

[If C is the Cantor set, then Df is (a constant multiple of) $\mathcal{H}^\gamma \llcorner C$, where $\gamma = \log 2/\log 3$.]

Assume henceforth that $n > 1$, where as usual $\Omega \subset R^n$ is nonempty and open.

9.2.8. NOTATION Given an $f \in BV(\Omega)$, put

$$\begin{cases} n_-(x) = \operatorname{ap} \liminf\limits_{\substack{y \to x \\ y \in \Omega}} f(y) \\ \\ n_+(x) = \operatorname{ap} \limsup\limits_{\substack{y \to x \\ y \in \Omega}} f(y) \end{cases} \quad (x \in \Omega).$$

9.2.9. LEMMA The functions

$$\begin{cases} x \to n_-(x) \\ x \to n_+(x) \end{cases} \quad (x \in \Omega)$$

are Borel measurable functions in Ω.

9.2.10. NOTATION

$$J_f = \{x \in \Omega : n_-(x) < n_+(x)\}.$$

[Accordingly, J_f is the set of points at which the approximate limit of f does not exist.]

9.2.11. N.B.

$$\mathcal{L}^n(J_f) = 0.$$

9.2.12. THEOREM J_f is $\mathcal{H}^{n-1}$-measurable.

9.2.13. THEOREM There exist countably many C^1-hypersurfaces S_k such that

$$\mathcal{H}^{n-1}(J_f - \bigcup_{k=1}^{\infty} S_k) = 0.$$

9.2.14. THEOREM

$$||Df|| \llcorner J_f = (n_+ - n_-)\mathcal{H}^{n-1} \llcorner J_f.$$

Let

$$Df = D^a f + D^s f$$

be the decomposition of Df into its absolutely continuous part w.r.t. Lebesgue measure L^n and its singular part $D^s f$. So

$$D^a f = f^a L^n,$$

where $f^a : \Omega \to R^n$ is the density of $D^a f$ and

$$D^s f \perp L^n.$$

9.2.15. DEFINITION

- The <u>jump part</u> of Df is

$$D^j f \equiv D^s f \llcorner J_f.$$

- The <u>Cantor part</u> of Df is

$$D^c f \equiv D^s f \llcorner (\Omega \backslash J_f).$$

Therefore

$$Df = D^a f + D^j f + D^c f.$$

9.2.16. THEOREM

$$D^j f = (n_+ - n_-)\nu_f \, H^{n-1} \llcorner J_f.$$

[Note: Here

$$\nu_f(x) = \frac{dD^j f}{d||D^j f||}(x)$$

for $||Df||$ almost every x in J_f.]

Therefore

$$Df = D^a f + (n_+ - n_-)\nu_f \, H^{n-1} \llcorner J_f + D^c f.$$

9.2.17. REMARK Earlier, under the assumption that n = 1, we exhibited a decomposition of a BV function but a result of this type does not obtain for BV functions of two or more variables.

§9.3. DIFFERENTIATION

Let Ω be a nonempty open subset of R^n.

9.3.1. RAPPEL Let $f \in W^{1,p}(\Omega)$ -- then f is approximately differentiable almost everywhere (cf. 7.1.6).

9.3.2. THEOREM Let $f \in BV(\Omega)$ -- then f is approximately differentiable almost everywhere.

[Note: Let

$$Df = D^a f + D^s f$$

be the decomposition of Df into its absolutely continuous part $D^a f$ w.r.t. Lebesgue measure L^n and its singular part $D^s f$ -- then

$$D^a f = f^a L^n,$$

where $f^a : \Omega \to R^n$ is the density of $D^a f$, and

$$\text{ap } df = f^a$$

almost everywhere.]

For the moment, take n = 1 and let $\Omega =]a,b[$. Suppose that $f \in BV(\Omega)$ -- then there is a $g \in BV(\Omega)$ such that g = f almost everywhere and g has an ordinary derivative almost everywhere.

[To see this, choose g admissible, thus

$$T_g]a,b[= e - T_f]a,b[< +\infty,$$

so g is of bounded variation in the traditional sense, thus has an ordinary derivative almost everywhere.]

These considerations can be extended to arbitrary n > 1.

9.3.3. THEOREM Let $f \in BV(\Omega)$ -- then there is a $g \in BV(\Omega)$ which is equivalent to f with the property that its ordinary partial derivative $\frac{\partial g}{\partial x_i}$ $(i = 1,\ldots,n)$ exists almost everywhere.

9.3.4. REMARK It follows that f has approximate partial derivatives almost everywhere, hence has an approximate differential almost everywhere.

[Note: Neither f nor any equivalent function need have an ordinary differential at any point.]

9.3.5. N.B. If in addition f is continuous, then f does have ordinary partial derivatives almost everywhere.

9.3.6. NOTATION Generically,

$$(x_1,\ldots,x_n) = (x_i^!,x_i),$$

where

$$x_i^! = (x_1,\ldots,x_{i-1}, x_{i+1},\ldots,x_n) \quad (i = 1,\ldots,n).$$

9.3.7. NOTATION An open rectangle

$$R =]a_1,b_1[\times \cdots \times]a_n,b_n[\subset R^n$$

can be viewed as the product of a rectangle

$$\begin{cases} R_i^! \subset R^{n-1} & \text{(variable } x_i^!\text{)} \\ R_i \subset R & \text{(variable } x_i\text{)} \end{cases}$$

and we write

$$R = R_i^! \times R_i.$$

Let $f \in BV(\Omega)$ -- then

$$Df = D^a f + D^s f,$$

where

$$D^a f = f^a L^n.$$

Here it is a question of R^n-valued vector measures:

$$f^a \in L^1(\Omega; R^n)$$

and for every Borel set $E \subset \Omega$,

$$Df(E) = \int_E f^a \, dL^n + D^s f(E).$$

9.3.8. THEOREM Let $f \in BV(\Omega)$, let $1 \leq i \leq n$, and let g be any function equivalent to f for which the ordinary partial derivative $\frac{\partial g}{\partial x_i}$ exists almost everywhere -- then

$$\frac{\partial g}{\partial x_i} = f_i^a.$$

almost everywhere.

The proof is on the lengthy side and will be broken up into 3 steps. Write for brevity ∂_i in place of $\frac{\partial}{\partial x_i}$.

<u>Step 1:</u> Consider a convex function $\Phi: R \to [0, +\infty[$ and let R be an open rectangle whose closure $\bar{R}$ is contained in Ω -- then the claim is that

$$\int_R \Phi(\partial_i g) \, dL^n \leq \liminf_{\varepsilon \downarrow 0} \int_{R_\varepsilon} \Phi(\partial_i f_\varepsilon) \, dL^n,$$

where

$$R_\varepsilon = \{x \in R : \text{dist}(x, \partial R) > \varepsilon\}.$$

To begin with, if ϕ is sufficiently smooth and if $h > 0$ is sufficiently small, then

$$\Phi\left(\frac{\phi(x_i', x_i + h) - \phi(x_i', x_i)}{h}\right)$$

$$= \Phi\left(\frac{1}{h} \int_{x_i}^{x_i + h} \partial_i \phi(x_i', t)\, dt\right)$$

$$\leq \frac{1}{h} \int_{x_i}^{x_i + h} \Phi(\partial_i \phi(x_i', t))\, dt \quad \text{(Jensen)}.$$

Now integrate this along $(R_i)_h$, hence

$$\int_{(R_i)_h} \Phi\left(\frac{\phi(x_i', x_i + h) - \phi(x_i', x_i)}{h}\right) dx_i$$

$$\leq \frac{1}{n} \int_{(R_i)_h} \int_{x_i}^{x_i + h} \Phi(\partial_i \phi(x_i', t))\, dt dx_i$$

$$\leq \frac{1}{h} \int_{R_i} \int_{t-h}^{t} \Phi(\partial_i \phi(x_i', t))\, dx_i dt$$

$$= \int_{R_i} \Phi(\partial_i \phi(x_i', t))\, dt$$

=>

$$\int_{R_h} \Phi\left(\frac{\phi(x_i', x_i + h) - \phi(x_i', x_i)}{h}\right) dx \leq \int_R \Phi(\partial_i \phi(x))\, dx.$$

Specialize and take $\phi = f_\varepsilon$, the mollification of f -- then almost everywhere

$$f_\varepsilon(x_i', x_i + h) - f_\varepsilon(x_i', x_i) \to f(x_i', x_i + h) - f(x_i', x_i) \quad (\varepsilon \downarrow 0),$$

thus by Fatou,

$$\int_{R_h} \Phi\left(\frac{f(x_i', x_i + h) - f(x_i', x_i)}{h}\right) dx$$

$$\leq \liminf_{\varepsilon \downarrow 0} \int_{R_\varepsilon} \Phi(\partial_i f_\varepsilon(x))\, dx$$

or still,

$$\int_{R_h} \Phi\left(\frac{g(x_i', x_i + h) - g(x_i', x_i)}{h}\right) dx$$

$$\leq \liminf_{\varepsilon \downarrow 0} \int_{R_\varepsilon} \Phi(\partial_i f_\varepsilon(x))\, dx.$$

To finish, it remains only to send h to 0.

<u>Step 2:</u> The next claim is that

$$\int_{R_\varepsilon} \Phi(\partial_i f_\varepsilon)\, dL^n \leq \int_R \Phi(f_i^a)\, dL^n + ||D_i^s||(R),$$

where the convex function Φ is subject to the condition

$$\Phi(s + t) \leq \Phi(s) + |t|$$

for all $s,t \in R$. First, for every $x \in \Omega_\varepsilon$ and any $i \in \{1,\ldots,n\}$,

$$\partial_i f_\varepsilon(x) = \frac{\partial f_\varepsilon}{\partial x_i}(x)$$

$$= \int_\Omega \frac{\partial \varphi_\varepsilon}{\partial x_i}(x - y) f(y)\, dy$$

$$= - \int_\Omega \frac{\partial \varphi_\varepsilon}{\partial y_i}(x - y) f(y)\, dy$$

$$= \int_\Omega \varphi_\varepsilon(x - y)\, dD_i f$$

$$= \int_\Omega \varphi_\varepsilon(x - y) f_i^a(y)\, dy + \int_\Omega \varphi_\varepsilon(x - y)\, dD_i^s f(y)$$

=>

$$\Phi(\partial_i f_\varepsilon(x)) \leq \Phi\left(\int_\Omega \varphi_\varepsilon(x - y)\, f_i^a(y)\, dy\right) + \left|\int_\Omega \varphi_\varepsilon(x - y)\, dD_i^s f(y)\right|.$$

Since

$$\int_\Omega \varphi_\varepsilon(x - y)\, dy = 1,$$

the first term can be estimated by Jensen, so

$$\Phi(\partial_i f_\varepsilon(x)) \le \int_\Omega \varphi_\varepsilon(x - y)\Phi(f_i^a(y))\, dy + \int_\Omega \varphi_\varepsilon(x - y) \quad d||D_i^s f||(y).$$

Therefore

$$\int_{R_\varepsilon} \Phi(\partial_i f_\varepsilon)\, dL^n \le \int_R \Phi(f_i^a)\, dL^n + \int_R d||D_i^s f||.$$

And

$$\int_R d||D_i^s f|| = ||D_i^s||(R).$$

9.3.9. <u>N.B.</u> Step 1 and Step 2

=>

$$\int_R \Phi(\partial_i g)\, dL^n \le \liminf_{\varepsilon \downarrow 0} \int_{R_\varepsilon} \Phi(\partial_i f_\varepsilon)\, dL^n$$

$$\le \int_R \Phi(f_i^a)\, dL^n + ||D_i^s f||(R).$$

<u>Step 3:</u> Work with

$$R(x_0, r) \equiv x_0 +]-\frac{r}{2}, \frac{r}{2}[^n.$$

Then

$$\lim_{r \to 0} \frac{||D_i^s f||(R(x_0, r))}{r^n} = 0$$

for almost all $x_0 \in \Omega$ (differentiation principle). Fix such an x_0 and take it to be a Lebesgue point for $\Phi(\partial_i g)$ and $\Phi(f_i^a)$. Since

$$L^n(R(x_0, r)) = r^n,$$

when $r \to 0$,

$$\left[\begin{array}{l} \frac{1}{r^n} \int_{R(x_0,r)} \Phi(\partial_i g)\, dL^n \to \Phi(\partial_i g)(x_0) \\ \\ \frac{1}{r^n} \int_{R(x_0,r)} \Phi(f_i^a)\, dL^n \to \Phi(f_i^a)(x_0) \end{array}\right. \qquad \to 0,$$

=>

$$\Phi(\partial_i g)(x_0) \leq \Phi(f_i^a(x_0)).$$

Choose now for Φ the function

$$\Psi(t) = \left[\begin{array}{ll} e^t & (t < 0) \\ \\ t + 1 & (t \geq 0). \end{array}\right.$$

Because Ψ is monotone increasing, it follows that $\partial_i g(x) \leq f_i^a(x)$ almost everywhere in Ω and consideration of $\Psi(-t)$ implies that $\partial_i g(x) \geq f_i^a(x)$ almost everywhere in Ω.

9.3.10. SCHOLIUM Start with an $f \in BV(\Omega)$, replace it by an equivalent $g \in BV(\Omega)$ with the property that the ordinary partial derivatives $\frac{\partial g}{\partial x_i}$ $(i = 1,\ldots,n)$ exist almost everywhere -- then

$$\nabla g = \left(\frac{\partial g}{\partial x_1},\ldots, \frac{\partial g}{\partial x_n}\right)$$

$$= (f_1^a,\ldots, f_n^a) = f^a$$

almost everywhere.

[Note:

$$f^a \in L^1(\Omega;R^n) \Rightarrow \nabla g \in L^1(\Omega;R^n).]$$

§9.4. BVL

Let us first review the situation when $n = 1$.

9.4.1. RAPPEL If $\Omega \subset R$ is open and nonempty and if $f \in L^1(\Omega)$, then the essential variation of f, denoted $e - T_f\Omega$, is the set

$$\inf \{T_g\Omega : g = f \text{ almost everywhere}\}.$$

Moreover $f \in BV(\Omega)$ iff

$$e - T_f\Omega < +\infty.$$

And then

$$e - T_f\Omega = ||Df||(\Omega).$$

[Note: Recall that Ω is the union of its connected components, these being intervals (finite or infinite).]

Let Ω be a nonempty open subset of R^n.

9.4.2. NOTATION Per $x_i' \in R^{n-1}$, put

$$\Omega_{x_i'} = \{x_i \in R : (x_i', x_i) \in \Omega\}$$

and if $\Omega_{x_i'} \neq \emptyset$ and if $f : \Omega \to R$ is Lebesgue measurable, denote by

$$e - T_{f(x_i', -)}\Omega_{x_i'}$$

the essential variation of the function

$$x_i \to f(x_i', x_i).$$

9.4.3. NOTATION If $f \in L^1(\Omega)$, write

$$\int_\Omega f\, dL^n = \int_{R^{n-1}} \int_{\Omega_{x_i'}} f(x_i', x_i)\, dx_i\, dx_i'.$$

[Put

$$\int_{\Omega_{x_i'}} f(x_i', x_i)\, dx_i = 0$$

if $\Omega_{x_i'} = \emptyset$.]

9.4.4. CRITERION Let $f \in L^1(\Omega)$. Suppose that there exists an equivalent function $g \in L^1(\Omega)$ and nonnegative functions $V_1, \ldots, V_n$ in $L^1(R^{n-1})$ such that

$$e - T_{g(x_i',\, -)}\, \Omega_{x_i'} \leq V_{i(x_i')}$$

for all $x_i' \in R^{n-1}$ such that $\Omega_{x_i'}$ is nonempty $(i = 1, \ldots, n)$ -- then $f \in BV(\Omega)$.

PROOF Fix $i \in \{1, \ldots, n\}$ and note that $V_i(x_i')$ is finite almost everywhere (being in $L^1(R^{n-1})$), hence

$$e - T_{g(x_i',\, -)}\, \Omega_{x_i'}$$

is finite almost everywhere. But

$$g(x_i', -) \in L^1(\Omega_{x_i'})$$

for almost all x_i' (Fubini), so the conclusion is that for almost all x_i',

$$g(x_i', -) \in BV(\Omega_{x_i'}),$$

from which a finite signed Radon measure $D_{x_i'}$ in $\Omega_{x_i'}$ of finite total variation such that

$$\partial_i g(x_i', -) = D_{x_i'}$$

with

$$||D_{x_i'}||(\Omega_{x_i'}) = e - T_{g(x_i', -)}^{\Omega_{x_i'}}.$$

Proceeding, define a linear functional Λ on $C_c^1(\Omega)$ by the rule

$$\Lambda(\phi) = \int_\Omega f\partial_i\phi \, dL^n.$$

Then Λ is continuous w.r.t. the topology of $C_0(\Omega)$. Proof:

$$\Lambda(\phi) = \int_{R^{n-1}} \int_{\Omega_{x_i'}} f\partial_i\phi \, dx_i dx_i'$$

$$= - \int_{R^{n-1}} \int_{\Omega_{x_i'}} \phi \, dD_{x_i'} \, dx_i$$

=>

$$|\Lambda(\phi)| \leq \max_\Omega |\phi| \int_{R^{n-1}} ||D_{x_i'}||(\Omega_{x_i'}) \, dx_i$$

$$= \max_\Omega |\phi| \int_{R^{n-1}} e - T_{g(x_i', -)}^{\Omega} {}_{x_i'}$$

$$\leq \max_\Omega |\phi| \int_{R^{n-1}} V_i(x_i') \, dx_i'.$$

Now extend it to a continuous linear functional on $C_0(\Omega)$ and use the "C_0" version of the RRT to get a finite signed Radon measure D_i such that for all $\phi \in C_0(\Omega)$,

$$\Lambda(\phi) = - \int_\Omega \phi \, dD_i,$$

or still, $\forall\ \phi \in C_c^1(\Omega)$,

$$\int_\Omega f\partial_i\phi\ dL^n = - \int_\Omega \phi\ dD_i.$$

Since

$$||D_i||(\Omega) = ||\Lambda||^* \le \int_{R^{n-1}} V_i(x_i') \ dx_i' < +\infty,$$

it follows that $f \in BV(\Omega)$.

9.4.5. REMARK Take $n = 1$ and suppose that $\Omega =]a,b[$ — then x_i' is just an abstract point, call it $*$, and Ω_* can be identified with $]a,b[$. Starting with $f \in L^1(]a,b[)$, the assumption above amounts to saying that there exists a $g \in L^1(]a,b[)$ and a constant $C \ge 0$ such that

$$e - T_{g(*,\ -)}\ ]a,b[\le C < +\infty.$$

But then

$$g \in BV(]a,b[)$$

which implies that

$$f \in BV(]a,b]).$$

Take

$$Q = [0,1] \times [0,1]$$

and let $f:Q \to R$ be a continuous function (hence $f \in L^1(Q)$).

9.4.6. NOTATION

$$\begin{cases} V_x(f;y) = T_{f(-,\ y)}\ [0,1] & (0 \le y \le 1) \\ V_y(f;x) = T_{f(x,\ -)}\ [0,1] & (0 \le x \le 1). \end{cases}$$

[Note: Since f is continuous,

$$\begin{cases} T_{f(-,\ y)}\ [0,1] = T_{f(-,\ y)}\ ]0,1[\\ \\ T_{f(x,\ -)}\ [0,1] = T_{f(x,\ -)}\ ]0,1[. \end{cases}$$

9.4.7. LEMMA $V_x(f;y)$ and $V_y(f;x)$ are Lebesgue measurable.

9.4.8. DEFINITION f is said to be of <u>bounded variation in the sense of Tonelli</u> if

$$\begin{cases} \int_0^1 V_x(f;y)\ dy < +\infty \\ \\ \int_0^1 V_y(f;x)\ dx < +\infty. \end{cases}$$

9.4.9. <u>N.B.</u> When dealing with essential variation on open subsets of the line, if the function in question is continuous, one can work instead with the usual variation, the reason being that the approximation via approximate points of continuity amounts to approximation via points of continuity.

9.4.10. SCHOLIUM If f is of bounded variation in the sense of Tonelli, then $f|Q^o \in BV(Q^o)$ and the ordinary partial derivatives

$$\frac{\partial f}{\partial x},\ \frac{\partial f}{\partial y}$$

exist almost everywhere in Q^o ($f|Q^o$ being continuous).

Relax the assumption that $f:Q \rightarrow R$ is continuous to merely that $f \in L^1(Q^o)$.

9.4.11. NOTATION

$$\begin{cases} e - V_x(f;y) = e - T_{f(-,\,y)}\ ]0,1[& (0 < y < 1) \\ e - V_y(f;x) = e - T_{f(x,\,-)}\ ]0,1[& (0 < x < 1). \end{cases}$$

[Note: The essential variations here are per]0,1[:

$$\begin{cases} f(-;\,y) \text{ is the function } x \to f(x,y) & (0 < x < 1) \\ f(x,\,-) \text{ is the function } y \to f(x,y) & (0 < y < 1).] \end{cases}$$

9.4.12. LEMMA $e - V_x(f;y)$ and $e - V_y(f;x)$ are Lebesgue measurable.

9.4.13. DEFINITION f is said to be of <u>bounded variation in the sense of Cesari</u> if

$$\begin{cases} \int_0^1 e - V_x(f;y)\, dy < +\infty \\ \int_0^1 e - V_y(f;x)\, dx < +\infty. \end{cases}$$

9.4.14. REMARK Under the preceding circumstances, it can be shown that there exists a function g equivalent to f such that

$$\begin{cases} \int_0^1 V_x(g;y)\, dy < +\infty \\ \int_0^1 V_y(g;x)\, dx < +\infty. \end{cases}$$

9.4.15. SCHOLIUM If f is of bounded variation in the sense of Cesari, then $f \in BV(Q^O)$ and there is a $g \in BV(Q^O)$ which is equivalent to f with the property that the ordinary partial derivatives

$$\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}$$

exist almost everywhere in Q°.

SECTION 10: ABSOLUTE CONTINUITY

10.1. RAPPEL Let Ω be a nonempty open subset of R -- then a function $f:\Omega \to R$ is absolutely continuous if $\forall\ \varepsilon > 0$, $\exists\ \delta > 0$ such that for every finite collection $[a_1,b_1],\ldots,[a_k,b_k]$ of pairwise disjoint closed intervals in Ω,

$$\sum_{j=1}^{k} L^1([a_j,b_j]) < \delta \Rightarrow \sum_{j=1}^{k} |f(a_j) - f(b_j)| < \varepsilon.$$

Here is one extrapolation from R to R^n $(n > 1)$, where now Ω is a nonempty open subset of R^n.

10.2. DEFINITION A function $f:\Omega \to R^n$ is <u>n-absolutely continuous</u> if $\forall\ \varepsilon > 0$, $\exists\ \delta > 0$ such that for every finite collection $B(x_1,r_1),\ldots,B(x_k,r_k)$ of pairwise disjoint closed balls in Ω,

$$\sum_{j=1}^{k} L^n(B(x_j,r_j)) < \delta \Rightarrow \sum_{j=1}^{k} \operatorname{osc}(f,B(x_j,r_j))^n < \varepsilon.$$

[Note: If E is a subset of R^n, then

$$\operatorname{osc}(f,E) = \operatorname{diam}(f(E)).]$$

Obviously,

n-absolute continuity => continuity.

10.3. NOTATION Put

$$V_n(f,\Omega) = \sup\{\sum_{j=1}^{k} (\operatorname{osc}(f,B(x_j,r_j)))^n\}.$$

10.4. DEFINITION f is of <u>bounded n-variation</u> in Ω if

$$V_n(f,\Omega) < +\infty.$$

10.5. NOTATION $BV^n(\Omega)$ is the set of all functions $f:\Omega \to R^n$ of bounded n-variation in Ω.

10.6. NOTATION $AC^n(\Omega)$ is the set of all functions $f \in BV^n(\Omega)$ that are n-absolutely continuous.

10.7. REMARK The notion of n-absolutely continuous uses closed balls. One could also work with closed cubes. Here, however, one has to be careful: Examples have been constructed which show that working with closed balls is not the same as working with closed cubes, thus that these two concepts are incomparable.

10.8. DEFINITION A function $f:\Omega \to R^n$ satsifies the condition RR if there is a nonnegative function $\varphi \in L^1(\Omega)$, a so-called weight, such that

$$(\mathrm{osc}(f,B(x,r)))^n \leq \int_{B(x,r)} \varphi \, dL^n$$

for every $B(x,r) \subset \Omega$.

10.9. NOTATION Denote by $RR^n(\Omega)$ the set of all functions $f:\Omega \to R^n$ which satisfy condition RR.

10.10. THEOREM

$$RR^n(\Omega) = AC^n(\Omega).$$

10.11. THEOREM Let $f \in BV^n(\Omega)$ -- then f is differentiable almost everywhere.

Matters can be generalized, thus suppose that $0 < \lambda \leq 1$.

10.12. DEFINITION A function $f:\Omega \to R^n$ is λ, n-absolutely continuous if

$\forall\ \varepsilon > 0$, $\exists\ \delta > 0$ such that for any finite collection $B(x_1,r_1),\ldots,B(x_k,r_k)$ of pairwise disjoint closed balls in Ω,

$$\sum_{j=1}^{k} L^n(B(x_j,r_j)) < \delta \Rightarrow \sum_{j=1}^{k} \operatorname{osc}(f,B(x_j,\lambda r_j))^n < \varepsilon.$$

10.13. <u>N.B.</u> 1, n-absolute continuity coincides with n-absolute continuity.

10.14. NOTATION Put

$$V_{\lambda,n}(f,\Omega) = \sup\{\sum_{j=1}^{k} (\operatorname{osc}(f,B(x_j,\lambda r_j)))^n\}.$$

10.15. DEFINITION f is of <u>bounded λ, n-variation</u> in Ω if

$$V_{\lambda,n}(f,\Omega) < +\infty.$$

10.16. NOTATION $BV^{\lambda,n}(\Omega)$ is the set of all functions $f:\Omega \to R^n$ of bounded λ, n-variation in Ω.

10.17. NOTATION $AC^{\lambda,n}(\Omega)$ is the set of all functions $f \in BV^{\lambda,n}(\Omega)$ which are λ, n-absolutely continuous.

10.18. <u>N.B.</u>

$$AC^n(\Omega) = AC^{1,n}(\Omega)$$

$$\subset AC^{\lambda,n}(\Omega) \qquad (0 < \lambda < 1)$$

and it can be shown that the containment is proper.

10.19. LEMMA Let $f:\Omega \to R^n$ and suppose that $0 < \lambda_1 < \lambda_2 < 1$ -- then f is λ_1, n-absolutely continuous iff f is λ_2, n-absolutely continuous.

10.20. THEOREM Suppose that $0 < \lambda_1 < \lambda_2 < 1$ -- then

$$BV^{\lambda_1,n}(\Omega) = BV^{\lambda_2,n}(\Omega).$$

10.21. SCHOLIUM There are but two classes of λ, n-absolutely continuous functions, viz. those corresponding to $\lambda = 1$ and to $0 < \lambda < 1$.

10.22. DEFINITION Let $f:\Omega \to R^n$ and suppose that $0 < \lambda < 1$ -- then f satisfies the <u>condition RR_λ</u> if there is a nonnegative function $\varphi \in L^1(\Omega)$, a so-called <u>weight</u>, such that

$$(\text{osc}(f,B(x,\lambda r)))^n \leq \int_{B(x,r)} \varphi \, dL^n$$

for every $B(x,r) \subset \Omega$.

[Note: Formally, $RR_1 = RR$.]

10.23. NOTATION Denote by $RR^{\lambda,n}(\Omega)$ the set of all functions $f:\Omega \to R^n$ which satisfy condition RR_λ.

10.24. THEOREM

$$RR^{\lambda,n}(\Omega) = AC^{\lambda,n}(\Omega).$$

10.25. THEOREM Let $f \in B^{\lambda,n}(\Omega)$ -- then f is differentiable almost everywhere.

Return now to the beginning.

10.26. LEMMA Suppose that Ω is bounded and $f:\Omega \to R^n$ is Lipschitz, say

$$||f(x_1) - f(x_2)|| \leq M||x_1 - x_2||$$

for all $x_1,x_2 \in \Omega$ -- then $f \in RR^n(\Omega)$.

[Define $\varphi:\Omega \to R$ by the rule

$$\varphi(x) = \frac{M^n 2^n}{\omega_n} \quad (\Rightarrow \varphi \in L^1(\Omega)).$$

Then for any $B(x,r) \subset \Omega$,

$$\int_{B(x,r)} \varphi \, dL^n = \frac{M^n 2^n}{\omega_n} \int_{B(x,r)} 1 \, dL^n$$

$$= \frac{M^n 2^n}{\omega_n} \omega_n r^n$$

$$= M^n 2^n r^n$$

$\Rightarrow$

$$\left(\int_{B(x,r)} \varphi \, dL^n\right)^{1/n} = M(2r).$$

But

$$x_1, x_2 \in B(x,r) \Rightarrow ||x_1 - x_2|| \leq 2r$$

$\Rightarrow$

$$||f(x_1) - f(x_2)|| \leq M||x_1 - x_2|| \leq M(2r)$$

$\Rightarrow$

$$\begin{aligned} \operatorname{osc}(f,B(x,r)) &\leq M(2r) \\ &= \left(\int_{B(x,r)} \varphi \, dL^n\right)^{1/n} \end{aligned}$$

$\Rightarrow$

$$(\operatorname{osc}(f,B(x,r)))^n \leq \int_{B(x,r)} \varphi \, dL^n.]$$

10.27. LEMMA Suppose that Ω is bounded and $f \in W^{1,p}(\Omega;R^n)$ $(p > n)$ is continuous -- then $f \in RR^n(\Omega)$.

[Upon consideration of components, one can take $n = 1$. This said, for any $B(x,r) \subset \Omega$ (cf. 7.4.11),

$$\operatorname{osc}(f,B(x,r)) \leq Cr^{1-n/p}\left(\int_{B(x,r)} ||\nabla f||^p \, dL^n\right)^{1/p}$$

=>

$$\operatorname{osc}(f,B(x,r))^n \leq Cr^{n(1-n/p)}\left(\int_{B(x,r)} ||\nabla f||^p \, dL^n\right)^{n/p}$$

$$\leq C\left(r^n + \int_{B(x,r)} ||\nabla f||^p \, dL^n\right)$$

$$\leq C\int_{B(x,r)} (1 + ||\nabla f||^p) \, dL^n.$$

So, for the weight, one can take

$$\varphi = C(1 + ||\nabla f||^p).]$$

[Note: The usual convention on the constant "C" is in force, i.e., it may change from line to line.]

SECTION 11: MISCELLANEA

§11.1. PROPERTY (N)

Let Ω be a nonempty open subset of R^n.

11.1.1. DEFINITION A continuous function $f:\Omega \to R^n$ is said to have <u>property (N)</u> if f sends sets of Lebesgue measure 0 to sets of Lebesgue measure 0:

$$L^n(E) = 0 \quad (E \subset \Omega) \Rightarrow L^n(f(E)) = 0.$$

11.1.2. SUBLEMMA If $E \in M_L^n$, then there exists an F_σ-set $F \subset E$ such that $L^n(E\backslash F) = 0$. Choose next a countable collection of compact sets C_j for which $F = \underset{j}{\cup} C_j$ and put $K_j = \underset{k=1}{\overset{j}{\cup}} C_k$, thus $\{K_j\}$ is an increasing sequence of compact sets with $\underset{j}{\cup} K_j = \underset{j}{\cup} C_j = F$. Finally, since E is the disjoint union of F and $E\backslash F$, we have

$$\begin{aligned} L^n(E) &= L^n(F) + L^n(E\backslash F) \\ &= L^n(F) = \lim_{j \to \infty} L^n(K_j). \end{aligned}$$

11.1.3. LEMMA Suppose that $f:\Omega \to R^n$ has property (N) -- then the implication

$$E \in M_L^n \quad (E \subset \Omega) \Rightarrow f(E) \in M_L^n$$

obtains.

PROOF As above, write

$$E = F \cup (E\backslash F) \quad (F = \underset{j}{\cup} K_j).$$

Then

$$f(E) = f(F) \cup f(E\backslash F)$$

$$= \underset{j}{\cup} f(K_j) + f(E/F).$$

Since f is continuous, the $f(K_j)$ are compact, hence measurable, so the union $\underset{j}{\cup} f(K_j)$ is measurable. On the other hand,

$$L^n(E\backslash F) = 0 \Rightarrow L^n(f(E\backslash F)) = 0$$

$$\Rightarrow f(E\backslash F) \in M_L^n.$$

All told therefore,

$$f(E) \in M_L^n.$$

11.1.4. EXAMPLE Take $n = 1$, $\Omega =]a,b[$, and suppose that $f:]a,b[\to R$ is absolutely continuous -- then f has property (N).

11.1.5. EXAMPLE If $f:\Omega \to R^n$ is locally Lipschitz, then f has property (N) (cf. 2.3.23).

[Note: In particular: A C'-function f has property (N) (being locally Lipschitz).]

11.1.6. N.B. The preceding consideration is false if f is merely continuous or even Hölder continuous with exponent $0 < \alpha < 1$.

[The Cantor function f sends the Cantor set C ($L^1(C) = 0$) to f(C) ($L^1(f(C)) = 1$). And f is Hölder continuous with exponent $\alpha = \frac{\ell og\ 2}{\ell og\ 3}$.]

11.1.7. RAPPEL (VITALI) Let $\mathcal{B}$ be a system of closed balls in R^n such that

$$\sup\{\mathrm{diam}(B) : B \in \mathcal{B}\} < +\infty.$$

Then there exists a pairwise disjoint, at most countable subsystem $\{B(x_i, r_i)\} \subset \mathcal{B}$ such that

$$\bigcup_{B \in \mathcal{B}} \subset \bigcup_i B(x_i, 5r_i).$$

11.1.8. THEOREM Suppose that $f:\Omega \to R^n$ is n-absolutely continuous -- then f has property (N).

PROOF Fix an $E \subset \Omega$ of Lebesgue measure 0. Given $\varepsilon > 0$, choose $\delta > 0$ per the definition of n-absolute continuity, subject to $\delta < \varepsilon$. Let $G \subset \Omega$ be an open set containing E with $L^n(G) < \delta$. Given an $x \in E$, choose $r(x) > 0$ such that

$$B(x,r(x)) \subset G,\ r(x) < \frac{\varepsilon}{10},\ \text{and}\ \rho(x) \equiv \mathrm{osc}(f,B(x,r(x))) < \frac{\varepsilon}{10}.$$

Using Vitali, determine a disjoint system

$$\{B(f(x_i),\rho(x_i))\} \subset \{B(f(x),\rho(x)) : x \in E\}$$

such that

$$f(E) \subset \bigcup_i B(f(x_i),\ 5\rho(x_i)).$$

Since

$$f(B(x_i,r(x_i))) \subset B(f(x_i),\rho(x_i)),$$

the $B(x_i,r(x_i))$ are pairwise disjoint, hence

$$\begin{aligned} H^n_\varepsilon(f(E)) &\le C \sum_i \rho(x_i)^n \\ &\le C \sum_i \mathrm{osc}(f,B(x_i,r(x_i)))^n \\ &\le C\ \varepsilon. \end{aligned}$$

Now let $\varepsilon \to 0$ to conclude that

$$\mathcal{H}^n(f(E)) = 0$$

or still, that

$$L^n(f(E)) = 0.$$

11.1.9. APPLICATION Suppose that $f \in W^{1,p}(\Omega;R^n)$ $(p > n)$ is continuous -- then f has property (N).

[In fact,

$$f \in RR^n(\Omega) \quad (\text{cf. } 10.27).$$

But

$$RR^n(\Omega) = AC^n(\Omega) \quad (\text{cf. } 10.10).]$$

11.1.10. REMARK There are continuous functions in $W^{1,n}(\Omega;R^n)$ $(n > 1)$ that do not have property (N).

[E.g., it is possible to construct a continuous $f \in W^{1,n}(R^n;R^n)$ $(n > 1)$ which sends $[0,1]$ onto $[0,1]^n$. Therefore f does not have property (N).]

11.1.11. THEOREM If $f \in W^{1,n}(\Omega;R^n)$ is continuous and open, then f has property (N).

11.1.12. N.B. There exists a homeomorphism

$$f \in W^p((]-1, 1[)^n, (]-1, 1[)^n) \quad (p < n)$$

which does not have property (N).

11.1.13. THEOREM If $f \in W^{1,n}(\Omega;R^n)$ is Hölder continuous, then f has property (N).

* * * * * * * * * * * *

APPENDIX

LEMMA Let $1 \leq k \leq n$, let Ω be a nonempty open subset of R^k, and let $T:\Omega \to R^n$ be continuous and one-to-one -- then

$$E \in \mathcal{B}(R^k) \ (E \subset \Omega) \Rightarrow T(E) \in \mathcal{B}(R^n).$$

PROOF Ω is a σ-compact subset of R^k, hence $T(\Omega)$ is a σ-compact subset of R^n (T being continous), hence $T(\Omega) \in \mathcal{B}(R^n)$. Let now

$$\mathcal{A} = \{E \subset \Omega : T(E) \in \mathcal{B}(R^n)\},$$

a σ-algebra of subsets of Ω (as regards complementation, note that $T(\Omega\backslash E) = T(\Omega)\backslash T(E)$, T being one-to-one). It is clear that $\mathcal{A}$ contains the open subsets of Ω, (per the initial observation), so $\mathcal{A}$ contains the Borel σ-algebra $\mathcal{B}(\Omega)$. But $\mathcal{B}(A) = \mathcal{B}(R^k) \cap \Omega$, thus $\forall\ E \in \mathcal{B}(R^k) \cap \Omega$, $T(E) \in \mathcal{B}(R^n)$.

§11.2. THE MULTIPLICITY FUNCTION

Let Ω be a nonempty open subset of R^n and let $f:\Omega \to R^n$ be a continuous function.

11.2.1. DEFINITION If $E \subset \Omega$, then

$$N(f,E,y) = \#\{x \in E:f(x) = y\}$$

or still,

$$N(f,E,y) = H^0(E \cap f^{-1}(y))$$

is the multiplicity function of f at $y \in R^n$ w.r.t. E.

[Note: $N(f,E,y)$ is the cardinality of $E \cap f^{-1}(y)$ and if this set is infinite, then we put

$$N(f,E,y) = +\infty.]$$

11.2.2. LEMMA

$$E_1 \subset E_2 \Rightarrow N(f,E_1,y) \leq N(f,E_2,y).$$

11.2.3. LEMMA If $\{E_k\}$ is an increasing sequence of subsets of Ω, then

$$N(f,E,y) = \lim_{k \to \infty} N(f,E_k,y),$$

where $E = \bigcup_{k=1}^{\infty} E_k$.

11.2.4. THEOREM Suppose that $f:\Omega \to R^n$ has property (N) -- then for any Lebesgue measurable set $E \subset \Omega$, the multiplicity function

$$y \to N(f,E,y)$$

is Lebesgue measurable in R^n.

PROOF Take E bounded and for every $m \in N$ construct a partition of E into pairwise nonintersecting measurable sets

$$E_1^{(m)}, \ldots, E_{k_m}^{(m)}$$

such that

$$\operatorname{diam}(E_i^{(m)}) \leq \frac{1}{m} \ (i = 1, \ldots, k(m)).$$

Put

$$N(f,m,—) = \chi_{f(E_1^{(m)})} + \cdots + \chi_{f(E_i^{(m)})}$$

and note that each of the sets $f(E_i^{(m)})$ is measurable (since f has property (N)), hence $N(f,m,—)$ is measurable. Accordingly it need only be shown that

$$\lim_{m \to \infty} N(f,m,y) = N(f,E,y)$$

to establish the contention. Given $y \in R^n$, there are two possibilities for $E \cap f^{-1}(y)$: It is either finite or it is infinite. To treat the first of these, say

$$E \cap f^{-1}(y) = \{x_1, \ldots, x_k\},$$

take

$$m_0 > \min_{i \neq j} \frac{1}{||x_i - x_j||}.$$

If $m > m_0$, then none of the $E_i^{(m)}$ contain two distinct x_r, x_s, so it can be assumed that

$$x_1 \in E_1^{(m)}, \ldots, x_k \in E_k^{(m)}.$$

Next, $\forall \ m > m_0$,

$$N(f,m,y) \geq N(f,E,y)$$

=>

$$\liminf_{m \to \infty} N(f,m,y) \geq N(f,E,y).$$

On the other hand,

$$N(f,m,y) \leq N(f,E,y)$$

=>

$$\limsup_{m \to \infty} N(f,m,y) \leq N(f,E,y).$$

Therefore

$$\lim_{m \to \infty} N(f,m,y) = N(f,E,y).$$

11.2.5. LEMMA If $f:\Omega \to R^n$ is continuous and open, then for every open set $G \subset \Omega$, the function $y \to N(f,G,y)$ is lower semicontinuous in R^n.

§11.3. JACOBIANS

Let Ω be a nonempty open subset of R^n.

11.3.1. DEFINITION Let $f = (f^1,\ldots,f^n) \in W^{1,n}(\Omega;R^n)$ -- then the Jacobian of f, denoted J_f, is the determinant

$$\det(\nabla f^1,\ldots,\nabla f^n).$$

11.3.2. N.B. The coordinate functions f^i $(1 \leq i \leq n)$ of f and their first order distributional derivatives belong to $L^n(\Omega)$.

[Note: Nevertheless, an element of $W^{1,n}(\Omega;R^n)$ may be nowhere continuous, hence nowhere differentiable.]

11.3.3. THEOREM If $f \in W^{1,n}(\Omega;R^n)$, then $J_f \in L^1(\Omega)$.

PROOF J_f is a sum of terms, each of which is (plus or minus) the product of n weak partial derivatives of the components of f and, as noted above, each of these is in $L^n(\Omega)$. The product of n $L^n(\Omega)$ functions is in $L^1(\Omega)$ (apply Hölder), hence $f \in L^1(\Omega)$.

11.3.4. FACT If $f:R^n \rightarrow R^n$ is Lipschitz, then for any $E \in M_L^n$,

$$\int_E |J_f| \, dL^n = \int_{R^n} N(f,E,y) \, dy.$$

[This will be established in §12.5.]

11.3.5. RAPPEL Let $f \in W^{1,p}(\Omega)$ -- then there is a partition

$$\Omega = (\bigcup_{k=1}^{\infty} E_k) \cup Z,$$

where the E_k are Lebesgue measurable sets such that $f|E_k$ is Lipschitz and Z has Lebesgue measure 0 (cf. 7.1.5).

11.3.6. THEOREM Suppose that $f \in W^{1,n}(\Omega;R^n)$ has property (N) -- then

$$\int_\Omega |J_f| \, dL^n = \int_{R^n} N(f,\Omega,y) \, dy.$$

PROOF The foregoing decomposition principle extends from $W^{1,n}(\Omega)$ to $W^{1,n}(\Omega;R^n)$ and the Lipschitz function $f|E_k$ extends to a Lipschitz function $f_k:R^n \to R^n$, hence per supra

$$\int_{E_k} |J_{f_k}| \, dL^n = \int_{R^n} N(f_k,E_k,y) \, dy.$$

Put $E_0 = \bigcup_{k=1}^{\infty} E_k$ -- then $\Omega = E_0 \cup Z$ $(L^n(Z) = 0)$, so

$$\begin{aligned}
\int_\Omega |J_f| \, dL^n &= \int_E |J_f| \, dL^n \\
&= \sum_{k=1}^{\infty} \int_{E_k} |J_f| \, dL^n \\
&= \sum_{k=1}^{\infty} \int_{E_k} |J_{f_k}| \, dL^n \\
&= \sum_{k=1}^{\infty} \int_{R^n} N(f_k,E_k,y) \, dy \\
&= \int_{R^n} N(f,E,y) \, dy \\
&\le \int_{R^n} N(f,\Omega,y) \, dy \\
&= \int_{R^n} N(f,E,y) \, dy + \int_{R^n} N(f,Z,y) \, dy
\end{aligned}$$

$$= \int_{R^n} N(f,E,y)\,dy + \int_{f(Z)} N(f,Z,y)\,dy$$

$$= \int_{R^n} N(f,E,y)\,dy \quad (L^n(f(Z)) = 0)$$

=>

$$\int_\Omega |J_f|\,dL^n = \int_{R^n} N(f,\Omega,y)\,dy.$$

[Note:

$$f(Z) = \{y:N(f,Z,y) \neq 0\}.]$$

11.3.7. N.B. The assumption that f has property (N) implies that the relevant multiplicity functions are Lebesgue measurable.

11.3.8. THEOREM If $f \in W^{1,n}(\Omega;R^n)$ is continuous and if $J_f > 0$ almost everywhere in Ω, then f has property (N).

11.3.9. REMARK Examples have been constructed of continuous functions f in $W^{1,n}(\Omega;R^n)$ such that $J_f = 0$ almost everywhere in Ω but such that f fails to have property (N).

On general grounds, an $f \in W^{1,p}(\Omega;R^n)$ $(1 \leq p < +\infty)$ is approximately differentiable almost everywhere in Ω. More is true if $p = n$, namely

$$ap - df$$

is "regular" (i.e., "E" can be written as a union of concentric spheres centered at x).

SECTION 12: AREA FORMULAS

§12.1. THE LINEAR CASE

12.1.1. RAPPEL Let $T:R^n \to R^n$ be a nonsingular linear transformation -- then

$$E \in M_L^n \Rightarrow T(E) \in M_L^n$$

and

$$L^n(T(E)) = |\det(M_T)| L^n(E).$$

12.1.2. N.B. This is the simplest instance of what is known as an "area formula". As will be shown below, it leads to a "change of variable formula".

Retaining T and E, suppose given a function $f:E \to [-\infty, +\infty]$.

12.1.3. LEMMA If f is Lebesgue measurable on E, then $f \circ T$ is Lebesgue measurable on $T^{-1}(E)$.

12.1.4. THEOREM

$$\int_E f \, dL^n = |\det(M_T)| \int_{T^{-1}(E)} f \circ T \, dL^n$$

in the sense that if one of the two integrals exists then so does the other and the two are equal.

PROOF Let $S \in M_L^n$ -- then

$$T^{-1}(E \cap S) = T^{-1}(E) \cap T^{-1}(S)$$

$\Rightarrow$

$$L^n(E \cap S) = L^n((T \circ T^{-1})(E \cap S))$$

$$= L^n(T(T^{-1}(E \cap S)))$$

$$= L^n(T(T^{-1}(E) \cap T^{-1}(S)))$$

$$= |\det(M_T)| \; L^n(T^{-1}(E) \cap T^{-1}(S)).$$

Take for f the characteristic function χ_S of S, hence

$$\int_E f \, dL^n = \int_E \chi_S \, dL^n$$

$$= L^n(E \cap S)$$

$$= |\det(M_T)| \; L^n(T^{-1}(E) \cap T^{-1}(S))$$

$$= |\det(M_T)| \int_{T^{-1}(E)} \chi_{T^{-1}(S)} \, dL^n$$

$$= |\det(M_T)| \int_{T^{-1}(E)} \chi_S \circ T \, dL^n$$

$$= |\det(M_T)| \int_{T^{-1}(E)} f \circ T \, dL^n.$$

One can then proceed from here to a nonnegative simple function on E and then to a nonnegative extended real valued Lebesgue measurable function on E and finally to the general case (write $f = f^+ - f^-$ and work separately with f^+ and f^-).

[Note: By way of a justification, monotone convergence is used when coupled with the fact that there exists an increasing sequence $\{f_j\}$ of nonnegative simple functions such that $f_j \uparrow f$.]

12.1.5. REMARK Matters can be restated, viz.

$$\int_{T(E)} f \, dL^n = |\det(M_T)| \int_E f \circ T \, dL^n,$$

the underlying supposition being that in this context, $f:T(E) \rightarrow [-\infty, +\infty]$ is Lebesgue measurable.

Assume: $k,n \in \mathbb{N}$, $1 \leq k \leq n$.

12.1.6. DEFINITION Suppose that $T:R^k \rightarrow R^n$ is a linear transformation -- then the adjoint of T is the linear transformation $T^*:R^n \rightarrow R^k$ characterized by the condition

$$< x, T^*y > = < Tx, y >$$

for all $x \in R^k$ and for all $y \in R^n$.

[Note: In terms of matrices,

$$M_{T^*} = M_T^T,$$

the transpose of M_T.]

12.1.7. NOTATION Given a linear transformation $T:R^k \rightarrow R^n$, put

$$J(T) = \sqrt{\det(T^*T)}.$$

[Note: $J(T)$ is nonzero iff T is nonsingular.]

12.1.8. N.B. If $k = n$, then

$$\begin{aligned}\det(T^*T) &= \det(M_{T^*T}) \\ &= \det(M_{T^*}M_T) \\ &= \det(M_T^T M_T) \\ &= \det(M_T^T)\det(M_T) \\ &= \det(M_T)\det(M_T)\end{aligned}$$

$$= \det(M_T)^2$$

=>

$$J(T) = \sqrt{\det(T^*T)}$$

$$= \sqrt{\det(M_T)^2} = |\det(M_T)| = |\det(T)|.$$

12.1.9. DEFINITION

- A linear map $U:R^k \to R^n$ is said to be <u>orthogonal</u> if $< Ux,Uy > = < x,y >$ for all $x,y \in R^k$.

- A linear map $S:R^k \to R^k$ is said to be <u>symmetric</u> if $< x,Sy > = < Sx,y >$ for all $x,y \in R^k$.

12.1.10. POLAR DECOMPOSITION Let $T:R^k \to R^n$ be an injective linear transformation -- then there exists a symmetric map $S:R^k \to R^k$ and an orthogonal map $U:R^k \to R^n$ such that $T = US$.

12.1.11. THEOREM If $T:R^k \to R^n$ is a injective linear transformation and if $E \in M_L^k$, then $T(E) \in M(\mathcal{H}^k)$ and

$$\mathcal{H}^k(T(E)) = J(T)L^k(E).$$

PROOF To establish the purported equality, consider first the case when $k = n$, thus

$$\mathcal{H}^n(T(E)) = L^n(T(E))$$

$$= |\det(M_T)|L^n(E)$$

$$= |\det(T)|L^n(E)$$

$$= J(T)L^n(E).$$

Supposing now that $k < n$, write

$$\begin{aligned} \mathcal{H}^k(T(E)) &= \mathcal{H}^k(US(E)) \\ &= \mathcal{H}^k(S(E)), \end{aligned}$$

U being an isometry. But

$$\begin{aligned} \mathcal{H}^k(S(E)) &= L^k(S(E)) \\ &= |\det(S)|L^k(E). \end{aligned}$$

And

$$\begin{aligned} T^*T &= S^*U^*US \\ &= S^*S \quad (U^*U = \mathrm{id}) \\ &= S^2 \quad (S^* = S) \end{aligned}$$

=>

$$\det(T^*T) = \det(S)^2$$

=>

$$J(T) = \sqrt{\det(T^*T)} = \sqrt{\det(S)^2} = |\det(S)|.$$

12.1.12. REMARK If $T:R^k \to R^n$ is Lipschitz, then (cf. 12.3.1)

$$E \in M_L^k \Rightarrow T(E) \in M(\mathcal{H}^k).$$

[Note:

$$T \text{ linear} \Rightarrow T \text{ Lipschitz}.$$

Proof:

$$||Tx - Ty|| \le ||T||\ ||x - y|| \quad (x,y \in R^k).]$$

12.1.13. SCHOLIUM

$$\int_E J(T)\, dL^k = \int_{T(E)} \mathcal{H}^0(E \cap T^{-1}(y))\, d\mathcal{H}^k(y).$$

To repeat: $k,n \in N$, $1 \le k \le n$.

12.1.14. NOTATION

$$\Lambda_{k,n} = \{\lambda \in N^k : 1 \le \lambda_1 < \cdots < \lambda_k \le n\}.$$

The matrix M_T associated with T is $n \times k$. Given $\lambda \in N^k$, let M_T^λ be the $k \times k$ submatrix of M_T made up of the rows $\lambda_1, \ldots, \lambda_k$ of M_T.

12.1.15. CAUCHY-BINET FORMULA

$$J(T)^2 = \sum_{\alpha \in \Lambda_{k,n}} (\det(M_T^\lambda))^2.$$

Therefore $J(T)$ is the square root of the sum of the squares of the $k \times k$ subdeterminants of $\det(M_T)$.

12.1.16. EXAMPLE Suppose that $k = 2$, $n = 3$, and

$$M_T = \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}.$$

Put

$$u = \begin{bmatrix} a \\ c \\ e \end{bmatrix}, \quad v = \begin{bmatrix} b \\ d \\ f \end{bmatrix}$$

and set

$$\begin{cases} E = ||u||^2 = a^2 + c^2 + e^2 \\ F = \langle u,v \rangle = ab + cd + ef \\ G = ||v||^2 = b^2 + d^2 + f^2. \end{cases}$$

Then

$$\det(M_T^* M_T) = \det \begin{bmatrix} E & F \\ F & G \end{bmatrix} = EG - F^2.$$

On the other hand,

$$\Lambda_{2,3} = \{(1,2),\ (2,3),\ (1,3)\},$$

so

$$M_T^{(1,2)} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv A$$

$$M_T^{(2,3)} = \begin{bmatrix} c & d \\ e & f \end{bmatrix} \equiv B$$

$$M_T^{(1,3)} = \begin{bmatrix} a & b \\ e & f \end{bmatrix} \equiv C,$$

and by Cauchy-Binet,

$$\det(M_T^* M_T) = \det(A)^2 + \det(B)^2 + \det(C)^2$$

$$= (ad - bc)^2 + (cf - ed)^2 + (af - be)^2.$$

Consider now $u \times v$, the vector cross product of u and v:

$$\begin{bmatrix} cf - ed \\ eb - af \\ ad - cb \end{bmatrix}$$

=>

$$||u \times v||^2 = \det(M_T^* M_T)$$

=>

$$||u \times v|| = J(T).$$

§12.2. THE C' CASE

It was shown in the previous § that if $T:R^k \to R^n$ $(k \le n)$ is an injective linear transformation, then

$$\begin{cases} E \in M_L^k \Rightarrow T(E) \in M(H^k) \\ H^k(T(E)) = J(T)L^k(E). \end{cases}$$

This conclusion can be generalized:

(1) Replace R^k by a nonempty open subset $\Omega \subset R^k$.

(2) Replace T by a one-to-one function $\Phi:\Omega \to R^n$ of class C'.

After a fair amount of effort, matters then will read

$$\begin{cases} E \in M_L^k \ (E \subset \Omega) \Rightarrow \Phi(E) \in M(H^k) \\ H^k(\Phi(E)) = \int_E J(\Phi)\, dL^k. \end{cases}$$

Setting aside the proof until later, we shall first deal with some preliminaries and consider some examples.

12.2.1. LEMMA Let $T:R^k \to R^n$ $(k \le n)$ be a linear transformation -- then

$$\text{rank } T \le k, \ \ker T = \ker T^*T,$$

and the following are equivalent:

(a) $J(T) \equiv \sqrt{\det(T^*T)} = 0$,

(b) $\ker T \neq \{0\}$,

(c) $\text{rank } T < k$.

[Note: If T is injective, then $\forall\, x \in R^k$,

$$||x|| = ||T^{-1}Tx||$$

$$\leq ||T^{-1}|| \ ||Tx||$$

=>

$$||x|| \frac{1}{||T^{-1}||} \leq ||Tx||.]$$

12.2.2. NOTATION Given $x_0 \in \Omega$, put

$$J(\Phi)(x_0) = J(d\Phi(x_0))$$

$$= \sqrt{\det(d\Phi(x_0)^* \, d\Phi(x_0))} \ ,$$

from which a function $J(\Phi):\Omega \to R$.

[Note: $J(\Phi)(x_0) \neq 0$ iff $d\Phi(x_0)$ is injective.]

12.2.3. RAPPEL If $\Phi:\Omega \to R^n$ is differentiable at a point $x_0 \in \Omega$, then the $n \times k$ matrix

$$D\Phi(x_0) = \nabla\Phi(x_0) = \begin{bmatrix} \nabla\Phi_1(x_0) \\ \cdot \\ \cdot \\ \cdot \\ \nabla\Phi_n(x_0) \end{bmatrix}$$

is the <u>Jacobian matrix</u> of Φ at x_0.

[Note: $D\Phi(x_0)$ is the matrix that represents $d\Phi(x_0)$.]

12.2.4. <u>N.B.</u> Suppose that $k = n$ -- then

$$\det(Df(x_0) = J_\Phi(x_0).$$

What follows are some particular cases of the relation

$$\mathcal{H}^k(\Phi(E)) = \int_E J(\Phi) \, dL^k.$$

12.2.5. EXAMPLE Take $k = 1$, $n > 1$, take $\Omega =]a,b[$, so $\Phi:]a,b[\rightarrow R^n$ is a curve:

$$\Phi(x) = (\Phi^1(x),\ldots,\Phi^n(x)) \quad (a < x < b).$$

And $D\Phi(x)$ is an $n \times 1$ matrix or still, upon switching the column vector to a row vector, $D\Phi(x)$ becomes a $1 \times n$ matrix, viz.

$$\left(\frac{d\Phi^1}{dx},\ldots,\frac{d\Phi^n}{dx}\right),$$

thus

$$J(\Phi) = \sqrt{\left(\frac{d\Phi^1}{dx}\right)^2 + \cdots + \left(\frac{d\Phi^n}{dx}\right)^2} = ||\dot{\Phi}||.$$

If therefore Φ is one-to-one, then

$$\mathcal{H}^1(\Phi(]a,b[)) = \int_a^b ||\dot{\Phi}||\, dt.$$

E.g.: Let

$$\Phi(x) = (\cos x, \sin x, x) \quad (0 < x < 1),$$

hence

$$\dot{\Phi}(x) = (-\sin x, \cos x, 1) \Rightarrow ||\dot{\Phi}|| = \sqrt{2},$$

hence

$$\mathcal{H}^1(\Phi(]0,1[)) = \int_0^1 ||\dot{\Phi}||\, dt = \sqrt{2}.$$

12.2.6. EXAMPLE The <u>graph</u> of a C' function $f:\Omega \rightarrow R$ $(\Omega \subset R^k)$ is the subset of R^{k+1} defined by

$$Gr_f = \{(x,f(x)) \in R^k \times R : x \in \Omega\},$$

i.e., Gr_f is the image of the injective map $\Phi(x) = (x,f(x))$ from Ω to R^{k+1}. Here

$$D\Phi(x) = \begin{bmatrix} \boxed{\text{id}} \\ \\ Df(x) \end{bmatrix},$$

thus by Cauchy-Binet,

$$J(D\Phi(x)) = \sqrt{1 + \left(\frac{\partial f}{\partial x_1}\right)^2 + \cdots + \left(\frac{\partial f}{\partial x_k}\right)^2}$$

or still,

$$J(D\Phi(x)) = \sqrt{1 + ||Df(x)||^2}$$

=>

$$\begin{aligned} H^k(Gr_f) &= H^k(\Phi(\Omega)) \\ &= \int_\Omega \sqrt{1 + ||Df||^2}\ dL^k. \end{aligned}$$

12.2.7. EXAMPLE Let $\Phi:\Omega \to R^{k+1}$ $(\Omega \subset R^k)$ be a one-to-one map of class C' -- then the Jacobian matrix of Φ has $k + 1$ rows and k columns and its $k \times k$ submatrices can be indexed by the missing row. If

$$\frac{\partial(\Phi_1,\ldots,\Phi_{i-1},\Phi_{i+1},\ldots,\Phi_k)}{\partial(x_1,\ldots,x_k)}$$

denotes the determinant of the submatrix obtained by removing the i^{th} row, it therefore follows that

$$H^k(\Phi(\Omega)) = \int_\Omega \left(\sum_{i=1}^{k} \left(\frac{\partial(\Phi_1,\ldots,\Phi_{i-1},\Phi_{i+1},\ldots\Phi_k)}{\partial(x_1,\ldots,x_k)}\right)^2\right)^{1/2} dL^k.$$

12.2.8. SCHOLIUM Take $k = n$ and let $\Phi:\Omega \to R^n$ be a one-to-one function of class C' -- then $\forall\ E \in M_L^n$ $(E \subset \Omega)$,

$$L^n(\Phi(E)) = \int_E J(\Phi)\ dL^n$$

$$= \int_\Omega \chi_E J(\Phi)\ dL^n$$

from which

$$\int_{\Phi(\Omega)} \chi_E\ dL^n = \int_\Omega (\chi_E \circ \Phi) J(\Phi)\ dL^n.$$

[The first point is a special case of the general theory and the second point follows from the first. To see this, assume to begin with that E is Borel, hence that $\Phi^{-1}(E) = \{x \in \Omega : \Phi(x) \in E\}$ is Borel (Φ being continuous), so

$$\int_{\Phi(\Omega)} \chi_E\ dL^n = L^n(E \cap \Phi(\Omega))$$

$$= L^n(\Phi(\Phi^{-1}(E) \cap \Omega))$$

$$= \int_\Omega \chi_{\Phi^{-1}(E)}\ J(\Phi)\ dL^n$$

$$= \int_\Omega (\chi_E \circ \Phi) J(\Phi)\ dL^n.$$

To proceed in general, let $E \in M_L^n$ $(E \subset \Omega)$ and write $E = F \cup N$, where $F \cap N = \emptyset$, F is an F_σ-set, and N is a subset of a G_δ-set G with $L^n(G) = 0$. Since F and G are Borel,

$$\int_{\Phi(\Omega)} \chi_F\ dL^n = \int_\Omega (\chi_F \circ \Phi) J(\Phi)\ dL^n$$

and

$$0 = \int_{\Phi(\Omega)} \chi_G\ dL^n = \int_\Omega (\chi_G \circ \Phi) J(\Phi)\ dL^n.$$

From here, it remains only to incorporate N... .]

12.2.9. THEOREM If $\Phi:\Omega \to R^n$ is one-to-one and if $f:\Phi(\Omega) \to [-\infty, +\infty]$ is Lebesgue measurable, then

$$\int_{\Phi(\Omega)} f\, dL^n = \int_{\Omega} (f \circ \Phi) J(\Phi)\, dL^n.$$

[This is true when $f = \chi_E$ and the general case follows by a standard approximation argument.]

12.2.10. <u>N.B.</u> The relation

$$L^n(\Phi(E)) = \int_E J(\Phi)\, dL^n$$

is an instance of a so-called "area formula".

12.2.11. EXAMPLE Work in R^2, take

$$\Omega =]0, +\infty[\times]-\pi, \pi[,$$

and for $(r,\theta) \in \Omega$, define $\Phi:\Omega \to R^2$ by the rule

$$\begin{cases} \Phi^1(r,\theta) = r\cos\theta = x \\ \Phi^2(r,\theta) = r\sin\theta = y. \end{cases}$$

Then Φ is one-to-one, of class C', and its range $\Phi(\Omega)$ is $R^2 \backslash \Lambda$, where

$$\Lambda =]-\infty, 0] \times \{0\} \subset R \times R \quad (\Rightarrow L^2(\Lambda) = 0).$$

The Jacobian matrix $D\Phi(r,\theta)$ is given by

$$\begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}$$

and the Jacobian $J_\Phi\,(r,\theta)$, i.e., $\det(D\Phi(r,\theta))$, equals r. So formally

$$\int_{R^2} f(x,y)\ dL^n(x,y)$$

$$= \int_{R^2\setminus\Lambda} f(x,y)\ dL^n(x,y)$$

$$= \int_{\Phi(\Omega)} f\ dL^n$$

$$= \int_{\Omega} (f \circ \Phi)\left|\det J_\Phi\right|\ dL^n$$

$$= \int_{]0,+\infty[\times]-\pi,\pi[} f(r\cos\theta,\ r\sin\theta)r\ dL^2(r,\theta).$$

§12.3. PROOF

12.3.1. SUBLEMMA If $T:R^k \to R^n$ $(1 \le k \le n)$ is Lipschitz continuous and if $E \subset R^k$ is Lebesgue measurable, then $T(E)$ is H^k-measurable.

PROOF It can be assumed that E is bounded. Accordingly let $\{K_j\}$ be a sequence of compact sets such that $K_j \subset E$, $K_j \subset K_{j+1}$, and

$$L^k(E) = L^k(\bigcup_j K_j) \quad (\Rightarrow L^k(E \backslash \bigcup_j K_j)) = 0.$$

Since T is continuous, it follows that

$$T(\bigcup_j K_j) = \bigcup_j T(K_j)$$

is Borel (being a countable union of compacta), thus is H^k-measurable. Now write

$$T(E) = T(\bigcup_j K_j) \cup T(E \backslash \bigcup_j K_j).$$

Then

$$\begin{aligned} &H^k(T(E \backslash \bigcup_j K_j)) \\ &\le \mathrm{Lip}(T)^k\, H^k(E \backslash \bigcup_j K_j) \\ &= \mathrm{Lip}(T)^k\, L^k(E \backslash \bigcup_j K_j) \\ &= \mathrm{Lip}(T)^k \cdot 0 = 0. \end{aligned}$$

Therefore $T(E)$ is the union of a Borel set and a set of zero H^k-measure, so $T(E)$ is H^k-measurable.

[Note: There are various easy variations on this theme.]

12.3.2. DATA $1 \leq k \leq n$, $\Omega \subset R^k$ a nonempty open set, $\Phi:\Omega \to R^n$ a one-to-one function of class C', $E \subset \Omega$ a Lebesgue measurable set.

12.3.3. THEOREM $\Phi(E)$ is $\mathcal{H}^k$-measurable and

$$\mathcal{H}^k(\Phi(E)) = \int_E J(\Phi)\, dL^k \quad \text{(area formula)}.$$

12.3.4. EXAMPLE Suppose that $G \subset \Phi(\Omega)$ is Borel, hence $\mathcal{H}^k$-measurable -- then

$$\Phi(\Phi^{-1}(G)) = G \cap \Phi(\Omega) = G,$$

so

$$\begin{aligned}\mathcal{H}^k(G) &= \mathcal{H}^k(\Phi(\Phi^{-1}(G))) \quad (E = \Phi^{-1}(G) \subset \Omega)\\ &= \int_{\Phi^{-1}(G)} J(\Phi)\, dL^k.\end{aligned}$$

LEMMA A Let $x_0 \in \Omega$ and assume that $d\Phi(x_0) \in \text{Hom}(R^k,R^n)$ is injective -- then $\forall\, \varepsilon > 0$ (< 1) there exists a neighborhood $U \subset \Omega$ of x_0 such that for all $x', x'' \in U$,

$$\begin{aligned}(1 - \varepsilon)\, &||d\Phi(x_0)x' - d\Phi(x_0)x''||\\ &\leq ||\Phi(x') - \Phi(x'')||\\ &\leq (1 + \varepsilon)\, ||d\Phi(x_0)x' - d\Phi(x_0)x''||.\end{aligned}$$

PROOF Fix $\varepsilon > 0$ (< 1) and choose $C > 0$:

$$||d\Phi(x_0)x|| \geq C||x|| \quad (x \in R^k).$$

Since Φ is class C', there exists $\delta > 0$ such that

$$||x - x_0|| < \delta \Rightarrow ||d\Phi(x) - d\Phi(x_0)|| \leq C\varepsilon.$$

So, for $x', x'' \in U \equiv B(x_0,\delta)^\circ$,

$$||\Phi(x') - \Phi(x'') - d\Phi(x_0)\ (x' - x'')||$$

$$= ||\int_0^1 \frac{d}{dt}\ (\Phi(x'' + t(x' - x'') - d\Phi(x_0)(x'' + t(x' - x''))) \ dt||$$

$$= ||\int_0^1 [d\Phi(x'' + t(x' - x'')) - d\Phi(x_0)]\ (x' - x'')\ dt||$$

$$\leq C\varepsilon||x' - x''|| \leq \varepsilon||d\Phi(x_0)(x' - x'')||.$$

Therefore

- $||\Phi(x') - \Phi(x'')||$

$$\leq ||d\Phi(x_0)(x' - x'')|| + ||\Phi(x') - \Phi(x'') - d\Phi(x_0)(x' - x'')||$$

$$\leq ||d\Phi(x_0)x' - d\Phi(x_0)x''|| + \varepsilon||d\Phi(x_0)x' - d\Phi(x_0)x''||$$

$$\leq (1 + \varepsilon)||d\Phi(x_0)x' - d\Phi(x_0)x''||.$$

- $||\Phi(x') - \Phi(x'')||$

$$\geq ||d\Phi(x_0)(x' - x'')|| - ||\Phi(x') - \Phi(x'') - d\Phi(x_0)(x' - x'')||$$

$$\geq ||d\Phi(x_0)x' - d\Phi(x_0)x''|| - \varepsilon||d\Phi(x_0)x' - d\Phi(x_0)x''||$$

$$\geq (1 - \varepsilon)||d\Phi(x_0)x' - d\Phi(x_0)x''||.$$

LEMMA B Let $x_0 \in \Omega$ and assume that $d\Phi(x_0) \in \text{Hom}(R^k,R^n)$ is injective — then $\forall\ \varepsilon > 0\ (< 1)$ there exists a neighborhood $U \subset \Omega$ of x_0 such that for each Lebesgue measurable set $E \subset U$, $\Phi(E)$ is $\mathcal{H}^k$-measurable and

$$(1 - \varepsilon)^{k+1} \int_E J(\Phi)\ dL^k$$

$$\leq \mathcal{H}^k(\Phi(E))$$

$$\leq (1+\varepsilon)^{k+1} \int_E J(\Phi)\, dL^k.$$

PROOF Since the linear transformation $d\Phi(x_0): R^k \to R^n$ is injective,

$$d\Phi(x_0)^{-1}: d\Phi(x_0)R^k \to R^k.$$

Given $\varepsilon > 0$ (< 1), choose $\delta > 0$ so small that the conclusion of LEMMA A holds, where as there $U \equiv B(x_0, \delta)^{\circ}$ and in addition

$$(1+\varepsilon)^{-1} J(\Phi)(x) \leq J(\Phi)(x_0) \leq (1+\varepsilon) J(\Phi)(x),$$

Φ being of class C' ($||x - x_0|| < \delta$). In the relation

$$||\Phi(x') - \Phi(x'')|| \leq (1+\varepsilon)||d\Phi(x_0)x' - d\Phi(x_0)x''||,$$

take

$$\left[\begin{array}{l} x' = d\Phi(x_0)^{-1}y' \\ \\ x'' = d\Phi(x_0)^{-1}y'' \end{array}\right. \qquad (y', y'' \in d\Phi(x_0)(U)).$$

Then

$$||(\Phi \circ d\Phi(x_0)^{-1})(y') - (\Phi \circ d\Phi(x_0)^{-1})(y'')||$$

$$= ||\Phi(x') - \Phi(x'')||$$

$$\leq (1+\varepsilon)||\, d\Phi(x_0)x' - d\Phi(x_0)x''||$$

$$= (1+\varepsilon)||y' - y''||.$$

Therefore

$$\Phi \circ d\Phi(x_0)^{-1}: d\Phi(x_0)(U) \to R^n$$

is Lipschitz continuous with

$$\mathrm{Lip}(\Phi \circ d\Phi(x_0)^{-1}) \leq 1 + \varepsilon.$$

Consequently,

$$\begin{aligned}
&\mathcal{H}^k(\Phi(E)) \\
&= \mathcal{H}^k((\Phi \circ d\Phi(x_0)^{-1})\ (d\Phi(x_0)(E)))) \\
&\leq (1 + \varepsilon)^k\ \mathcal{H}^k(d\Phi(x_0)(E)) \qquad (\text{cf. } 2.3.10) \\
&= (1 + \varepsilon)^k\ J(d\Phi(x_0))\mathcal{L}^k(E) \qquad (\text{cf. } 12.1.11) \\
&= (1 + \varepsilon)^k\ J(\Phi)(x_0)\mathcal{L}^k(E) \qquad (\text{cf. } 12.2.2) \\
&= (1 + \varepsilon)^k\ J(\Phi)(x_0)\ \int_E 1\ d\mathcal{L}^k \\
&= (1 + \varepsilon)^k \int_E J(\Phi)(x_0)\ d\mathcal{L}^k \\
&\leq (1 + \varepsilon)^{k+1} \int_E J(\Phi)(x)\ d\mathcal{L}^k\ \ (||x - x_0|| < \delta),
\end{aligned}$$

the sought for estimate from above. To arrive at the estimate from below, in the relation

$$(1 - \varepsilon)||d\Phi(x_0)x' - d\Phi(x_0)x''|| \leq ||\Phi(x') - \Phi(x'')||$$

take

$$\left[\begin{array}{l} x' = \Phi^{-1}(y') \\ \\ x'' = \Phi^{-1}(y'') \end{array}\right. \qquad (y',y'' \in \Phi(U))$$

to get

$$||(d\Phi(x_0) \circ \Phi^{-1})(y') - (d\Phi(x_0) \circ \Phi^{-1})(y'')||$$

$$\leq (1 - \varepsilon)^{-1} ||\Phi(x') - \Phi(x")||$$

$$= (1 - \varepsilon)^{-1} ||y' - y"||.$$

Therefore

$$d\Phi(x_0) \circ \Phi^{-1} : \Phi(U) \to R^n$$

is Lipschitz continuous with

$$\mathrm{Lip}(d\Phi(x_0) \circ \Phi^{-1}) \leq (1 - \varepsilon)^{-1}.$$

Now manipulate as before:

$$(1 + \varepsilon)^{-1} \int_E J(\Phi)(x) \; dL^k$$

$$\leq \int_E J(d\Phi(x_0)) \; dL^k$$

$$= H^k(d\Phi(x_0)(E))$$

$$= H^k((d\Phi(x_0) \circ \Phi^{-1})(\Phi(E)))$$

$$\leq (1 - \varepsilon)^{-k} H^k(\Phi(E))$$

=>

$$(1 - \varepsilon)^k (1 + \varepsilon)^{-1} \int_E J(\Phi)(x) \; dL^k \leq H^k(\Phi(E))$$

=>

$$(1 - \varepsilon)^{k+1} \int_E J(\Phi)(x) \; dL^k \leq H^k(\Phi(E)) \; ((1 + \varepsilon)^{-1} \geq 1 - \varepsilon).$$

12.3.5. N.B. $E \subset U$ is Lebesgue measurable and the claim is that $\Phi(E)$ is H^k-measurable.

[To see this, let $T = \Phi \circ d\Phi(x_0)^{-1}$, thus by construction $T: d\Phi(x_0)(U) \to R^n$ is

Lipschitz continuous. And

$$\Phi(E) = T(df(x_0)(E)).$$

But $df(x_0)$ is Lipschitz continuous, so $df(x_0)(E)$ is $\mathcal{H}^k$-measurable, thus the same is true of $T(df(x_0)(E)$.]

LEMMA C Suppose that $\forall\ x \in \Omega$, $d\Phi(x)$ is injective -- then

$$\mathcal{H}^k(\Phi(E)) = \int_E J(\Phi)\ dL^k.$$

PROOF Fix $\varepsilon > 0$ (< 1) and cover Ω with countably many $U_i \subset \Omega$ such that for any $E \subset U_i$, LEMMA B is in force. Given now an $E \subset \Omega$, define inductively

$$E_1 = E \cap U_1, \ldots, E_i = (E \ \cap U_i) \backslash \bigcup_{j=1}^{i-1} B_j.$$

Then the E_i are pairwise disjoint and

$$E = \bigcup_{i=1}^{\infty} E_i.$$

Proceeding, apply LEMMA B to each E_i, thus

$$\begin{aligned}(1 - \varepsilon)^{k+1} \int_{E_i} J(\Phi)\ dL^k &\\ &\leq \mathcal{H}^k(\Phi(E_i)) \\ &\leq (1 + \varepsilon)^{k+1} \int_{E_i} J(\Phi)\ dL^k\end{aligned}$$

or still, upon summing over i and bearing in mind that Φ is one-to-one,

$$\begin{aligned}(1 - \varepsilon)^{k+1} \int_E J(\Phi)\ dL^k &\\ &\leq \mathcal{H}^k(\Phi(E))\end{aligned}$$

$$\leq (1+\varepsilon)^{k+1} \int_E J(\Phi)\, dL^k.$$

Finish by sending ε to 0.

12.3.6. N.B. $\forall$ i, $\Phi(E_i)$ is H^k-measurable, thus the same is true of $\Phi(E)$.

LEMMA D Let $\Sigma \subset \Omega$ be the set of $x \in \Omega$ with the property that $d\Phi(x)$ is not injective, hence that $J(\Phi)(x) = 0$ -- then

$$H^k(\Phi(\Sigma)) = 0.$$

PROOF Since the matter is local, it can be assumed that Ω is bounded and that $d\Phi$ is bounded in Ω, say $||d\Phi(x)|| \leq M$ for all $x \in \Omega$. Given $\varepsilon > 0$, consider the function

$$\Phi_\varepsilon : \Omega \to R^n \times R^k$$

defined by the rule

$$\Phi_\varepsilon(x) = (\Phi(x), \varepsilon x) \quad (x \in \Omega),$$

so

$$\Phi = \Pi \circ \Phi_\varepsilon,$$

where

$$\Pi : R^n \times R^k \to R^n$$

is the projection operator given by $\Pi(y,x) = y$, a Lipschitz continuous function with Lipschitz constant 1 (i.e., $\text{Lip}(\Pi) = 1$). Since $\forall$ $x \in \Omega$, $d\Phi_\varepsilon(x)$ is injective, it follows from LEMMA C that

$$\begin{aligned} H^k(\Phi(\Sigma)) &= H^k(\Pi(\Phi_\varepsilon(\Sigma))) \\ &\leq (\text{Lip } \Pi)^k\, H^k(\Phi_\varepsilon(\Sigma)) \\ &= H^k(\Phi_\varepsilon(\Sigma)) \\ &= \int_\Sigma J(\Phi_\varepsilon)\, dL^k. \end{aligned}$$

To estimate this integral, use Cauchy-Binet to produce a constant $C > 0$ such that $\forall\ x \in \Omega$,

$$((J(\Phi_\varepsilon)(x))^2 \leq (J(\Phi)(x))^2 + C^2||d\Phi(x)||^2\ \varepsilon^2.$$

In particular, if $x \in \Sigma$, then

$$\begin{aligned}(J(\Phi_\varepsilon)(x))^2 &\leq C^2||d\Phi(x)||^2\ \varepsilon^2 \\ &\leq C^2M^2\varepsilon^2\end{aligned}$$

or still,

$$J(\Phi_\varepsilon)(x) \leq CM\varepsilon.$$

Therefore

$$\begin{aligned}\mathcal{H}^k(\Phi(\Sigma)) &\leq \int_\Sigma J(\Phi_\varepsilon)\ dL^k \\ &\leq CM\varepsilon L^k(\Sigma).\end{aligned}$$

Now let $\varepsilon \to 0$ to get

$$\mathcal{H}^k(\Phi(\Sigma)) = 0.$$

PROOF OF THEOREM Given a Lebesgue measurable set $E \subset \Omega$, write

$$E = E\backslash\Sigma \cup E \cap \Sigma.$$

Then

$$\Phi(E) = \Phi(E\backslash\Sigma) \cup \Phi(E \cap \Sigma).$$

Owing to LEMMA D, $\Phi(E \cap \Sigma)$ is a set of zero $\mathcal{H}^k$measure. On the other hand, $E\backslash\Sigma \subset \Omega\backslash\Sigma$ (an open set), hence $\Phi(E\backslash\Sigma)$ is $\mathcal{H}^k$-measurable (cf. LEMMA C and 12.3.6). Therefore $\Phi(E)$ is $\mathcal{H}^k$-measurable, And finally

$$\begin{aligned}\mathcal{H}^k(\Phi(E)) &= \mathcal{H}^k(\Phi(E\backslash\Sigma)) \\ &= \int_{E\backslash\Sigma} J(\Phi)\ dL^k \\ &= \int_E J(\Phi)\ dL^k.\end{aligned}$$

The supposition that Φ is one-to-one can be dropped.

12.3.7. THEOREM (AREA FORMULA) If $\Phi:\Omega \to R^n$ is a function of class C' and if $E \subset \Omega$ is a Lebesgue measurable set, then $\Phi(E)$ is $\mathcal{H}^k$-measurable and

$$\int_{R^n} \mathcal{H}^0(E \cap \Phi^{-1}(y))\, d\mathcal{H}^k(y) = \int_E J(\Phi)\, dL^k.$$

[Note: If Φ is one-to-one, then matters reduce to

$$\mathcal{H}^k(\Phi(E)) = \int_E J(\Phi)\, dL^k.]$$

12.3.8. <u>N.B.</u> The arrow

$$y \to \mathcal{H}^0(E \cap \Phi^{-1}(y))$$

$$= \#\{x \in E:\Phi(x) = y\}$$

$$\equiv N(\Phi,E,y)$$

is the multiplicity function of Φ at $y \in R^n$ w.r.t. E and the assignment

$$y \to N(\Phi,E,y)$$

defines an $\mathcal{H}^k$-measurable function (cf. 11.2.4.) (recall that Φ has property (N)).

[Note:

$$\Phi(E) = \{y \in R^n:N(\Phi,E,y) \neq 0\},$$

so the integral over R^n can be replaced by an integral over $\Phi(E)$.]

12.3.9. CHANGE OF VARIABLES Suppose that $\Phi:\Omega \to R^n$ is class C' and $u:\Omega \to [0, +\infty]$ is Lebesgue measurable -- then the assignment

$$y \to \sum_{x\in\Phi^{-1}(y)} u(x)$$

defines an $\mathcal{H}^k$-measurable function and

$$\int_\Omega uJ(\Phi)\, dL^k = \int_{R^n} \Big(\sum_{x\in\Phi^{-1}(y)} u(x)\Big)\, d\mathcal{H}^k(y).$$

[The proof is canonical, given what we know. Thus start with $u = \chi_E$ ($E \subset \Omega$ Lebesgue measurable) and note that

$$\sum_{x\in\Phi^{-1}(y)} \chi_E(x) = \mathcal{H}^0(E \cap \Phi^{-1}(y)).$$

Therefore

$$\begin{aligned}
&\int_\Omega \chi_E\, J(\Phi)\, dL^k \\
&\quad = \int_E J(\Phi)\, dL^k \\
&\quad = \int_{R^n} \mathcal{H}^0(E \cap \Phi^{-1}(y))\, d\mathcal{H}^k(y) \quad \text{(area formula)} \\
&\quad = \int_{R^n} \Big(\sum_{x\in\Phi^{-1}(y)} \chi_E(x)\Big)\, d\mathcal{H}^k(y).
\end{aligned}$$

By linearity, this settles the contention for simple functions, thence... .]

12.3.10. SCHOLIUM Suppose that $\Phi:\Omega \to R^n$ is class C' and $v:R^n \to [0, +\infty]$ is $\mathcal{H}^k$-measurable -- then for every Lebesgue measurable set $E \subset \Omega$,

$$\int_E (v \circ \Phi)J(\Phi)\, dL^k = \int_{R^n} v(y)N(\Phi,E,y)\, d\mathcal{H}^k(y).$$

<u>Step 1:</u> Take $E = \Omega$ and $v = \chi_V$ ($V \subset R^n$), V open -- then

$$\begin{aligned}
&\int_\Omega (\chi_V \circ \Phi)J(\Phi)\, dL^k \\
&\qquad = \int_{\Phi^{-1}(V)} J(\Phi)\, dL^k \\
&\qquad = \int_{R^n} N(\Phi,\Phi^{-1}(V),y)\, d\mathcal{H}^k(y) \quad \text{(area formula)}.
\end{aligned}$$

And

$$\int_{R^n} = \int_{\Phi(\Phi^{-1}(V))} = \int_{V\cap\Phi(\Omega)}$$

=>

$$\int_{R^n} N(\Phi,\Phi^{-1}(V),y)\ dH^k(y)$$

$$= \int_V N(\Phi,\Phi^{-1}(V),y)\ dH^k(y)$$

$$= \int_V N(\Phi,\Omega,y)\ dH^k(y)$$

$$= \int_{R^n} \chi_V(y) N(\Phi,\Omega,y)\ dH^k(y).$$

[Note:

$$\int_{R^n} = \int_{V\cap\Phi(\Omega)}.$$

Meanwhile

$$\int_{R^n} \geq \int_V \geq \int_{V\cap\Phi(\Omega)}$$

$$\Rightarrow \int_{R^n} = \int_V.]$$

<u>Step 2:</u> Take $E \subset \Omega$ compact and v a simple function constant on open sets.

<u>Step 3:</u> Take $E \subset \Omega$ compact and v an arbitrary simple function.

<u>Step 4:</u> Take $E \subset \Omega$ compact and $v \geq 0$ an arbitrary measurable function.

<u>Step 5:</u> Take $E \subset \Omega$ Lebesgue measurable and $v \geq 0$ an arbitrary measurable function.

§12.4. THE DIFFERENTIABLE CASE

The central conclusion of the preceding § is the fact that $\Phi(E)$ is H^k-measurable and

$$H^k(\Phi(E)) = \int_E J(\Phi)\, dL^k.$$

Here $1 \leq k \leq n$, $\Omega \subset R^k$ is a nonempty open set, $\Phi:\Omega \to R^n$ is a one-to-one function of class C', and $E \subset \Omega$ is a Lebesgue measurable set.

It turns out that one can drop the assumption that Φ is class C', it being enough to suppose that Φ is merely differentiable (as well as one-to-one).

12.4.1. WHITNEY APPROXIMATION PRINCIPLE There exists a sequence of disjoint closed sets $F_j \subset \Omega$ and a sequence of C' functions $\Phi_j:R^k \to R^n$ such that in F_j,

$$\Phi = \Phi_j \text{ and } J(\Phi) = J(\Phi_j).$$

Moreover

$$L^k(\Omega\backslash F) = 0,$$

where $F = \bigcup_j F_j$.

12.4.2. LEMMA

$$H^k(\Phi(\Omega\backslash F)) = 0.$$

PROOF Write

$$\Omega\backslash F = \bigcup_{\ell=1}^{\infty} E_\ell,$$

E_ℓ being the set of all points $x \in \Omega\backslash F$ such that

$$\frac{||\Phi(x) - \Phi(x')||}{||x - x'||} \leq \ell$$

for all $x' \in \Omega$ with

$$0 < ||x - x'|| \leq \frac{1}{\ell}.$$

Claim: The restriction of Φ to E_ℓ is locally Lipschitz. For suppose that x, x' belong to a compact $K \subset E_\ell$ and $||x - x'|| \leq \frac{1}{\ell}$ — then

$$||\Phi(x) - \Phi(x')|| \leq \ell \, ||x - x'||$$

by the very definition of E_ℓ. On the other hand, if $||x - x'|| > \frac{1}{\ell}$, then

$$\begin{aligned} ||\Phi(x) - \Phi(x')|| &\leq 2 \max_K ||\Phi|| \\ &\leq 2 \max_K ||\Phi|| \cdot 1 \\ &\leq 2 \max_K ||\Phi|| \ell ||x - x'||. \end{aligned}$$

Hence the claim. Consequently

$$\begin{aligned} H^k(\Phi(K)) &\leq \operatorname{Lip}(\Phi|K;K) H^k(K) \\ &= \operatorname{Lip}(\Phi|K;K) L^k(K), \end{aligned}$$

where

$$\operatorname{Lip}(\Phi|K;K) \leq \ell(1 + 2 \max_K ||\Phi||).$$

But

$$K \subset E_\ell \subset \Omega \backslash F$$

$$\Rightarrow$$

$$L^k(K) \leq L^k(E_\ell) \leq L^k(\Omega \backslash F) = 0.$$

Therefore

$$H^k(\Phi(K)) = 0.$$

Now let $K \uparrow E_\ell$ invade E_ℓ to get

$$\mathcal{H}^k(\Phi(E_\ell)) = 0,$$

so in the end

$$\begin{aligned}\mathcal{H}^k(\Phi(\Omega \backslash F)) &\le \sum_{\ell=1}^{\infty} \mathcal{H}^k(\Phi|E_\ell)) \\ &= 0.\end{aligned}$$

12.4.3. APPLICATION

$$\Omega = F \cup \Omega \backslash F$$

$$\Rightarrow$$

$$\Phi(\Omega) = \Phi(F) \cup \Phi(\Omega \backslash F)$$

$$\Rightarrow$$

$$\Phi(E) = \Phi(E) \cap \Phi(\Omega) = \Phi(E) \cap \Phi(F) \cup \Phi(E) \cap \Phi(\Omega \backslash F)$$

$$\Rightarrow$$

$$\mathcal{H}^k(\Phi(E)) = \mathcal{H}^k(\Phi(E) \cap \Phi(F)) + \mathcal{H}^k(\Phi(E) \cap \Phi(\Omega \backslash F)).$$

And

$$\begin{aligned}\mathcal{H}^k(\Phi(E) \cap \Phi(\Omega \backslash F)) &\le \mathcal{H}^k(\Phi(\Omega \backslash F)) \\ &= 0.\end{aligned}$$

Now compute:

$$\begin{aligned}\mathcal{H}^k(\Phi(E)) &= \mathcal{H}^k(\Phi(E) \cap \Phi(F)) \\ &= \sum_j \mathcal{H}^k(\Phi(E) \cap \Phi(F_j)) \\ &= \sum_j \mathcal{H}^k(\Phi(E) \cap \Phi_j(F_j)) \\ &= \sum_j \int_{\Phi_j^{-1}(\Phi(E)) \cap F_j} J(\Phi_j)\, dL^k\end{aligned}$$

$$= \sum_j \int_{\Phi^{-1}(\Phi(E)) \cap F_j} J(\Phi)\, dL^k$$

$$= \sum_j \int_{E \cap F_j} J(\Phi)\, dL^k$$

$$= \int_{E \cap F} J(\Phi)\, dL^k$$

$$= \int_{E \cap F \cup E \cap \Omega\setminus F} J(\Phi)\, dL^k$$

$$= \int_E J(\Phi)\, dL^k.$$

§12.5. *THE LIPSCHITZ CASE*

12.5.1. DATA $1 \le k \le n$, $\Phi : R^k \to R^n$ a Lipschitz continuous function, $E \subset R^k$ a Lebesgue measurable set.

12.5.2. RAPPEL Owing to Rademacher, $J(\Phi)$ is defined L^k-almost everywhere.

12.5.3. THEOREM (AREA FORMULA) $\Phi(E)$ is $\mathcal{H}^k$-measurable and

$$\int_{R^n} \mathcal{H}^0(E \cap \Phi^{-1}(y))\, d\mathcal{H}^k(y) = \int_E J(\Phi)\, dL^k.$$

12.5.4. N.B. There is an a priori estimate

$$\int_{R^n} \mathcal{H}^0(E \cap \Phi^{-1}(y))\, d\mathcal{H}^k(y) \le (\mathrm{Lip}(\Phi))^k L^k(E).$$

12.5.5. REMARK Φ has property (N), thus the assignment

$$y \to N(\Phi,E,y)$$

defines an $\mathcal{H}^k$-measurable function (cf. 11.2.4).

12.5.6. LEMMA $\forall\, \varepsilon > 0$, there exists a closed set $F_\varepsilon \subset R^k$ and a C' function $\Phi_\varepsilon : R^k \to R^n$ such that in F_ε,

$$\Phi = \Phi_\varepsilon \text{ and } D\Phi = D\Phi_\varepsilon.$$

Moreover

$$L^k(R^k \backslash F_\varepsilon) < \varepsilon.$$

Granted this and bearing in mind the C' version of the area formula, we have

$$\int_E J(\Phi)\, dL^k = \int_{F_\varepsilon \cap E} J(\Phi)\, dL^k + \int_{E \backslash F_\varepsilon} J(\Phi)\, dL^k$$

$$= \int_{F_\varepsilon \cap E} J(\Phi_\varepsilon)\, dL^k + \int_{E \setminus F_\varepsilon} J(\Phi)\, dL^k$$

$$= \int_{R^n} H^0(F_\varepsilon \cap E \cap \Phi_\varepsilon^{-1}(y))\, dH^k(y) + \int_{E \setminus F_\varepsilon} J(\Phi)\, dL_k$$

$$= \int_{R^n} H^0(F_\varepsilon \cap E \cap \Phi^{-1}(y))\, dH^k(y) + \int_{E \setminus F_\varepsilon} J(\Phi)\, dL^k.$$

Now send ε to 0, noting that $L^k(E \setminus F_\varepsilon) \to 0$ (use monotone convergence).

12.5.7. EXAMPLE Given a Lipschitz continuous function $f: R^k \to R$, put

$$\Phi(x) = (x, f(x)) \qquad (x \in R^k).$$

Then $\Phi: R^k \to R^{k+1}$ is Lipschitz continuous, one-to-one, and

$$H^k(\Phi(E)) = \int_E J(\Phi)\, dL^k = \int_E \sqrt{1 + ||Df||^2}\, dL^k.$$

§12.6. THE SOBOLEV CASE

Let Ω be a nonempty open subset of R^n. Given a continuous function $f \in W^{1,p}(\Omega)$ $(1 \leq p < +\infty)$ and a Lebesgue measurable set $E \subset \Omega$, put

$$\mathrm{Gr}_f(E) = \{(x,f(x)):x \in E\} \subset R^{n+1}.$$

12.6.1. THEOREM

$$\mathcal{H}^n(\mathrm{Gr}_f(E)) = \int_E \sqrt{1 + ||\nabla f||^2}\, dL^n.$$

Per 7.1.5., write

$$\Omega = (\bigcup_{k=1}^{\infty} E_k) \cup Z,$$

where the E_k are Lebesgue measurable sets such that $f|E_k$ is Lipschitz and Z has Lebesgue measure 0. Extend $f|E_k$ to a Lipschitz function $f_k:R^n \to R$ -- then $||\nabla f_k|| = ||\nabla f||$ almost everywhere in E_k. Now apply 12.5.7. to get

$$\mathcal{H}^n(\mathrm{Gr}_f(E \cap E_k)) = \int_{E \cap E_k} \sqrt{1 + ||\nabla f||^2}\, dL^n.$$

Put $E_0 = \bigcup_{k=1}^{\infty} E_k$ and sum over k, hence

$$\begin{aligned}\mathcal{H}^n(\mathrm{Gr}_f(E \cap E_0)) &= \int_{E \cap E_0} \sqrt{1 + ||\nabla f||^2}\, dL^n \\ &= \int_E \sqrt{1 + ||\nabla f||^2}\, dL^n.\end{aligned}$$

It remains to pass from

$$\mathcal{H}^n(\mathrm{Gr}_f(E \cap E_0)) \text{ to } \mathcal{H}^n(\mathrm{Gr}_f(E))$$

and for this, it need only be shown that

$$\mathcal{H}^n(\mathrm{Gr}_f(E\backslash E_0)) = 0.$$

12.6.2. LEMMA Let $f \in W^{1,p}(\Omega)$ -- then $\mathcal{H}^n(\mathrm{Gr}_f(S)) = 0$ if $S \subset \Omega$ is a set of Lebesgue measure 0.

[It suffices to make the verification in

$$W^{1,1}_{\ell oc}(\Omega) \quad (\supset W^{1,p}(\Omega)).]$$

§12.7. THE APPROXIMATE CASE

Suppose that $\Phi:\Omega \to R^n$ is approximately differentiable almost everywhere in Ω -- then using approximate partial derivatives, one can form $J_{ap}(\Phi)$.

12.7.1. LEMMA (cf. 12.4.1) There exists an increasing sequence $K_1 \subset K_2 \subset \ldots$ of compact subsets of $\Omega (\subset R^k)$ for which

$$L^k(\Omega \backslash K) = 0 \quad (K = \bigcup_j K_j)$$

and a sequence of C' functions $\Phi_j:R^k \to R^k$ such that in K_j,

$$\Phi = \Phi_j \text{ and } J_{ap}(\Phi) = J(\Phi_j).$$

12.7.2. NOTATION Given $y \in R^n$, let $m^j(y)$ be the cardinality of $\Phi^{-1}(y) \cap K_j$.

12.7.3. LEMMA $m^j(-)$ is Borel measurable and $\forall$ y,

$$m^1(y) \leq m^2(y) \leq \ldots .$$

Put

$$m(y) \equiv \lim_{y \to \infty} m^j(y).$$

Then $m(-)$ is Borel measurable.

12.7.4. THEOREM

$$\int_{R^n} m(y)\, dH^k(y) = \int_\Omega J_{ap}(\Phi)\, dL^k.$$

It suffices to show that

$$\int_{R^n} m(y)\, dH^k(y) \geq \int_\Omega J_{ap}(\Phi)\, dL^k$$

and

$$\int_{R^n} m(y)\ d\mathcal{H}^k(y) \leq \int_\Omega J_{ap}(\Phi)\ dL^k.$$

The second point being the easier of the two, note that

$$\begin{aligned}\int_\Omega J_{ap}(\Phi)\ dL^k & \\ &\geq \int_{K_j} J_{ap}(\Phi)\ dL^k \\ &= \int_{K_j} J(\Phi_j)\ dL^k \\ &= \int_{R^n} N(\Phi_j, K_j, y)\ d\mathcal{H}^k(y) \\ &= \int_{R^n} \mathcal{H}^0(\Phi_j^{-1}(y) \cap K_j)\ d\mathcal{H}^k(y) \\ &= \int_{R^n} m^j(y)\ d\mathcal{H}^k(y) \\ &\xrightarrow[(j \to \infty)]{} \int_{R^n} m(y)\ d\mathcal{H}^k(y).\end{aligned}$$

12.7.5. N.B. Under the supposition that

$$\int_\Omega J_{ap}(\Phi)\ dL^k < +\infty,$$

the "m" is independent of the choice of data, i.e., if

$$\left[\begin{array}{l} m_1(-) \longleftrightarrow \{K_j^1\} \\ \\ m_2(-) \longleftrightarrow \{K_j^2\}, \end{array}\right.$$

then $m_1 = m_2$ $\mathcal{H}^k$ almost everywhere.

[Let

$$m_3(-) \longleftrightarrow \{K_j^1 \cap K_j^2\}.$$

Then

$$m_3(y) \le \left[\begin{array}{l} m_1(y) \\ \\ m_2(y) \end{array}\right. \qquad (y \in R^n).$$

But

$$\int_\Omega J_{ap}(\Phi)\, d\mathcal{L}^k$$

$$= \int_{R^n} m_3(y)\, d\mathcal{H}^k(y)$$

$$= \left[\begin{array}{l} \int_{R^n} m_1(y)\, d\mathcal{H}^k(y) \\ \\ \\ \int_{R^n} m_2(y)\, d\mathcal{H}^k(y) \end{array}\right.$$

=>

$$\left[\begin{array}{l} \int_{R^n} (m_1 - m_3)\, d\mathcal{H}^k = 0 \\ \\ \int_{R^n} (m_2 - m_3)\, d\mathcal{H}^k = 0 \end{array}\right.$$

=>

$$\left[\begin{array}{l} m_1 - m_3 = 0 \\ \\ m_2 - m_3 = 0 \end{array}\right. \qquad \mathcal{H}^k \text{ almost everywhere}$$

=>

$$m_1 = m_2 \ \mathcal{H}^k \text{ almost everywhere.}$$