LUZIN TYPE APPROXIMATION OF FUNCTIONS OF BOUNDED VARIATION

by

Gregory P. Francos

B.S. in Mathematics, Lafayette College, 2005

Submitted to the Graduate Faculty of

the Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

ARTS AND SCIENCES

This dissertation was presented

by

Gregory P. Francos

It was defended on

May 6th, 2011

and approved by

Piotr Hajłasz, Department of Mathematics

Juan Manfredi, Department of Mathematics

Frank Beatrous, Department of Mathematics

Giovanni Leoni, Department of Mathematics, Carnegie Mellon University

Dissertation Director: Piotr Hajłasz, Department of Mathematics

LUZIN TYPE APPROXIMATION OF FUNCTIONS OF BOUNDED VARIATION

Gregory P. Francos, PhD

University of Pittsburgh, 2011

This paper is divided into two sections:

(I) Consider the function space

$$BV^m = \{u \in W^{m-1,1} : D^\alpha u \text{ is a measure for } |\alpha| = m\}.$$

Such functions are called mth order functions of bounded variation. We will show that a given function $u \in BV^m(\mathbb{R}^n)$ possesses the so-called C^m-Luzin property; that is, u coincides with a $C^m(\mathbb{R}^n)$ function outside a set of arbitrarily small Lebesgue measure.

(II) Consider a set of Lebesgue measurable functions $f^\alpha : \mathbb{R}^N \to \mathbb{R}$ indexed by the multi-indices in $\mathbb{R}^N$ of order $|\alpha| = m$. We will prove that for any such collection, there is $g \in C^{m-1}(\mathbb{R}^N)$ which is m times differentiable almost everywhere, and such that

$$D^\alpha g(x) = f^\alpha(x) \text{ a.e. for all } |\alpha| = m.$$

Keywords: Functions of Bounded Variation, Calculus of Variations, Luzin Property, Whitney Extension Theorem, Real Analysis, Distributions .

TABLE OF CONTENTS

PREFACE

I would like to thank my parents and my sister for their support. I would like to thank my advisor, Piotr Hajłasz, for his attention to this work.

1.0 INTRODUCTION

As a starting point for this thesis, recall a well-known result of Luzin:

Let $f : \mathbb{R}^n \to \mathbb{R}$ be Lebesgue measurable. For each $\epsilon > 0$, there exists a closed set C such that $|\mathbb{R}^n \backslash C| < \epsilon$ and $f|_C$ is continuous.

Federer ([8], p. 442) proved that if $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable almost everywhere, then f coincides with a C^1 function outside a set of arbitrarily small measure. The proof utilized a result of Whitney [25], which provided a characterization of functions that are restrictions of C^m functions to closed subsets of $\mathbb{R}^n$. Federer showed that an almost everywhere differentiable functions satisfies the assumptions of Whitney's result for $m = 1$ on closed sets whose complement has arbitrarily small measure.

Following a similar idea, namely a version of Whitney's result for L^p functions, Calderón and Zygmund [6] proved that a function in the Sobolev space $W^{m,p}(\mathbb{R}^n)$ coincides with a C^m function outside an open set of arbitrarily small measure. Recall that $W^{m,p}$ is the space of functions whose weak derivatives of orders less than or equal to m are L^p functions.

More precise variations of Calderón and Zygmund's results for $W^{m,p}$ functions have been the subject of [16], [20], [27], [28], [13], and [24]. For instance, Liu [16] showed that in addition to coinciding with a C^m function outside a set of small measure, one could also estimate the error in terms of the Sobolev norm. Michael and Ziemer [28] showed that the approximation can be made to coincide outside a set of small Bessel capacity. Bojarski, Hajłasz, and Strzelecki [12] fine-tuned these results by replacing the Whitney Extension theorem with Whitney Smoothing to obtain improved results in norm and capacity.

Our aim is to generalize the result to the function space BV^m. These functions' weak derivatives of orders less that m are L^1 functions and the mth order distributional derivatives are Radon measures of finite total variation. Our main result is that a function in BV^m coincides with a C^m function outside a set of arbitrarily small measure.

It is well known that the space BV^m is strictly larger than the space $W^{m,1}$. As an important example, convex functions are BV^2_{loc}, and our result shows that a convex function coincides with a C^2 function outside a set of arbitrarily small measure. Alberti provided a sketch of a proof of this result in [3]. However, our proof is based on a different method. We utilize pointwise estimates for BV^m functions similar to those developed in [12], [13] to show that our function satisfies the conditions of Whitney's result on a large closed set.

$BV(\mathbb{R}^n)$ functions ($m = 1$) are known to possess the C^1-Luzin property [9], i.e. for any $\epsilon > 0$ the function coincides with some $C^1(\mathbb{R}^n)$ function outside a set of measure ϵ. The ideas used in the proof do not generalize to higher order derivatives, so a different technique must be utilized.

In the thesis we also discuss another independent, but related result based on another theorem of Luzin from 1917 [23]:

Let $f : \mathbb{R} \to \mathbb{R}$ be Lebesgue measurable. Then there exists $g \in C(\mathbb{R})$ which is differentiable almost everywhere and for which $g'(x) = f(x)$ almost everywhere.

This result can be looked at two ways. On one hand, Lebesgue measurable functions on $\mathbb{R}$ possess enough structure to be derivatives of continuous functions outside a set of measure zero. On the other hand, one can view this result as saying that classical derivatives have a certain lack of structure.

Recently Moonens and Pfeffer [19] generalized Luzin's result to vector fields on $\mathbb{R}^N$; any measurable vector field is a gradient of an a.e. differentiable continuous function. This is interesting because it allows us to find a.e. solutions to non-integrable systems:

$$\frac{\partial f}{\partial x_i} = p_i, \qquad \frac{\partial p_i}{\partial x_j} \neq \frac{\partial p_j}{\partial x_i}.$$

A construction of this kind was used by Balogh [5] to construct surfaces in the Heisenberg group with large characteristic sets, while the horizontal distribution in the Heisenberg group is not integrable.

Our second main result is the following generalization of Moonens and Pfeffer's result:

Theorem 1. *Let U be open in $\mathbb{R}^n$ and let $f = \{f^\alpha\}_{|\alpha|=m}$ be a Lebesgue measurable function defined on U. Then for any $\sigma > 0$, there is $u \in C^{m-1}(\mathbb{R}^n)$ which is m-times differentiable almost everywhere and such that*

$$D^m u(x) = f(x) \quad \text{for a.e. } x \in U.$$

$$\|D^\gamma u\|_\infty \leq \sigma \qquad \text{for each } |\gamma| < m.$$

2.0 BACKGROUND

2.1 VECTOR-VALUED MEASURES

Our goal is to study functions of bounded variation on $\mathbb{R}^n$. The distributional derivative of a function of bounded variation is a vector-valued measure. Therefore, we begin with some basic results on this topic. The results here generally follow from those in ([22], Chapter 6) involving complex measures. We denote by $\mathcal{B}$ the Borel σ-algebra on $\mathbb{R}^n$. A vector-valued Radon measure on $\mathbb{R}^n$ is any set function $\mu : \mathcal{B} \to \mathbb{R}^m$ whose components are signed Radon measures, i.e. for each $i = 1, ..., m$, μ^i is countably additive, $\mu^i(\emptyset) = 0$, and for any compact subset $K \subset \mathbb{R}^n$, $|\mu^i|(K)$ is finite.

The total variation of the measure μ defined by

$$|\mu|(E) := \sup \left\{ \sum_{i=0}^{\infty} |\mu(E_i)| : E_i \in \mathcal{B} \text{ are pairwise disjoint}, E = \bigcup_{i=0}^{\infty} E_i \right\}$$

is a positive Radon measure on $\mathcal{B}$. One can easily show that $|\mu|(E)$ is finite if and only if $|\mu^i|(E) < \infty$ for each $i = 1, \ldots, n$. If μ is a signed measure and $f \in L^1_{\text{loc}}(\mu)$, we will denote by $f\mu$ the measure given by

$$(f\mu)(A) := \int_A f \, d\mu.$$

The notion of absolute continuity for vector measures is as follows: Let μ be a positive Radon measure and let ν be a vector-valued Radon measure. We say ν is *absolutely continuous* with respect to μ, and write $\nu << \mu$, if for any Borel set B,

$$\mu(B) = 0 \Rightarrow |\nu|(B) = 0.$$

For any Borel set $E \subset \mathbb{R}^n$, we say the measure μ is *concentrated on* E if $\mu(B) = \mu(B \cap E)$ for each Borel set B. We say ν is *singular* with respect to μ if there exist Borel sets X_ν, X_μ such that

$$X_\nu \cup X_\mu = \mathbb{R}^n,\ X_\nu \cap X_\mu = \emptyset,$$

$$|\nu| \text{ is concentrated on } X_\nu, \text{and } |\mu| \text{ is concentrated on } X_\mu.$$

We present some fundamental results for vector-valued measures:

Theorem 2 (Radon-Nikodym)**.** *Let μ be a positive Radon measure and let ν be a vector-valued Radon measure on $\mathbb{R}^n$. Assume $\nu << \mu$. Then there exists a unique function $f \in L^1_{\text{loc}}(\mathbb{R}^n, \mathbb{R}^m)$ such that $\nu = f\mu$, i.e. for all μ-measurable sets A,*

$$\nu(A) = \int_A f \, d\mu$$

The function f is called the density of ν with respect to μ.

Indeed, this follows from ([22],Theorem 6.10) applied to the components of ν.

Every complex measure has a polar decomposition, i.e. given a complex measure μ, there exists a complex function σ such that $\mu = \sigma\,|\mu|$. The same holds true for vector-valued measures:

Theorem 3. *Let μ be a vector-valued measure on $\mathbb{R}^n$. Then there exists a unique function $f \in L^1_{\text{loc}}(\mathbb{R}^n, \mathbb{R}^m)$ such that $\mu = f|\mu|$. Also $|f(x)| = 1$ for $\mu -$ almost every $x \in \mathbb{R}^n$.*

Proof. Existence of a density f follows trivially from Theorem 2 since $\mu << |\mu|$. It remains to show that $|f| = 1$ $\mu-$a.e. Fix $r < 1$, and consider the set $A_r = \{|f| < r\}$. Then if $\{E_j\}_{j=1}^\infty$ is a partition of A_r,

$$\sum_{j=1}^{\infty} |\mu(E_j)| = \sum_{j=1}^{\infty} \Big| \int_{E_j} f \ d|\mu| \Big| \leq r \sum_{j=1}^{\infty} |\mu|(E_j) = r|\mu|(A_r)$$

which implies $|\mu|(A_r) \leq r|\mu|(A_r)$, so $|\mu|(A_r) = 0$. Thus $|f| \geq 1$ μ-almost everywhere.

On the other hand, for any measurable set E with $|\mu|(E) > 0$, define

$$f_E = \frac{1}{|\mu|(E)} \int_E f \, d|\mu|$$

and note that

$$|f_E| = \Big|\frac{1}{|\mu|(E)}\int_E f\, d|\mu|\Big| = \frac{|\mu(E)|}{|\mu|(E)} \le 1.$$

Let $B(a,r) \subset \{|x| > 1\}$ be arbitrary, and let $A = f^{-1}(B(a,r))$. Note that for $x \in A$, $|f(x) - a| < r$. Suppose that $|\mu|(A) > 0$. Then since $|f_A| \le 1$,

$$r < |f_A - a| \le \frac{1}{|\mu|(A)}\int_A |f - a|\, d|\mu| \le r,$$

which is a contradiction, so $|\mu|(A) = 0$. Since $\{|f| > 1\}$ is the countable union of such sets, $|f| \le 1$ μ-a.e. □

For an $\mathbb{R}^m$-valued vector measure μ and a μ-measurable $\varphi : \mathbb{R}^n \to \mathbb{R}^m$, define the integral

$$\int \varphi \cdot d\mu = \int \varphi_1\, d\mu^1 + ... + \int \varphi_m\, d\mu^m.$$

We derive the following useful formula for the total variation of an open set from Theorem 3.

Theorem 4. *Let μ be an R^m-valued Radon measure on $\mathbb{R}^n$. Then for every open set $\Omega \subset \mathbb{R}^n$,*

$$|\mu|(\Omega) = \sup\{\int_\Omega \varphi \cdot d\mu : \varphi \in C_c(\Omega, \mathbb{R}^m), \|\varphi\|_\infty \le 1\}$$

Remark: In the above theorem, $C_c(\Omega; \mathbb{R}^m)$ can be replaced with $C_c^k(\Omega; \mathbb{R}^n)$, for any $k \in \mathbb{N} \cup \{\infty\}$.

Proof. Let $\mu = \sigma|\mu|$, where $|\sigma(x)| = 1$ μ-almost everywhere. Assume first that $|\mu|(\Omega)$ is finite. Then

$$\int_\Omega \varphi \cdot d\mu = \sum_{i=1}^m \int_\Omega \varphi^i\, \sigma^i\, d|\mu| = \int_\Omega \varphi \cdot \sigma\, d|\mu| \le \|\varphi\|_\infty |\mu|(\Omega).$$

then taking the supremum over all $\varphi \in C_c(\Omega, \mathbb{R}^m)$ with $\|\varphi\|_\infty \le 1$ yields the inequality

$$|\mu|(\Omega) \ge \sup\{\int_\Omega \varphi \cdot d\mu : \varphi \in C_c(\Omega, \mathbb{R}^m), |\varphi| \le 1\}$$

By density of compactly supported functions in L^1, for each $i = 1, ..., m$, we can select a sequence $\{\varphi_j^i\}_{j=1}^\infty \subset C_c(\Omega, \mathbb{R})$ such that

$$\|\varphi_j^i - \sigma^i\|_{L^1(\Omega, |\mu|)} \to 0$$

For $j \in \mathbb{N}$, let $\varphi_j = \left(\varphi_j^1, ..., \varphi_j^m\right)$. We can assume that $\|\varphi_j\|_\infty \leq 1$.

Since $|\mu|(\Omega) < \infty$, we can apply the dominated convergence theorem.

$$\lim_{j\to\infty} \int_\Omega \varphi_j \cdot d\mu = \lim_{j\to\infty} \int_\Omega \varphi_j \cdot \sigma \, d|\mu| = \int_\Omega |\sigma|^2 \, d|\mu| = |\mu|(\Omega).$$

If $|\mu|(\Omega) = \infty$, let $U_n = B(0, n) \cap \Omega$. There exists $\varphi_n \in C_c(U_n)$ such that

$$\int_{U_n} \varphi_n \cdot d\mu \ > |\mu|(U_n) - 1.$$

Since $|\mu|(U_n) \to \infty$

$$\int_{U_n} \varphi_n \cdot d\mu = \int_\Omega \varphi_n \cdot d\mu \to \infty.$$

□

For positive Radon measures μ, ν on $\mathbb{R}^n$, define

$$D_\mu \nu(x) = \begin{cases} \lim\limits_{r\to 0} \dfrac{\nu(B(x,r))}{\mu(B(x,r))} & \text{if } \mu(B(x,r)) > 0 \text{ for all } r > 0. \\ 0 & \text{if } \mu(B(x,r)) = 0 \text{ for some } r > 0. \end{cases}$$

The following non-trivial result requires the Besicovitch Covering Theorem ([9], p. 40).

Theorem 5 (Besicovitch)**.** *$D_\mu \nu(x)$ exists and is finite for μ-almost every $x \in \mathbb{R}^n$. In addition if $\nu << \mu$, then $D_\mu \nu$ is the density of ν with respect to μ, i.e. for each Borel set $A \subset \mathbb{R}^n$,*

$$\nu(A) = \int_A D_\mu \nu \, d\mu$$

Theorem 6 (Lebesgue Decomposition Theorem)**.** *Let μ, ν be (positive) Radon measures on $\mathbb{R}^n$. Then there exists a unique pair of Radon measures ν_s, ν_{ac} such that*

$$\nu = \nu_s + \nu_{ac}, \qquad \nu_{ac} << \mu, \qquad \nu_s \perp \mu \tag{2.1}$$

Moreover $D_\mu \nu$ is the density of ν_{ac} with respect to μ and

$$D_\mu \nu_s = 0 \quad \mu\text{-a.e.} \tag{2.2}$$

Proof. The proof of (2.1) can be found in ([22], Theorem 6.10). For (2.2), let μ be concentrated on the set C, $\nu_s(C) = 0$. Fix $\alpha > 0$, and let

$$D = \{x : \limsup_{r\to 0} \frac{\nu_s(B(x,r))}{\mu(B(x,r))} > \alpha\}.$$

Let $E = C \cap D$, and let U be any open set containing E. Then the collection of closed balls

$$\mathcal{F} = \{B(x,r) \subset U : x \in E,\, r < 1, \nu_s(B(x,r)) > \alpha\mu(B(x,r))\}$$

is a fine Besicovitch covering for E. By the Besicovitch covering theorem, there exists a disjoint countable subfamily of closed balls $B(x_i, r_i) \in \mathcal{F}$ such that

$$\mu(E \backslash \bigcup_{i=1}^{\infty} B(x_i, r_i)) = 0.$$

Then

$$\alpha\mu(E) \leq \alpha \sum_{i=1}^{\infty} \mu(B(x_i, r_i)) \leq \sum_{i=1}^{\infty} \nu_s(B(x_i, r_i)) \leq \nu_s(U).$$

Taking inf over all such U,

$$\alpha\mu(D) = \alpha\mu(E) \leq \nu_s(E) \leq \nu_s(C) = 0.$$

So $D_\mu \nu_s = 0$ μ-a.e. The equality $D_\mu \nu_{ac} = D_\mu \nu$ μ-a.e. follows immediately from the equation $\nu = \nu_{ac} + \nu_s$ and hence $D_\mu \nu$ is the density of ν_{ac} by Theorem 5. □

Remark: We can analogously define $D_\mu \nu$ when ν is any vector-valued Radon measure. Note that the above result continues to hold if we replace ν by any signed (Radon) measure by separating ν into its positive and negative parts, and also to the case where ν is a vector-valued measure on $\mathbb{R}^n$.

Corollary 1. *Let μ be a positive Radon measure and ν a vector-valued Radon measure on $\mathbb{R}^n$. For any Borel set $A \subset \mathbb{R}^n$,*

$$\nu(A) = \int_A D_\mu \nu \, d\mu + \nu_s(A).$$

Recall that $C_0(\mathbb{R}^n)$ is the space of continuous functions on $\mathbb{R}^n$ such that $\lim_{|x|\to\infty} |f(x)| = 0$. The following important result states that $[C_0(\mathbb{R}^n, \mathbb{R}^m)]^*$ is isometrically isomorphic to the space of $\mathbb{R}^m$-valued Radon measures of finite total variation on $\mathbb{R}^n$.

Theorem 7 (Riesz Representation Theorem)**.** *Let $L : C_0(\mathbb{R}^n; \mathbb{R}^m) \to \mathbb{R}$ be a bounded linear functional. Then there exist a unique finite positive measure μ and a function $\sigma : \mathbb{R}^n \to \mathbb{R}^m$ such that $|\sigma(x)| = 1$ μ-a.e. and*

$$L(\varphi) = \int_{\mathbb{R}^n} \varphi \cdot \sigma \, d\mu \quad \text{for every } \varphi \in C_0(\mathbb{R}^n; \mathbb{R}^m).$$

In addition, $|\mu|(\mathbb{R}^n) = \|L\|_{[C_0(\mathbb{R}^m;\mathbb{R}^n)]^*}$.

Proof. Let L satisfy the assumptions of the above theorem, and define for $i = 1, ..., m$ the functional

$$L_i : C_0(\mathbb{R}^n; \mathbb{R}) \to \mathbb{R} \qquad \text{by} \qquad L_i(\psi) = L(\psi \, e_i).$$

Then $\|L_i\| \leq \|L\|$, so by the Riesz Representation Theorem for complex measures ([22], Theorem 6.19), there exists a signed measure μ^i of finite total variation on $\mathbb{R}^n$ such that

$$L_i(\psi) = \int_{\mathbb{R}^n} \psi \, d\mu^i \qquad \text{for each } \psi \in C_c(\mathbb{R}^n, \mathbb{R}).$$

Let $\vec{\mu} = \{\mu^1, ..., \mu^m\}$. Then

$$\begin{aligned} L(\varphi) &= L(\varphi_1 \, e_1 + ... + \varphi_m \, e_m) \\ &= L_1(\varphi_1) + ... + L_m(\varphi_m) \\ &= \int_{\mathbb{R}^n} \varphi_1 d\mu^1 + ... + \int_{\mathbb{R}^n} \varphi_m d\mu^m = \int_{\mathbb{R}^m} \varphi \cdot d\vec{\mu}. \end{aligned}$$

It follows immediately from Theorem 3 that we have the polar decomposition $\vec{\mu} = \sigma \, \mu$ where μ is the total variation of $\vec{\mu}$. Moreover, Theorem 4 immediately implies the last statement of the theorem. □

The following result ([22], Theorem 6.12) relates a finite measure with density g to its total variation.

Theorem 8. *Let μ be a finite positive measure and $g \in L^1(\mu)$ be a vector valued function on $\mathbb{R}^n$. Let $\lambda = g\mu$. Then $|\lambda| = |g|\mu$.*

2.2 SOBOLEV SPACES

Let $\Omega \subset \mathbb{R}^n$ be an open set, $p \in [1, \infty]$, and m be any positive integer. The Sobolev Space $W^{m,p}(\Omega)$ consists of functions $f \in L^p(\Omega)$ such that for all multi-indices α, $0 \le |\alpha| \le m$, the distributional derivatives are represented by functions $v^\alpha \in L^p(\Omega)$. That is, there exists a function $v^\alpha \in L^p(\Omega)$ such that for every $\phi \in C_c^\infty(\Omega)$,

$$\int_\Omega f(x) D^\alpha \phi(x) dx = (-1)^{|\alpha|} \int_\Omega v^\alpha(x) \phi(x) dx.$$

The functions v^α are called weak derivatives of f. Whenever weak derivatives exist, it is natural (and common practice) to denote weak derivative v^α more simply by $D^\alpha f$. However, one must note that in the context of Sobolev spaces this is not the classical derivative.

The space of functions with locally p-integrable weak derivatives of orders less than or equal to m is denoted $W^{m,p}_{\text{loc}}(\mathbb{R}^n)$. A *Sobolev function* is a function in the space $W^{1,1}_{\text{loc}}(\Omega)$.

$C^1(\Omega) \subset W^{1,p}_{\text{loc}}(\Omega)$, since if $f \in C^1(\Omega)$ then integration by parts is valid, and continuity of f and ∇f guarantees local integrability. Also, it is easy to see that the weak and classical derivatives of f coincide.

$W^{m,p}(\Omega)$ is a Banach space ([10], Chap. 6) with the norm

$$\|f\|_{m,p} = \sum_{|\alpha| \le m} \|D^\alpha f\|_p.$$

Note: Where unambiguous we will drop the dependence on Ω in $\|\cdot\|_{L^p(\Omega)}$ and simply write $\|\cdot\|_p$. Similarly $\|\cdot\|_{W^{m,p}(\Omega)}$ will be replaced with $\|\cdot\|_{m,p}$.

Clearly a sequence $f_k \to f$ in $W^{m,p}(\Omega)$ if and only if

$$\|D^\alpha f_k - D^\alpha f\|_{L^p(\Omega)} \to 0 \text{ as } k \to \infty \text{ for each } |\alpha| \le m.$$

We present some key results in the theory of Sobolev spaces.

Theorem 9 (Density of Smooth Functions). *([9], p. 125) Let $\Omega \subset \mathbb{R}^n$ be any open set. Let $f \in W^{m,p}(\Omega)$, $p \in [1, \infty)$. Then there exists a sequence of functions $\{f_k\}$ in $W^{m,p}(\Omega) \cap C^\infty(\Omega)$ such that $f_k \to f$ in $W^{m,p}(\Omega)$.*

In particular, $W^{m,p}(\Omega)$ is the closure of $C^\infty(\Omega)$ in the space $(W^{m,p}(\Omega), \|\cdot\|_{m,p})$.

We say that a bounded open set Ω is a *Lipschitz domain* if its boundary is locally the graph of a Lipschitz function. If Ω is a Lipschitz domain, then $C^1(\overline{\Omega})$ is dense in $W^{1,p}(\Omega)$ for $1 \leq p < \infty$.

Theorem 10 (Traces). *([9], p. 133) Let $p \in [1, \infty)$ and let Ω be a bounded Lipschitz domain. Then there exists a bounded linear 'trace' operator*

$$T : W^{1,p}(\Omega) \to L^p(\partial\Omega, d\sigma)$$

such that for every function $u \in C^1(\overline{\Omega})$, $Tu = u|_{\partial\Omega}$. Moreover for every $\phi \in C^1(\mathbb{R}^n; \mathbb{R}^n)$ and $f \in W^{1,p}(\Omega)$,

$$\int_\Omega f \, \operatorname{div} \phi \, dx = -\int_\Omega \phi \cdot Df \, dx + \int_{\partial\Omega} (\phi \cdot \nu) \, Tf \, d\sigma.$$

Tf is called a trace of f on $\partial\Omega$ and provides a well-defined notion of "boundary values" of f. Here $Df = \{D^\alpha f\}_{|\alpha|=1}$ is the weak gradient of f. The existence of a trace allows us to integrate by parts against smooth functions which do not vanish on $\partial\Omega$. It also makes Sobolev functions excellent candidates for weak solutions of PDE with given boundary data.

Theorem 11 (Extensions). *([1], p.91) Let $p \in [1, \infty]$, and assume Ω is a bounded Lipschitz domain. Let V be any open set with $\Omega \subset\subset V$. Then there exists a bounded linear extension operator*

$$E : W^{m,p}(\Omega) \to W^{m,p}(\mathbb{R}^n)$$

such that $\operatorname{supp}(Ef) \subset V$ *for every $f \in W^{m,p}(\Omega)$.*

By an extension operator we mean that $Ef|_\Omega = f$, and by boundedness we mean existence of a constant C such that

$$\|Ef\|_{W^{m,p}(\mathbb{R}^n)} \leq C\|f\|_{W^{m,p}(\Omega)} \text{ for every } f \in W^{m,p}(\Omega).$$

Sobolev functions $f \in W^{1,p}_{\mathrm{loc}}(\Omega)$ have certain smoothness properties. If $p = \infty$, the function has a locally Lipschitz representative. If $p > n$, the function has a representative which is Holder continuous with exponent $1 - \frac{n}{p}$. In both these cases, the function is almost everywhere differentiable in the classical sense (for $n < p < \infty$ this is an extension of the classical Rademacher Theorem).

For $p \leq n$ we still have some regularity. Indeed, every Sobolev function is absolutely continuous on almost all lines parallel to the coordinate axes. This excludes the characteristic functions of a large class of sets from being Sobolev functions. In addition, the set of 'measure-theoretically significant' discontinuities of a Sobolev function has Hausdorff dimension less than $n - 1$.

2.3 FUNCTIONS OF BOUNDED VARIATION

Let $\Omega \subset \mathbb{R}^n$ be open. We say that a function f has bounded variation in Ω, or $f \in BV(\Omega)$, if $f \in L^1(\Omega)$ and the first order distributional derivatives of f are Radon measures of finite total variation in Ω. That is, there exist Radon measures $\mu^i,\ i = 1, ..., m$ such that $|\mu^i|(\Omega) < \infty$ and for each $\varphi \in C^1_c(\Omega)$,

$$\int_\Omega f\, \partial_i \varphi\, dx = -\int_\Omega \varphi\, d\mu^i.$$

The derivative of f is the vector-valued measure $Df := \{\mu^i\}_{i=1}^m$. The total variation of Df is a finite positive measure and is denoted $\|Df\|$.

Similarly we define $BV_{\mathrm{loc}}(\Omega)$ as the space of locally integrable functions whose distributional derivatives are (not necessarily finite) Radon measures.

Functions of bounded variation are a natural generalization of Sobolev functions. A larger space than $W^{1,1}(\mathbb{R}^n)$, the set of 'measure theoretically significant' discontinuities of $BV(\mathbb{R}^n)$ functions may have positive $n-1$-dimensional Hausdorff dimension, although the Hausdorff dimension of this set may not exceed $n-1$. A fundamental and surprising result is that for $\mathcal{H}^{n-1}$-a.e. point at which there is a jump, the jump is across a hyperplane (see [9], p. 213).

$BV(\Omega)$ is a Banach space with the norm:

$$\|f\|_{BV(\Omega)} = \|f\|_{L^1(\Omega)} + \|Df\|(\Omega).$$

For the proof, see Corollary 16.

A simple example illustrates the vast difference between BV spaces and Sobolev spaces. Consider the function $f = \chi_{(0,1)}$ on $(-1,1)$. This is not a Sobolev function, since a Sobolev function on $\mathbb{R}$ must be equal almost everywhere to an absolutely continuous function, and f has an unremovable jump discontinuity at 0. However, f is in $BV(-1,1)$. Indeed, for any $\varphi \in C_c^\infty(-1,1)$,

$$\int_{-1}^{1} f\varphi' dx = -\varphi(0) = -\int_{-1}^{1} \varphi \, d\delta_0,$$

where δ_0 is the Dirac measure on $\mathbb{R}$ concentrated at 0. So the distributional derivative of f is the measure δ_0. One can easily see that total variation $\|Df\|(-1,1) = 1$.

In fact, we have the following result for *any* nonempty open set $\Omega \subset \mathbb{R}^n$:

Theorem 12. $W^{1,1}(\Omega) \subsetneq BV(\Omega)$. *In addition* $W^{1,1}(\Omega)$ *is a closed isometrically embedded subspace of* $BV(\Omega)$.

Proof. If $f \in W^{1,1}(\Omega)$ then the distributional derivative is the measure $(\nabla f)\mathcal{L}^n$, which by Theorem 8 has total variation $|\nabla f|\mathcal{L}^n$, so $f \in BV(\Omega)$ and

$$\|f\|_{BV(\Omega)} = \|f\|_{L^1(\Omega)} + \|Df\|(\Omega) = \|f\|_{L^1(\Omega)} + \|\nabla f\|_{L^1(\Omega)} = \|f\|_{W^{1,1}(\Omega)}.$$

Since $W^{1,1}(\Omega)$ is complete, it is a closed subspace of $BV(\Omega)$.

If $B \subset\subset \Omega$ is an open ball, then

$$\int_\Omega \chi_B \ \operatorname{div} \varphi \ dx = \int_B \operatorname{div} \varphi \ dx = -\int_{\partial B} \varphi \cdot \nu \ d\mathcal{H}^{n-1},$$

from which it easily follows that $\|D\chi_B\|(\Omega) = \mathcal{H}^{n-1}(\partial B) < \infty$, so $\chi_B \in BV(\Omega)$. However, the function χ_B is discontinuous on each line passing through the interior of B, so it is not in $W^{1,1}(\Omega)$. □

Remark: As a consequence, the limit of any convergent sequence of smooth functions in $BV(\Omega)$ must lie in $W^{1,1}(\Omega)$. Therefore we have no hope of approximating functions in $BV(\Omega)\backslash W^{1,1}(\Omega)$ by smooth functions with respect convergence in BV norm. However, we are guaranteed convergence of a slightly weaker type, see Theorem 17.

We now proceed to give an equivalent characterization of the space $BV(\Omega)$.

For $f \in L^1_{\text{loc}}(\Omega)$, define

$$V(f, \Omega) := \sup \left\{ \int_\Omega f \operatorname{div} \varphi \, dx : \varphi \in C_c(\Omega, \mathbb{R}^m), \|\varphi\|_\infty \leq 1 \right\}.$$

Theorem 13. *$f \in BV(\Omega)$ if and only if $f \in L^1(\Omega)$ and $V(f, \Omega) < \infty$. In case either of these holds, $\|Df\|(\Omega) = V(f, \Omega)$.*

Proof. Let $f \in BV(\Omega)$. Recall that if $\varphi \in C^1_c(\Omega, \mathbb{R}^m)$, then

$$\int_\Omega \varphi \cdot d\, Df = -\int_\Omega f \ \operatorname{div} \varphi \ dx.$$

Then the fact that $\|Df\|(\Omega) = V(f, \Omega)$ is an immediate consequence of the fact that $\|Df\|$ is a finite measure and Theorem 4.

For the other implication, fix $f \in L^1(\Omega)$ and define the linear functional L on $C^1_c(\Omega; \mathbb{R}^m)$ by

$$L(\varphi) = -\int_\Omega f \operatorname{div} \varphi \ dx.$$

Assuming $V(f,\Omega) < \infty$, L is a bounded linear functional on $C^1_c(\Omega;\mathbb{R}^m)$, so it extends uniquely to a bounded linear functional (also denoted L) on $C_c(\Omega;\mathbb{R}^m)$. By the Riesz Representation Theorem, there exists a finite Radon measure μ on Ω and a unique $\sigma : \Omega \to \mathbb{R}^m$, $|\sigma| = 1$ μ−a.e. such that

$$L(\varphi) = \int_\Omega \varphi \cdot \sigma \, d\mu.$$

Let $Df := \sigma\,\mu$. Measurability of σ implies that Df is a vector-valued Radon measure on Ω. For each $\varphi \in C^1_c(\Omega;\mathbb{R}^n)$,

$$\int_\Omega f \operatorname{div} \varphi \, dx = -\int_\Omega \varphi \cdot d\, Df.$$

Therefore Df is the distributional derivate of f. Since $V(f,\Omega) < \infty$, Df has finite total variation, and therefore $f \in BV(\Omega)$. □

Theorem 7 easily implies that for $f \in BV_{\text{loc}}(\Omega)$, its distributional derivative is a measure Df defined on the Borel sets of Ω, i.e. a measure satisfying

$$\int_\Omega f \, \operatorname{div} \varphi \, dx = -\int_\Omega \varphi \cdot d\, Df \quad \text{for every } \varphi \in C^1_c(\Omega)$$

However, for $f \in BV_{\text{loc}}(\Omega)$ the variation $V(f,\Omega)$ is not, in general, finite.

Let's make a brief note about BV functions of one variable. For an interval $(a,b) \subset \mathbb{R}$, define the variation of a function $f : (a,b) \to \mathbb{R}$ by

$$\bigvee_a^b f = \left\{ \sum_{i=0}^{k-1} |f(t_{i+1}) - f(t_i)| : t_0 < t_1 < \ldots < t_k \in (a,b) \right\}.$$

A function $f \in L^1(a,b)$ is in $BV(a,b)$ if and only if it has a pointwise representative f^* such that

$$\bigvee_a^b f^* < \infty.$$

See [4], Theorem 3.27.

As a further remark, functions of bounded variation on $\Omega \subset \mathbb{R}^n$ are in $BV(\ell \cap \Omega)$ for $\mathcal{L}^{n-1}$-a.e. line ℓ parallel to the coordinate axes, and in fact may be characterized in terms of these one-dimensional variations (see [28], Theorem 5.3.4). The property of being absolutely continuous on lines for Sobolev functions is analagous to the property of being of bounded variation on lines for BV functions. In fact, one can recover the measure Df from its restrictions to almost all one dimensional slices, see ([4], Theorem 3.107).

Decomposition of the measure Df:

Let $f \in BV_{\text{loc}}(\Omega)$. To each component μ^i of Df, apply the Lebesgue Decomposition Theorem (Theorem 6). There exist Radon measures μ^i_s and μ^i_{ac} such that

$$\mu^i = \mu^i_s + \mu^i_{ac}, \qquad \mu^i_{ac} << \mathcal{L}^n, \qquad \mu^i_s \perp \mathcal{L}^n.$$

By the Radon-Nikodym Theorem, $\mu^i_{ac} = (g^i)\mathcal{L}^n$ for some $g^i \in L^1_{\text{loc}}(\mathbb{R}^n)$. The vector field $\nabla f := \{g^i\}_{i=1}^n$ is called the density of the absolutely continuous part of the measure Df with respect to the Lebesgue measure, and the measure $Df_s = \{\mu^i_s\}_{i=1}^n$ is called the singular part of the measure. Then

$$Df = Df_s + \nabla f \, \mathcal{L}^n$$

and

$$\|Df\| = \|Df\|_s + |\nabla f| \, \mathcal{L}^n.$$

As an immediate consequence of the decomposition we have the following result:

Theorem 14. *Let $f \in BV_{\text{loc}}(\Omega)$. Then $f \in W^{1,1}_{\text{loc}}(\Omega)$ if and only if $Df_s \equiv 0$.*

Remark: The main difficulty in the study of BV functions is accounting for the singular part of the measure Df.

We now provide some fundamental results for BV functions (see [9], Chapter 5):

Theorem 15 (Weak Lowersemicontinuity)**.** *Let $\{f_k\}$ be a sequence in $BV(\Omega)$ such that $f_k \to f$ in $L^1_{\text{loc}}(\Omega)$. Then $\|Df\|(\Omega) \leq \liminf_{k\to\infty} \|Df_k\|(\Omega)$.*

Proof. Fix $\varphi \in C^1_c(\Omega; \mathbb{R}^n)$ with $|\varphi| \leq 1$. Then

$$\int_\Omega f \, \operatorname{div} \varphi \, dx = \lim_{k\to\infty} \int_\Omega f_k \, \operatorname{div} \varphi \, dx = \liminf_{k\to\infty} - \int_\Omega \varphi \cdot \sigma \, d\, Df_k \leq \liminf_{k\to\infty} \|Df_k\|(\Omega)$$

Now take the supremum over all such φ. □

As a corollary, we can now prove that $(BV(\Omega), \|\cdot\|_{BV})$ is a Banach space:

Theorem 16. *$BV(\Omega)$ is a Banach space with the norm*

$$\|f\|_{BV} = \|f\|_1 + \|Df\|(\Omega).$$

Proof. Let $\{f_n\}$ be a Cauchy sequence in $BV(\Omega)$. Then the sequence is Cauchy in $L^1(\Omega)$, so there exists $f \in L^1(\Omega)$ with $f_n \to f$ in $L^1(\Omega)$. By the weak lowersemicontinuity (15), $\|Df\|(\Omega) \leq \liminf_{k\to\infty} \|Df_k\|(\Omega)$, so $f \in BV(\Omega)$. For any $\epsilon > 0$, there exists n sufficiently large such that

$$\|D(f_n - f)\|(\Omega) \leq \liminf_{m\to\infty} \|D(f_n - f_m)\| < \epsilon.$$

□

We noted earlier that functions in $BV(\Omega)\backslash W^{1,1}(\Omega)$ cannot be approximated in BV norm by smooth functions. However, we have the following weaker type approximation.

Theorem 17 (Approximation by of Smooth Functions)**.** *Let $f \in BV(\Omega)$. Then there exists a sequence of functions $\{f_k\}$ in $BV(\Omega) \cap C^\infty(\Omega)$ such that*

(i) $f_k \to f$ *in* $L^1(\Omega)$.

(ii) $\|Df_k\|(\Omega) \to \|Df\|(\Omega)$.

In addition, we have the weak- convergence*

(iii) $\|Df_k\| \xrightharpoonup{*} \|Df\|$ *in* $[C_c(\Omega)]^*$.

In fact, for any $g \in C(\Omega)$,

$$\int_\Omega g\, d\|Df_k\| \to \int_\Omega g\, d\|Df\| \text{ as } k \to \infty.$$

We will prove a more general result for BV^m functions so we omit the proof here.

Theorem 18 (Traces)**.** *([9],p. 177) Assume $\partial\Omega$ is Lipschitz and Ω is bounded. Then there exists a bounded linear trace operator $T : BV(\Omega) \to L^1(\partial\Omega)$ such that for every function $u \in C^1(\overline{\Omega})$, $Tu = u|_{\partial\Omega}$. Moreover for every $\phi \in C^1(\mathbb{R}^n;\mathbb{R}^n)$ and $f \in BV(\Omega)$,*

$$\int_\Omega f \operatorname{div} \phi\, dx = -\int_\Omega \phi \cdot d\,Df + \int_{\partial\Omega} \phi \cdot \nu\, Tf\, d\sigma$$

Remark: This trace operator and the trace defined for Sobolev functions agree on $W^{1,1}(\Omega)$.

Theorem 19 (Extensions)**.** *([9], p. 183) Assume Ω is a bounded Lipschitz domain. Let $f_1 \in BV(\Omega)$ and $f_2 \in BV(\mathbb{R}^n \backslash \overline{\Omega})$. Then $f \in BV(\mathbb{R}^n)$, where*

$$f(x) = \begin{cases} f_1(x) & \text{if} \quad x \in \Omega \\ f_2(x) & \text{if} \quad x \in \mathbb{R}^n / \bar{\Omega} \end{cases} \tag{2.3}$$

Moreover,

$$\|Df\|(\mathbb{R}^n) = \|Df_1\|(\Omega) + \|Df_2\|(\mathbb{R}^n \backslash \overline{\Omega}) + \int_{\partial\Omega} |Tf_1 - Tf_2| d\sigma$$

It follows that if Ω is a bounded Lipschitz domain, then any function in $BV(\Omega)$ can be extended to a function in $BV(\mathbb{R}^n)$ by setting the extension equal to 0 outside Ω.

Theorem 20 (Compactness)**.** *([9], p. 176) Let Ω be a Lipschitz domain. Then the embedding $BV(\Omega) \subset\subset L^1(\Omega)$ is compact, i.e. if $\{g_k\}_{k=1}^{\infty} \subset BV(\Omega)$ is a sequence such that*

$$\sup_k \|g_k\|_{BV(\Omega)} < \infty$$

then there is are a subsequence $g_{k'}$ and a function $f \in L^1(\Omega)$ such that

$$\|f - g_{k'}\|_{L^1(\Omega)} \to 0$$

In addition, it follows from weak lowersemicontinuity that $f \in BV(\Omega)$.

Let $f \in L^1_{\text{loc}}(\mathbb{R}^n; \mathbb{R}^m)$. We say f is *approximately differentiable* at x if there is a linear mapping

$$L : \mathbb{R}^n \to \mathbb{R}^m$$

such that for each $\epsilon > 0$,

$$\lim_{r \to 0} \frac{|B(x,r) \cap \{y : \frac{|f(x) - f(y) - L(y-x)|}{|x-y|} > \epsilon\}|}{|B(x,r)|} = 0.$$

We denote

$$\text{ap}\, Df(x) := L.$$

The following characterization of functions that are approximately differentiable almost everywhere is due to Whitney [26].

Theorem 21 (Whitney)**.** *Let $E \subset \mathbb{R}^n$ be measurable and let $f : E \to \mathbb{R}$ be a measurable function. The following are equivalent:*

(i) *f is approximately differentiable at almost every $x \in E$.*

(ii) *For every $\epsilon > 0$, there is a closed set $C \subset E$ and a locally Lipschitz function $g : \mathbb{R}^n \to \mathbb{R}$ such that $g|_C = f|_C$ and $|E \backslash C| < \epsilon$.*

(iii) *For every $\epsilon > 0$, there is a closed set $C \subset E$ and a function $g \in C^1(\mathbb{R}^n)$ such that $g|_C = f|_C$ and $|E \backslash C| < \epsilon$.*

Now the Luzin-type approximation of BV functions by C^1 functions follows from approximate differentiability:

Theorem 22. *Let $f \in BV_{\text{loc}}(\Omega)$. Then for almost every $x \in \Omega$, f is approximately differentiable at x and $\operatorname{ap} Df(x) = \nabla f(x)$, where ∇f denotes the density of the absolutely continuous part of the derivative of f.*

Proof. For the proof we cite the following result from ([9],p. 228):

Lemma 1. *Let $f \in BV(\mathbb{R}^n)$. Then for $\mathcal{L}^n$-a.e. $x \in \mathbb{R}^n$,*

$$\frac{1}{r} \fint_{B(x,r)} |f(y) - f(x) - \nabla f(x) \cdot (y - x)| \, dy \to 0 \quad \text{as } r \to 0.$$

Fix any such point x. We show that for any $\epsilon > 0$,

$$\lim_{r \to 0} \frac{|B(x,r) \cap \{|f(y) - f(x) - \nabla f(x) \cdot (y - x)| > \epsilon |y - x|\}|}{|B(x,r)|} = 0.$$

Suppose not. Then there are $\epsilon > 0$, $\gamma > 0$ and $r_j \to 0$ such that

$$\frac{|B(x,r_j) \cap \{|f(y) - f(x) - \nabla f(x) \cdot (y - x)| > \epsilon |y - x|\}|}{|B(x,r_j)|} > \gamma.$$

Hence there exists $\sigma > 0$ small such that,

$$\frac{|B(x,r_j) \backslash B(x,\sigma r_j) \cap \{|f(y) - f(x) - \nabla f(x) \cdot (y - x)| > \epsilon |y - x|\}|}{|B(x,r_j)|} > \frac{\gamma}{2}.$$

Since in this annulus, $|y - x| \geq \sigma r_j$, we have

$$\frac{|B(x,r_j) \cap \{|f(y) - f(x) - \nabla f(x) \cdot (y - x)| > \epsilon \sigma r_j\}|}{|B(x,r_j)|} > \frac{\gamma}{2},$$

but the left hand side is less than or equal to

$$\frac{1}{r_j}\fint_{B(x,r_j)} \frac{|f(y) - f(x) - \nabla f(x)\cdot(y-x)|}{\epsilon\sigma} \to 0 \quad \text{as } j\to\infty,$$

which is a contradiction. Therefore at any such x, f is approximately differentiable and

$$\operatorname{ap} Df(x) = \nabla f(x).$$

□

As an immediate consequence of Theorem 21 we obtain:

Corollary 2. *Let $f \in BV(\mathbb{R}^n)$. Given $\epsilon > 0$, there exists $g \in C^1(\mathbb{R}^n)$ and $C \subset \mathbb{R}^n$ closed such that $f = g$ on C and $|\mathbb{R}^n \backslash C| < \epsilon$.*

We will now prove an easy corollary of the lemma:

Corollary 3. *Let $f \in BV_{\mathrm{loc}}(\Omega)$ and suppose there exists a Borel set $E \subset \Omega$ and a function g that is differentiable at almost every point of E. Suppose $f = g$ on E. Then $\nabla f = \nabla g$ a.e. on E.*

Proof. Fix $x \in E$ which is a density point of E such that g is differentiable at x. Define $H : \Omega\backslash\{x\} \to \mathbb{R}$ by

$$H(y) = \frac{|f(x) - f(y) - \nabla g(x)\cdot(y-x)|}{|x-y|}.$$

Then fix $\epsilon > 0$.

$$\begin{aligned}
&\frac{|B(x,r)\cap\{y : H(y) > \epsilon\}|}{|B(x,r)|}\\
&\quad= \frac{|B(x,r)\cap E\cap\{y : H(y) > \epsilon\}|}{|B(x,r)|} + \frac{|B(x,r)\cap E^c\cap\{y : H(y) > \epsilon\}|}{|B(x,r)|}\\
&\quad\le \frac{|B(x,r)\cap\{y : |g(x) - g(y) - \nabla g(x)\cdot(y-x)| > \epsilon|x-y|\}|}{|B(x,r)|} + \frac{|B(x,r)\backslash E|}{|B(x,r)|}\\
&\qquad\to 0 \text{ as } r\to 0.
\end{aligned}$$

The convergence of the first term to zero is due to the differentiability of g at x and convergence of the second term to 0 follows from the assumption that x is a density point of E. Thus $\operatorname{ap} Df(x) = \nabla g(x)$. Since $\operatorname{ap} Df(x) = \nabla f(x)$ at almost every x in Ω, the result follows. □

2.4 THE WHITNEY EXTENSION THEOREM

The Whitney Extension Theorem provides a necessary and sufficient condition for a family $\{f^\alpha\}_{|\alpha|\le m}$ of continuous functions defined on a compact set $K \subset \mathbb{R}^n$ to be the restrictions $D^\alpha f|_K$ of *some* function $f \in C^m(\mathbb{R}^n)$. It is easy to obtain a necessary condition on the set $\{f^\alpha\}_{|\alpha|\le m}$ by applying the Taylor formula to $f \in C^m(\mathbb{R}^n)$. Indeed, consider the Taylor remainder

$$R_x^m f(y) := f(y) - \sum_{|\alpha|\le m} \frac{(y-x)^\alpha}{\alpha!} D^\alpha f(x).$$

For each β, $0 \le |\beta| \le m$, $D^\beta f \in C^{m-|\beta|}(\mathbb{R}^n)$. Therefore the Taylor formula implies

$$R_x^{m-|\beta|} D^\beta f(y) = D^\beta f(y) - \sum_{|\gamma|\le m-|\beta|} \frac{(y-x)^\gamma}{\gamma!} D^{\beta+\gamma} f(x) = o(|x-y|^{m-|\beta|}) \qquad (2.4)$$

uniformly as $|x-y| \to 0$ on compact subsets of $\mathbb{R}^n$.

Now fix a compact set $K \subset \mathbb{R}^n$, and suppose $f^\alpha := D^\alpha f|_K$. Let $F := \{f^\alpha\}_{|\alpha|\le m}$. Define for each $|\beta| \le m$,

$$(R_x^m F)^\beta(y) := f^\beta(x) - \sum_{|\gamma|\le m-|\beta|} \frac{(y-x)^\gamma}{\gamma!} f^{\beta+\gamma}(y).$$

Then necessarily

$$\frac{(R_x^m F)^\beta(y)}{|x-y|^{m-|\beta|}} \rightrightarrows 0 \quad \text{uniformly as } |x-y| \to 0,\ x, y \in K. \qquad (2.5)$$

Indeed, since $f^\alpha = D^\alpha f|_K$, this is just a restatement of (2.4).

Now we let $F := \{f^\alpha\}_{|\alpha|\le m}$ be an arbitrary collection of continuous functions on K. We have just seen that in order for there to exist $f \in C^m(\mathbb{R}^n)$ such that $f^\alpha = D^\alpha f|_K$ for each $|\alpha| \le m$, the condition (2.5) is necessary.

The surpising fact is that it is also sufficient. This is the celebrated theorem of Whitney. Although proving necessity is straightforward, proving sufficiency requires an explicit construction of the extension and this is quite a difficult task. However, having an explicit formula for an extension is useful in many applications. Here we only provide a sketch of the proof. The interested reader can find complete details in [25], ([18], pp. 1 - 8).

Let's set some notation. For a compact set $K \subset \mathbb{R}^n$, define

$$J^m(K) = \{ \{f^\alpha\}_{|\alpha|\le m} \,\big|\, f^\alpha : K \to \mathbb{R} \text{ are continuous}\},$$

the set of all *jets* of order m on K. For a fixed $|\beta| \le m$, we define an operator

$$D^\beta : J^m(K) \to J^{m-|\beta|}(K) \quad \text{by} \quad \{f^\alpha\}_{|\alpha|\le m} \longmapsto \{f^{\beta+\alpha}\}_{|\alpha|\le m-|\beta|}.$$

We will write $F = \{f^\alpha\}_{|\alpha|\le m}$ for a jet of order m. For each $a \in K$, define the operator

$$T^m_a : J^m(K) \to C^\infty(\mathbb{R}^n) \quad \text{by} \quad \{f^\alpha\}_{|\alpha|\le m} \longmapsto \sum_{|\alpha|\le m} f^\alpha(a)\frac{(\cdot - a)^\alpha}{\alpha!}.$$

$T^m_a F$ is a *formal* Taylor polynomial constructed from the jet F and centered at a.

Denote by $\mathcal{E}^m(K)$ the space of all jets F in $J^m(K)$ such that

$$(R^m_x F)^\alpha(y) = o(|x-y|^{m-|\alpha|}) \text{ as } |x-y| \to 0 \text{ for every } x, y \in K \text{ and } |\alpha| \le m. \tag{2.6}$$

Each $F \in \mathcal{E}^m(K)$ is called a *Whitney jet* of class C^m on K. Now we state the theorem.

Theorem 23 (Whitney Extension Theorem)**.** *There exists a linear extension operator $W : \mathcal{E}^m(K) \to C^m(\mathbb{R}^n)$ such that for every $F \in \mathcal{E}^m(K)$ and every $x \in K$, $D^\alpha WF(x) = f^\alpha(x)$ for $|\alpha| \le m$.*

For a Whitney jet F, the construction of the extension WF relies on the well known Whitney decomposition of the open set $\mathbb{R}^n \backslash K$. The main feature of the decomposition is that it partitions this set into diadic cubes whose diameter is comparable to the distance of the cube to the boundary of K. Here is a brief outline of the construction: First divide $\mathbb{R}^n$ into diadic cubes S of side length 1. Collect all such cubes with $d(S,K) > \sqrt{n}$. Now partition the remaining space into diadic cubes of side length 1/2, and collect all cubes with $d(S,K) > \frac{\sqrt{n}}{2}$. Continue with cubes of side length $1/4, 1/8, \ldots$ ad finitum. Let I be the collection of all the cubes. This is a partition of $\mathbb{R}^n \backslash K$. For each $S \in I$, define $\widetilde{S} := \frac{3}{2}S$ to be the concentric cube of 3/2 side length. Let $\{\phi_S\}_{S\in I}$ be a $C^\infty(\mathbb{R}^n)$ partition of unity with

(a) $0 \le \phi_S \le 1$

(b) $\text{supp}(\phi_S) \subset \text{int}(\widetilde{\text{S}})$

Then the following hold:

1. $\text{int(S)} \cap \text{int(T)} = \emptyset$ whenever S, T $\in$ I.
2. Adjacent cubes differ in size by at most a factor of 2 (linear dilation).
3. For each $S \in I$, $\#\{T \in I : \text{supp}(\phi_S) \cap \text{supp}(\phi_T) \neq \emptyset\} \leq 4^n$.
4. $\widetilde{S} \cap \widetilde{T} \neq \emptyset$ iff S is adjacent to T. Therefore the partition of unity is locally finite.
5. If S is not side length 1 and $x \in \widetilde{S}$, the $d(x, K) \leq (17/4)\text{diam(S)}$.
6. There exists $C(\beta, n)$ only such that $|D^\beta \phi_S(x)| \leq C(1 + \frac{1}{d(x,K)^{|\beta|}})$; $|\beta| \leq m$, $x \in \mathbb{R}^n \backslash K$.

Now we can give an explicit formula for the extension WF: For every $S \in I$, choose a point a_S in K that is nearest to $\text{supp}(\phi_S)$. Let $f^0 = f$.

$$WF(x) = \begin{cases} f^0(x) & \text{if } x \in K, \\ \displaystyle\sum_{S \in I} \phi_S(x) T^m_{a_S} F(x) & \text{if } x \in \mathbb{R}^n \backslash K. \end{cases} \tag{2.7}$$

First note that this is a finite sum at each x since the partition of unity is locally finite. To define the function outside K, we take a weighted sum (over all cubes in the partition) of formal Taylor polynomials of the jet centered at nearest points on the boundary of K. As x approaches K, the points at which the polynomials are centered become clustered near x. This limiting behavior as x approaches K is the crucial part of the construction. A priori we cannot differentiate the function at any point in K. To prove the extension is in fact a C^m function, we define a jet $\{\tilde{f}^\alpha\}_{|\alpha| \leq m}$ and show inductively that the functions in this jet are in fact classical derivatives of WF. Let

$$\tilde{f}^\alpha(x) = \begin{cases} f^\alpha(x) & \text{if x} \in K, \\ D^\alpha WF(x) & \text{if x} \in \mathbb{R}^n \backslash K. \end{cases} \tag{2.8}$$

Now since $T^m_{a_S} F$ is C^∞, it is immediate that the function WF is smooth on $\mathbb{R}^n \backslash K$. We need to prove inductively that for each $|\alpha| < m$, $a \in K$ and $x \in \mathbb{R}^n$,

$$\tilde{f}^\alpha(x) - f^\alpha(a) - \nabla \tilde{f}^\alpha(a) \cdot (x - a) = o(|x - a|) \tag{2.9}$$

where $\nabla \tilde{f}^\alpha = \{\tilde{f}^{\alpha+e_i}\}_{i=1}^n$ is the formal gradient of $\tilde{f}^\alpha$ consisting of functions in the jet. Indeed, if (2.9) holds then implies that $\tilde{f}^\alpha$ is differentiable at a and $\nabla \tilde{f}^\alpha$ coincides with the classical gradient of $\tilde{f}^\alpha$. If $a \in \text{int}(K)$, then (2.9) is satisfied for points x near a. This follows immediately from (2.6). So WF is part C^m in the interior of K. Therefore the interesting case is when a is a boundary point of K.

The key estimate is the following:

Let L be a cube with $K \in \text{int}(L)$ and $\lambda = \sup_{x\in L} d(x, K)$. Then there exists $C(m, n, \lambda)$ such that for each β with $|\beta| \le m$, $a \in K$ and $x \in L$,

$$\left|\tilde{f}^\alpha(x) - D^\alpha T_a^m F(x)\right| \le C\omega(|x-a|)|x-a|^{m-|\alpha|} \tag{2.10}$$

where ω is a modulus of continuity whose existence is guaranteed by (2.6). This is the most difficulty part of the proof and we will not include the details.

To finish the proof using (2.10), for $|\alpha| < m$, add and subtract terms to get

$$\frac{|\tilde{f}^\alpha(x) - f^\alpha(a) - \nabla \tilde{f}^\alpha(a)\cdot(x-a)|}{|x-a|}$$
$$\le \frac{\left|\tilde{f}^\alpha(x) - D^\alpha T_a^m F(x)\right|}{|x-a|} + \sum_{1<|\beta|\le m-|\alpha|} \frac{|x-a|^{|\beta|}|D^{\alpha+\beta}f(a)|}{|x-a|}$$

Letting $|x-a| \to 0$, the first term converges to zero by (2.10) and the second since $|\beta| > 1$. Inductively, this proves that WF has derivatives up to order m, and $D^\alpha WF = \tilde{f}^\alpha$. When $|\alpha| = m$, (2.10) reduces to

$$|\tilde{f}^\alpha(x) - \tilde{f}^\alpha(a)| \le C\omega(|x-a|)$$

which proves continuity of the mth order derivatives, and completes the proof.

Corollary 4. *Let C be a closed set on $\mathbb{R}^n$ and $F = \{f^\beta\}_{|\beta|\le m} \in J^m(C)$. Suppose the Taylor remainder-type estimates (2.6) hold on each compact set $K \subset C$. Then there is a function $g \in C^m(\mathbb{R}^n)$ such that for each $x \in C$, $|\beta| \le m$,*

$$D^\beta g(x) = f^\beta(x).$$

Moreover the mapping $F \mapsto g$ is linear.

Proof. If x in not in C, there exists a nearest point $y \in C$ such that $\operatorname{dist}(x, C) = |x - y|$. Decompose $\mathbb{R}^n \setminus C$ using the Whitney Decomposition. The same inequalities hold with K replaced by C.

We define the extension by

$$g = \begin{cases} f(x) & \text{if } \ \mathrm{x} \in C, \\ \sum_{S \in I} \phi_S(x) T^m_{a_S} F(x) & \text{if } \ \mathrm{x} \in \mathbb{R}^n \setminus C, \end{cases} \tag{2.11}$$

where $a_S \in C$ is such that $\operatorname{dist}(S, C) = \operatorname{dist}(a_S, C)$. g is clearly C^∞ on $\mathbb{R}^n \backslash C$.
On each compact $K \subset C$ we get that

$$\sup_{x,y \in K, |x-y| < \rho} \frac{|(R^m_x F)^\beta(y)|}{|x - y|^{m-|\beta|}} \to 0 \tag{2.12}$$

as $\rho \to 0$, which guarantees existence of a modulus of continuity ω_K for the compact set K ([18], Thm 2.2.1).

Fix $a \in C$, and let $K = \bar{B}(0, 1) \cap C$. If $a \in \operatorname{int}(C)$, this is not a problem because $|R^m_a F(y)| \to 0$ uniformly in K.

For $a \in \partial C$, we get the estimate $|(R^m_a F)^\beta(y)| \leq C\omega_K(x - a)|x - a|^{m-|\beta|}$, which implies as in the original proof that the function is C^m is a neighborhood of a. $\square$

3.0 APPROXIMATION BY LIPSCHITZ FUNCTIONS

We proved, see Corollary 2, that $f \in BV$ equals to a C^1, and in particular, to a locally Lipschitz function outside a set of arbitrarily small measure. We will provide two alternative proofs of this result that will use certain pointwise estimates of f. The motivation stems from the fact that in the following sections we will use a similar method to prove a higher order analogue for BV^m functions.

3.1 FIRST POINTWISE ESTIMATE FOR BV FUNCTIONS

Henceforth, we identify $f \in BV_{\mathrm{loc}}(\Omega)$ with a pointwise representative

$$\tilde{f}(x) := \limsup_{r \to 0} \fint_{B(x,r)} f\,dy \tag{3.1}$$

defined at <u>every</u> point of Ω.

For $f \in L^1(\mathbb{R}^n)$ and a Radon measure $\mu \in \mathcal{B}(\mathbb{R}^n)$, we define the following *maximal functions*:

$$\text{(i) } \mathcal{M}f(x) = \sup_{r>0} \fint_{B(x,r)} |f|\,dy, \qquad \text{(ii) } \mathcal{M}\mu(x) = \sup_{r>0} \frac{|\mu|(B(x,r))}{|B(x,r)|},$$

$$\text{(iii) } \mathcal{M}_\delta \mu(x) = \sup_{0<r<\delta} \frac{|\mu|(B(x,r))}{|B(x,r)|}.$$

For Lipschitz functions, we want to utilize the following pointwise estimate:

Theorem 24. *Let $f \in BV_{\mathrm{loc}}(\mathbb{R}^n)$. Then for each $x, y \in \mathbb{R}^n$,*

$$|f(x) - f(y)| \le C(n)|x-y|(\mathcal{M}_{|x-y|}\|Df\|(x) + \mathcal{M}_{|x-y|}\|Df\|(y)). \tag{3.2}$$

This material in this section is devoted to proving the above estimate. We start with some well-known weak-type estimates (see [27], Sec. 2.8).

Theorem 25. *Let $f \in L^1(\mathbb{R}^n)$, $\mu \in \mathcal{B}(\mathbb{R}^n)$ be as above, and let $t > 0$. Then*

$$t|\{|f| > t\}| \leq \int_{\{|f|>t\}} |f|\, dy \leq \|f\|_1, \tag{3.3}$$

$$t|\{\mathcal{M}f > t\}| \leq 2 \cdot 5^n \int_{\{|f|>t/2\}} |f|\, dy \leq 2 \cdot 5^n \|f\|_1, \tag{3.4}$$

$$t|\{\mathcal{M}\mu > t\}| \leq 5^n |\mu|(\mathbb{R}^n). \tag{3.5}$$

By *Chebychev's inequality* (3.3), $|\{|f| > t\}| \to 0$ as $t \to \infty$. We also know that for $f \in L^1(\mu)$,

$$\int_E |f|\, d\mu \to 0 \qquad \text{as } \mu(E) \to 0.$$

These facts along with inequalities (3.3),(3.4) give us the follow result:

Corollary 5. *If $f \in L^1(\mathbb{R}^n)$, then*

$$t|\{|f| > t\}| \to 0 \quad \textit{and} \quad t|\{\mathcal{M}f > t\}| \to 0 \quad \textit{as } t \to \infty.$$

Remark: The convergence $t|\{\mathcal{M}\mu > t\}| \to 0$ as $t \to \infty$ is not true for arbitrary $\mu \in \mathcal{B}(\mathbb{R}^n)$. In fact, a simple example where this fails is δ_0, the Dirac measure concentrated at zero. This presents a major difficulty in estimates involving BV functions, because it may happen that $t|\{\mathcal{M}\|Df\| > t\}|$ does not converge to 0 as $t \to \infty$.

To prove our pointwise estimate, we use the following well-known potential estimate for smooth functions. For a proof see ([11], Theorem 7.12):

Lemma 2. *Let $Q \subset \mathbb{R}^n$ be an open cube, and let $f \in C^1(Q)$. Then for each $x \in Q$*

$$|f(x) - f_Q| \leq C(n) \int_Q \frac{|\nabla f(y)|}{|x-y|^{n-1}} dy \tag{3.6}$$

We will also need the following lemma ([13], Lemma 3.2):

Lemma 3. *Fix $x \in \mathbb{R}^n$, $m < n$. Then there exists $C_{m,n}$ independent of r such that*

$$⨍_{B(x,r)} \frac{1}{|y-z|^{n-m}}\, dy \leq \frac{C_{m,n}}{|x-z|^{n-m}} \text{ for each } r > 0,\ z \in \mathbb{R}^n.$$

These lemmas along with approximation of $BV(Q)$ functions by C^∞ functions yield the following result:

Theorem 26. *Let $f \in BV_{\mathrm{loc}}(\mathbb{R}^n)$ be defined pointwise as in (3.1), and let $Q \subset \mathbb{R}^n$ be an open cube. Then for each $x \in Q$,*

$$|f(x) - f_Q| \leq C(n) \int_Q \frac{d\|Df\|(y)}{|x-y|^{n-1}} \tag{3.7}$$

Proof. Note that $f \in BV(Q)$. By Theorem 17, we can find a sequence $\{f_k\} \subset BV(Q) \cap C^\infty(Q)$ such that

(i) $f_k \to f$ in $L^1(Q)$,

(ii) $\|Df_k\|(Q) \to \|Df\|(Q)$,

and moreover, for each function u which is continuous and bounded on Q,

$$\int_Q u\, d\|Df_k\| \to \int_Q u\, d\|Df\|.$$

Now we will apply this to the function $u(z) = \int_{B(x,\epsilon)} \frac{dy}{|z-y|^{n-1}}$. By first applying inequality (3.6) and Fubini's Theorem, we find:

$$\begin{aligned} \left| ⨍_{B(x,\epsilon)} f_k(y) - (f_k)_Q \right| &\leq C(n) ⨍_{B(x,\epsilon)} \left(\int_Q \frac{|\nabla f_k(z)|}{|y-z|^{n-1}} dz \right) dy \\ &\leq C(n) \int_Q \left(⨍_{B(x,\epsilon)} \frac{dy}{|y-z|^{n-1}} \right) |\nabla f_k(z)|\, dz. \end{aligned}$$

Letting $k \to \infty$, applying Fubini's Theorem and Lemma 3 yields

$$\left| ⨍_{B(x,\epsilon)} f(y) - f_Q \right| \leq C(n) \int_Q \left(⨍_{B(x,\epsilon)} \frac{dy}{|y-z|^{n-1}} \right) d\|Df\|(z) \tag{3.8}$$

$$\leq C(n) \int_Q \frac{d\|Df\|(z)}{|x-z|^{n-1}} \tag{3.9}$$

which is a bound independent of ϵ. Now taking the limsup as $\epsilon \to 0$ on the left hand side yields the result.

□

Now we prove the following lemma due to Hedberg [14], see also [12] Lemma 3.4:

Lemma 4 (Hedberg)**.** *If $m > 0$ and $Q \subset \mathbb{R}^n$ be a cube, then there exists a constant $C(n,m)$ such that*

$$\int_Q \frac{d|\mu|(y)}{|x-y|^{n-m}} dy \le C(\mathrm{diam}(Q))^m \mathcal{M}_{\mathrm{diam}(Q)}\mu(x)$$

for each $\mu \in \mathcal{B}(\mathbb{R}^n)$ and $x \in Q$.

Proof. Let $\delta = \mathrm{diam}(Q)$. Let $A_k = Q \cap B(x, \delta/2^k) \backslash B(x, \delta/2^{k+1})$ for $k = 0, 1, 2, ...$

On A_k, $\frac{\delta}{2^{k+1}} \le |x-y| \le \frac{\delta}{2^k}$,

$$\begin{aligned} \int_Q \frac{d|\mu|(y)}{|x-y|^{n-m}} &= \sum_k \int_{A_k} \frac{d|\mu|(y)}{|x-y|^{n-m}} \\ &\le \sum_k (\frac{2^{k+1}}{\delta})^{n-m} \int_{A_k} d|\mu|(y) \\ &\le C \sum_k (\frac{2^{k+1}}{\delta})^{n-m} (\frac{\delta}{2^k})^n ⨍_{A_k} d|\mu|(y) \\ &\le C 2^{n-m} \sum_k (\frac{1}{2^m})^k \delta^m \mathcal{M}_\delta \mu(x). \end{aligned}$$

Summing the geometric series gives the result. □

Now we can prove Theorem 24.

Proof. By Theorem 26 and Hedberg's lemma,

$$|f(x) - f_Q| \le \int_Q \frac{d\|Df\|(z)}{|x-z|^{n-1}} \le C(\mathrm{diam}(Q)) M_{\mathrm{diam}(Q)} \|Df\|(x).$$

Applying the Triangle Inequality, for any cube Q containing x, y, we have

$$|f(x) - f(y)| \le C(\mathrm{diam}(Q))(M_{\mathrm{diam}(Q)}\|Df\|(x) + M_{\mathrm{diam}(Q)}\|Df\|(y)). \tag{3.10}$$

To complete the argument, note that for any $\epsilon > 0$ we can find a cube containing x and y whose diameter is less that $|x-y| + \epsilon$. □

3.2 FIRST APPROXIMATION BY LIPSCHITZ FUNCTIONS

We provide a first proof of the main result of this section:

Theorem 27. *Let $f \in BV(\Omega)$. For every $\epsilon > 0$, there exists a locally Lipschitz function g_ϵ such that*

$$\text{(i) } |\{f \neq g_\epsilon\}| < \epsilon \quad \textit{and} \quad \text{(ii) } \|g_\epsilon - f\|_1 < \epsilon.$$

Proof. Using a partition of unity we may assume that $\Omega = \mathbb{R}^n$ and that f has compact support. Let

$$E_t = \{|f| \leq t\} \cap \{\mathcal{M}\|Df\| \leq t\}. \tag{3.11}$$

Note $f|_{E_t}$ is a Lipschitz mapping with constant $2Ct$ by (3.2). Let $\widetilde{f_t}$ be the Mcshane (Lipschitz) extension of $f|_{E_t}$ to all of $\mathbb{R}^n$. Then let

$$f_t = \begin{cases} \widetilde{f_t} & \text{if } |\widetilde{f_t}| \leq t, \\ t & \text{if } \widetilde{f_t} > t, \\ -t & \text{if } \widetilde{f_t} < -t. \end{cases}$$

f_t is Lipschitz with the same Lipschitz constant as $f|_{E_t}$, since the Mcshane extension and the truncation both preserve the Lipschitz constant.

(i): On E_t, $f = \widetilde{f_t}$ and $|\widetilde{f_t}| = |f| \leq t$. So $f_t = \widetilde{f_t} = f$ on E_t. Then to prove f_t coincides with f outside a set of small measure, it suffices to prove that $|E_t^c| \to 0$. Note

$$|E_t^c| \leq |\{|f| > t\}| + |\{\mathcal{M}\|Df\| > t\}|.$$

Since $f \in BV(\mathbb{R}^n)$, $|E_t^c| \to 0$ as $t \to \infty$ by (3.3) and (3.5).

(ii) We begin by finding a uniform bound on $\|Df_t\|(\mathbb{R}^n)$. Note that by (3.3),(3.5),

$$t|E_t^c| \leq t|\{\mathcal{M}f > t\}| + t|\{\mathcal{M}\|Df\| > t\}| \leq 2 \cdot 5^n(\|f\|_1 + \|Df\|(\mathbb{R}^n)).$$

By Rademacher Theorem([9], p. 81), f_t is differentiable almost everywhere, and the Lipschitz condition implies $\|\nabla f_t\|_\infty \le C\,t$. By Corollary 3, $|\nabla f_t| = |\nabla f|$ on E_t, where ∇f is the density of the absolutely continuous part of Df. Then

$$\begin{aligned}\int_{\mathbb{R}^n} |\nabla f_t| dx &\le \int_{E_t} |\nabla f| dx + \int_{E_t^c} C\,t \\ &\le \|Df_{ac}\|(E_t) + C\,t|E_t^c| \\ &\le C(\|f\|_1 + \|Df\|(\mathbb{R}^n)).\end{aligned}$$

We also have the estimate

$$\int_{\mathbb{R}^n} |f_t| \le \int_{E_t^c} t\,dx + \int_{E_t} |f| \le t|E_t^c| + \|f\|_1 \le C(\|f\|_1 + \|Df\|(\mathbb{R}^n)).$$

Hence

$$\sup_{t>0} \|f_t\|_{BV} \le C(\|f\|_1 + \|Df\|(\Omega)).$$

Since f has compact support, f_t has compact support for t large. Indeed, if $d = \operatorname{dist}(x, \operatorname{supp}(f))$, then

$$\mathcal{M}\|Df\|(x) \le \frac{\|Df\|(\mathbb{R}^n)}{\omega_n d^n},$$

so for

$$t > \frac{\|Df\|(\mathbb{R}^n)}{\omega_n d^n},$$

$x \in E_t$ because $|f(x)| = 0 < t$ and $\mathcal{M}\|Df\|(x) \le t$, so $f_t(x) = f(x) = 0$.

This implies that family f_t all have support in the same compact ball, so we can apply Theorem 20 to guarantee existence of $g \in L^1(\mathbb{R}^n)$ such that $\|f_t - g\|_{L^1(\mathbb{R}^n)} \to 0$. Since $|\{f \ne f_t\}| \to 0$ and E_t are increasing sets, $f_t \to f$ almost everywhere, so f is the $L^1(\mathbb{R}^n)$ limit of f_t. □

3.3 SECOND APPROXIMATION BY LIPSCHITZ FUNCTIONS

Now we provide an alternative construction for a Lipschitz approximation of a BV function. Our first approximation was constructed by restricting the function f to the set E_t, whereby we lost all information about f on E_t^c. This is a significant problem because despite the fact that the measure of E_t^c is small, $\|Df\|(E_t^c)$ may be large.

Recall again our pointwise estimate (3.2). Now instead let $E_t = \{\mathcal{M}\|Df\| \leq t\}$, and consider the Whitney decomposition of E_t^c, which is an open set by lowersemicontinuity of the maximal function $\mathcal{M}\|Df\|$. Recall the Whitney decomposition is made up of pairwise disjoint cubes $\{Q_i\}_{i\in I}$, and for each i we let $\widetilde{Q_i}$ be a concentric cube linearly dilated by a factor of $3/2$. Finally recall the associated locally finite partition of unity $\{\xi_i\}$ subordinate to $\{\widetilde{Q_i}\}_{i\in I}$ has the property that $\|\nabla\xi_i\|_\infty \leq C(n)\text{diam}(\widetilde{Q_i})^{-1}$. For the relevant details, see section 2.4 of this thesis.

Now for $t > 0$, let

$$g_t = \begin{cases} f(x) & \text{if } x \in E_t, \\ \sum_i \xi_i(x) f_{\widetilde{Q_i}} & \text{if } x \in E_t^c. \end{cases} \tag{3.12}$$

Clearly since $g_t = f$ on E_t, $|\{g_t \neq f\}| \to 0$ as $t \to \infty$. We must show that g_t is Lipschitz.

Let $x \in E_t^c$. Let $\overline{x} \in E_t$ such that $\text{dist}\,(x, E_t) = |x - \overline{x}|$.

<u>Claim:</u> $|g_t(x) - f(\overline{x})| \leq C(n)t|x - \overline{x}|$.

$$\begin{aligned} |g_t(x) - f(\overline{x})| &= \left|\sum_i \xi_i(x) f_{\widetilde{Q_i}} - f(\overline{x})\right| \\ &= \left|\sum_i \xi_i(x)(f_{\widetilde{Q_i}} - f(\overline{x}))\right| \\ &\leq \sum_i |f_{\widetilde{Q_i}} - f(\overline{x})| \end{aligned}$$

Now let $\widetilde{T_i}$ be a cube containing $\widetilde{Q_i}$ and also the point $\overline{x}$. Note that since $|x-\overline{x}| < C\text{diam}(\widetilde{Q_i})$, we can ensure that $\text{diam}(\widetilde{T_i}) < C\text{diam}(\widetilde{Q_i})$. Since $\overline{x} \in \widetilde{T_i}$ and $\widetilde{Q_i} \subset \widetilde{T_i}$, a simple variation of the inequality (3.6) and the same approximation argument as in section 3.1, along with

Lemma (4) implies that

$$|f(\overline{x}) - f_{\widetilde{Q_i}}| \le C(n)\frac{|\widetilde{T_i}|}{|\widetilde{Q_i}|}\int_{\widetilde{T_i}} \frac{\|Df\|(z)}{|\overline{x}-z|^{n-1}} \le C(n)\mathrm{diam}(\widetilde{T_i})\mathcal{M}\|Df\|(\overline{x})$$
$$\le C\mathrm{diam}(\widetilde{Q_i})t,$$

because $x \in E_t$ implies $\mathcal{M}\|Df\|(\overline{x}) < t$.

Combining these inequalities, and noting that $\mathrm{diam}(\widetilde{Q_i}) \le C|x-\overline{x}|$, we see that

$$|g_t(x) - f(\overline{x})| \le \sum_i |f(\overline{x}) - f_{\widetilde{Q_i}}| \le \sum_i C\,t\,\mathrm{diam}(\widetilde{Q_i})$$
$$\le 4^n C\,t\,|x-\overline{x}|.$$

Now let $x, y \in E_t^c$, and suppose $|x-y| \ge |x-\overline{x}|$. Then since $|y-\overline{x}| \ge |y-\overline{y}|$, a simple application of the triangle inequality shows

$$\begin{aligned}|g_t(x) - g_t(y)| &\le |g_t(x) - f(\overline{x})| + |f(\overline{x}) - f(\overline{y})| + |f(\overline{y}) - g_t(\overline{y})| \\ &\le Ct(|x-\overline{x}| + |\overline{x}-\overline{y}| + |\overline{y}-y|) \\ &\le Ct|x-y|.\end{aligned}$$

If $|x-\overline{x}|, |y-\overline{y}| \ge |x-y|$, then noting that $f(\overline{x}) = \sum_i \xi_i(x)f(\overline{x}) = \sum_i \xi_i(y)f(\overline{x})$, we have

$$\begin{aligned}|g_t(x) - g_t(y)| &= \left|\sum_i \xi_i(x)f_{\widetilde{Q_i}} - \sum_i \xi_i(y)f_{\widetilde{Q_i}}\right| \\ &\le \left|\sum_i (\xi_i(x) - \xi_i(y))(f_{\widetilde{Q_i}} - f(\overline{x}))\right| \\ &\le \sum_i \|\nabla\xi_i\|_\infty |x-y||f(\overline{x}) - f_{\widetilde{Q_i}}| \\ &\le \sum_i C\mathrm{diam}(\widetilde{Q_i})^{-1}|x-y|\mathrm{diam}(\widetilde{Q_i})t \\ &\le C\,t\,|x-y|.\end{aligned}$$

Therefore g_t is Lipschitz.

The advantage to this approximation is that the information about f on E_t^c is built into the function g_t. We will see that this yields stronger bounds than the previous approximation.

Claim 1: $\|g_t - f\|_1 \to 0$ as $t \to \infty$.

$$\int_{E_t^c} |g_t| dx = \int_{E_t^c} \sum_i |\xi_i(x)||f_{\widetilde{Q_i}}| dx \le \int_{E_t^c} \sum_i \chi_{\widetilde{Q_i}}(x)|f_{\widetilde{Q_i}}| dx$$
$$= \sum_i \int_{\widetilde{Q_i}} |f| dx \to 0 \text{ as } t \to \infty$$

since $f \in L^1(\mathbb{R}^n)$.

Claim 2: $\|Dg_t\|(\mathbb{R}^n) \le C(n)\|Df\|(\mathbb{R}^n)$.

It suffices to show that $\|Dg_t\|(E_t^c) \le C(n)\|Df\|(E_t^c)$. Since $\nabla(\sum_i \xi_i(x)) = \nabla(1) = 0$,

$$\begin{aligned}\int_{E_t^c} |\nabla g_t|\, dx \le \sum_i \int_{E_t^c} |\nabla \xi_i(x) f_{\widetilde{Q_i}}| dx = & \int_{E_t^c} |\nabla \xi_i(x)(f_{\widetilde{Q_i}} - f(x))| dx \\ & \le \sum_i \|\nabla \xi_i\|_\infty \int_{\widetilde{Q_i}} |f_{\widetilde{Q_i}} - f(x)| dx \\ & \le \sum_i C \mathrm{diam}(\widetilde{Q_i})^{-1} \mathrm{diam}(\widetilde{Q_i}) \|Df\|(\widetilde{Q_i}) \\ & \le C \sum_i \|Df\|(\widetilde{Q_i}).\end{aligned}$$

In the second to last step we employed the Poincaré inequality for BV functions on cubes ([4],Remark 3.50). Since each $\widetilde{Q_i}$ is covered by finitely many Q_i and each $\widetilde{Q_i}$ intersects at most 4^n other cubes, we get that

$$\int_{E_t^c} |\nabla g_t| dx \le C(n) \sum_i \|Df\|(\widetilde{Q_i}) \le C(n)\|Df\|(E_t^c).$$

I believe this bound is not optimal. In fact, the way we have defined g_t on E_t^c is by somewhat crude approximation of f by averaging over cubes in the Whitney decomposition. I believe that by modifying the definition of g_t on E_t^c, we can, in fact find a Lipschitz function g_t which approximates the given function f in the Luzin sense, in terms of L^1 convergence, and also $\|Dg_t\|(\mathbb{R}^n) \to \|Df\|(\mathbb{R}^n)$ as $t \to \infty$.

Such an approximation will be more difficult to construct when dealing with C^1 approximations, and ever moreso for higher order functions of bounded variation, but I believe it is possible and I am currently working on this issue.

3.4 APPROXIMATION BY C^1 FUNCTIONS

The approximation of BV functions by C^1 functions will result from an application of the Whitney Extension Theorem (Theorem 23), and this requires formulation of certain Taylor remainder type estimates. We will provide a sketch of the proof, which will motivate and clarify the proof in the higher order case. We start by proving a potential estimate similar to that developed for Lipschitz functions.

Theorem 28. *Let $f \in BV_{\text{loc}}(\mathbb{R}^n)$. Then for any cube Q,*

$$\left| f(x) - f_Q - \fint_Q (x-z) \cdot d\, Df(z) \right| \leq C(n) \int_Q \frac{d\|Df - \vec{a}\mathcal{L}^n\|(z)}{|x-z|^{n-1}} \tag{3.13}$$

holds for each $x \in Q$ and each vector $\vec{a} \in \mathbb{R}^n$.

Proof. Fix $x \in \mathbb{R}^n$, a cube Q containing x, and a vector $\vec{a} \in \mathbb{R}^n$.

Note that by Lemma 5 with $m = 1$ (proven in section 4.2) that for any $f_k \in C^1(Q)$,

$$\left| f_k(y) - (f_k)_Q - \fint_Q (y-z) \cdot \nabla f_k(z) dz \right| \leq C \int_Q \frac{|\nabla f_k(z) - \vec{a}|}{|y-z|^{n-1}}\, dz$$

holds for each $y \in Q$.

We want to find a sequence $\{f_k\} \in C^\infty(Q) \cap BV(Q)$ such that the vector measures and total variations

$$Df_k \overset{*}{\rightharpoonup} Df \quad \text{and} \quad \|Df_k - \vec{a}\mathcal{L}^n\| \overset{*}{\rightharpoonup} \|Df - \vec{a}\mathcal{L}^n\| \tag{3.14}$$

weakly in the sense of Radon measures.

Let A be a linear function with $DA = \vec{a}$. Since A is smooth and locally integrable, $A \in BV(Q)$. By Theorem 4.4, there exists a sequence $\{g_k\} \subset C^\infty(Q) \cap BV(Q)$ such that

$$g_k \to (f - A) \text{ in } \mathcal{L}^1(Q) \qquad \text{and } \|Dg_k\|(Q) \to \|D(f-A)\|(Q) = \|Df - \vec{a}\mathcal{L}^n\|(Q).$$

Let $f_k = g_k + A$. Then $f_k \in C^\infty(Q)$ and $D(f_k - A) = D(g_k)$. So

$$f_k \to f \text{ in } L^1(\Omega) \qquad \text{and } \|Df_k - \vec{a}\mathcal{L}^n\|(Q) \to \|Df - \vec{a}\mathcal{L}^n\|(Q).$$

This guarantees the weak-* convergence (3.14).

Note that the functions

$$z \mapsto \int_Q \frac{dy}{|y-z|^{n-1}} \qquad \text{and } z \mapsto \int_Q (y-z)dy$$

are continuous on Q and can be extended to continuous compactly supported functions on $\mathbb{R}^n$.

For any $\epsilon > 0$, average the inequality (3.13) over a ball $B(x,\epsilon)$

$$\left| \fint_{B(x,\epsilon)} \left(f_k(y) - (f_k)_Q - \fint_Q (y-z) \cdot \nabla f_k(z) dz \right) dy \right| \le C \int_Q \left(\fint_{B(x,\epsilon)} |y-z|^{1-n} \, dy \right) |\nabla f_k(z) - \vec{a}| dz$$

Applying the weak-* convergence of the measures and Fubini Theorem implies

$$\begin{aligned} \left| \fint_{B(x,\epsilon)} \left(f(y) - f_Q - \fint_Q (y-z) \cdot Df(z) \right) dy \right| &\le C \int_Q \left(\fint_{B(x,\epsilon)} \frac{1}{|y-z|^{n-1}} dy \right) d\|Df - \vec{a}\mathcal{L}^n\|(z) \\ &\le C \int_Q \frac{d\|Df - \vec{a}\mathcal{L}^n\|(z)}{|x-z|^{n-1}}, \end{aligned}$$

the last step applying Lemma 3.

Taking the limsup as $\epsilon \to 0$ on the left hand size, and noting that

$$\fint_{B(x,\epsilon)} (y-z) \, dy \to (x-z) \text{ as } \epsilon \to 0$$

completes the argument. □

Define

$$T_Q^1 f(y) = f_Q + \fint_Q (y-z) \cdot d\, Df(z),$$

$$T_x^1 f(y) = f(x) + \nabla f(x) \cdot (y-x),$$

recalling that ∇f is the density of Df_{ac} with respect to Lebesgue measure.

Theorem 29. *If $f \in BV(\mathbb{R}^n)$ then for any $x, y \in \mathbb{R}^n$,*

$$\frac{|f(y) - T_x^1 f(y)|}{|x-y|} \le C(n) \big(\mathcal{M}_{|x-y|} \|Df - \nabla f(x)\mathcal{L}^n\|(x) + \mathcal{M}_{|x-y|} \|Df - \nabla f(y)\mathcal{L}^n\|(y) \big)$$

Proof. $T^1_Q f(y)$ is a polynomial in y. Indeed, differentiating under the integral sign, the gradient is constant:

$$\partial_i T^1_Q f = \partial_i \fint_Q (y-z)\cdot d\,Df(z) = \sum_j \fint_Q \delta_{ij} d\mu^j(z) = \fint_Q d\mu^i(z).$$

So $T^1_Q f$ is equal to its first degree Taylor polynomial centered at x.

$$\begin{aligned} T^1_Q f(y) &= T^1_Q f(x) + \sum_{i=1}^n \big(\fint_Q d\mu^i(z)\big)(y_i - x_i) \\ &= T^1_Q f(x) + \frac{Df(Q)}{|Q|}\cdot (y-x). \end{aligned}$$

By the triangle inequality,

$$\begin{aligned} |f(y) - T^1_x f(y)| &\le |f(y) - T^1_Q f(y)| + |T^1_Q f(y) - T^1_x f(y)| \\ &\le |f(y) - T^1_Q f(y)| + |f(x) - T^1_Q f(x)| + \left|\Big(\frac{Df(Q)}{|Q|} - \nabla f(x)\Big)\cdot (y-x)\right| \\ &\le C\int_Q \frac{d\|Df - \vec{a}\mathcal{L}^n\|(z)}{|x-z|^{n-1}} + C\int_Q \frac{d\|Df - \vec{b}\mathcal{L}^n\|(z)}{|y-z|^{n-1}} dy \\ &\quad + \left|\frac{Df(Q)}{|Q|} - \nabla f(x)\right| |y-x|. \end{aligned}$$

Now select the vectors $\vec{a} = \nabla f(x)$ and $\vec{b} = \nabla f(y)$. It follows from Lemma 4 that

$$\int_Q \frac{d\|Df - \vec{a}\mathcal{L}^n\|(z)}{|x-z|^{n-1}} \le C(\mathrm{diam}(Q))\mathcal{M}_{\mathrm{diam}Q}\|Df - \vec{a}\mathcal{L}^n\|(x).$$

Divide both sides of the inequality by $|y-x|$ and note for any $\epsilon > 0$ we can choose Q containinig x, y with $\mathrm{diam}(Q) < |x-y| + \epsilon$. To complete the proof, estimate

$$\begin{aligned} \left|\frac{Df(Q)}{|Q|} - \nabla f(x)\right| &= \frac{|Df(Q) - \nabla f(x)\mathcal{L}^n(Q)}{|Q|} \\ &\le \frac{\|Df - \nabla f(x)\mathcal{L}^n\|(Q)}{|Q|} \\ &\le C\mathcal{M}_{\mathrm{diam}(Q)}\|Df - \nabla f(x)\mathcal{L}^n\|(x). \end{aligned}$$

□

We will now use our pointwise inequality to show that for any $\epsilon > 0$, there is a closed set C with $|\mathbb{R}^n \backslash C| < \epsilon$ and such that the jet $\{f, \nabla f\}$ is a Whitney jet on C. This implies the existence of C^1 extension, say g, of $f|_C$ with $g = f$ and $\nabla g = \nabla f$ on C.

Theorem 30. *Let $f \in BV(\mathbb{R}^n)$. Then for each $\epsilon > 0$, there is a function $g \in C^1(\mathbb{R}^n)$ such that*

$$|\{f \neq g\}| < \epsilon.$$

Moreover $|\{\nabla g \neq \nabla f\}| < \epsilon$, where ∇f is the absolutely continuous density of the derivative measure Df.

Proof. Define

$$\eta_k(z) = \sup_{r<\frac{1}{k}} \mathcal{M}_r \|Df - \nabla f(z)\mathcal{L}^n\|(z)$$

$$\text{Claim 1: } \eta_k(z) \to 0 \text{ for almost every } z \in \mathbb{R}^n. \tag{3.15}$$

To prove this, note that since the maximal function is subadditive,

$$\begin{aligned}\mathcal{M}_\delta\|Df - \nabla f(x)\mathcal{L}^n\|(x) &= \mathcal{M}_\delta\|Df_s + Df_{ac} - \nabla f(x)\mathcal{L}^n\|(x) \\ &\leq \mathcal{M}_\delta\|Df_s\|(x) + \mathcal{M}_\delta\|Df_{ac} - \nabla f(x)\mathcal{L}^n\|(x) \to 0 \text{ as } \delta \to 0\end{aligned}$$

for almost every $x \in \mathbb{R}^n$. Indeed, convergence of the first term is an immediate consequence of the Besicovitch Differentiation Theorem. The second term converges to zero whenever x is a Lebesgue point of ∇f.

Fix $\epsilon > 0$. By Lusin Theorem, there is a closed set $\widetilde{C} \subset \mathbb{R}^n$ with $|\mathbb{R}^n \setminus \widetilde{C}| < \epsilon/2$ such that $f|_{\widetilde{C}}$, $Df|_{\widetilde{C}}$ are continuous.

Claim (3.15) implies that by Egorov Theorem, there exists a closed set $C \subset \widetilde{C}$ such that $|\widetilde{C} \setminus C| \leq \epsilon/2$ and

$$\eta_k \rightrightarrows 0 \text{ uniformly on compact subsets of } C$$

Claim 2: Let K be a compact subset of C.

$$\sup_{x,y\in K, |x-y|<\rho} \frac{|f(y) - T_x^1 f(y)|}{|x-y|} \to 0 \text{ as } \rho \to 0.$$

Fix $0 < \rho < 1$. Choose $k(\rho)$ such that $\frac{1}{k+1} < \rho < \frac{1}{k}$. Fix $x, y \in K$ with $|x - y| < \rho$. Let Q contain x, y with $\operatorname{diam}(Q) = r < \frac{1}{k}$. Then

$$\begin{aligned}\mathcal{M}_r \|Df - \nabla f(x)\mathcal{L}^n\|(x) + \mathcal{M}_r \|Df - \nabla f(y)\mathcal{L}^n\|(y) \\ \leq 2 \sup_{z \in K, r < \frac{1}{k}} \mathcal{M}_r \|Df - \nabla f(z)\mathcal{L}^n\|(z).\end{aligned}$$

Hence

$$\frac{|f(y) - T_x^1 f(y)|}{|x - y|} \leq C \sup_{z \in K} \eta_k(z)$$

Take the supremum over all $x, y \in K$, $|x - y| < \rho$, and let $\rho \to 0$. Then $k \to \infty$, and uniform convergence of η_k to 0 on K implies Claim 2.

Claim 2 simply says that the formal Taylor remainder of degree 1 for the jet $\{f, \nabla f\}$ converges uniformly to 0 on compact subsets $K \subset C$. Thus the jet $\{f, \nabla f\}$ is a Whitney jet on C. By the Whitney Extension Theorem, there exists a function $g \in C^1(\mathbb{R}^n)$ such that $g|_C = f|_C$ and $\nabla g|_C = \nabla f|_C$. Since $|\mathbb{R}^n \backslash C| < \epsilon$, this proves the theorem.

$\square$

4.0 HIGHER ORDER FUNCTIONS OF BOUNDED VARIATION

Let $\Omega \subset \mathbb{R}^n$ be open. We say a function f is in $BV^m(\Omega)$ if f is in $W^{m-1,1}(\Omega)$ and in addition the mth order distributional derivatives of f are Radon measures of finite total variation. This is a natural definition because in order for the mth order distibutional derivatives to be measures, we must require that the $(m-1)$st order derivatives exist in the weak sense and be locally integrable. When $f \in BV^m(\Omega)$, for each multi-index α of order m, there exists a signed Radon measure μ^α such that

$$\int_\Omega f\, D^\alpha \varphi\, dx = (-1)^m \int_\Omega \varphi\, d\mu^\alpha \qquad \text{for each } \varphi \in C_c^m(\Omega) \tag{4.1}$$

and $|\mu^\alpha|(\Omega) < \infty$.

When $f \in BV^m$ we will denote the distributional derivative μ^α by $D^\alpha f$.

Similarly we define $BV^m_{\text{loc}}(\Omega)$ to be the space of functions $f \in W^{m,1}_{\text{loc}}(\Omega)$ whose mth order distributional derivatives are signed Radon measures (whose total variation may not in general be finite).

The mth derivate of f is a vector-valued measure $D^m f$ defined by

$$D^m f := \{D^\alpha f\}_{|\alpha|=m}.$$

Its total variation will be denoted $\|D^m f\|$, and as for any vector-valued measure,

$$\|D^m f\|(A) := \sup \left\{ \sum_{k=0}^{\infty} |D^m f(A_k)| : A_k \text{ are disjoint and } A = \bigcup_{k=0}^{\infty} A_k \right\}$$

for any Borel set A.

$BV^m(\Omega)$ is a Banach space with the norm

$$\|f\|_{BV^m(\Omega)} = \|f\|_{W^{m-1,1}(\Omega)} + \|D^m f\|(\Omega).$$

The proof follows from weak lowersemicontinuity of the total variation with respect to L^1 convergence, and is analogous to the proof of Theorem 16.

$\|D^m f\|$ has a polar decomposition $D^m f = \sigma \|D^m f\|$, where $|\sigma(x)| = 1$ for $\|D^m f\|$-a.e. $x \in \Omega$. If we let $M := \#\{\alpha \in \mathbb{N}^n : |\alpha| = m\}$, from Theorem 4

$$\|D^m f\|(\Omega) = \sup \left\{ \sum_{|\alpha|=m} \int_\Omega \phi_\alpha d(D^\alpha f) : \phi \in C_c^m(\Omega, \mathbb{R}^M), |\phi| \leq 1 \right\}, \tag{4.2}$$

or equivalently,

$$\|D^m f\|(\Omega) = \sup \left\{ \int_\Omega f \Big(\sum_{|\alpha|=m} D^\alpha \phi_\alpha \Big)\, dx : \phi \in C_c^m(\Omega, \mathbb{R}^M), |\phi| \leq 1 \right\}. \tag{4.3}$$

Analogous to Theorem 13 for the case $m = 1$, $f \in BV^m(\Omega)$ if and only if $f \in W^{m-1,1}(\Omega)$ and the right hand side of equation (4.3) is finite.

4.1 BASIC PROPERTIES

Theorem 31. *Let $f_k \in BV^m(\Omega)$, and suppose $f_k \to f$ in $L^1(\Omega)$.*
Then $\|D^m f\|(\Omega) \leq \liminf_{k \to 0} \|D^m f_k\|(\Omega)$.

Proof. Let $\psi \in C_c^m(\Omega, \mathbb{R}^M)$ with $\|\psi\|_\infty \leq 1$. Then for $f, \{f_k\}$ as above,

$$\begin{aligned}
\int_\Omega f \big(\sum_{|\alpha|=m} D^\alpha \psi_\alpha \big) dx &= \lim_{k\to\infty} \int_\Omega f_k \big(\sum_{|\alpha|=m} D^\alpha \psi_\alpha \big) dx \\
&= \lim_{k\to\infty} (-1)^m \int_\Omega D^m f_k \cdot \varphi dx \\
&\leq \liminf_{k\to\infty} \|D^m f_k\|(\Omega)
\end{aligned}$$

Then taking the supremum over all such ψ, the result follows from (4.3). □

It is easy to check that

$$W^{m,1}(\Omega) \subsetneq BV^m(\Omega)$$

is an closed isometrically embedded subspace. So as in the case $m = 1$, we cannot approximate BV^m functions by C^∞ functions in $(BV^m, \|\cdot\|_{BV^m})$. However, we have a result similar to Theorem 17 for BV^m functions, which says that we can approximate by smooth functions in the following weaker sense:

Theorem 32. *Let $f \in BV^m(\Omega)$. There exists a sequence $\{f_k\}$ in $C^\infty(\Omega) \cap BV^m(\Omega)$ such that*

$$f_k \to f \text{ in } W^{m-1,1}(\Omega) \qquad \text{and} \qquad \|D^m f_k\|(\Omega) \to \|D^m f\|(\Omega) \tag{4.4}$$

Proof. Fix $\epsilon > 0$. Let $U_{i,\ell} = \{x \in \Omega : dist(\partial\Omega, x) > \frac{1}{\ell+i}\} \cap B(0, i+\ell)$.

$\bigcup_{\ell=1}^{\infty} U_{1,\ell} = \Omega$, and $U_{1,\ell} \subset U_{1,\ell+1}$, so $\|D^m f\|(\Omega) = \lim_{l\to\infty} \|D^m f\|(U_{1,\ell})$.

Since $\|D^m f\|(\Omega) < \infty$, we can choose ℓ large so that $\|D^m f\|(\Omega \backslash U_{1,\ell}) < \epsilon$. Fix such an l, and let $U_i := U_{i,\ell}$. Then

$$\|D^m f\|(\Omega \backslash U_1) < \epsilon, \qquad U = \bigcup_{i=1}^{\infty} U_i$$

Let $V_1 = \emptyset$, and let $V_k = U_{k+1} \backslash \bar{U}_k$ for $k > 1$. Let ζ_k be a partition of unity subordinate to V_k.

Let η be a standard mollifier on $\mathbb{R}^n$, i.e. η is a function symmetric about the origin such that

$$\text{supp}(\eta) \subset B(0,1), \qquad \int_{B(0,1)} \eta\, dx = 1, \qquad 0 \le \eta \le 1,$$

and let $\eta_\delta = (\delta)^{-n}\eta(x/\delta)$. Let $\gamma(\epsilon) > 0$ be a constant to be later determined. It follows from standard theory for mollifiers that

(i) $\text{supp}(\eta_{\delta_k} * f\zeta_k) \subset V_k$ for δ_k small.

(ii) Since $f\zeta_k \in W^{m-1,1}(\Omega)$, for every β with $|\beta| \le m-1$,

$$\|D^\beta(f\zeta_k) * \eta_{\delta_k} - D^\beta(f\zeta_k)\|_{L^1(\Omega)} < \frac{\epsilon}{2^k} \tag{4.5}$$

for δ_k sufficiently small.

(iii) Since $fD^{\alpha-\beta}\zeta_k \in W^{m-1,1}(\Omega)$, for every α with $|\alpha| = m$, and every $\beta < \alpha$,

$$\|D^\beta(fD^{\alpha-\beta}\zeta_k) * \eta_{\delta_k} - D^\beta(fD^{\alpha-\beta}\zeta_k)\|_{L^1(\Omega)} \leq \frac{\gamma(\epsilon)}{2^k} \tag{4.6}$$

for δ_k sufficiently small.

For each integer k, choose δ_k small so that all of the finitely many conditions (i)-(iii) are satisfied. For each k, we will write $\eta_k := \eta_{\delta_k}$

Let

$$f^\epsilon = \sum_{k=0}^{\infty} f\zeta_k * \eta_k$$

Since the partition of unity is locally finite, this is a finite sum in some neighborhood of each $x \in \Omega$, so $f^\epsilon \in C^\infty(\Omega)$.

We have $\|f^\epsilon - f\|_{W^{m-1,1}} < \epsilon$. Indeed,

$$\begin{aligned}
\|D^\beta f^\epsilon - D^\beta f\|_1 &= \|D^\beta\big(\sum_{k=1}^{\infty} \eta_k * f\zeta_k\big) - D^\beta\big(\sum_{k=1}^{\infty} f\zeta_k\big)\|_1 \\
&\leq \sum_{k=1}^{\infty} \|\eta_k * D^\beta(f\zeta_k) - D^\beta(f\zeta_k)\|_1 \\
&< \sum_{k=1}^{\infty} \frac{\epsilon}{2^k} = \epsilon \text{ by } (4.5)
\end{aligned}$$

By Theorem 31, this implies

$$\|D^m f\|(\Omega) \leq \liminf_{\epsilon \to 0} \|D^m f^\epsilon\|(\Omega) \tag{4.7}$$

We will frequently use the following fact: If $g \in W^{k,p}(\Omega)$ and $|\beta| \leq k$, then

$$D^\beta(g * \eta_k) = D^\beta g * \eta_k = g * D^\beta \eta_k \tag{4.8}$$

Fix $\Psi = \{\psi_\alpha\}_{|\alpha|=m} \in C_c^\infty(\Omega; \mathbb{R}^M)$ with $\|\Psi\|_\infty \leq 1$.

$$
\begin{aligned}
I^\epsilon := \int f^\epsilon\Big(\sum_{|\alpha|=m} D^\alpha \psi_\alpha\Big) dx &= \sum_{k=1}^{\infty}\sum_{|\alpha|=m}\int (f\zeta_k * \eta_k)(x)\, D^\alpha\psi_\alpha(x)\, dx \\
&= \sum_{k=1}^{\infty}\sum_{|\alpha|=m}\int D^\alpha\psi_\alpha(x)\Big(\int f\zeta_k(y)\eta_k(y-x)\, dy\Big)\, dx \\
&= \sum_{k=1}^{\infty}\sum_{|\alpha|=m}\int f\zeta_k(y)\Big(\int \eta_k(y-x) D^\alpha\psi_\alpha(x) dx\Big)\, dy \\
&= \sum_{k=1}^{\infty}\sum_{|\alpha|=m}\int f\zeta_k(y)\, \big(\eta_k * D^\alpha\psi_\alpha\big)(y)\, dy \\
&= \sum_{k=1}^{\infty}\sum_{|\alpha|=m}\int f\zeta_k(y)\, D^\alpha\big(\eta_k * \psi_\alpha\big)(y)\, dy
\end{aligned}
$$

Here we have moved sums outside the integral, used Fubini Theorem, symmetry of η, and used property (4.8).

Now by Leibniz Rule, if $g \in W^{m-1,1}(\Omega)$ and $h_\alpha \in C^\infty(\Omega)$

$$
gD^\alpha h_\alpha = D^\alpha\big(gh_\alpha\big) - \sum_{\beta<\alpha} C_{\alpha,\beta} D^{\alpha-\beta} g D^\beta h_\alpha
$$

Substituting $h_\alpha = \eta_k * \psi_\alpha$ and $g = \zeta_k$, we have

$$
\begin{aligned}
I^\epsilon &= \sum_{k=1}^{\infty}\sum_{|\alpha|=m}\int f\zeta_k\, D^\alpha\big(\eta_k * \psi_\alpha\big)\, dy \\
&= \sum_{k=1}^{\infty}\sum_{|\alpha|=m}\Big[\int fD^\alpha\big(\zeta_k(\eta_k*\psi_\alpha)\big)\, dy - \sum_{\beta<\alpha} C_{\alpha,\beta}\int fD^{\alpha-\beta}\zeta_k D^\beta(\eta_k*\psi_\alpha)\, dy\Big]
\end{aligned}
$$

Note also that since η is symmetric, $D^\beta\eta_k(y-x) = (-1)^{|\beta|}D^\beta\eta_k(x-y)$, so

$$
\begin{aligned}
\int fD^{\alpha-\beta}\zeta_k D^\beta(\eta_k*\psi_\alpha)\, dy &= \int fD^{\alpha-\beta}\zeta_k(D^\beta\eta_k*\psi_\alpha)\, dy \\
&= \int (-1)^{|\beta|}\psi_\alpha(D^\beta\eta_k * fD^{\alpha-\beta}\zeta_k)\, dy \\
&= (-1)^{|\beta|}\int \psi_\alpha(\eta_k * D^\beta(fD^{\alpha-\beta}\zeta_k))\, dy
\end{aligned}
$$

For $\beta < \alpha$, $\sum_{k=1}^{\infty} D^{\alpha-\beta}\zeta_k = D^{\alpha-\beta}(1) = 0$. Therefore

$$\sum_{k=1}^{\infty} \psi_\alpha D^\beta(fD^{\alpha-\beta}\zeta_k) = \psi_\alpha D^\beta\Big(\sum_{k=1}^{\infty} fD^{\alpha-\beta}\zeta_k\Big) = 0$$

in which case

$$\begin{aligned} I^\epsilon = &\sum_{k=1}^{\infty}\sum_{|\alpha|=m}\Big(\int fD^\alpha\big(\zeta_k(\eta_k * \psi_\alpha)\big)\, dy\Big) \\ &\quad - \sum_{k=1}^{\infty}\sum_{|\alpha|=m}\sum_{\beta<\alpha} C_{\alpha,\beta}(-1)^{|\beta|}\int \psi_\alpha\big(\eta_k * D^\beta(fD^{\alpha-\beta}\zeta_k) - D^\beta(fD^{\alpha-\beta}\zeta_k)\big)\, dy\Big) \\ = &\; I_1^\epsilon + I_2^\epsilon \end{aligned}$$

Since $\|\zeta_k\big(\eta_k * \Psi\big)\|_\infty \leq 1$, $\zeta_k\big(\eta_k * \Psi\big) \in C^\infty(\Omega, \mathbb{R}^M)$, and each point in Ω belongs to at most three of the sets $\{V_k\}_{k=1}^{\infty}$,

$$\begin{aligned} |I_1^\epsilon| &\leq \int f \sum_{|\alpha|=m} D^\alpha\big(\zeta_1(\eta_1 * \psi_\alpha)\big)\, dy + \sum_{k=2}^{\infty}\int f \sum_{|\alpha|=m} D^\alpha\big(\zeta_k(\eta_k * \psi_\alpha)\big)\, dy \\ &\leq \|D^m f\|(\Omega) + 3\|D^m f\|(\Omega\backslash U_1) \\ &\leq \|D^m f\|(\Omega) + 3\epsilon. \end{aligned}$$

Also by condition (iii) above, selecting $\gamma(\epsilon) < \epsilon\Big(\sum_{|\alpha|=m}\sum_{\beta<\alpha} C_{\alpha,\beta}\Big)^{-1}$,

$$\begin{aligned} |I_2^\epsilon| &\leq \sum_{k=1}^{\infty}\sum_{|\alpha|=m}\sum_{\beta<\alpha} C_{\alpha,\beta}\int \big|\eta_k * D^\beta(fD^{\alpha-\beta}\zeta_k) - D^\beta(fD^{\alpha-\beta}\zeta_k)\big|\, dy \\ &\leq \sum_{k=1}^{\infty}\sum_{|\alpha|=m}\sum_{\beta<\alpha} C_{\alpha,\beta}\frac{\gamma(\epsilon)}{2^k} < \epsilon. \end{aligned}$$

Hence

$$\int_\Omega f^\epsilon\Big(\sum_{|\alpha|=m} D^\alpha\psi_\alpha\Big)dx \leq \|D^m f\|(\Omega) + 4\epsilon$$

and therefore

$$\|D^m f^\epsilon\|(\Omega) \leq \|D^m f\|(\Omega) + 4\epsilon$$

Now the proof is complete comparing with (4.7). □

The following theorems assert weak-$*$ convergence of the distributional measures and total variation measures when f_k is an smooth approximating sequence of $f \in BV^m(\Omega)$.

Theorem 33 (Weak Convergence of Total Variation)**.** *Let $f \in BV^m(\Omega)$ and let f_k be a smooth approximating sequence satisfying equation (4.4). Let μ_k and μ be the extensions of the measures $\|D^m f_k\|$ and $\|D^m f\|$ by zero, i.e. for any Borel set A*

$$\mu_k(A) = \|D^m f_k\|(A \cap \Omega)$$

$$\mu(A) = \|D^m f\|(A \cap \Omega).$$

Then

$$\mu_k \overset{*}{\rightharpoonup} \mu$$

weakly in the sense of Radon measures on $\mathbb{R}^n$. In particular, for every $\phi \in C_c(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} \phi \, d\mu_k \to \int_{\mathbb{R}^n} \phi \, d\mu$$

First we prove the following simple proposition:

Proposition 1. *Let $\{a_i\}$, $\{b_i\}$, $\{c_i\}$ be sequences in $\mathbb{R}$ with $a_i = b_i + c_i$ for each i, and assume $\{a_i\}$ is convergent and $\{c_i\}$ bounded below. Then*

$$\lim_{i\to\infty} a_i \geq \limsup_{i\to\infty} b_i + \liminf_{i\to\infty} c_i.$$

Proof. Let $L = \lim_{i\to\infty} a_i$. Fix $\epsilon > 0$, and choose k large so $n \geq k$ implies $a_n < L + \epsilon$. Then

$$b_n = a_n - c_n < L + \epsilon - c_n < L + \epsilon - \inf_{n\geq k} c_n \quad \text{for each } n \geq k.$$

So $\sup_{n\geq k} b_n < L + \epsilon - \inf_{n\geq k} c_n$. Letting $k \to \infty$,

$$\limsup_{i\to\infty} b_i + \liminf_{i\to\infty} c_i < L + \epsilon.$$

Since ϵ is arbitrary, this proves the proposition. □

Now we prove the theorem:

Proof. Let $f_k \in C^\infty$ be a sequence satisfying (4.4). Let $C \subset \mathbb{R}^n$ be closed.

Claim: $\limsup_{k\to\infty} \|D^m f_k\|(C\cap\Omega) \le \|D^m f\|(C\cap\Omega)$.

Let $V = \Omega\backslash C$. Then V is open, so $\|D^m f\|(V) \le \liminf_{k\to\infty} \|D^m f_k\|(V)$, hence

$$\begin{aligned}\|D^m f\|(C\cap\Omega) &= \|D^m f\|(\Omega) - \|D^m f\|(V)\\ &= \lim_{k\to\infty} \|D^m f_k\|(\Omega) - \|D^m f\|(V)\\ &\ge \liminf_{k\to\infty} \|D^m f_k\|(V) + \limsup_{k\to\infty} \|D^m f_k\|(C\cap\Omega) - \|D^m f\|(V)\\ &\ge \limsup_{k\to\infty} \|D^m f_k\|(C\cap\Omega). \end{aligned} \tag{4.9}$$

Let $K \subset \mathbb{R}^n$ be compact, and let $U \subset \mathbb{R}^n$ be open. The previous calculation shows

$$\limsup_{k\to\infty} \mu_k(K) = \limsup_{k\to\infty} \|D^m f_k\|(K\cap\Omega) \le \|D^m f\|(K\cap\Omega) = \mu(K)$$

Since $f_k \to f$ in $L^1(\Omega)$

$$\mu(U) = \|D^m f\|(U\cap\Omega) \le \liminf_{k\to\infty} \|D^m f_k\|(U\cap\Omega) = \liminf_{k\to\infty} \mu_k(U)$$

by weak lowersemicontinuity.

By Theorem 1.9.1 in [9], this is sufficient to prove the convergence $\mu_k \overset{*}{\rightharpoonup} \mu$.

□

Corollary 6. *Let Ω as above, and suppose Ω is bounded. Let $g \in C(\bar{\Omega})$. Then*

$$\int_\Omega g\, d\|D^m f_k\| \to \int_\Omega g\, d\|D^m f\|$$

This follows easily by extending g to a $C_c(\mathbb{R}^n)$ function.

Theorem 34 (Weak Convergence of Derivatives)**.** *Let $f \in BV^m(\Omega)$, and let $\{f_k\}$ be as above. Define μ_k and μ to be the extensions of the vector-valued measures $D^m f_k$ and $D^m f$, respectively. Then $\mu_k \overset{*}{\rightharpoonup} \mu$ in the sense of weak convergence of vector-valued measures, i.e. for each $\varphi \in C_c(\mathbb{R}^n, \mathbb{R}^M)$,*

$$\int \varphi \cdot d\mu_k \to \int \varphi \cdot d\mu \quad \textit{as } k \to \infty$$

Proof. Fix $\epsilon > 0$. Choose $U_1 \subset \Omega$ large enough so that $\|D^m f\|(\Omega \backslash U_1) < \epsilon$.

Let $\zeta \in C_c^\infty(\Omega)$ with $0 \le \zeta \le 1$ and $\zeta = 1$ on U_1. Let $\varphi \in C_c^\infty(\mathbb{R}^n; \mathbb{R}^M)$. The idea is that $(\varphi\zeta)$ will have compact support in Ω, and $(1-\zeta)|\varphi| \le |\varphi|\chi_{\Omega\backslash U_1}$.

$$\begin{aligned}
&\left| \int_\Omega \varphi \cdot D^m f_k \, dx - \int_\Omega \varphi \cdot d(D^m f) \right| \\
= &\left| \int_\Omega (\varphi\zeta) \cdot D^m f_k \, dx - \int_{\Omega\backslash U_1} (1-\zeta)\varphi \cdot D^m f_k \, dx - \int_\Omega (\varphi\zeta) \cdot d(D^m f) - \int_{\Omega\backslash U_1} (1-\zeta)\varphi \cdot d(D^m f) \right| \\
\le &\left| \int_\Omega \sum_{|\alpha|=m} D^\alpha(\varphi\zeta) f_k \, dx - \int_\Omega \sum_{|\alpha|=m} D^\alpha(\varphi\zeta)\, f \, dx \right| + \|\varphi\|_\infty \big(\|D^m f_k\|(\Omega\backslash U_1) + \|D^m f\|(\Omega\backslash U_1) \big)
\end{aligned}$$

Note that $\limsup_{k\to\infty} \|D^m f_k\|(\Omega\backslash U_1) \le \|D^m f\|(\Omega\backslash U_1)$ by (4.9). Also, $f_k \to f$ in $L^1(\Omega)$, so for k large, we have

$$\left| \int_\Omega \sum_{|\alpha|=m} D^\alpha(\varphi\zeta) f_k \, dx - \int_\Omega \sum_{|\alpha|=m} D^\alpha(\varphi\zeta)\, f \, dx \right| < \epsilon$$

and

$$\|D^m f_k\|(\Omega\backslash U_1) + \|D^m f\|(\Omega\backslash U_1) < 2\|D^m f\|(\Omega\backslash U_1) + \epsilon < 3\epsilon.$$

Hence

$$\left| \int_\Omega \varphi \cdot D^m f_k \, dx - \int_\Omega \varphi \cdot d(D^m f) \right| \le C\epsilon \|\varphi\|_\infty, \tag{4.10}$$

which completes the proof in the case where φ is smooth.

Now an approximation argument shows this is true for $\varphi \in C_c(\mathbb{R}^n, \mathbb{R}^M)$. Let $\varphi \in C_c(\mathbb{R}^n)$, let $\operatorname{supp}(\varphi) \subset K$, and let $\varphi_k = \varphi * \eta_{(1/k)}$. Then $\operatorname{supp}(\varphi_k) \subset K + \bar{B}(0,1)$, which is compact, and

$$\varphi_k \to \varphi \text{ uniformly in } K + \bar{B}(0,1) \text{ as } k \to \infty$$

For $k \ge N$ large,

$$\|D^m f_k\|(\Omega) \le \|D^m f\|(\Omega) + \epsilon \quad \text{and} \quad \|D^m f_k\|(\Omega\backslash U_1) \le \|D^m f\|(\Omega\backslash U_1) + \epsilon.$$

Then

$$\begin{aligned}
&\left|\int_\Omega \varphi \cdot D^m f_k \, dx - \int_\Omega \varphi \cdot d(D^m f)\right| \\
&\leq \left|\int_\Omega (\varphi - \varphi_j) \cdot D^m f_k \, dx\right| + \left|\int_\Omega (\varphi - \varphi_j) \cdot d(D^m f)\right| + \left|\int_\Omega \varphi_j D^m f_k \cdot dx - \int_\Omega \varphi_j \cdot d(D^m f)\right| \\
&\leq \|\varphi - \varphi_j\|_\infty \big(\|D^m f_k\|(\Omega) + \|D^m f\|(\Omega)\big) + \left|\int_\Omega \varphi_j \cdot D^m f_k \, dx - \int_\Omega \varphi_j \cdot d(D^m f)\right|
\end{aligned}$$

Now by (4.10), for $k \geq N$ large, we have the bound

$$\left|\int_\Omega \varphi_j D^m f_k \, dx - \int_\Omega \varphi_j \, d(D^m f)\right| \leq C\|\varphi_j\|_\infty \|D^m f\|(\Omega)$$

So for $k \geq N$,

$$\begin{aligned}
\left|\int_\Omega \varphi \, D^m f_k \, dx - \int_\Omega \varphi \, d(D^m f)\right| &\leq \lim_{j\to\infty} C\|\varphi - \varphi_j\|_\infty \|D^m f\|(\Omega) + C\|\varphi_j\|_\infty \|D^m f\|(\Omega \backslash U_1) \\
&\leq C\|\varphi\|_\infty \|D^m f\|(\Omega \backslash U_1) \leq C\epsilon \|\varphi\|_\infty
\end{aligned}$$

□

4.2 POINTWISE ESTIMATES

To develop pointwise estimates similar to the case $m = 1$, we will apply Lemma 5 and a similar approximation argument. However, the computations are more involved.

For notational simplicity, we will denote the density of the absolutely continuous part of $D^m f$ with respect to the Lebesgue measure by $\nabla^m f = \{\nabla^\beta f\}_{|\beta|=m}$.

Henceforth, we will utilize the following pointwise representatives at <u>every</u> x:

$$D^\beta f(x) = \limsup_{r\to 0} \fint_{B(x,r)} D^\beta f(y) dy \quad \text{for } |\beta| \leq m-1.$$

$$\nabla^\beta f(x) = \limsup_{r\to 0} \fint_{B(x,r)} \nabla^\beta f(y) dy \quad \text{for } |\beta| = m.$$

For $x \in \mathbb{R}^n$, define

$$T_x^m f(y) := \sum_{|\beta|\leq m-1} \frac{(y-x)^\beta}{\beta!} D^\beta f(x) + \sum_{|\beta|=m} \nabla^\beta f(x) \frac{(y-x)^\beta}{\beta!} \tag{4.11}$$

For any cube $Q \subset \mathbb{R}^n$, define

$$T_Q^m f(y) := \sum_{|\beta| \le m-1} \fint_Q D^\beta f(z) \frac{(y-z)^\beta}{\beta!} \, dz + \sum_{|\beta|=m} \fint_Q \frac{(y-z)^\beta}{\beta!} \, d(D^\beta f)(z) \tag{4.12}$$

$T_x^m f(y)$ is a formal Taylor expansion of f of degree m centered at x, constructed from weak derivatives of f for $|\beta| < m$ and the densities $\nabla^\beta f$ for $|\beta| = m$.

$T_Q^m f(y)$ is *nearly* an averaging of the Taylor expansion $T_z^m f(y)$ over $z \in Q$, except that we integrate the mth degree terms with respect to $D^m f$, hence creating a disparity arising from the singular part $(D^m f)_s$ of the measure. Luckily when approximating up to sets of small Lebesgue measure, the singular part is, for our purposes, easily dealt with.

Let us introduce an important lemma for C^m functions that is the basis for the pointwise estimates used in this paper. This is a generalization of inequality (3.6) and was first proven in [13]:

Lemma 5. *Let $f \in C^m(Q)$ where $Q \subset \mathbb{R}^n$ is an open cube. Then*

$$\text{(i) } |T_Q^{m-1} f(y) - f(y)| \le C \int_Q \frac{|D^m f(z)|}{|y-z|^{n-m}} \, dz$$

$$\text{(ii) } |T_Q^m f(y) - f(y)| \le C \int_Q \frac{|D^m f(z) - \vec{a}|}{|y-z|^{n-m}} \, dz$$

for each $y \in Q$ and any constant vector $\vec{a} = \{a_\alpha\}_{|\alpha|=m}$.

Proof. Fix $x \in Q$ and, for $y \in Q$, define

$$\varphi_x(y) = \sum_{|\beta|<m} D^\beta f(y) \frac{(x-y)^\beta}{\beta!} + \sum_{|\alpha|=m} a_\alpha \frac{(x-y)^\alpha}{\alpha!}$$

Then clearly $\varphi_x \in C^1(Q)$ and

$$\begin{aligned} \partial_i \varphi_x(y) &= \sum_{|\beta|<m} D^{\beta+e_i} f(y) \frac{(x-y)^\beta}{\beta!} - \sum_{|\beta|<m} D^\beta f(y) \frac{(x-y)^{\beta-e_i}}{(\beta-e_i)!} + \sum_{|\alpha|=m} a_\alpha \frac{(x-y)^{\alpha-e_i}}{(\alpha-e_i)!} \\ &= \sum_{|\alpha|=m, \alpha \ge e_i} (D^\alpha f(y) - a_\alpha) \frac{(x-y)^{\alpha-e_i}}{(\alpha-e_i)!} \end{aligned}$$

Note that

$$T_y^m f(x) = \varphi_x(y) + \sum_{|\alpha|=m} (D^\alpha f(y) - a_\alpha) \frac{(x-y)^\alpha}{\alpha!},$$

so by the inequality 3.6

$$\begin{aligned}|f(x)-T_Q^m f(x)| &\le |\varphi_x(x)-(\varphi_x)_Q| + \sum_{|\alpha|=m} \fint_Q (D^\alpha f(y)-a_\alpha)\frac{(x-y)^\alpha}{\alpha!}dy \\ &\le C(n)\int_Q \frac{|\nabla\varphi_x(y)|dy}{|x-y|^{n-1}} + \sum_{|\alpha|=m}\fint_Q (D^\alpha f(y)-a_\alpha)\frac{(x-y)^\alpha}{\alpha!}dy \\ &\le C(m,n)\int_Q \frac{|\nabla^m f(y)-\vec{a}||x-y|^{m-1}dy}{|x-y|^{n-1}} + \fint_Q |\nabla^m f(y)-\vec{a}||x-y|^m dy \\ &\le C(m,n)\int_Q \frac{|\nabla^m f(y)-\vec{a}|dy}{|x-y|^{n-m}}\end{aligned}$$

since $|x-y|^n < C(n)|Q|$. Note that $\nabla^m f(y)-\vec{a} = \{D^\alpha f(y)-a_\alpha\}_{|\alpha|=m}$.

□

Now we apply the previous result to get the following potential estimates for BV^m functions.

Theorem 35. *Let $f \in BV^m(Q)$ where $Q \subset \mathbb{R}^n$ is an open cube. Then*

$$\text{(i) } |T_Q^{m-1} f(y) - f(y)| \le \int_Q \frac{d\|D^m f\|(z)}{|y-z|^{n-m}}$$

$$\text{(ii) } |T_Q^m f(y) - f(y)| \le \int_Q \frac{d\|[D^m f]-\vec{a}\mathcal{L}^n\|(z)}{|y-z|^{n-m}} \tag{4.13}$$

for each $y \in Q$ and any constant vector $\vec{a} = \{a_\alpha\}_{|\alpha|=m}$.

Proof. We will prove (ii), from which (i) follows easily.

Let f_k be a sequence in $C^\infty(Q) \cap BV^m(Q)$ such that

$$\|f_k - f\| \to 0 \text{ in } W^{m-1,1}(Q), \qquad \|D^m f_k\|(Q) \to \|D^m f\|(Q)$$

Such a sequence exists by a Theorem 17. Then Theorems 33 & 34 imply the weak-$*$ convergence

$$D^m f_k \overset{*}{\rightharpoonup} D^m f \qquad \qquad \|D^m f_k - \vec{a}\mathcal{L}^n\| \overset{*}{\rightharpoonup} \|D^m f - \vec{a}\mathcal{L}^n\|.$$

Fix $x \in Q$. Since $m \geq 1$, for each $\epsilon > 0$, the function $z \mapsto \int_{B(x,\epsilon)} \frac{dy}{|y-z|^{n-m}}$ is continuous. As $k \to \infty$, by the weak-* convergence and Lemma 3,

$$\int_{B(x,\epsilon)} \Big(\int_Q \frac{|\nabla D^m f_k(z) - \vec{a}|}{|y-z|^{n-1}} dz \Big) dy \to \int_{B(x,\epsilon)} \Big(\int_Q \frac{d\|D^m f - \vec{a}\mathcal{L}^n\|(z)}{|y-z|^{n-1}} \Big) dy$$
$$\leq C(m,n) \int_Q \frac{d\|D^m f - \vec{a}\mathcal{L}^n\|(z)}{|x-z|^{n-1}} \tag{4.14}$$

Similarly for $|\beta| \leq m-1$, $D^\beta f_k \to D^\beta f$ in $L^1(Q)$, so

$$\sum_{|\beta|<m} \fint_{B(x,\epsilon)} \left(\int_Q D^\beta f_k(z) \frac{(y-z)^\beta}{\beta!} \, dy \right) dz \to \sum_{|\beta|<m} \fint_{B(x,\epsilon)} \left(\int_Q D^\beta f(z) \frac{(y-z)^\beta}{\beta!} \, dy \right) dz.$$

For $|\beta| = m$, the weak-* convergence above implies that

$$\sum_{|\beta|=m} \fint_{B(x,\epsilon)} \left(\int_Q \frac{(y-z)^\beta}{\beta!} \, dy \right) dD^\beta f_k(z) \to \sum_{|\beta|=m} \fint_{B(x,\epsilon)} \left(\int_Q D^\beta f(z) \frac{(y-z)^\beta}{\beta!} \, dy \right) dD^\beta f(z).$$

Therefore

$$\left| \fint_{B(x,\epsilon)} T_Q^m f_k(y) - f(y) dx \right| \to \left| \fint_{B(x,\epsilon)} T_Q^m f(y) - f(y) dx \right|.$$

Now we recall Lemma 5 holds for each k. Letting $k \to \infty$ and applying the estimate (4.14), which is independent of ϵ, we have

$$\left| \fint_{B(x,\epsilon)} T_Q^m f(y) - f(y) dx \right| \leq C(m,n) \int_Q \frac{d\|D^m f - \vec{a}\mathcal{L}^n\|(z)}{|x-z|^{n-1}}.$$

Taking the limsup as $\epsilon \to 0$ implies the result. □

Now we can prove the crucial pointwise estimate foe BV^m functions:

Theorem 36. *Let $f \in BV^m(\mathbb{R}^n)$. Then there is a constant $C(m,n)$ such that for all $x, y \in \mathbb{R}^n$,*

$$\frac{|f(y) - T_x^m f(y)|}{|x-y|^m} \leq C(\mathcal{M}_{|x-y|} \|D^m f - \nabla^m f(x)\|(x) + \mathcal{M}_{|x-y|} \|D^m f - \nabla^m f(y)\|(y)) \tag{4.15}$$

Proof. First note $T_Q^m f(y)$ is an mth degree polynomial in y. Indeed, recall

$$T_Q^m f(y) := \sum_{|\beta|\le m-1} \fint_Q D^\beta f(z) \frac{(y-z)^\beta}{\beta!}\, dz + \sum_{|\beta|=m} \fint_Q \frac{(y-z)^\beta}{\beta!}\, d(D^\beta f)(z).$$

Differentiating under the integral sign, we see that $D^m(T_Q^m f)(y) = \dfrac{D^m f(Q)}{|Q|}$. Therefore $T_Q^m f(y)$ is equal to its Taylor polynomial centered at x:

$$T_Q^m f(y) = \sum_{|\alpha|\le m} D^\alpha\big(T_Q^m f\big)(x) \frac{(y-x)^\alpha}{\alpha!} \tag{4.16}$$

Compute

$$\begin{aligned}
D^\alpha T_Q^m f(x) &= D^\alpha\Big(\sum_{|\beta|\le m-1} \fint_Q D^\beta f(z) \frac{(x-z)^\beta}{\beta!}\, dz + \sum_{|\beta|=m} \fint_Q \frac{(x-z)^\beta}{\beta!}\, d(D^\beta f)(z)\Big) \\
&= \sum_{|\beta|\le m-1} \fint_Q D^\beta f(z) \frac{(x-z)^{\beta-\alpha}}{(\beta-\alpha)!}\, dz + \sum_{|\beta|=m} \fint_Q \frac{(x-z)^{\beta-\alpha}}{(\beta-\alpha)!}\, d(D^\beta f)(z) \\
&= \sum_{|\gamma|\le m-|\alpha|-1} \fint_Q D^{\gamma+\alpha} f(z) \frac{(x-z)^\gamma}{\gamma!}\, dz + \sum_{|\gamma|=m-|\alpha|} \fint_Q \frac{(x-z)^\gamma}{(\gamma)!}\, d(D^{\gamma+\alpha} f)(z) \\
&= T_Q^{m-|\alpha|} D^\alpha f(x)
\end{aligned}$$

Employing (4.16), we compute

$|T_Q^m f(y) - T_x^m f(y)|$

$$\begin{aligned}
&\le \sum_{|\alpha|<m} |T_Q^{m-|\alpha|} D^\alpha f(x) - D^\alpha f(x)| \frac{|y-x|^{|\alpha|}}{\alpha!} + \sum_{|\alpha|=m} |(D^\alpha f)_Q - \nabla^\alpha f(x)| \frac{|y-x|^m}{m!} \\
&\le \sum_{|\alpha|<m} \Big(\int_Q \frac{d\|D^{m-|\alpha|}(D^\alpha f) - \vec{a}\mathcal{L}^n\|(z)}{|x-z|^{n-(m-|\alpha|)}} \Big) \frac{|y-x|^{|\alpha|}}{\alpha!} + C \left| \frac{D^m f(Q)}{|Q|} - \nabla^m f(x) \right| \frac{|y-x|^m}{m!} \\
&= \sum_{|\alpha|<m} \Big(\int_Q \frac{d\|D^m f - \vec{a}\mathcal{L}^n\|(z)}{|x-z|^{n-(m-|\alpha|)}} \Big) \frac{|y-x|^{|\alpha|}}{\alpha!} + C \left| \frac{D^m f(Q)}{|Q|} - \nabla^m f(x) \right| \frac{|y-x|^m}{m!} \qquad (4.17)
\end{aligned}$$

where in the second inequality we applied our potential estimate (4.13) to $D^\alpha f \in BV^{m-|\alpha|}(\mathbb{R}^n)$ for $|\alpha| < m$.

By Hedberg's Lemma,

$$\int_Q \frac{d\|D^m f - \vec{a}\mathcal{L}^n\|(z)}{|x-z|^{n-(m-|\alpha|)}} \le C \mathrm{diam}(Q)^{m-|\alpha|} M_{\mathrm{diam}(Q)}(\|D^m f - \vec{a}\mathcal{L}^n\|)(x)$$

Therefore using the fact that $|x-y| < \text{diam}(Q)$ and (4.17) we have

$$|T_Q^m f(y) - T_x^m f(y)| \le C\text{diam}(Q)^m \mathcal{M}_{\text{diam}(Q)} \|D^m f - \vec{a}\mathcal{L}^n\|(x) + C\left|\frac{D^m f(Q)}{|Q|} - \nabla^m f(x)\right| \frac{|y-x|^m}{m!}.$$

Also by (4.13),

$$|f(y) - T_Q^m f(y)| \le \int_Q \frac{d\|D^m f - \vec{b}\mathcal{L}^n\|(z)}{|y-z|^{n-m}} \le C\text{diam}(Q)^m \mathcal{M}_{\text{diam}(Q)} \|D^m f - \vec{b}\mathcal{L}^n\|(y)$$

Since $\vec{a}$, $\vec{b} \in \mathbb{R}^M$ are arbitrary vectors, we choose $\vec{a} = \nabla^m f(x)$, $\vec{b} = \nabla^m f(y)$.
Applying the triangle inequality, $|f(y) - T_x^m f(y)| \le |f(y) - T_Q^m f(y)| + |T_Q^m f(y) - T_x^m f(y)|$, yields

$$|f(y) - T_x^m f(y)| \le C\text{diam}(Q)^m \Big(\mathcal{M}_{\text{diam}(Q)} \|D^m f - \nabla^m f(x)\mathcal{L}^n\|(x) + \mathcal{M}_{\text{diam}(Q)} \|D^m f - \nabla^m f(y)\mathcal{L}^n\|(y)\Big) + \left|\frac{D^m f(Q)}{|Q|} - \nabla^m f(x)\right| |y-x|^m$$

Note that by the same calculation as (3.15),

$$\left|\frac{D^m f(Q)}{|Q|} - \nabla^m f(x)\right| \le C\mathcal{M}_{\text{diam}(Q)} \|D^m f - \nabla^m f(x)\mathcal{L}^n\|(x).$$

The statement follows since we can select Q containing x, y such that $\text{diam}(Q) < |x-y| + \epsilon$. □

For $|\beta| < m$, $D^\beta f \in BV^{m-|\beta|}(\mathbb{R}^n)$, so replacing f with $D^\beta f$ in the previous inequality yields the following theorem:

Theorem 37. *Let $f \in BV^m(\mathbb{R}^n)$. Then the exists a constant $C(m,n)$ such that for all $x, y \in \mathbb{R}^n$ and any multiindex β with $|\beta| < m$,*

$$\frac{|D^\beta f(y) - T_x^{m-|\beta|} D^\beta f(y)|}{|x-y|^{m-|\beta|}} \le C\mathcal{M}_{|x-y|} \|D^m f - \nabla^m f(x)\|(x) + \mathcal{M}_{|x-y|} \|D^m f - \nabla^m f(y)\|(y) \tag{4.18}$$

The proof follows immediately from (36), noting that $D^m f$ has more terms than $D^{m-|\beta|} D^\beta f$ and $\nabla^m f$ has more terms than $\nabla^{m-|\beta|} \nabla^\beta f$, and that all nonzero terms coincide with terms in $D^m f$ and $\nabla^m f$ respectively.

4.3 LUZIN TYPE APPROXIMATION BY SMOOTH FUNCTIONS

Now we can prove our main result: Our goal is to apply the Whitney Extension Theorem to an appropriate jet of functions.

Theorem 38. *Let $f \in BV^m(\mathbb{R}^n)$. For each $\epsilon > 0$, there exists a function $g \in C^m(\mathbb{R}^n)$ such that*

$$|\{D^\beta f \neq D^\beta g\}| < \epsilon \text{ for each } \beta,\, |\beta| \leq m-1, \quad |\{\nabla^m f \neq \nabla^m g\}| < \epsilon$$

Proof. Fix $\epsilon > 0$. By Luzin Theorem, there is a closed set $\widetilde{C} \subset \mathbb{R}^n$ with $|\mathbb{R}^n \backslash \widetilde{C}| < \epsilon/2$ such that $D^\beta f|_{\widetilde{C}}$ and $\nabla^m f|_{\widetilde{C}}$ are continuous functions on $\widetilde{C}$. Define

$$\eta_k(z) = \sup_{r<\frac{1}{k}} M_r \|D^m f - \nabla^m f(z)\mathcal{L}^n\|(z) \tag{4.19}$$

Claim: $\eta_k(z) \to 0$ for almost every $z \in \mathbb{R}^n$.

Indeed,

$$\begin{aligned} M_r\|D^m f - \nabla^m f(z)\mathcal{L}^n\|(z) &= M_r\|D^m f_s + D^m f_{ac} - \nabla^m f(z)\mathcal{L}^n\|(z) \\ &\leq M_r\|D^m f\|_s(z) + M_r\|D^m f_{ac} - \nabla^m f(z)\mathcal{L}^n\|(z) \to 0 \end{aligned}$$

as $r \to 0$, just as in the proof for the C^1 case. (see 3.15)

Applying Egorov theorem to the sequence η_k, we find a closed set $C \subset \widetilde{C}$ such that $|\widetilde{C} \backslash C| \leq \epsilon/2$ and

$$\eta_k \rightrightarrows 0 \text{ uniformly on compact subsets of } C$$

Then we define the jet $F = \{f^\beta\}_{|\beta| \leq m}$ on $\widetilde{C}$ by

$$f^\beta = \begin{cases} D^\beta f|_C & \text{if } |\beta| < m \\ \nabla^\beta f|_C & \text{if } |\beta| = m \end{cases}$$

Note that

$$T^{m-|\beta|} D^\beta F = (T^{m-|\beta|} D^\beta f)|_C$$

Let K be a compact subset of C. Fix a multiindex β with $|\beta| < m$

$$\underline{\text{Claim:}} \quad \sup_{x,y \in K, |x-y|<\delta} \frac{|f^\beta(y) - T_x^{m-|\beta|} D^\beta F(y)|}{|x-y|^{m-|\beta|}} \to 0 \quad \text{as} \quad \delta \to 0 \tag{4.20}$$

Fix $k(\delta)$ so that $\frac{1}{k+1} < \delta < \frac{1}{k}$. Let $x, y \in K$ be any pair of points such that $|x-y| < \delta$.

$$\mathcal{M}_{|x-y|}\|D^m f - \nabla^m f(x)\mathcal{L}^n\|(x) + \mathcal{M}_{|x-y|}\|D^m f - \nabla^m f(y)\mathcal{L}^n\|(y) \\ \leq \sup_{z\in K, r<\frac{1}{k}} \mathcal{M}_r\|D^m f - \nabla^m f(x)\mathcal{L}^n\|(z).$$

Employing inequality (4.18),

$$\frac{|f^\beta(y) - T_x^{m-|\beta|}D^\beta F(y)|}{|x-y|^{m-|\beta|}} \leq C \sup_{z\in K} \eta_k(z) \to 0 \text{ as } k \to \infty$$

Taking the supremum over all $x, y \in K$ with $0 < |x-y| < \delta$, and noting that $k \to \infty$ as $\delta \to 0$ implies the claim.

For $|\beta| = m$, equation (4.20) reduces to

$$\sup_{x,y\in K, |x-y|<\delta} |\nabla^\beta f(y) - \nabla^\beta f(x)| \to 0 \text{ as } \delta \to 0 \tag{4.21}$$

which is true because $\nabla^m f$ is uniformly continuous on the compact set K.

The above estimates imply that F is a Whitney jet on the closed set C. Therefore by the Whitney Extension Theorem, there exists a function $g \in C^m(\mathbb{R}^n)$ such that

$$g(x) = f(x), \quad D^\beta g(x) = D^\beta f(x) \ (|\beta| < m), \quad D^m g(x) = \nabla^m f(x) \quad \forall x \in C$$

Since $|\mathbb{R}^n \backslash C| < \epsilon$, this proves the theorem. □

4.4 BOUNDING THE ERROR IN VARIATION

Given a function $f \in BV^m(\mathbb{R}^n)$, we want to construct a function $g \in C^m(\mathbb{R}^n)$ such that g coincides with f on large set, say F, whose complement has measure less that ϵ, and moreover, such that we have estimates on the errors in $W^{m-1,1}$ norm and the variation:

$$\|f - g\|_{W^{m-1,1}(\mathbb{R}^n)} < \epsilon \qquad \Big| \|D^m f\|(\mathbb{R}^m) - \|D^m g\|(\mathbb{R}^n) \Big| < \epsilon.$$

The approximation given by the Whitney Extension Theorem in the previous section will not, in general, provide us with good estimates. At issue is the fact that the extension provided by the Whitney Theorem is determined using only the information about f on the set F. However, since the mth order derivative may have a singular part, the information about f outside F is crucial, despite the fact that F has small measure. Therefore we need to modify our approximation to account for this, and one way to do so is by a kind of averaging process over cubes in the Whitney decomposition. This technique is called *Whitney smoothing* and was first introduced in [12]. In general, the Whitney smoothing will not produce a C^m function even provided we have good estimates on C; however, BV^m functions have enough structure to ensure the desired regularity.

Start by recalling that the function

$$\eta_k(z) = \sup_{r<\frac{1}{k}} M_r \|D^m f - \nabla^m f(z) \mathcal{L}^n\|(z) \to 0$$

as $k \to \infty$ for almost every z in $\mathbb{R}^n$.

So for $\epsilon > 0$, by Luzin and Egorov Theorems we can find a closed set F with $|F^c| < \epsilon$, such that $D^\beta f|_F$ for $|\beta| < m$ and $\nabla^\beta f|_F$ for $|\beta| = m$ are continuous, and

$$\eta_k \rightrightarrows 0 \text{ uniformly on compact subsets of } F.$$

For each $a \in F$, let $K_a = F \cap \overline{B}(a, 1)$, and define the function

$$\eta_a(t) = \sup_{z \in K_a} \sup_{r<t} M_r \|D^m f - \nabla^m f(z) \mathcal{L}^n\|(z)$$

By the uniform convergence of η_k, it is easy to check $\eta_a(t) \to 0$ as $t \to 0$. We will consider the concave envelope of this function, which we will also denote η_a.

Our goal is to show that the Whitney smoothing

$$g(x) = \begin{cases} f(x) & \text{if } \mathrm{x} \in F \\ \sum_{i\in I} \varphi_i(x) T^m_{Q_i} f(x) & \text{if } \mathrm{x} \notin F \end{cases} \tag{4.22}$$

is a C^m function on $\mathbb{R}^n$. Here Q_i are the enlarged cubes in the Whitney decomposition and $\{\varphi_i\}$ is the partition of unity on F^c subordinate to $\{\widetilde{Q_i}\}$.

Proof. Recall the definitions of $T^m_{\widetilde{Q_i}} f$ and $T^m_a f(x)$ (see (4.12) and (4.11)).

Let's include a few important estimates: Let $x \in \mathbb{R}^n$, let $\overline{Q}$ be a cube containing x and $S \subset \overline{Q}$ be any measurable subset of positive measure. Let $f \in BV^m(\overline{Q})$. Then

$$|f(x) - T^m_{\overline{S}} f(x)| \leq C(n) \frac{|\overline{Q}|}{|S|} \int_{\overline{Q}} \frac{d\|D^m f - \vec{a}\mathcal{L}^n\|(z)}{|x-z|^{n-m}}.$$

This follows from a variant of (3.6) and an approximation argument.

If Q is a cube containing both x and a, then by (36),

$\frac{|f(x)-T^m_a f(x)|}{|x-a|^m}$

$$\leq C(m,n) \left[\mathcal{M}_{|x-y|}\|D^m f - \nabla^m f(x)\mathcal{L}^n\|(x) + \mathcal{M}_{|x-y|}\|D^m f - \nabla^m f(a)\mathcal{L}^n\|(a)\right]$$

In particular if $a \in F$ and $x \in K_a$, then we can easily find a cube Q of diameter less than $2|x-a|$ containing x and a, which yields the estimate

$$|f(x) - T^m_a f(x)| \leq 2\eta_a(2|x-a|)|x-a|^m \tag{4.23}$$

immediately from definition of η_a.

Clearly $g|_{F^c} \in C^m(F^c)$, as $T^m_{\widetilde{Q_i}} f$ is a polynomial for each i and the sum is locally finite. To prove it is C^m in a neighborhood of each point $a \in F$, we will show that

$$|D^\alpha g(x) - D^\alpha g(a) - \nabla(D^\alpha f)(a) \cdot (x-a)| = o(|x-a|).$$

where $D^\alpha g(a)$ is equal to $D^\alpha f(a)$ for $|\alpha| \leq m-1$ and is equal to $\nabla^\alpha f(a)$ for $|\alpha| = m$.

The idea is to prove that for $a \in F$ and x in a sufficiently small neigbhood of a, that

$$|D^\alpha g(x) - T_a^{m-|\alpha|} D^\alpha f(x)| \leq C(m,n)\eta_a(|x-a|)|x-a|^{m-|\alpha|} \tag{4.24}$$

which by the argument in section 2.4 shows that that g is C^m in a neigborhood of a; moreover, $D^k g|_F = D^k f|_F$ for $k < m$ and $D^m g|_F = \nabla^m f|_F$.

First consider the case $a \in \text{int}\,(F)$. Then for $x \in B(a,\delta) \subset K_a$ for δ sufficiently small, and $|\alpha| < m$ we have the estimate

$$|D^\alpha f(x) - T_a^{m-|\alpha|} D^\alpha f(x)| \leq \eta_a(2|x-a|)|x-a|^{m-|\alpha|} \qquad \text{for each } x \in B(a,\delta),$$

which is true by (4.23). Concavity of η_a implies that $\eta_a(c|x-a|) \leq c\eta_a(|x-a|)$ for $c \geq 1$, which implies (4.24). By the induction, g is m-times differentiable at a with derivative $\nabla^m f(a)$. When $|\alpha| = m$, it is enough to note

$$|\nabla^\alpha f(x) - \nabla^\alpha f(a)| \to 0 \text{ as } |x-a| \to 0,\ x \in K_a,$$

which proves continuity of the mth order derivative.

Now assume $a \in \partial F$. Fix $x \in F^c$ with $|x-a| < 1$, and let b be the nearest point to x in K_a, q_i be the nearest point to $\widetilde{Q_i}$ in K_a.
Expand $\left|D^\alpha g(x) - T_a^{m-|\alpha|} D^\alpha f(a)\right|$ as

$$\sum_{\beta+\gamma=\alpha} \binom{\alpha}{\beta} S^{\beta,\gamma}(x)$$

where

$$S^{\beta,\gamma}(x) = \sum_i D^\beta \varphi_i(x) D^\gamma (T^m_{\widetilde{Q_i}} f(x) - T_a^m f(x)).$$

For $|\beta| > 0$, we can replace a by b.

$$\begin{aligned}
&|D^\gamma (T^m_{\widetilde{Q_i}} f(x) - T_a^m f(x))| \\
&\qquad \leq |T^{m-|\gamma|}_{\widetilde{Q_i}} f(x) - T^{m-|\gamma|}_{q_i} f(x)| + |T^{m-|\gamma|}_{q_i} f(x) - T^{m-|\gamma|}_a f(x)| \\
&\qquad = I + II
\end{aligned}$$

where again we replace a by b if $|\gamma| < |\alpha|$.

To estimate II, we need to use a lemma ([18],Lemma 2.1.5), which states that if η_a is a modulus of continuity such that for each $|\alpha| \le m$, $x, y \in K_a$

$$|D^\alpha f(y) - T_x^{m-|\alpha|} D^\alpha f(y)| \le \eta_a(|x-y|)|x-y|^{m-|\alpha|}$$

then for some constant C, for every $x, y \in K_a$ and $z \in \mathbb{R}^n$,

$$|T_x^{m-|\alpha|} D^\alpha f(z) - T_y^{m-|\alpha|} D^\alpha f(z)| \le C\eta_a(|x-y|)(|x-z|^{m-|\alpha|} + |y-z|^{m-|\alpha|}).$$

Now the conditions of this Lemma hold with q_i and a immediately from our above estimates, and therefore

$$|T_{q_i}^{m-|\gamma|} D^\gamma f(x) - T_a^{m-|\gamma|} D^\gamma f(x)| \le C\eta_a(|q_i - a|)(|x-a|^{m-|\gamma|} + |x-q_i|^{m-|\gamma|}),$$

noting that for $|\beta| > 0$ we can replace a by b. Since all these are comparable to $|x-a|$ we get that

$$II \le C\eta_a(|x-a|)|x-a|^{m-|\gamma|}.$$

To estimate I, we note that $T_{\widetilde{Q_i}}^{m-|\gamma|} D^\gamma f(x)$ is a polynomial of degree m, so is equal to its Taylor expansion centered at q_i. Therefore

$$I = \left| \sum_{\sigma \le m-|\gamma|} \left[T_{\widetilde{Q_i}}^{m-|\gamma+\sigma|} D^{\gamma+\sigma} f(q_i) - D^{\gamma+\sigma} f(q_i) \right] \frac{(x-q_i)^\sigma}{\sigma!} \right|.$$

The left and right terms are just the two Taylor expansions written out.

First we will consider the case $|\gamma| < m$. Let $\overline{Q}$ be a cube containing Q and q_i such that $\mathrm{diam}(\overline{Q}) < 10\mathrm{diam}(\widetilde{Q_i})$. Then for $|\alpha| < m$, we estimate

$|T_{\widetilde{Q_i}}^{m-|\alpha|} D^\alpha f(q_i) - D^\alpha f(q_i)|$

$$\begin{aligned}
&\le C(n)\frac{|\overline{Q}|}{|Q_i|} \int_{\overline{Q}} \frac{\|D^m f - \nabla^m f(z)\mathcal{L}^n\|(z)}{|q_i - z|^{n-(m-|\alpha|)}} \\
&\le C(n)\mathrm{diam}(\widetilde{Q_i})^{m-|\alpha|} \mathcal{M}_{10\mathrm{diam}(\widetilde{Q_i})} \|D^m f - \nabla^m f(q_i)\mathcal{L}^n\|(q_i) \\
&\le C(n)|x-q_i|^{m-|\alpha|} \mathcal{M}_{C|x-q_i|} \|D^m f - \nabla^m f(q_i)\mathcal{L}^n\|(q_i) \\
&\le C(n)|x-q_i|^{m-|\alpha|} \eta_a(|x-q_i|)
\end{aligned}$$

because $x \in \widetilde{Q_i}$ and $\text{diam}(\widetilde{Q_i}) < C'\text{dist}\,(Q_i, K_a) = C'\text{dist}\,(Q_i, q_i) \leq C|x - q_i|$. We also apply concavity of η_a in the last step.

Now for $|\alpha| = m$, we cannot apply this estimate because $D^\alpha f$ is a measure. Instead we have

$$
\begin{aligned}
||T^0_{\widetilde{Q_i}} D^\alpha f(q_i) - \nabla^\alpha f(q_i)| \\
&= \left| \frac{D^\alpha f(\widetilde{Q_i})}{|\widetilde{Q_i}|} - \nabla^\alpha f(q_i) \right| = \frac{|D^\alpha f(\widetilde{Q_i}) - \nabla^\alpha f(q_i)\mathcal{L}^n(\widetilde{Q_i})|}{|\widetilde{Q_i}|} \\
&\leq \frac{\|D^m f - \nabla^m f(q_i)\mathcal{L}^n\|(\widetilde{Q_i})}{|\widetilde{Q_i}|} \\
&\leq C(n)\mathcal{M}_{C|x-q_i|}\|D^m f - \nabla^m f(q_i)\mathcal{L}^n\|(q_i) \\
&\leq C(n)\eta_a(|x - q_i|)
\end{aligned}
$$

since $q_i \in K_a$.

Note that when $|\alpha| = |\gamma+\sigma| = m$, that $|\sigma| = m-|\gamma|$. The above estimates for $|\gamma+\sigma| \leq m$ imply that

$$I \leq C\eta_a(|x - q_i|)|x - q_i|^{m-|\gamma|} \leq C\eta_a(|x - a|)|x - a|^{m-|\gamma|}.$$

Now we compute

$$
\begin{aligned}
|S^{\beta,\gamma}(x)| &\leq C(n) \sum_i D^\beta \varphi_i(x)\eta_a(|x - a|)|x - a|^{m-|\gamma|} \\
&\leq C(n)\text{diam}(Q_i)^{-|\beta|}\eta_a(|x - a|)|x - a|^{m-|\gamma|} \\
&\leq C(n)\eta_a(|x - a|)|x - a|^{m-|\beta+\gamma|} = C\eta_a(|x - a|)|x - a|^{m-|\alpha|}
\end{aligned}
$$

This proves the estimate for each $a \in F$, and thus proves that $g \in C^m(\mathbb{R}^n)$. □

Now we can sketch some partial results involving the above estimates in the case $m = 1$.

Let g_t be an approximation such that $|\{f \neq g_t\}| < 1/t$, and let F_t be the set $\{g_t = f\}$. Recall that for $x \in F_t^c$,

$$g_t(x) = \sum_i \varphi_i T^1_{\widetilde{Q_i}} f(x) = \sum_i \varphi_i(x)\Big(f_{\widetilde{Q_i}} + \fint_{\widetilde{Q_i}} (x - z) \cdot dDf(z)\Big)$$

Then

$$\int_{F_t^c} |g_t| dx = \sum_i \int_{F_t^c} \left| \varphi_i(x)\, f_{\widetilde{Q_i}} + \fint_{\widetilde{Q_i}} \varphi_i(x)(x-z)\cdot dDf(z) \right| dx$$
$$\leq \sum_i \int_{F_t^c} \chi_{\widetilde{Q_i}}(x)|f_{\widetilde{Q_i}}|\, dx + \int_{F_t^c} \chi_{\widetilde{Q_i}}(x) \fint_{\widetilde{Q_i}} |x-z| d\|Df\|(z)\, dx$$
$$\leq \sum_i \int_{\widetilde{Q_i}} |f|\, dx + \int_{\widetilde{Q_i}} \frac{\operatorname{diam}(\widetilde{Q_i})}{|\widetilde{Q_i}|} \|Df\|(\widetilde{Q_i})$$
$$\leq C(n)\|f\|_1 + \sum_i \operatorname{diam}(\widetilde{Q_i})\|Df\|(\widetilde{Q_i})$$
$$\leq C(n)(\|f\|_1 + \|Df\|(\mathbb{R}^n))$$

where the last step follows from the fact that $\operatorname{diam}(\widetilde{Q_i})$ are uniformly bounded and the fact that each $x \in F_t^c$ lies in at most 4^n of the cubes $\widetilde{Q_i}$.

Now we will get an upper estimate on the error in variation. Note that since g_t is locally Lipschitz,

$$\|Dg_t\|(\mathbb{R}^n) = \int_{\mathbb{R}^n} |\nabla g_t(x)|\, dx.$$

Differentiating, we see that

$$\nabla g_t = \sum_i \nabla \varphi_i(x) T^1_{Q_i} f(x) + \sum_i \varphi_i(x) \frac{Df(Q_i)}{|Q_i|}.$$

Then recalling $|\nabla \varphi_i(x)| \leq C(n)\operatorname{diam}(Q_i)^{-1}$, and the fact that $\sum_i \nabla\varphi_i(x) f(x) = 0$, we estimate

$$\|Dg_t\|(F_t^c) \leq \sum_i \int_{F_t^c} \left| \nabla\varphi_i(x)\Big(f_{\widetilde{Q_i}} + \fint_{\widetilde{Q_i}} |x-z| \|Df\|(z) \Big) \right| + \int_{F_t^c} \chi_{\widetilde{Q_i}}(x) \frac{|Df(\widetilde{Q_i})|}{|\widetilde{Q_i}|} dx$$
$$\leq \sum_i C(n) \int_{F_t^c} \operatorname{diam}(\widetilde{Q_i})^{-1} \chi_{\widetilde{Q_i}} |f_{\widetilde{Q_i}} - f(x)| dx + \fint_{\widetilde{Q_i}} \operatorname{diam}(\widetilde{Q_i})^{-1} \operatorname{diam}(\widetilde{Q_i}) \|Df\|(\widetilde{Q_i}) dx$$
$$+ \fint_{\widetilde{Q_i}} |Df(\widetilde{Q_i})| dx$$
$$\leq C(n) \sum_i \Big(\int_{\widetilde{Q_i}} |f_{\widetilde{Q_i}} - f(x)| dx + 2\|Df\|(\widetilde{Q_i})\Big)$$

Now we apply the Poincaré inequality for BV functions and the fact that $\widetilde{Q_i} \subset F_t^c$ and each $x \in F_t^c$ lies in at most 4^n of the cubes $\widetilde{Q_i}$ to recover

$$\|Dg_t\|(F_t^c) \leq C(n)\|Df\|(F_t^c) \tag{4.25}$$

Then the sequence $\{g_t\}$ is uniformly bounded in $BV(\mathbb{R}^n)$. Using the same argument as in the proof for Lipschitz functions (see p. 31), we can assume that all the functions g_t are supported in a compact ball. Therefore they converge in $L^1(\mathbb{R}^n)$ to some function $\widetilde{f}$. Since $g_t(x) \to f(x)$ for almost every $x \in \mathbb{R}^n$, it follows that

$$\|g_t - f\|_{L^1(\mathbb{R}^n)} \to 0 \text{ as } t \to \infty.$$

By weak lowersemicontinuity (Theorem 15) this implies that

$$\|Df\|(\mathbb{R}^n) \leq \liminf_{t\to\infty} \|Dg_t\|(\mathbb{R}^n).$$

In addition, inequality (4.25) gives us a bound on the error in variation; however, it does not imply the convergence

$$|\|Dg_t\|(\mathbb{R}^n) - \|Df\|(\mathbb{R}^n)| \to 0 \text{ as } t \to \infty.$$

I believe that these estimates can be significantly improved, and in fact that by modifying the construction of g_t on the set $\{g_t \neq f\}$, one can recover

$$\|Dg_t\|(\mathbb{R}^n) \to \|Df\|(\mathbb{R}^n) \text{ as } t \to \infty.$$

However, this is a nontrivial result and I am currently working on the details.

5.0 THE LUZIN THEOREM FOR HIGHER ORDER DERIVATIVES

In 1917, N. Luzin proved [17] [23] the following surprising result: For any Lebesgue measurable function $f : \mathbb{R} \to \mathbb{R}$ there is a continuous almost everywhere differentiable function g such that $g' = f$ almost everywhere. This is surprising even for the function $f(x) = 1/x$, because the antiderivative of f is discontinuous and, in fact, unbounded at 0. In this case, we correct the antiderivative by adding continuous functions which are differentiable almost everywhere with derivative equal to zero but which are not constant (one example of such a function is a Cantor staircase).

The original proof due to Luzin is purely one-dimensional, and offers no insight into a proof in higher dimensions. However, in 2008 Moonens and Pfeffer [19] proved the following generalization:

Let U be an open subset of $\mathbb{R}^N$. Given any Lebesgue measurable function $f : U \to \mathbb{R}^N$, there is an almost everywhere differentiable function $g \in C(\mathbb{R}^N)$ such that $\nabla g = f$ almost everywhere.

The goal of this paper is to extend the results to include higher order derivatives.

For an m-times differentiable function g defined in an open subset $U \subset \mathbb{R}^N$ we write

$$D^m g = (D^\alpha g)_{|\alpha|=m}$$

to denote the collection of all partial derivatives of order m. Our main result reads as follows:

Theorem 39. *Let* $f = (f^\alpha)_{|\alpha|=m}$ *be a Lebesgue measurable function defined in an open set* $U \subset \mathbb{R}^n$*. Then there is a function* $g \in C^{m-1}(\mathbb{R}^n)$ *which is m-times differentiable a.e. and such that*

$$D^m g = f \quad a.e. \text{ in } U,$$

i.e.

$$D^\alpha g = f^\alpha \quad a.e. \text{ in } U \text{ for } |\alpha| = m.$$

Moreover, for any $\sigma > 0$*, the function* g *may be chosen such that*

$$\|D^\gamma g\|_\infty < \sigma \qquad \text{for every } |\gamma| < m.$$

The outline of the proof is as follows: If $f = (f^\alpha)_{|\alpha|=m}$ is continuous and bounded on an open set U of finite measure, then we can find and a function $g \in C^m$ such that $D^m g$ *approximates* f on a large compact set. Using this approximation and a suitable limiting process, we can find $g \in C^m$ such that $D^m g$ is *equal to* f on a large compact set. We then show that the same holds more generally for the class of Lebesgue measurable functions, since such functions are continuous and bounded when restricted to a large compact set. The final construction involves piecing together approximations of f using a compact exhaustion of $\mathbb{R}^n$, taking care to avoid any overlap which would cause the resulting approximation to lose its desired form. The proof requires careful estimates for the approximation, which is the main difficulty.

5.1 SMOOTH FUNCTIONS WITH PRESCRIBED HIGHER ORDER DERIVATIVES ON LARGE SETS

Throughout the paper $|U|$ denotes the N-dimensional Lebesgue measure of a set U.

Given a continuous function $f = (f^\alpha)_{|\alpha|=m}$ defined in an open set $U \subset \mathbb{R}^n$ with $|U| < \infty$, our first task is to construct a compactly supported function $u \in C_c^m(U)$ such that $D^m u = f$ on a large compact subset of U. To this end we need the following approximation result. For the case $m = 1$, see [2].

Lemma 6. *Fix $m \in \mathbb{N}$ and let $U \subset \mathbb{R}^N$ be open with $|U| < \infty$. Let $f = (f^\alpha)_{|\alpha|=m}$ be a continuous and bounded function on U. Then for any $\epsilon, \eta, \sigma > 0$, there exists a function $u \in C_c^\infty(U)$ and a compact set $K \subset U$ such that for each $p \in [1, \infty]$ the following hold:*

(i) $|U \backslash K| < \epsilon$,

(ii) $|D^m u(x) - f(x)| < \eta$ *for each* $x \in K$,

(iii) $\|D^m u\|_p \leq C(m, N)\epsilon^{1/p-m}\|f\|_p$,

(iv) $\|D^\gamma u\|_\infty \leq \sigma$ *for every* $|\gamma| < m$.

Proof. Fix $\epsilon, \eta, \sigma > 0$. By $Q(x, r)$ we will denote the closed cube centered at x with side length r. Select a compact set $K' \subset U$ such that $|U \backslash K'| < \epsilon/2$. Choose $\delta > 0$ so small that

$$Q(x, 4\delta) \subset U \text{ for all } x \in K'$$

and

$$(Q(z, \delta) \cap K' \neq \emptyset, (x, y) \in Q(z, \delta)) \Rightarrow |f(x) - f(y)| < \eta \tag{5.1}$$

Cover $\mathbb{R}^N$ with a lattice of closed cubes of side length δ. Let $\{T_i\}_{i \in I}$ be the finite subcollection of cubes whose intersection with K' is nonempty. Clearly

$$K' \subset \bigcup_{i \in I} T_i \subset U.$$

For each i, let Q_i be a closed cube concentric with T_i and side length $(1 - \epsilon/2N)\delta$. Denote the center of the cube by c_i.

If δ is sufficiently small, then the set

$$K = \bigcup_{i \in I} Q_i$$

satisfies

$$|U \backslash K| = |U \backslash \bigcup_{i \in I} Q_i| < \epsilon.$$

Defining

$$a_i^\alpha = ⨍_{T_i} f^\alpha(y)\ dy,\ |\alpha| = m,$$

the function

$$g_i(x) = \sum_{|\alpha|=m} \frac{a_i^\alpha}{\alpha!}(x - c_i)^\alpha$$

is a polynomial such that

$$D^\alpha g_i(x) = a_i^\alpha \quad \text{for } |\alpha| = m,$$

and hence (5.1) yields that for all $x \in T_i$ and $|\alpha| = m$,

$$|D^\alpha g_i(x) - f^\alpha(x)| \leq \fint_{T_i} |f^\alpha(y) - f^\alpha(x)|\, dy < \eta.$$

Now if $\Phi_i \in C_c^\infty(T_i)$ with $\Phi_i \equiv 1$ on Q_i is a cut-off function,

$$u = \sum_{i \in I} \Phi_i g_i \in C_c^\infty(U)$$

satisfies

$$|D^m u(x) - f(x)| < \eta \text{ for all } x \in K.$$

We only need to choose Φ_i carefully to guarantee the estimates (iii) and (iv). Let

$$T = \left[\frac{-1}{2}, \frac{1}{2}\right]^N \text{ and } Q = \left[\frac{-1}{2} + \frac{\epsilon}{4N}, \frac{1}{2} - \frac{\epsilon}{4N}\right]^N,$$

i.e. Q is the cube concentric with T with side length $(1 - \epsilon/2N)$.

Let $\zeta \in C_c^\infty(B^N(0,1))$, $\zeta \geq 0$, $\int_{\mathbb{R}^N} \zeta = 1$ and let $\zeta_\epsilon(x) := \epsilon^{-N}(x/\epsilon)$ be a standard mollifier.

For

$$\widetilde{Q} = \left[\frac{-1}{2} + \frac{\epsilon}{8N}, \frac{1}{2} - \frac{\epsilon}{8N}\right]^N,$$

we define

$$\Phi = \chi_{\widetilde{Q}} * \zeta_{(\epsilon/16N)}.$$

Clearly $\Phi \in C_c^\infty(T)$, $\Phi = 1$ on Q and

$$|D^\alpha \Phi(x)| \leq C(m, N)\epsilon^{-|\alpha|} \quad \text{for } |\alpha| \leq m \text{ and } x \in T_i.$$

Finally we define

$$\Phi_i(x) = \Phi\left(\frac{x - c_i}{\delta}\right)$$

and

$$u = \sum_{i \in I} \Phi_i g_i.$$

Observe that for $x \in T_i$ and $|\beta|, |\gamma| \leq m$ we have

$$\left|D^\beta g_i(x)\right| \leq C(m,N)\|f\|_\infty \delta^{m-|\beta|}$$
$$|D^\gamma \Phi_i(x)| \leq C(m,N)\epsilon^{-|\gamma|}\delta^{-|\gamma|}.$$

Hence for any $|\alpha| \leq m$ and $x \in Q_i$,

$$\begin{aligned} |D^\alpha u(x)| &= |D^\alpha(g_i\Phi_i)(x)| \\ &\leq \sum_{\beta+\gamma=\alpha} \frac{\alpha!}{\beta!\,\gamma!} \left|D^\beta g_i(x)\right| |D^\gamma \Phi_i(x)| \\ &\leq C(m,N)\|f\|_\infty \epsilon^{-|\alpha|}\delta^{m-|\alpha|} \end{aligned}$$

Note that by choosing δ small enough we can ensure

$$C(m,N)\|f\|_\infty \sup_{|\alpha|<m} \delta^{m-|\alpha|}\epsilon^{-|\alpha|} < \sigma,$$

which proves (iv). Considering the case $|\gamma| = m$, we see that the proof of (iii) is complete for the case $p = \infty$.

We are left with the case $1 \leq p < \infty$ of (iii). Observe that for $|\gamma| > 0$

$$\text{supp } D^\gamma \Phi_i \subset \overline{T_i \backslash Q_i}$$

and

$$\frac{|\overline{T_i \backslash Q_i}|}{|T_i|} = 1 - (1-\epsilon/2N)^N < \epsilon/2$$

by Bernoulli's inequality.

Hence for $|\alpha| = m$ and $x \in T_i$

$$\begin{aligned} |D^\alpha u(x)| &= |D^\alpha(g_i\Phi_i)(x)| \\ &\leq |D^\alpha g_i(x)||\Phi_i(x)| + C \sum_{\substack{\beta+\gamma=\alpha \\ |\gamma|>0}} |D^\beta g_i(x)||D^\gamma \Phi_i(x)| \\ &\leq \left(⨍_{T_i} |f|\right) \chi_{T_i}(x) + C \sum_{\substack{\beta+\gamma=\alpha \\ |\gamma|>0}} \left(⨍_{T_i} |f|\right) \delta^{m-|\beta|}\epsilon^{-|\gamma|}\delta^{-|\gamma|}\chi_{\overline{T_i\backslash Q_i}}(x) \\ &\leq \left(\int_{T_i} |f|^p\right)^{1/p} |T_i|^{-1/p}\chi_{T_i}(x) \\ &\quad + C\epsilon^{-m}\left(\int_{T_i} |f|^p\right)^{1/p} |T_i|^{-1/p}\chi_{\overline{T_i\backslash Q_i}}(x). \end{aligned}$$

Thus

$$
\begin{aligned}
\|D^\alpha u\|_p &\le \|f\|_p + C\epsilon^{-m}\left(\sum_{i\in I}\left(\int_{T_i}|f|^p\right)\frac{|\overline{T_i\backslash Q_i}|}{|T_i|}\right)^{1/p}\\
&\le \|f\|_p(1+C\epsilon^{1/p-m}) \le C'\|f\|_p\epsilon^{1/p-m}
\end{aligned}
$$

□

Let $V \subset \mathbb{R}^n$ be open with $|V| < \infty$. Recall that if $u_i \in C_c^m(V)$ satisfies

$$u = \sum_{i=1}^{\infty} u_i \quad \text{converges uniformly in } V$$

and

$$\sum_{i=1}^{\infty} \|D^m u_i\|_\infty < \infty,$$

then $u \in C^m(V)$.

In the next lemma we will exhaust U by compact sets K_i and build a series $u_i \in C_c^m(U)$,

$$u = \sum_{i=1}^{\infty} u_i \in C_c^m(U).$$

The functions u_i will be constructed with the help of Lemma 6, so the partial sums of the series $D^m u_i$ will approximate a given continuous function $f = (f^\alpha)_{|\alpha|=m}$ on U. This will result in the fact that for the limiting function u, $D^m u$ will coincide with f on a large compact set.

Lemma 7. *Fix $m \in \mathbb{N}$ and let $U \subset \mathbb{R}^N$ be open with $|U| < \infty$. Let $f = (f^\alpha)_{|\alpha|=m}$ be a continuous and bounded function on U. For any $\epsilon, \sigma > 0$, there exists a function $u \in C_c^m(U)$ and a compact set $K \subset U$ such that the following hold:*

(i) $|U\backslash K| < \epsilon$,

(ii) $D^m u(x) = f(x)$ *for each* $x \in K$,

(iii) $\|D^m u\|_p \le C(m,N)\epsilon^{1/p-m}\|f\|_p$ *for all* $1 \le p \le \infty$,

(iv) $\|D^\gamma u\|_\infty < \sigma$ *for* $|\gamma| < m$.

Proof. We can assume that $f \neq 0$. Then the function

$$\varphi(p) = |U|^{1/p} \|f\|_p^{-1}, \quad p \in [1, \infty)$$

is continuous and $\varphi(p) \to \|f\|_\infty^{-1}$ as $p \to \infty$, so φ is bounded and hence

$$0 < A := \sup_{1 \le p < \infty} |U|^{1/p} \|f\|_p^{-1} < \infty.$$

Let $\eta_0 = \|f\|_\infty$ and $\eta_i = 2^{-(m+1)i} A^{-1},\ i = 1, 2, \ldots$ Then

$$\sum_{i=1}^{\infty} 2^{mi} \eta_i = A^{-1}.$$

Let $V \subset\subset U$ be open with $|V \backslash U| < \epsilon/2$. Let $f_1 = f|_V$. Applying Lemma 6, we select a compact subset K_1 of V and $u_1 \in C_c^m(V)$ such that

$$|V \backslash K_1| < 2^{-2}\epsilon,$$

$$|D^m u_1(x) - f_1(x)| < \eta_1 \quad \text{for } x \in K_1,$$

$$\|D^m u_1\|_p \le C(m, N)\epsilon^{1/p-m} \|f_1\|_p,$$

$$\|D^\gamma u_1\|_\infty \le 2^{-1}\sigma \quad \text{for } |\gamma| < m.$$

We will recursively construct sequences f_n, $K_n \subset U$ compact, $u_n \in C_c^m(V)$ such that

(I) $|V \backslash K_n| < 2^{-(n+1)}\epsilon$

(II) $|D^m u_n(x) - f_n(x)| < \eta_n$ for each $x \in K_n$

(III) $\|D^m u_n\|_p \le C(m, N)(2^{-n}\epsilon)^{1/p-m} \|f_n\|_p$

(IV) $\|D^\gamma u_n\|_\infty < 2^{-n}\sigma \quad$ for $|\gamma| < m$.

Assume that f_{n-1}, K_{n-1}, and u_{n-1} have been selected to satisfy (I) - (IV). Define a function $\widetilde{f}_n$ by

$$\widetilde{f}_n(x) = f_{n-1}(x) - D^m u_{n-1}(x), \qquad x \in \bigcap_{i=1}^{n-1} K_i$$

Applying Teizte extension theorem to $\widetilde{f}_n$ yields a continuous function f_n on U, which by (II) satisfies

$$\|f_n\|_\infty \leq \eta_{n-1}.$$

By Lemma 6, there is a compact set K_n and $u_n \in C_c^m(V)$ satisfying (I) - (IV).

Define $K = \bigcap_{i=1}^\infty K_i$. Clearly K is compact and

$$|U \backslash K| \leq |U \backslash V| + |V \backslash K| < \epsilon.$$

Define $u = \sum_{i=1}^\infty u_i$. To show (iii), for $p \in [1, \infty)$ we estimate

$$\begin{aligned}
\sum_{i=1}^\infty \|D^m u_i\|_p &\leq C(m,N)\epsilon^{\frac{1}{p}-m} \sum_{i=1}^\infty (2^{m-\frac{1}{p}})^i \|f_i\|_p \\
&\leq 2^m C(m,N)\epsilon^{\frac{1}{p}-m}\|f\|_p \left(1 + \|f\|_p^{-1} \sum_{i=2}^\infty 2^{m(i-1)} \|f_i\|_p\right) \\
&\leq 2^m C(m,N)\epsilon^{\frac{1}{p}-m}\|f\|_p \left(1 + \frac{|U|^{1/p}}{\|f\|_p} \sum_{i=2}^\infty 2^{m(i-1)} \|f_i\|_\infty\right) \\
&\leq 2^m C(m,N)\epsilon^{\frac{1}{p}-m}\|f\|_p \left(1 + A \sum_{i=2}^\infty (2^m)^{i-1} \eta_{i-1}\right) \\
&\leq 2^{m+1} C(m,N)\epsilon^{\frac{1}{p}-m}\|f\|_p.
\end{aligned} \tag{5.2}$$

Now we claim that $u \in C_c^m(U)$. By (IV), for $|\gamma| < m$ we have

$$\sum_{i=1}^\infty \|D^\gamma u_i\|_\infty < \sigma,$$

which implies the uniform convergence of the series in U. Moreover, note that since $|U| < \infty$, we can let $p \to \infty$ in (5.2) and hence

$$\sum_{i=1}^\infty \|D^m u_i\|_\infty \leq C'(m,N)\epsilon^{-m}\|f\|_\infty.$$

As we remarked preceeding the proof, this implies $u \in C^m(U)$. Since each u_i is supported in V and $V \subset\subset U$, this implies $u \in C_c^m(U)$, and (iii) and (iv) follows.

We are left with the proof of (ii). Fix $x \in K$. An easy inductive argument shows that

$$f_n(x) = f(x) - \sum_{i=1}^{n-1} D^m u_i(x).$$

Hence for every n,

$$|f(x) - \sum_{i=1}^{n} D^m u_i(x)| = |f_n(x) - D^m u_n(x)| < \eta_n.$$

Thus

$$\begin{aligned} |f(x) - D^m u(x)| &\leq |f(x) - \sum_{i=1}^{n} D^m u_i(x)| + \sum_{i=n+1}^{\infty} \|D^m u_i\|_\infty \\ &\leq \eta_n + \sum_{i=n+1}^{\infty} \|D^m u_i\|_\infty \to 0 \text{ as } n \to \infty. \end{aligned}$$

□

Any Lebesgue measurable function on a set $U \subset \mathbb{R}^n$ with $|U| < \infty$ is continuous and bounded outside a set of small measure. This fact allows us to prove a similar result to Lemma (7) without the restrictions that f be bounded or continuous. We simply isolate the region where f is badly behaved.

Lemma 8. *Let f be a Lebesgue measurable function on an open set $U \subset \mathbb{R}^N$ with $|U| < \infty$. Then for any $\epsilon > 0$, there is a compact set $K \subset U$ and a continuous, bounded function $\widetilde{f}$ on U such that*

(i) $|U \backslash K| < \epsilon$,

(ii) $f = \widetilde{f}$ *on* K,

(iii) $\|\widetilde{f}\|_p \leq 2\|f\|_p$ *for all* $p \in [1, \infty]$.

Proof. Fix $\epsilon > 0$. Suppose first that f is essentially unbounded. Then there exists an $R > 0$ such that

$$0 < |\{|f| > R\}| < \epsilon/2.$$

Let $K \subset \{|f| \leq R\}$ be a compact set such that $f|_K$ is continuous and

$$|\{|f| \leq R\} \backslash K| < |\{|f| > R\}| < \epsilon/2.$$

Let $\widetilde{f}$ be the Tietze extension of $f|_K$. Clearly $\|\widetilde{f}\|_\infty \leq R$. We have

$$U \backslash K \subset (\{|f| \leq R\} \backslash K) \cup \{|f| > R\}.$$

Hence $|U \backslash K| < \epsilon$. Also $\widetilde{f} = f$ on K. We are left with the estimate for the L^p norm.

$$\begin{aligned} \int_U |\widetilde{f}|^p &\leq \int_K |f|^p + \int_{\{|f| \leq R\} \backslash K} R^p + \int_{\{|f| > R\}} R^p \\ &\leq \int_K |f|^p + 2 \int_{\{|f| > R\}} R^p \leq 2 \int_U |f|^p. \end{aligned}$$

Now suppose that f is essentially bounded, say $\|f\|_\infty = M > 0$. If $|\{|f| = M\}| = 0$, then the proof follows from the previous argument since we can find $0 < R < M$ with

$$0 < |\{|f| > R\}| < \epsilon/2.$$

Thus we may suppose that $|\{|f| = M\}| > 0$. Let $K \subset U$ be compact such that $f|_K$ is continuous and

$$|U \backslash K| < \min \{\epsilon, |\{|f| = M\}|\}.$$

Let $\widetilde{f}$ be the Tietze extension of $f|_K$. Clearly $\|\widetilde{f}\|_\infty \leq M$.

As above, we estimate the L^p norm:

$$\begin{aligned} \int_U |\widetilde{f}|^p &\leq \int_K |f|^p + \int_{U \backslash K} M^p \\ &\leq \int_K |f|^p + \int_{\{|f| = M\}} M^p \\ &= \int_K |f|^p + \int_{\{|f| = M\}} |f|^p \leq 2 \int_U |f|^p \end{aligned}$$

□

As a consequence we have the following immediate result:

Theorem 40. *Fix $m \in \mathbb{N}$ and let $U \subset \mathbb{R}^n$ be open with $|U| < \infty$. Let $f = (f^\alpha)_{|\alpha|=m}$ be Borel. Then for any ϵ, $\sigma > 0$, there exists a function $u \in C_c^m(U)$ and an compact set $K \subset U$ such that for each $p \in [1, \infty]$ the following hold:*

(i) $|U \backslash K| < \epsilon$,

(ii) $D^m u(x) = f(x)$ *for each* $x \in K$,

(iii) $\|D^m u\|_p \leq C(m, N)\epsilon^{1/p-m}\|f\|_p$,

(iv) $\|D^\gamma u\|_\infty < \sigma$ *for* $|\gamma| < m$.

To prove Theorem 40, we simply note that by the lemma we can replace f by $\widetilde{f}$ which is bounded and continuous and apply Lemma 7, noting that $\widetilde{f} = f$ on a large compact set and $\|\widetilde{f}\|_p \leq 2\|f\|_p$.

5.2 MAIN RESULT 2

Now we come to the main result. We no longer require that the open set U has finite measure.

Theorem 41. *Let U be open in $\mathbb{R}^n$ and let $f = (f^\alpha)_{|\alpha|=m}$ be a Lebesgue measurable function defined on U. Then for any $\sigma > 0$, there is $u \in C^{m-1}(\mathbb{R}^n)$ which is m-times differentiable almost everywhere and such that*

$$D^m u(x) = f(x) \quad \text{for a.e. } x \in U.$$

$$\|D^\gamma u\|_\infty \leq \sigma \qquad \text{for each } |\gamma| < m.$$

Proof. Let $U_1 = U \cap B(0, 1)$. We claim that there is a compact set $K_1 \subset U_1$ and $u_1 \in C_c^m(U_1)$ such that:

$$D^m u_1(x) = f(x) \text{ for } x \in K_1,$$

$$|U_1 \backslash K_1| < 2^{-1},$$

$$|D^\gamma u_1(x)| < 2^{-1}\sigma \min\{1, \operatorname{dist}^2(x, U_1^c)\} \qquad x \in \mathbb{R}^N, |\gamma| < m. \tag{5.3}$$

Indeed, let $V_1 \subset\subset U_1$ with $|U_1 \backslash V_1| < 1/4$. According to Theorem 40, for any $\eta > 0$ there is a compact set $K_1 \subset V_1$ and $u_1 \in C_c^m(V_1)$ such that

$$|V_1 \backslash K_1| < 1/4 \ \text{ (and hence } |U_1 \backslash K_1| < 1/2).$$

$$D^m u_1 f(x) = f(x) \quad \text{for } x \in K_1$$

$$|D^\gamma u_1(x)| < \eta, \ \ x \in \mathbb{R}^N, |\gamma| < m. \tag{5.4}$$

Since $\operatorname{dist}(\overline{V_1}, U_1) > 0$, by taking η small enough, (5.4) implies (5.3).

Now we will construct a sequence of compact sets K_n and functions u_n by induction. Suppose that $K_1, \cdots, K_{n-1}$ and $u_1, \cdots, u_{n-1}$ have been defined. Let $U_n = U \cap B(0, n) \backslash (K_1 \cup \cdots \cup K_{n-1})$. Using a similar argument as above we may find a compact set $K_n \subset U_n$ and $u_n \in C_c^m(U_n)$ such that

$$D^m u_n(x) = f(x) - \sum_{i=1}^{n-1} D^m u_i(x) \quad \text{for } x \in K_n \tag{5.5}$$

$$|U_n \backslash K_n| < 2^{-n},$$

$$|D^\gamma u_n(x)| < 2^{-n} \min\{1, \operatorname{dist}^2(x, U_n^c)\} \quad x \in \mathbb{R}^N,\ |\gamma| < m.$$

Now let $C = \bigcup_{n=1}^\infty K_n$. It is easy to see that $|U \backslash C| = 0$. We will show that

$$u = \sum_{n=1}^\infty u_n$$

satisfies the claim of the theorem.

First, note that clearly $\operatorname{supp}(u) \subset \overline{U}$. Since for $|\gamma| < m$,

$$\sum_{n=1}^\infty \|D^\gamma u_n\|_\infty < \sigma,$$

it follows that $u \in C^{m-1}(\mathbb{R}^n)$ and

$$\|D^\gamma u\|_\infty \leq \sigma \quad \text{for } |\gamma| < m.$$

It remains to show that for $x \in C$, u is m-times differentiable at x and $D^m u(x) = f(x)$.

Let $x \in C$. Then $x \in K_n$ for some n. Observe that (5.5) implies that

$$\sum_{j=1}^{n} D^m u_j(x) = f(x).$$

Thus it remains to show that the function

$$\sum_{j>n} u_j \tag{5.6}$$

is m-times differentiable at x and the m-th derivative at x is 0. Since the function (5.6) is clearly of class C^{m-1} it suffices to show that for $|\gamma| = m - 1$

$$Dg(x) := D(\sum_{j>n} D^\gamma u_j)(x) = 0.$$

Since the functions $D^\gamma u_j$ are supported in U_j, and $x \notin U_j$ for $j > n$, we have

$$g(x) = \sum_{j>n} D^\gamma u_j(x) = 0.$$

Let $h \in \mathbb{R}^N$. If $x + h \notin U_j$, then $|D^\gamma u_j(x + h)| = 0$. On the other hand, if $x + h \in U_j$, since $x \notin U_j$ for $j > n$,

$$|D^\gamma u_j(x + h)| < 2^{-j}\sigma \min\{1, \operatorname{dist}^2(x + h, U_j^c)\} \leq 2^{-j}\sigma|h|^2.$$

Thus

$$\begin{aligned}|g(x + h) - g(x)| &= |g(x + h)| \\ &\leq \sum_{j>n} |D^\gamma u_j(x + h)| \\ &\leq \sigma|h|^2 \sum_{j>n} 2^{-j} < \sigma|h|^2.\end{aligned}$$

Then $Dg(x) = 0$, which completes the proof. □

6.0 CURRENT AND RELATED RESEARCH PROBLEMS

As a corollary of the Luzin-type approximation of BV^m functions, it was noted that this implies in the case $m = 2$ that

$$\textit{Given } f : \mathbb{R}^n \to \mathbb{R} \textit{ convex and } \epsilon > 0, \textit{ there exists } g \in C^2(\mathbb{R}^n) \textit{ such that } |\{f \neq g\}| < \epsilon.$$

A related question is whether convex functions can be approximated in the Luzin sense by C^2 *convex* functions. The difficulty is that the Whitney extension, or Whitney smoothing, does not preserve the convexity. This seems a natural question because we approximate a convex function by a C^2 function, and then 'smooth out' the corners.

In a 1988 paper, S. Imomkulov [15] provided a proof that subharmonic functions on $\mathbb{R}^n$ have the C^2-Luzin property. We are working on verifying this claim. His argument was based on the representation of any subharmonic function u as

$$u = K * \mu + v,$$

where v is harmonic, μ is a finite Borel measure, and K is the fundamental solution to the Laplace equation:

$$K(x) = \Phi(|x|) = \begin{cases} \frac{1}{2\pi} ln|x| & \text{n} = 2 \\ \frac{-\Gamma(n/2-1)}{4\pi^{n/2}} \frac{1}{|x|^{n-2}} & \text{n} \geq 3 \end{cases}$$

If we let $K_\delta(x) = \Phi(\sqrt{|x|^2 + \delta^2})$ be smoothed kernels, the claim is that for almost every $x \in \mathbb{R}^n$

$$\lim_{\delta \to 0} \int \frac{\partial^2}{\partial_{x_i} \partial_{x_j}} K_\delta(x-y) \mathrm{d}\mu(y) = \frac{1}{n} \delta_{i,j} \mu'(x) + \lim_{\delta \to 0} \int_{|x-y|>\delta} \frac{\partial^2}{\partial_{x_i} \partial_{x_j}} K(x-y) \mathrm{d}\mu(y),$$

where

$$\mu'(x) = \lim_{r \to 0} \frac{\mu(B(x,r))}{|B(x,r)|}.$$

This result is presented without proof and references a paper of Calderón and Zygmund [7]; however, the result is not proven explicitly in this paper either. Using the Calderón-Zygmund decomposition of $\mathbb{R}^n$ for the finite measure μ, we are attempting to verify this claim and the result.

Another direction is to improve the final result of this thesis. Specifically, we would like to show that given $f = \{f^\alpha\}_{|\alpha|=m}$, there exists a function $u \in \bigcap_{0<\lambda<1} C^{m-1,\lambda}(\mathbb{R}^n)$ differentiable almost everywhere, with $D^m u = f$ almost everywhere. This has applications to the paper of Balogh [5] related to the Heisenberg group.

BIBLIOGRAPHY

[1] R. Adams, *Sobolev Spaces.* Academic Press, New York (1975).

[2] G. Alberti. *A Lusin Type Theorem for Gradients,* J. Funct. Anal. 100 (1991), 100-118.

[3] G. Alberti, *On the Structure of Singular Sets of Convex Functions,* Calc. Var. Partial Differential Equations 2 (1994), 17-27.

[4] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems.* Oxford University Press, New York (2000).

[5] Z. Balogh, *Size of Characteristic Sets and Functions with Prescribed Gradient.* Journal Reine Agnew. Math. 564 (2003), 63-83.

[6] A.P. Calderón, A. Zygmund *Local Properties of Solutions of Partial Differential Equations.* Studia Math. 20 (1961), 171-225.

[7] A.P. Calderón, A. Zygmund *On the existence of certain singular integrals.* Acta Math. 88 (1952) 85-139.

[8] H. Federer. *Surface area II.* Trans. Amer. Math. Soc. 55 (1944), 438-456.

[9] L. C. Evans, R. Gareipy, *Measure Theory and Fine Properties of Functions.* CRC Press, Boca Raton (1992).

[10] L. C. Evans *Partial Differential Equations.* American Mathematical Society, USA (1998).

[11] D. Gilbard and N. Trudinger, *Elliptic Partial Differential Equations of Second Order.* Fundamental Principles of Mathematical Sciences Volume 224, Springer-Verlag, Berlin (1983).

[12] B. Bojarski, P. Hajłasz, P. Strzelecki, *Improved* $C^{k,\lambda}$ *approximation of higher order Sobolev functions in norm and capacity.* Indiana Univ. Mat. J. 51 (2002), 507-540.

[13] B.Bojarski, P.Hajłasz, *Pointwise inequalities for Sobolev functions and some applications.* Studia Math. 106 (1993), 77-92.

[14] L. Hedberg, *On Certain Convolution Inequalities,* Proc. Amer. Math Soc. 36 (1972) 505-510.

[15] S. Imomkulov, *Twice differentiability of subharmonic functions* Izv. RAN. Ser. Mat. 56 (1992) 877-888.

[16] F. Liu, *A Lusin Type Property for Sobolev Functions* Indiana Univ. Math J. 26 (1977) 645-651.

[17] N. Luzin, *Sur la notion de l'integrale,* Annali Mat. Pura e Appl. 26 (1917) 77-129.

[18] B. Malgrange, *Ideals of Differentiable Functions.* Oxford Univ. Press, London (1966).

[19] L. Moonens and W. Pfeffer, *The Multidimensional Luzin Theorem.* J. Math. Anal. Appl. 339 (2008) 746-752.

[20] J. Michael and W.P. Ziemer, *A Lusin Type Approximation of Sobolev Functions by Smooth Functions,* Contemp. Math 42 (1985), 135-167.

[21] O. Nikodym, *Sur une classe de fonctions considerees dans le probleme de Dirichlet.* Fund. Math 21 (1933) 129-150.

[22] W. Rudin, *Real & Complex Analysis.* McGraw-Hill (1986).

[23] S. Saks, *Theory of the Integral.* G.E. Stechert and Company, New York (1937).

[24] D. Swanson, *Pointwise Inequalities and Approximation in Fractional Sobolev Spaces,* Studia Math. 149 (2002) 147-174.

[25] H. Whitney, *Analytic Extensions of Differentiable Functions Defined in Closed Sets.* Trans. Amer. Math Soc. 36 (1934), 63-89.

[26] H. Whitney, *On totally differentiable and smooth functions.* Pacific J. Math. Volume 1, Number 1 (1951), 143-159.

[27] W. P. Zeimer, *Weakly Differentiable Functions.* Springer, New York (1989).

[28] W. P. Zeimer, *Uniform Differentiability of Sobolev Functions, Sobolev Spaces and Functions of Bounded Variation,* Indiana Univ. Math. J., 37 (1988), 789-799.