BULLETIN OF THE
INSTITUTE OF MATHEMATICS
ACADEMIA SINICA
Volume 1, Number 1, June 1973

ON A THEOREM OF WHITNEY

BY

FON-CHE LIU[1]

Abstract. Whitney's theorem, for real-valued measurable functions of n variables, that approximate total differentiability is equivalent to the existence of approximate partial derivatives is generalized to differentiability of higher order.

1. Let S be a bounded measurable subset of R^n with $|S|>0$, where as usual $|A|$ denotes the Lebesgue measure of A in R^n; and let f be a real-valued measurable function defined on S. (All the functions considered in this paper are real-valued.) If for a point x_0 of S, there is a real vector $L=(l_1,\cdots,l_n)$ such that the set $R_\varepsilon =\{x\in S: |f(x)-f(x_0)-L\cdot(x-x_0)|\leq\varepsilon|x-x_0|\}$ has density one at x_0 for every given $\varepsilon>0$, then f is said to be approximately totally differentiable at x_0 (or a.t.d. at x_0). If R_ε has linear density one at x_0 in each direction of the coordinate axes for every $\varepsilon>0$, then f is said to have approximate partial derivatives at x_0 (or a.p.d.'s at x_0). The following theorem of Whitney [7] gives the connection of these concepts with smoothness.

THEOREM 1 (WHITNEY). *The following conditions are equivalent:*

(a) *f has a.p.d.'s at almost all points of S.*

(b) *f is a.t.d. at almost all points of S.*

(c) *Given $\varepsilon>0$, there is a C^1-function g on R^n such that*
$$|\{x\in S: f(x)\neq g(x)\}|<\varepsilon.$$

We note that the equivalence of (a) and (b) is Stepanov's theorem. It is the purpose of this paper to give a generalization of Theorem 1 to situations of higher order, which is motivated by questions of characterization of approximate extension type for functions of Sobolev spaces corresponding to a result of Michael for $\overset{\circ}{W}{}^1_1$ (see [3] or [2]).

Received by the editors March 1, 1973.

([1]) Communicated by T. C. Sun.

An application of the result of this paper to these questions will be given in a forthcoming paper.

Now let S, f, and x_0 be as before. Naturally, the a. p. d.'s of f at x_0 of order greater than one can be defined, if the a. p. d.'s of lower order are defined on a set with linear density one at x_0 in each direction of the coordinate axes. Since there will be no confusion, the usual notation for partial derivatives will be employed to denote approximate partial derivatives.

Function f is said to have unbiased a. p. d.'s up to order α at x_0 if the a. p. d.'s of f of order $\leq \alpha$ are defined at x_0 and if all the mixed a. p. d.'s do not depend on the permutations of the coordinates. In this case any a. p. d. of f at x_0 can be written in the form $\partial^\beta f(x_0) / \partial x_1^{j_1} \cdots \partial x_n^{j_n}$, $\beta = j_1 + \cdots + j_n \leq \alpha$. We shall write $\partial^\beta f(x_0) / \partial x_1^{j_1} \cdots \partial x_n^{j_n}$ as $f_J(x_0)$, put $|J| = \beta$, and denote $j_1! \cdots j_n!$ by $J!$, where $J = (j_1, \cdots, j_n)$. Sometimes J will be written as $(j_i, \tilde{j}_i)$, $i = 1, \cdots, n$. For convenience, if $x = (x_1, \cdots, x_n)$, then x^J will stand for $x_1^{j_1} \cdots x_n^{j_n}$. As is usual, if $x_i \leq y_i$, $i = 1, \cdots, n$, and $x \neq y$, we write $x < y$.

Suppose f has unbiased a.p.d.'s up to order α at x_0. Then we say that f is approximately partially totally differentiable of order α (or a.p.t.d. of order α) at x_0 if for every J with $|J| < \alpha$, and $i = 1, \cdots, n$, the set $\{(x_i^0 + x_t, \tilde{x}_i^0) : |f_J(x_i^0 + x_t, \tilde{x}_i^0) - \sum_{k=0}^{\alpha-|J|} f_{(j_i+k, \tilde{j}_i)}(x_0) x_i^k / k!| \leq \varepsilon |x_i|^{\alpha - |J|}\}$ has linear density one at x_0 for every $\varepsilon > 0$, where $x_0 = (x_1^0, \cdots, x_n^0)$. Of course, when $\alpha = 1$, this is equivalent to the condition that f is a. p. d. at x_0.

A function f will be said to be approximately totally differentiable of order α (or a. t. d. of order α) at x_0 if corresponding to every J with $|J| < \alpha$, a measurable function f^J is defined for almost all points of S in a neighborhood of x_0 and if $f^J(x_0)$ is defined for $|J| \leq \alpha$ in such a way that $f^0 = f$ and the set $\{x \in S : |f^J(x) - \sum_{|J'+J| \leq \alpha} (1/J'!) f^{J'+J}(x_0)(x - x_0)^{J'}| \leq \varepsilon |x - x_0|^{\alpha - |J|}\}$ has density one at x_0 for every $\varepsilon > 0$, and $|J| < \alpha$.

To clarify the meaning of local existence of functions f^J in the definition of a. t. differentiability of order α, we remark that if f is a. t. d. of order α at almost all points of S, then it is not hard to see by induction on α that the f^J's can be so chosen as to be defined on

S and to serve as the local functions in the definition of a. t. differentiability of order α at almost all points of S, and also that each f^J is the same as f_J up to a set of measure zero, and consequently is measurable on S. (See also the equivalence of (a) and (b) in Theorem 2.)

We shall prove the following theorem.

THEOREM 2. *Let S be a bounded measurable subset of R^n with $|S|>0$, and let f be a measurable function defined on S. The following three conditions are equivalent:*

(a) *f is a.p.t.d. of order α at almost all points of S.*

(b) *f is a.t.d. of order α at almost all points of S.*

(c) *Given $\varepsilon>0$, there is a C^α-function g defined on R^n such that*

$$|\{x\in S: f(x)\neq g(x)\}|<\varepsilon.$$

That (c) implies (a) is obvious. The equivalence between (a) and (b) can be proved by elaborating the method of proof for Stepanov's theorem; we omit the proof.

It remains to prove that (b) implies (c).

2. We recall Whitney's extension method, which is essential to our purpose. The references for facts quoted in this section can be found in [6], [1, pp. 225–229], and [5, pp. 170–180].

Let F be a closed subset of R^n, and let us fix a family $\{Q_k\}_{k=1}^\infty$ of closed cubes each of which is disjoint from F, satisfying the following conditions:

(2.1) $\bigcup Q_k = F^c$;

(2.2) there are constants $c_2>c_1>0$ such that $c_1 d(Q_k, F) \leq \operatorname{diam} Q_k \leq c_2 d(Q_k, F)$, where $d(A,B)$ denotes the distance between A and B, and $\operatorname{diam} A$ denotes the diameter of A; and

(2.3) there is a positive integer N such that every point of F^c is in at most N elements of $\{Q_k\}$.

We note that c_1, c_2, and N can be chosen to depend only on n, not on F, but this fact does not concern us in this paper.

Related to the family $\{Q_k\}$ is a family $\{\psi_k\}$ of nonnegative C^∞-functions with the following properties:

(2.4) ψ_k has its support in Q_k, $k=1, 2,\cdots$;

(2.5) for every x of F^c, $\sum \psi_k(x) = 1$ (we recall from (2.3) and (2.4) that $\sum \psi_k(x)$ is actually a finite sum); and

(2.6) there is a constant $M>0$ such that

$$|\psi_{kJ}(x)| = \left| \frac{\partial^{|J|}}{\partial x_1^{j_1}\cdots \partial x_n^{j_n}} \psi_k(x) \right| \le Md(x, F)^{|J|},$$

for $k=1, 2,\cdots$, $x \in F^c$, and $J=(j_1,\cdots, j_n)$.

Now for each k, we choose and fix a point p_k in F such that

$$d(p_k, Q_k) = d(F, Q_k)\,. \tag{2.7}$$

From (2.2) and (2.7), we see (by changing c_1 and c_2 if necessary) that

$$c_1 d(x, F) \le d(p_k, x) \le c_2 d(x, F)\,, \qquad x \in Q_k\,. \tag{2.8}$$

Suppose $\{f^J\}_{0\le|J|\le\beta}$ is a C^β-function defined on F in the sense of Whitney [6], i.e. each f^J is defined on F and for $x\in F$, $\varepsilon>0$ given, there is a $\delta>0$ such that if $y, z\in B_\delta(x)\cap F$ ($B_\delta(x)$ is the ball with center at x and radius δ), the following inequality is true for each J with $|J|\le\beta$,

$$\left| f^J(y) - \sum_{J\le J', |J'|\le\beta} \frac{1}{(J'-J)!} f^{J'}(z)(y-z)^{J'-J} \right| \le \varepsilon |y-z|^{\beta-|J|}.$$

Whitney [6] shows that f^0 can be extended to be a C^β-function h on R^n such that $h_J(x) = f^J(x)$ for $x\in F$ and $|J|=\beta$. Actually, h can be defined by

$$\begin{aligned} h(x) &= f^0(x)\,, \qquad \text{if} \qquad x\in F \\ &= \sum_k \psi_k(x) \sum_{0\le|J|\le\beta} \frac{1}{J!} f^J(p_k)(x-p_k)^J\,. \end{aligned}$$

We shall call this method of extending $\{f^J\}$ from F to R^n the Whitney extension method.

We shall need the following lemma.

LEMMA 1. *Let y be a density point of F which is not an interior point of F. For any $\varepsilon>0$, there is a $\delta>0$ such that if $x\in F^c\cap B_\delta(y)$, then $d(x, p_k)\le\varepsilon|x-y|$ for all k with $x\in Q_k$.*

Lemma 1 follows directly from (2.8) and the definition of density point.

3. Now we are ready to complete the proof of Theorem 2. Suppose (b) holds. We must show (c) holds. We proceed by induction on α. When $\alpha=1$, the desired result is contained in Theorem 1. Suppose now that (b) implies (c) when $1\leq\alpha\leq\beta$, and let us show that (b) implies (c) for $\alpha=\beta+1$. Given $\varepsilon>0$, by the induction hypothesis there is a compact subset Q of S with the properties:

(3.1) $|S-Q|<\varepsilon/3$;

(3.2) when restricted to Q, $\{f_J\}_{0\leq|J|\leq\beta}$ is a C^β-function in the sense of Whitney; and $\{f_{J+J'}\}_{0\leq|J'|\leq 1}$ is a C^1-function on Q, also in the sense of Whitney, for each J with $|J|=\beta$.

Let $\tilde{S}$ be the set of all those points of S at which f is a.t.d. of order $\beta+1$. For $x\in S$, and $i=1,2,\cdots$, let

$$Q(x)=\bigcap_{0\leq|J|\leq\beta}\{y:|f_J(y)-\pi(f_J;x;y)|\leq|y-x|^{\beta+1-|J|}\}$$

and

$$\phi_i(x)=|B_{1/i}(x)\setminus Q(x)|\cdot|B_{1/i}(x)|^{-1},$$

where $\pi(f_J;x;y)$ denotes the Taylor polynomial centered at x,

$$\sum_{0\leq|J'|\leq\beta+1-|J|}\frac{1}{J'!}f_{J+J'}(x)\,(y-x)^{J'}.$$

Then it is a routine matter to check that ϕ_i is a measurable function for each i. As $\lim_{i\to\infty}\phi_i(x)=0$ for $x\in\tilde{S}$, we infer from Egoroff's theorem that there is a compact subset $\tilde{Q}$ of $\tilde{S}$ such that

(3.3) $|\tilde{S}\setminus\tilde{Q}|<\varepsilon/3$; and

(3.4) $\phi_i(x)\to 0$ uniformly on $\tilde{Q}$.

Put $F=\tilde{Q}\cap Q$. Since $\{f_J\}_{0\leq|J|\leq\beta}$ is C^β on F, we can extend f outside F to obtain a C^β-function h on R^n by Whitney's extension method, i.e.

$$\begin{aligned} h(x)&=f(x), && \text{if } x\in F,\\ &=\sum_k\psi_k(x)\,\pi'(f;p_k;x)\equiv\sum_k\psi_k(x)\,\pi'(p_k;x), && \text{if } x\in F^c, \end{aligned}$$

where $\pi'(f;\, p_k;\, x)$ is the polynomial which consists of those terms of $\pi(f;\, p_k;\, x)$ with order $\leq \beta$, and $\{\psi_k\}$, $\{p_k\}$ are related to F as described in §2. We first show that for each J with $|J| = \beta$, the function h_J is totally differentiable almost everywhere. Since h_J is totally differentiable at every point not in F, it is sufficient to show that h_J is totally differentiable at every density point of F. So let y be a density point of F, where by (3.2) we may assume that y is not an interior point of F. We have to show that $h_J(x) - h_J(y) = \nabla f_J(y) \cdot (x - y) + o(|x - y|)$, as $x \to y$. For this purpose, as is obvious in view of (3.2), we need only consider the case $x \to y$ via F^c. For such x,

$$h_J(x) - h_J(y) = \sum_k \sum_{0 \leq J' \leq J} \psi_{kJ'}(x)\, \pi'_{J-J'}(p_k;\, x) - f_J(y) \equiv r_1 + r_2\,,$$

where

$$r_1 = \sum_k \psi_k(x)\, \pi'_J(p_k;\, x) - f_J(y)\,,$$

and

$$r_2 = \sum_k \sum_{0 < J' \leq J} \psi_{kJ'}(x)\, \pi'_{J-J'}(p_k;\, x)\,.$$

First, since $\sum_k \psi_k(x) = 1$, we have

$$\begin{aligned}
r_1 &= \sum_k \psi_k(x)\,[\pi'_J(p_k;\, x) - f_J(y)] \\
&= \sum_k \psi_k(x)\,[f_J(p_k) - f_J(y)] \\
&= \sum_k \psi_k(x)\,[\nabla f_J(y) \cdot (p_k - y) + o(|p_k - y|)] \\
&= \sum_k \psi_k(x)\,[\nabla f_J(y) \cdot (x - y) + \nabla f_J(y) \cdot (p_k - x) + o(|p_k - y|)] \\
&= \sum_k \psi_k(x)\,[\nabla f_J(y) \cdot (x - y)] + O(|p_k - x|) + o(|p_k - y|) \\
&= \nabla f_J(y) \cdot (x - y) + o(|x - y|)\,.
\end{aligned}$$

In the above equalities, the third line follows from (3.2), and the last line follows from the fact $\sum_k \psi_k(x) = 1$ and Lemma 1.

Now we have to show that $r_2 = o(|x - y|)$ as $x \to y$ via F^c. Let i_0 be such that $\phi_i(z) \leq \sigma/N$, for all $z \in \widetilde{Q}$ whenever $i \geq i_0$, where N is as in §2 and $\sigma > 0$ is a number chosen so that for any set A with

$\operatorname{diam} A \leq r$ we have $|\bigcap_{z\in A} B_{2r}(z)| > \sigma\, |B_{2r}(0)|$. If x is sufficiently near to y, then Lemma 1 shows that $\operatorname{diam}\{p_k : \psi_k(x) \neq 0\} \leq 1/2 i_0$. For such x, let i, which depends on x, be the largest integer with $\operatorname{diam}\{p_k : \psi_k(x) \neq 0\} \leq 1/2 i$. With i so chosen, we have

(3.5) $1/i < 8 d(x, F)$;

(3.6) $\phi_i(z) \leq \sigma/N \ \forall\ z \in \tilde{Q}$;

(3.7) $|\bigcap_{k, \psi_k(x)\neq 0} B_{1/i}(p_k)| > \sigma\, |B_{1/i}(0)|$.

It follows from the choice of i and (3.6) that

(3.8) $|\bigcup_{k, \psi_k(x)\neq 0} \{B_{1/i}(p_k) \setminus Q(p_k)\}| \leq \sigma\, B_{1/i}(0)$.

Combining (3.7) and (3.8), we infer that there is p with the properties that $p \in Q(p_k)$ for all k with $\psi_k(x) \neq 0$ and $|p - p_k| \leq 1/i < 8d(x, F)$. Consequently,

(3.9) $|f_{J''}(p) - \pi(f_{J''}; p_k; p)| \leq |p - p_k|^{\beta+1-|J''|} \leq M' d(x, F)^{\beta+1-|J''|}$,

whenever $0 \leq |J''| \leq \beta$ and $\psi_k(x) \neq 0$, where M' is a constant independent of x.

Now let x be in F^c and sufficiently near to y that we can choose i, and then p, such that (3.5) and (3.9) hold. Then

$$
\begin{aligned}
r_2 &= \sum_{0<J'\leq J} \sum_k \psi_{kJ'}(x)\, \pi'_{J-J'}(p_k;\ x) \\
&= \sum_{0<J'\leq J} \sum_k \psi_{kJ'}(x)\, \{\pi'_{J-J'}(p_k;\ x) - \pi'_{J-J'}(p;\ x)\},
\end{aligned} \tag{3.10}
$$

because $\sum_k \psi_{kJ'}(x) = 0$ when $J' > 0$. But since $\pi'_{J-J'}(p_k;\ x) - \pi'_{J-J'}(p;\ x)$ is a polynomial in x of degree $\leq \beta - |J - J'|$, it can be expressed as a Taylor polynomial centered at p as follows,

$$
\begin{aligned}
&\pi'_{J-J'}(p_k;\ x) - \pi'_{J-J'}(p,\ x) \\
&\quad = \sum_{\substack{J-J'\leq J'' \\ |J''|\leq \beta}} \left\{ \frac{1}{(J'' - J + J')!}\, \pi'_{J''}(p_k;\ p) - f_{J''}(p) \right\} \\
&\qquad \cdot (x - p)^{J''-J+J'}.
\end{aligned} \tag{3.11}
$$

From (3.9) and the fact that there is an $M'' > 0$ independent of x such that $|x - p| \leq M'' d(x, F)$ (see the definition of i, (3.5) and (2.2)), it follows that each term of (3.11) is less than a constant multiplied by $d(x, F)^{|J'|+1}$. Keeping this in mind, and observing

(2.6) of §2, we infer from (3.10) that there is a constant C independent of x such that

$$|r_2| \leq cd(x, F),$$

and therefore $r_2 = o(|x-y|)$ as $x \to y$ via F^c.

Thus we have succeeded in proving that h_J is totally differentiable a.e. for each J with $|J| = \beta$. Consequently from a theorem of Whitney (see [7] or [1]), there is a $C^{\beta+1}$-function g on R^n such that $|\{x \in F : g(x) \neq h(x)\}| < \varepsilon/3$. Obviously, $|\{x \in S : f(x) \neq g(x)\}| < \varepsilon$. We have completed the proof of Theorem 2.

REFERENCES

1. H. Federer, *Geometric measure theory*, Springer-Verlag, New York and Berlin, 1969.
2. C. Goffman, *Approximation of non-parametric surfaces of finite area*, J. Math. Mech. **12** (1963), 737–745.
3. J. H. Michael, *The equivalence of two areas for non-parametric discontinuous surfaces*, Illinois J. Math. **7** (1963), 59–78.
4. S. Saks, *The theory of integrals*, Second Revised Edition, Warzzawa-Luov, 1937.
5. E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Univ. Press, Princeton, N. J., 1970.
6. H. Whitney, *Analytic extensions of differentiable functions defined in closed sets*, Trans. Amer. Math. Soc. **36** (1934), 63–89.
7. ———, *On totally differentiable and smooth functions*, Pacific J. Math. **1** (1951), 143–159.

DEPARTMENT OF MATHEMATICS, WAYNE STATE UNIVERSITY, DETROIT, MICHIGAN 48202, U. S. A.

INSTITUTE OF MATHEMATICS, ACADEMIA SINICA, NANKANG, TAIPEI, TAIWAN, CHINA