Indian Statistical Institute, Kolkata

The Generalized Divergence Theorem for Sobolev Functions and the Co-Area Formula for BV Functions: An Exposition

End-Semester Report in Geometric Measure Theory (Third Semester)

Author: Jishnu Parayil Shibu

Supervisor: Partha Sarathi Chakraborty

Date: December 10, 2025

0 Abstract

This report provides an exposition of two central results in geometric measure theory: the generalized divergence theorem for Sobolev functions and the co-area formula for functions of bounded variation. These theorems extend classical results from smooth calculus to non-smooth settings that arise naturally in the study of PDEs, variational problems, and the theory of minimal surfaces. We first develop the theory of Sobolev functions and establish the properties needed to state and prove the generalized divergence theorem. We then present the theory of BV functions, which provides a rigorous framework for defining the variation of a function in the weakest possible sense, and culminate with the co-area formula for BV functions. This expresses the total variation of a function as an integral of the perimeters of its superlevel sets, generalizing both Fubini's theorem and the classical co-area formula for Lipschitz functions.

Contents

1 Introduction

The aim of this report is to present an exposition of two fundamental results in geometric measure theory: the *generalized divergence theorem for Sobolev functions*, and the *co-area formula for functions of bounded variation.*

Classically, the divergence theorem provides a relation between the integral of the divergence of a smooth vector field over a domain and the flux of the vector field across the boundary. However, in many applications, one encounters functions that are not smooth everywhere but lie in a Sobolev space $W^{1,p}(U)$, i.e. functions in L^p functions whose distributional derivative also lies in L^p (see Definition 2.5). In this setting, the classical theorem no longer applies directly. The central question is:

Question 1.1. Given a bounded domain $U \subset \mathbb{R}^n$ with suitable regularity conditions (if need be) and a Sobolev function $f \in W^{1,p}(U)$, can we still relate the integral of the divergence of a vector field weighted by f to a boundary integral in a meaningful way?

The *generalized divergence theorem* (see Theorem 2.16) answers this question affirmatively by assuming U has Lipschitz boundary (see Definition 2.11) and introducing the notion of the *trace* (see Theorem 2.14) of a Sobolev function on the boundary. Note that making sense of a function on the boundary is not obvious: if $f \in C^\infty(\overline{U})$, the restriction $f|_{\partial U}$ is well-defined, but for $f \in W^{1,p}(U)$ or even $f \in L^p(U)$, the function is defined only a.e. w.r.t $\mathcal{L}^n$, and usually (but not always) $\mathcal{L}^n(\partial U) = 0$, so a priori the values on the boundary are meaningless. The trace operator provides a canonical way to assign boundary values to such functions, and we shall see that this is indeed a meaningful generalization.

After studying Sobolev functions, we then turn to a larger class of functions, i.e. functions of bounded variation $BV_{loc}(U)$ (see Definition 3.2). These functions allow us to make sense of the variation $\mathrm{Var}(f, U)$ of a function f in the weakest possible setting, namely when the derivative is a signed Radon vector-valued measure (see Theorem 3.7). Recall that in the classical setting, when $f \in C^1(U)$, variation of f is given by $\int_U |Df|(x)$ and we shall see that the newly defined notion of variation $\mathrm{Var}(f, U)$ indeed coincides with this.

BV functions also provide a meaningful notion of *perimeter* for highly irregular sets. For instance, if $E \subset \mathbb{R}^n$ satisfies suitable conditions, the characteristic function χ_E belongs to BV, even though it may not lie in any Sobolev space. The total variation of χ_E then provides a rigorous definition of the perimeter of E. We shall see that in the case of sets with Lipschitz boundary, this coincides with $\mathcal{H}^{n-1}(\partial E)$ (see Proposition 3.9).

The co-area formula will give us a meaningful way of understanding $\mathrm{Var}(f, U)$ in terms of an integral over the perimeter of its superlevel sets. We first recall that for a Lipschitz function $f : U \to \mathbb{R}$, the classical co-area formula gives

$$\int_U |Df|(x)\, dx = \int_{-\infty}^{\infty} \mathcal{H}^{n-1}(\partial E_t)\, dt,$$

where $|Df|(x)$ exists $\mathcal{L}^n$-a.e. by Rademacher's theorem, and $E_t := \{x \in U \mid f(x) > t\}$. This indeed expresses the total variation of f as an integral over the perimeters of its superlevel sets E_t.

A natural question arises: can the same formula be generalized to BV functions, where the derivative exists only as a measure and the level sets may be even more irregular? In other words, we ask ourselves:

Question 1.2. For a function $f \in BV(U)$, can we express its total variation $\mathrm{Var}(f, U)$ in terms of the geometry of its level sets?

As we will see, the natural generalization of the co-area formula to BV functions will provide an affirmative answer (see Theorem 3.16).

It is worth noting that BV functions admit their own notions of trace and satisfy a corresponding divergence theorem, results that are even more general than those in the Sobolev setting. For the sake of brevity, we omit these developments here, although their proofs follow the similar pattern as in the Sobolev case.

With this in mind, the report is organized as follows. In the first part, we develop the theory of Sobolev functions, covering basic properties, density theorems, and traces, and culminating in the generalized divergence theorem. In the second part, we introduce BV functions, establish analogous density results, explore the notion of perimeter, and conclude with a proof of the co-area formula for BV functions. The main reference for this report is [2]

2 Generalized Divergence Theorem for Sobolev Functions

Throughout this report, $U \subset \mathbb{R}^n$ denotes an open set.

2.1 Approximate Identities and Mollifiers

Approximate identities and mollifiers are central tools in establishing the fundamental properties of Sobolev functions. In this section, we recall their basic properties and introduce the standard mollifier that will be used throughout this report.

Definition 2.1. An *approximate identity* is a family $\{k_\lambda\}_{\lambda>0} \subset L^1(\mathbb{R}^n)$ such that:

(a) $\displaystyle\int_{\mathbb{R}^n} k_\lambda(x)\,dx = 1$ for every $\lambda > 0$;

(b) $\displaystyle\sup_{\lambda>0} \|k_\lambda\|_1 < \infty$;

(c) for every $\delta > 0$, $\lim_{\lambda\to 0} \int_{|x|\geq\delta} |k_\lambda(x)|\,dx = 0$.

Proposition 2.2. *Recall the following fundamental results:*

(a) If $f \in L^p_{\text{loc}}(U)$, $1 \leq p < \infty$, then $f * k_t \to f$ in $L^p_{\text{loc}}(U)$. meaning that for every open set $V \subset\subset U$ (i.e., V compactly contained in U), $\|f * k_t - f\|_{L^p(V)} \to 0$.

(b) If f is continuous, then $f * k_t(x) \to f(x)$ pointwise for every x at which f is continuous. If f is uniformly continuous, then the convergence is uniform on compact sets.

Definition 2.3 (Mollifier)**.** A *mollifier* is an approximate identity $\{\eta_\epsilon\}_{\epsilon>0}$ on $\mathbb{R}^n$ such that each η_ϵ is non-negative and belongs to $C_c^\infty(\mathbb{R}^n)$, i.e.,

$$\eta_\epsilon \geq 0, \quad \eta_\epsilon \in C_c^\infty(\mathbb{R}^n), \quad \int_{\mathbb{R}^n} \eta_\epsilon(x)\,dx = 1.$$

The key advantage of mollifiers is that for any $f \in L^1_{\text{loc}}(U)$, the convolution $f * \eta_\epsilon$ is defined and smooth on $U_\epsilon = \{x \in U : \text{dist}(x, \partial U) > \epsilon\}$, a

Notation 2.4. We fix the following throughout the rest of the notes:

(i) For $\epsilon > 0$, define the ϵ-interior of U by $U_\epsilon := \{x \in U \mid \text{dist}(x, \partial U) > \epsilon\}$.

(ii) Define a smooth, compactly supported function $\eta : \mathbb{R}^n \to \mathbb{R}$ by

$$\eta(x) := \begin{cases} c\exp\left(\frac{1}{|x|^2-1}\right), & |x| < 1, \\ 0, & |x| \geq 1, \end{cases}$$

where the constant $c > 0$ is chosen so that $\int_{\mathbb{R}^n} \eta(x)\,dx = 1$.

(iii) For $\epsilon > 0$, define the scaled mollifier

$$\eta_\epsilon(x) := \frac{1}{\epsilon^n}\eta\left(\frac{x}{\epsilon}\right), \quad x \in \mathbb{R}^n.$$

Unless otherwise stated, η will always refer to this choice. While we could develop the theory by considering general mollifiers with additional properties such as symmetry etc, we use this standard one for simplicity.

(iv) For $f \in L^1_{\mathrm{loc}}(U)$, define the mollification

$$f_\epsilon(x) := (\eta_\epsilon * f)(x) = \int_U \eta_\epsilon(x-y)\, f(y)\, dy, \quad x \in U_\epsilon.$$

2.2 Sobolev Space $W^{1,p}_{\mathrm{loc}}(U)$ and Basic Properties

In this sections, we defined Sobolov spaces and prove a few basic properties.

Definition 2.5 (Sobolev Space $W^{1,p}$)**.** Let $1 \le p \le \infty$ and let $U \subset \mathbb{R}^n$ be open.

(i) A function f belongs to the Sobolev space $W^{1,p}(U)$ if $f \in L^p(U)$ and for each $i = 1, \dots, n$, the weak partial derivative $\partial_{x_i} f$ (which exists in the sense of distributions) additionally satisfies $\partial_{x_i} f \in L^p(U)$.

(ii) If $f \in W^{1,p}(U)$, we define norms

$$\|f\|_{W^{1,p}(U)} := \left(\int_U |f(x)|^p + |Df(x)|^p\, dx\right)^{1/p}, \quad 1 \le p < \infty,$$

$$\|f\|_{W^{1,\infty}(U)} := \operatorname{ess\,sup}_{x\in U}\Big(|f(x)| + |Df(x)|\Big).$$

This makes $W^{1,p}(U)$ a Banach space. It also follows immediately from the definition of the norm that a sequence $\{f_k\} \subset W^{1,p}(U)$ converges to $f \in W^{1,p}(U)$ if and only if

$$f_k \to f \text{ in } L^p(U) \quad \text{and} \quad \partial_{x_i} f_k \to \partial_{x_i} f \text{ in } L^p(U) \text{ for each } i = 1, \dots, n.$$

(iii) A function f belongs to the local Sobolev space $W^{1,p}_{\mathrm{loc}}(U)$ if $f \in W^{1,p}(V)$ for every open set $V \subset\subset U$. We refer to f as a *Sobolev function.*

Moreover, a sequence $\{f_k\}$ is said to converge to f in $W^{1,p}_{\mathrm{loc}}(U)$ if for every open set $V \subset\subset U$ we have $f_k \longrightarrow f \quad \text{in } W^{1,p}(V)$.

Notation 2.6. For locally integrable functions f, g on an open set $U \subset \mathbb{R}^n$, we denote

$$\langle f, g\rangle := \int_U f(x)\, g(x)\, dx,$$

whenever the integral is well defined.

Remark 2.7. In classical L^p theory, one usually works with functions defined on all of $\mathbb{R}^n$, and when a function is originally given in $L^p(U)$, where U open subset, it can be extended to $\mathbb{R}^n$ by declaring it to be 0 outside U. This causes no difficulty, since the extended functions is still in L^p.

For Sobolev functions, however, the situation is different. A function $f \in W^{1,p}(U)$ need not admit an extension to a function in $W^{1,p}(\mathbb{R}^n)$ unless U satisfies additional geometric conditions (for example, being a Lipschitz domain). Consequently, when we mollify f by convolving with a standard mollifier η_ε, the expression $f * \eta_\varepsilon(x)$ is defined only when the ball $B_\varepsilon(x)$ is contained in U. Thus, in general, the mollification $f * \eta_\varepsilon$ is well defined only on the inner domain

$$U_\varepsilon := \{x \in U : \operatorname{dist}(x, \partial U) > \varepsilon\},$$

and not necessarily on all of U.

This point is important to keep in mind when proving density and approximation results in Sobolev spaces: one often needs to combine mollification with cut-off functions or use other techniques such partitions of unity etc.

We now prove a few basic properties for Sobolov functions.

Lemma 2.8. *Let $U \subset \mathbb{R}^n$ be open, let $1 \le p < \infty$, and suppose $f \in W^{1,p}_{\mathrm{loc}}(U)$. Let $\{\eta_\varepsilon\}_{\varepsilon>0}$ be the standard mollifiers (Notation 2.4) and set $f_\varepsilon := f * \eta_\varepsilon$. Then for every open V with $V \subset\subset U$ there exists $\varepsilon_0 > 0$ such that f_ε is well defined on V for $0 < \varepsilon < \varepsilon_0$ and*

$$f_\varepsilon \longrightarrow f \quad \text{in } W^{1,p}(V) \qquad (\varepsilon \to 0).$$

Thus $f_\varepsilon \to f$ in $W^{1,p}_{\mathrm{loc}}(U)$.

Proof. Fix V with $V \subset\subset U$ and choose $\varepsilon_0 > 0$ with $\operatorname{dist}(V, \partial U) > \varepsilon_0$. For $0 < \varepsilon < \varepsilon_0$ the convolution $f_\varepsilon = f * \eta_\varepsilon$ is well defined and smooth on V.
We first show that for each $i = 1, \ldots, n$,

$$\partial_i (f * \eta_\varepsilon) = (\partial_i f) * \eta_\varepsilon \qquad \text{on } V,$$

where $\partial_i f$ denotes the weak partial derivative of f.

Indeed, let $\varphi \in C_c^\infty(V)$ be arbitrary. Using the definition of distributional derivative and Fubini/Tonelli,

$$\begin{aligned}
\Big\langle \partial_i (f * \eta_\varepsilon), \varphi \Big\rangle &= -\Big\langle f * \eta_\varepsilon, \partial_i \varphi \Big\rangle = -\Big\langle f,\, \eta_\varepsilon^{\sim} * \partial_i \varphi \Big\rangle \\
&= -\Big\langle f,\, \partial_i (\eta_\varepsilon^{\sim} * \varphi) \Big\rangle = \Big\langle \partial_i f,\, \eta_\varepsilon^{\sim} * \varphi \Big\rangle \\
&= \Big\langle (\partial_i f) * \eta_\varepsilon, \varphi \Big\rangle,
\end{aligned}$$

where $\eta_\varepsilon^{\sim}(x) := \eta_\varepsilon(-x)$. Since this identity holds for every test function $\varphi \in C_c^\infty(V)$ we obtain the claimed equality.

Convergence in $W^{1,p}(V)$ is now immediate. Because $f \in L^p_{\mathrm{loc}}(U)$ and $\partial_i f \in L^p_{\mathrm{loc}}(U)$, by Proposition 2.2,

$$f_\varepsilon \longrightarrow f \quad \text{in } L^p(V), \qquad \partial_i(f * \eta_\varepsilon) = (\partial_i f) * \eta_\varepsilon \longrightarrow \partial_i f \quad \text{in } L^p(V)$$

as $\varepsilon \to 0$. Thus $f_\varepsilon \to f$ in $W^{1,p}(V)$.
Since $V \subset\subset U$ was arbitrary, this proves convergence $f_\varepsilon \to f$ in $W^{1,p}_{\mathrm{loc}}(U)$. □

The following lemma provides a sufficient condition under which the product of two Sobolev functions is again a Sobolev function, and ensures that the Leibniz rule holds.

Lemma 2.9. *Let $1 \leq p < \infty$ and let $U \subset \mathbb{R}^n$ be open. If $f \in W^{1,p}(U)$ and $g \in W^{1,p}(U) \cap L^\infty(U)$, then $fg \in W^{1,p}(U)$ and for each $i = 1, \ldots, n$,*

$$\partial_i(fg) = (\partial_i f)\, g + f\, (\partial_i g),$$

Proof. Fix $\varphi \in C_c^\infty(U)$. Then using Lemma 2.8, DCT (which we can use as $g \in L_\infty(U)$), and Leibniz rule for smooth functions, we get:

$$\begin{aligned}
\left\langle \partial_i(fg), \varphi \right\rangle &= -\left\langle fg, \partial_i \varphi \right\rangle \\
&= \lim_{\varepsilon \to 0} -\left\langle f_\varepsilon g_\varepsilon,\, \partial_i \varphi \right\rangle \\
&= \lim_{\varepsilon \to 0} \left\langle \partial_i(f_\varepsilon g_\varepsilon),\, \varphi \right\rangle \\
&= \lim_{\varepsilon \to 0} \left\langle \partial_i(f_\varepsilon) g \,+ f\, \partial_i(g_\varepsilon),\, \varphi \right\rangle \\
&= \left\langle (\partial_i f) g \,+ f\, (\partial_i g),\, \varphi \right\rangle
\end{aligned}$$

Thus $\partial_i(fg) = (\partial_i f)\, g + f\, (\partial_i g)$, as distributions, and the right-hand side belongs to $L^p(U)$, so $fg \in W^{1,p}(U)$. □

2.3 Density Theorems for Sobolev Spaces

Theorem 2.10 (Density of Smooth Functions)**.** *Let $1 \leq p < \infty$ and let $U \subset \mathbb{R}^n$ be open. Then $C^\infty(U)$ is dense in $W^{1,p}(U)$*

Proof. The idea of the proof is as follows. Fix $1 \leq p < \infty$ and let $f \in W^{1,p}(U)$. A first attempt would be to approximate f by its mollifications f_ε, which are smooth and converge to f in $W^{1,p}$ on domains where the convolution is well-defined. However, as noted earlier, the mollification f_ε is not defined on all of U unless f is extended beyond the boundary. To overcome this, we work locally: we cover U by coordinate patches on which convolution is admissible, and then glue these local approximations together using a partition of unity.

With this in mind, let $\varepsilon > 0$. Choose a locally finite covering of U by relatively compact sets $\{V_i\}_{i=1}^{\infty}$. Then, there exists a partition of unity $\{\varphi_i\}_{i=1}^{\infty} \subset C_c^{\infty}(U)$ subordinate to $\{V_i\}_{i=1}^{\infty}$: each $0 \leq \varphi_i \leq 1$, $\operatorname{supp} \varphi_i \subset V_i$, the family $\{\operatorname{supp} \varphi_i\}$ is locally finite, and $\sum_{i=1}^{\infty} \varphi_i(x) = 1$ for all $x \in U$.

Define $g_k := \varphi_k f \in W^{1,p}(U)$. and by Lemma 2.9, $g_k \in W^{1,p}(U)$ and the Leibniz rule holds for g_k. By Lemma 2.8, for each k there exists $\varepsilon_k > 0$ small enough so that $\operatorname{supp}(g_{k,\varepsilon_k}) \subset V_i$, where $g_{k,\varepsilon_k} := g_k * \eta_{\varepsilon_k}$ and

$$\|g_{k,\varepsilon_k} - g_k\|_{L^p(U)} < \frac{\varepsilon}{2^k}, \qquad \|Dg_{k,\varepsilon_k} - Dg_k\|_{L^p(U)} < \frac{\varepsilon}{2^k}$$

(We can do this because mollifiers approximate in $W^{1,p}$ on any fixed relatively compact set containing $\operatorname{supp} g_i$.)

Set

$$g_\varepsilon := \sum_{i=1}^{\infty} g_{k,\varepsilon_k} \in C^{\infty}(U).$$

and this is well defined as for each $x \in U$, $g_{k,\varepsilon_k}(x) = 0$ except for a finite number of k (as V_i is locally finite and $\operatorname{supp}(g_{k,\varepsilon_k}) \subset V_i$). Thus, as $f = \sum_{k=1}^{\infty} g_k$ by definition, we get:

$$\|g_\varepsilon - f\|_{L^p(U)} = \left\| \sum_{k=1}^{\infty} \left(g_{k,\varepsilon_k} - g_k \right) \right\|_{L^p(U)} \leq \sum_{k=1}^{\infty} \left\| g_{k,\varepsilon_k} - g_k \right\|_{L^p(U)} \leq \varepsilon.$$

Similarly,

$$\|Dg_\varepsilon - Df\|_{L^p(U)} = \left\| \sum_{k=1}^{\infty} \left(Dg_{k,\varepsilon_k} - Dg_k \right) \right\|_{L^p(U)} \leq \sum_{k=1}^{\infty} \left\| Dg_{k,\varepsilon_k} - Dg_k \right\|_{L^p(U)} \leq \varepsilon.$$

Thus, $g_\varepsilon \in C^{\infty}(U)$ converges to f in $W^{1,p}(U)$, completing the proof. □

Next, we introduce domains U with Lipschitz boundary, a class of domains frequently used in analysis. Such domains enjoy the key property that an outward-pointing unit normal vector to ∂U exists for $\mathcal{H}^{n-1}$-almost every point, as we shall see. We then establish a stronger density result in this case, showing that functions in $W^{1,p}(U)$ can be approximated by functions $f_k \in C^{\infty}(\overline{U}) \cap W^{1,p}(U)$. Here, $C^{\infty}(\overline{U})$ consists of all functions $f \in C^{\infty}(U)$ such that for every multiindex α, the derivative $D^\alpha f$ extends continuously to $\overline{U}$.

Definition 2.11 (Lipschitz Boundary)**.** We say that an open set $U \subset \mathbb{R}^n$ has a *Lipschitz boundary* if for every point $x \in \partial U$, there exist $r > 0$ and a Lipschitz continuous function $\gamma : \mathbb{R}^{n-1} \to \mathbb{R}$ such that, after a suitable rotation and relabeling of coordinates, denoting the rotated coordinates by $(y_1, ..., y_n)$, we have:

$$U \cap Q(x,r) = \{y \in Q(x,r) \mid \gamma(y_1, \dots, y_{n-1}) < y_n\}$$

where

$$Q(x,r) := \{y \in \mathbb{R}^n \mid |y_i - x_i| < r,\ i = 1, \dots, n\}$$

Theorem 2.12 (Smooth Approximation on Lipschitz Domains)**.** *Let U be a bounded domain with Lipschitz boundary.*
For every $f \in W^{1,p}(U)$, $1 \le p < \infty$, there exists a sequence $\{f_k\} \subset W^{1,p}(U) \cap C^\infty(\overline{U})$ such that

$$f_k \longrightarrow f \quad \text{in } W^{1,p}(U) \text{ as } k \to \infty.$$

Here, $C^\infty(\overline{U})$ consists of all functions $f \in C^\infty(U)$ such that for every multiindex α, the derivative $D^\alpha f$ extends continuously to $\overline{U}$.

Proof. The idea of the proof is as follows. For each $x \in \partial U$, choose a cube $Q(x,r)$ as in Definition 2.11. Since U is bounded, ∂U is compact, and therefore there exist finitely many such cubes $Q(x_1,r_1), \dots, Q(x_N,r_N)$ whose union covers ∂U. Choose a smooth partition of unity $\{\psi_k\}_{k=0}^N$ subordinate to the cover $\{U, Q(x_1,r_1), ..., Q(x_n,r_n)\}$ such that

$$\operatorname{supp}\psi_0 \subset U, \qquad \operatorname{supp}\psi_k \subset U \cap Q(x_k,r_k) \quad \text{for } k = 1, \dots, N.$$

Define

$$f_k := f\,\psi_k \text{ and thus } f = \sum_{k=0}^N f_k$$

By Theorem 2.10, the function f_0 can be approximated in $W^{1,p}(U)$ by smooth functions g_k s.t. $\operatorname{supp}(g_k) \subset V$ where V is some pre-compact set in U s.t. $\operatorname{supp}(f_o) \subset V \subset \overline{V} \subset U$. Thus, each of these approximants g_k may be extended by 0 up to ∂U, giving functions in $C^\infty(\overline{U})$.

What remains is to treat f_k for $k = 1, \dots, N$. If this can be done, since f is a finite sum of these f_k, we will be done. Note that each such f_k is supported in $U \cap Q(x_k,r_k)$, i.e., in a region that intersects ∂U. Thus we must construct smooth approximations that also extend smoothly to ∂U, and this is where the Lipschitz structure of the boundary will be used. We now explain how to do this in the following steps.

Step 1. Work in the rotated coordinates $(y_1, \dots, y_n)$ associated to $x \in \partial U$ as in Definition 2.11, so that

$$U \cap Q(x,r) = \Big\{ y \in Q(x,r) : y_n > \gamma(y_1, \dots, y_{n-1}) \Big\},$$

where γ is Lipschitz. Let e_n denote the n-th coordinate unit vector in these $(y_1, \dots, y_n)$ coordinates.
Fix a constant $\alpha > \mathrm{Lip}(\gamma) + 1$. For any $y \in \overline{U \cap Q(x,r/2)}$ and $\varepsilon > 0$, define

$$y_\varepsilon := y + \varepsilon\alpha\, e_n.$$

As $y \in \overline{Q(x, r/2)}$, there exists $\varepsilon_0 > 0$ small enough s.t. $B(y_{\varepsilon_0}, \varepsilon_0) \subset Q(x, r)$

Claim. By our choice of ε_0 and α, we in fact get the stronger statement

$$B(y_{\varepsilon_0}, \varepsilon_0) \subset Q(x, r) \cap U \;\; \textit{for} \;\; y \in \overline{Q(x, r/2)}$$

Proof. To show $B(y_{\varepsilon_0}, \varepsilon_0) \subset U$, let $z \in B(y_{\varepsilon_0}, \varepsilon_0)$. Write $z = (z', z_n)$ and $y = (y', y_n)$ with $z', y' \in \mathbb{R}^{n-1}$. Then

$$\|z' - y'\| \leq \varepsilon, \qquad |\, z_n - (y_n + \alpha\varepsilon) \,| \leq \varepsilon.$$

By Lipschitz continuity of γ,

$$\gamma(z') \leq \gamma(y') + \mathrm{Lip}(\gamma) \, \|z' - y'\| \leq \gamma(y') + \mathrm{Lip}(\gamma) \, \varepsilon.$$

Moreover,

$$z_n \geq y_n + \alpha\varepsilon - \varepsilon = y_n + (\alpha - 1)\varepsilon.$$

Hence

$$z_n - \gamma(z') \geq (y_n - \gamma(y')) + \Big(\alpha - 1 - \mathrm{Lip}(\gamma)\Big)\varepsilon.$$

Since $y \in \overline{U}$ implies $y_n \geq \gamma(y')$, and $\alpha > \mathrm{Lip}(\gamma) + 1$ implies the second part is strictly positive, we get $z_n - \gamma(z') > 0$. Thus $z \in U$. This proves our claim and completes step 1. □

Step 2. Let $x \in \partial U$ and work in the rotated coordinates $(y_1, \ldots, y_n)$ as in Step 1. Fix a radius $r > 0$ and let $f \in W_{1,p}(U)$ s.t.

$$\operatorname{supp} f \subset Q(x, r/2) \cap U,$$

Choose $\alpha > \mathrm{Lip}(\gamma) + 1$ and, by Step 1, select $\varepsilon_0 > 0$ such that for every $0 < \varepsilon < \varepsilon_0$ and every $y \in Q(x, r/2)$ we have

$$B\Big(y_\varepsilon, \varepsilon\Big) \subset Q(x, r) \cap U \;\; \text{where} \;\; y_\varepsilon = y + \alpha\varepsilon e_n$$

For $0 < \varepsilon < \varepsilon_0$ define the translated mollification

$$f_\varepsilon(y) := (f * \eta_\varepsilon)\Big(y_\varepsilon\Big) \;\; = \;\; \int_U f(z) \, \eta_\varepsilon\Big(y + \alpha\varepsilon e_n - z\Big) \, dz, \qquad y \in \overline{Q(x, r/2) \cap U}$$

Fix a multi-index α and for $y \in \overline{Q(x, r/2) \cap U}$, we have the formula

$$D^\alpha f_\varepsilon(y) = \int_U f(z) \, D^\alpha \eta_\varepsilon\Big(y + \alpha\varepsilon e_n - z\Big) \, dz,$$

obtained by DCT and differentiating under the integral. The integrand is continuous in y for each fixed z, and is dominated by $|f(z)| \sup_w |D^\alpha \eta_\varepsilon(w)|$, which is integrable on the compact set $\operatorname{supp} f$. Hence, by the dominated convergence theorem, the map $y \mapsto D^\alpha f_\varepsilon(y)$ is continuous on $\overline{Q(x, r/2) \cap U}$. Since α was arbitrary, every derivative of f_ε (of any order) extends continuously to the boundary of the local chart; in particular $f_\varepsilon \in C^\infty(\overline{Q(x, r/2) \cap U})$.

We now show $f_\varepsilon \to f$ in $W^{1,p}(U)$ as $\varepsilon \downarrow 0$. Write the translation operator T_h by $T_h g(\cdot) = g(\cdot + h)$. Then

$$f_\varepsilon = T_{\alpha\varepsilon e_n}(f * \eta_\varepsilon).$$

Fix any $1 \le p < \infty$. Using the triangle inequality,

$$\|f_\varepsilon - f\|_{L^p} \le \left\|T_{\alpha\varepsilon e_n}(f * \eta_\varepsilon) - f * \eta_\varepsilon\right\|_{L^p} + \left\|f * \eta_\varepsilon - f\right\|_{L^p}.$$

The second term tends to 0 as $\varepsilon \downarrow 0$ by the standard approximation-by-mollifiers theorem. The first term also tends to 0 because translations are continuous in L^p: for any fixed integrable function g, $\|T_h g - g\|_{L^p} \to 0$ as $h \to 0$. Hence $\|f_\varepsilon - f\|_{L^p} \to 0$.

For the gradients we note that:

$$Df_\varepsilon = T_{\alpha\varepsilon e_n}\Big(Df * \eta_\varepsilon\Big).$$

The same two-term estimate and the same two facts (mollification in L^p and continuity of translations in L^p) imply

$$\|Df_\varepsilon - Df\|_{L^p} \longrightarrow 0 \qquad (\varepsilon \downarrow 0).$$

Combining the two convergences shows

$$f_\varepsilon \longrightarrow f \quad \text{in } W^{1,p}(U).$$

Step 3. Combining everything, we conclude the proof. Recall that we wrote

$$f = \sum_{k=0}^{N} f_k, \qquad f_k := f\,\psi_k.$$

By construction,

$$\operatorname{supp} f_0 \subset U, \qquad \operatorname{supp} f_k \subset U \cap Q(x_k, r_k) \quad (k \ge 1).$$

For f_0, Theorem 2.10 yields a sequence $\{g_j^{(0)}\} \subset C^\infty(\overline{U})$ with $g_j^{(0)} \to f_0$ in $W^{1,p}(U)$. For each $k = 1, \dots, N$, Step 2 gives a family of smooth functions $\{g_j^{(k)}\} \subset C^\infty(\overline{U})$such that

$$g_j^{(k)} \to f_k \quad \text{in } W^{1,p}(U).$$

Define

$$G_j := \sum_{k=0}^{N} g_j^{(k)}.$$

Each G_j lies in $C^\infty(\overline{U})$, and by the triangle inequality,

$$\|G_j - f\|_{W^{1,p}(U)} \le \sum_{k=0}^{N} \|g_j^{(k)} - f_k\|_{W^{1,p}(U)} \longrightarrow 0.$$

Thus $G_j \to f$ in $W^{1,p}(U)$, and the proof is complete.

□ □

2.4 Trace and a Generalized Divergence Theorem

We begin by proving the existence of an outward-pointing unit normal vector $\nu(x)$ for $\mathcal{H}^{n-1}$-almost every $x \in \partial U$, when U has Lipschitz boundary.

Proposition 2.13. *Let $U \subset \mathbb{R}^n$ be a bounded open set with Lipschitz boundary. Then the outer unit normal ν to ∂U exists for $\mathcal{H}^{n-1}$–almost every point of ∂U.*
Moreover, wherever $\nabla\gamma(y')$ exists, ν is given by:

$$\nu(y', \gamma(y')) \;=\; \frac{1}{\sqrt{1 + |\nabla\gamma(y')|^2}} \big(-\nabla\gamma(y'),\, 1\big).$$

Proof. By Definition 2.11, for each $x_0 \in \partial U$ there is a cube $Q(x_0, r)$ and rotated coordinates $(y', y_n) = (y_1, \dots, y_{n-1}, y_n)$ in which

$$U \cap Q(x_0, r) = \{y \in Q(x_0, r) : y_n > \gamma(y')\},$$

for some Lipschitz function $\gamma : \mathbb{R}^{n-1} \to \mathbb{R}$. Because ∂U is compact we may cover it by finitely many such charts; it suffices therefore to prove the claim in a single chart.

Fix one such chart and write

$$\Gamma := \{(y', \gamma(y')) : y' \in W\},$$

where $W \subset \mathbb{R}^{n-1}$ is the projection of $Q(x_0, r)$ onto the first $n-1$ coordinates. The boundary $\partial U \cap Q(x_0, r)$ is exactly Γ in these coordinates.

Since γ is Lipschitz, Rademacher's theorem implies that γ is differentiable for Lebesgue–a.e. $y' \in W$. Let $S \subset W$ denote the measure-zero set of non-differentiability points.
Define the parametrisation

$$\phi : W \longrightarrow \Gamma, \qquad \phi(y') := (\, y',\, \gamma(y')\,).$$

Since γ is Lipschitz, the map ϕ is Lipschitz between W and Γ, and therefore

$$\mathcal{H}^{n-1}(\phi(S)) \leq \mathcal{L}^{n-1}(S) = 0,$$

Thus, ϕ is differentiable for $\mathcal{H}^{n-1}$ a.e. Whenever ϕ is differentiable, the outward unit normal to U at the boundary point $\phi(y') = (y', \gamma(y'))$ is given directly by the standard formula for a graph:

$$\nu\big(\phi(y')\big) \;=\; \frac{1}{\sqrt{1+|\nabla\gamma(y')|^2}}\big(\nabla\gamma(y'),\, -1\big).$$

Since this argument applies to almost every point of Γ and the boundary is covered by finitely many such graphs, we conclude that the outer unit normal ν_U exists for $\mathcal{H}^{n-1}$-a.e. point of ∂U, as required. $\square$

In many applications—especially in the study of boundary value problems for PDEs—it is essential to give meaning to the boundary values of a Sobolev function. If $f \in C(\overline{U})$, its restriction $f|_{\partial U}$ is well defined. However, for $f \in W^{1,p}(U)$, the situation is more subtle: such functions are only defined up to sets of Lebesgue measure zero, and since in many cases, $\mathcal{L}^n(\partial U) = 0$, the pointwise restriction to ∂U is generally meaningless. The trace theorem provides a systematic way to assign boundary values to Sobolev functions, extending the classical notion of restriction in a consistent and useful way.

Theorem 2.14 (Traces of Sobolev functions)**.** *Assume $U \subset \mathbb{R}^n$ is bounded with Lipschitz boundary and let $1 \leq p < \infty$. There exists a bounded linear operator*

$$T : W^{1,p}(U) \longrightarrow L^p(\partial U; \mathcal{H}^{n-1})$$

(the trace operator*) such that*

$$Tf = f|_{\partial U} \qquad \textit{for all } f \in W^{1,p}(U) \cap C(\overline{U}).$$

Proof. We first give an idea of the proof. By Theorem 2.12, the subspace $W^{1,p}(U)\cap C^\infty(\overline{U})$ is dense in $W^{1,p}(U)$. Thus it suffices to define the trace operator T on $W^{1,p}(U) \cap C^\infty(\overline{U})$ by

$$Tf := f|_{\partial U} \qquad (f \in W^{1,p}(U) \cap C^\infty(\overline{U}))$$

and to show that this map is bounded from $W^{1,p}(U) \cap C^\infty(\overline{U}) \to L^p(\partial U; \mathcal{H}^{n-1})$. Once boundedness on is established, T extends uniquely (by continuity) to a bounded linear operator on all of $W^{1,p}(U)$.

Since ∂U is compact we may cover it by finitely many local charts $\{Q(x_j, r_j)\}_{j=1}^N$ (as in Definition 2.11). Let $\{\psi_j\}_{j=1}^N$ be a partition of unity subordinate to the cover $\{U, Q(x_1, r_1), ..., Q(x_N, r_N)\}$

this partition of unity we reduce the global boundedness to a finite collection of local estimates: it is enough to prove that for each chart $Q = Q(x_j, r_j)$ there exists a constant C_Q such that

$$\|f\|_{L^p(\partial U\cap Q;\mathcal{H}^{n-1})} \le C_Q \|f\|_{W^{1,p}(U)}$$

for every $f \in \mathcal{W}^{1,p}(U) \cap C^\infty(\overline{U})$ with $\operatorname{supp} f \subset U \cap Q$. (Indeed, writing $f = \sum_{j=0}^N f\psi_j$, one handles the boundary pieces by the local inequalities above and the interior piece f_0 can be safely ignored as $f_0|_{\partial U} = 0$)

Therefore the proof reduces to the following single local statement. □

Lemma 2.15 (Local trace estimate)**.** *Let $Q = Q(x, r)$ be a boundary cube from Definition 2.11, and let $1 \le p < \infty$. There exists a constant C with the following property: for every $f \in \mathcal{W}^{1,p}(U) \cap C^\infty(\overline{U})$ satisfying $\operatorname{supp} f \subset U \cap Q$ we have*

$$\|f\|_{L^p(\partial U\cap Q;\mathcal{H}^{n-1})} \le C\,\|f\|_{W^{1,p}(U)}.$$

Proof. Fix a boundary cube $Q = Q(x, r)$ and work in the rotated coordinates (y', y_n) from Definition 2.11, so that $U \cap Q = \{(y', y_n) \in Q : y_n > \gamma(y')\}$ with γ Lipschitz. By the expression of the outer normal ν in Proposition 2.13 there is a constant $c > 0$ (depending on $\operatorname{Lip}(\gamma)$) such that

$$-e_n\cdot\nu(y) \ge c > 0 \qquad \text{for } \mathcal{H}^{n-1}\text{-a.e. } y \in \partial U \cap Q. \tag{*}$$

Let $1 \le p < \infty$ and let $f \in W^{1,p}(U) \cap C^\infty(\overline{U})$ satisfy $\operatorname{supp} f \subset U \cap Q$. To treat the power $|f|^p$ smoothly we approximate the map $t \mapsto |t|^p$ by smooth functions. For $\varepsilon > 0$ define

$$\beta_\varepsilon(t) := \left(t^2 + \varepsilon^2\right)^{p/2} - \varepsilon^p, \qquad t \in \mathbb{R}.$$

Then $\beta_\varepsilon \in C^\infty(\mathbb{R})$, $\beta_\varepsilon \ge 0$, $\beta_\varepsilon(t) \uparrow |t|^p$ pointwise as $\varepsilon \downarrow 0$, and

$$\beta_\varepsilon'(t) = p\,t\left(t^2 + \varepsilon^2\right)^{\frac{p-2}{2}}$$

satisfies the growth bound

$$|\beta_\varepsilon'(t)| \le C_p\left(|t|^{p-1} + \varepsilon^{p-1}\right) \le C_p'\left(1 + |t|^{p-1}\right),$$

From this, one can derive that $|\beta_\varepsilon'(f)|^{p'} \lesssim 1 + |f|^p$ and is integrable on U.
Choose a fixed smooth vector field $\varphi \in C_c^1(Q; \mathbb{R}^n)$ with $\varphi \equiv e_n$ on a neighbourhood of $\partial U \cap Q$ (this is possible since $\partial U \cap Q$ is compactly contained in Q). Apply the classical divergence theorem (Gauss–Green) to the compactly supported smooth function $\beta_\varepsilon(f)$

and the vector field φ to obtain

$$\int_{\partial U\cap Q} \beta_\varepsilon(f)(\varphi\cdot\nu)\, d\mathcal{H}^{n-1} = \int_{U\cap Q} \beta_\varepsilon(f)\ \operatorname{div}\varphi\, dy + \int_{U\cap Q} \varphi\cdot\nabla\Big(\beta_\varepsilon(f)\Big)\, dy.$$

By the choice of φ and $(*)$ there is a constant $C_1 > 0$ with

$$\int_{\partial U\cap Q} \beta_\varepsilon(f)\, d\mathcal{H}^{n-1} \le C_1\left(\int_{U\cap Q} \beta_\varepsilon(f)\, dy + \int_{U\cap Q}\Big|\nabla(\beta_\varepsilon(f))\Big|\, dy\right).$$

Using the chain rule,

$$\nabla(\beta_\varepsilon(f)) = \beta_\varepsilon'(f)\,\nabla f,$$

and hence by Hölder's inequality (with conjugate exponent p') we get

$$\int_{U\cap Q}\Big|\beta_\varepsilon'(f)\,\nabla f\Big|\, dy \le \left(\int_{U\cap Q} |\beta_\varepsilon'(f)|^{p'}\, dy\right)^{1/p'} \|\nabla f\|_{L^p(U)}.$$

As noted above $|\beta_\varepsilon'(f)|^{p'} \lesssim 1 + |f|^p$, so the first factor on the right is bounded by a constant multiple of $1 + \|f\|_{L^p(U)}$. Combining the preceding estimates yields

$$\int_{\partial U\cap Q} \beta_\varepsilon(f)\, d\mathcal{H}^{n-1} \le C\Big(\|f\|^p_{L^p(U)} + \|f\|^{p-1}_{L^p(U)}\|\nabla f\|_{L^p(U)} + \|\nabla f\|_{L^p(U)}\Big)$$

for some constant $C = C(Q,p)$. By Young's inequality the right-hand side is controlled by a constant times $\|f\|^p_{L^p(U)} + \|\nabla f\|^p_{L^p(U)}$. Therefore

$$\int_{\partial U\cap Q} \beta_\varepsilon(f)\, d\mathcal{H}^{n-1} \le C'\Big(\|f\|^p_{L^p(U)} + \|\nabla f\|^p_{L^p(U)}\Big).$$

Finally let $\varepsilon \downarrow 0$. By monotone convergence $\beta_\varepsilon(f) \uparrow |f|^p$ pointwise and the integrable bound just obtained, so we may pass to the limit and obtain

$$\int_{\partial U\cap Q} |f|^p\, d\mathcal{H}^{n-1} \le C'\Big(\|f\|^p_{L^p(U)} + \|\nabla f\|^p_{L^p(U)}\Big).$$

Taking p-th roots gives the desired estimate

$$\|f\|_{L^p(\partial U\cap Q;\mathcal{H}^{n-1})} \le C\,\|f\|_{W^{1,p}(U)}.$$

□

With the trace theorem in place, we can now extend the classical divergence theorem to Sobolev functions. Even though functions in $W^{1,p}(U)$ need not possess pointwise derivatives or well-defined boundary values, the trace operator allows the familiar boundary term to be recovered in a meaningful way. The following result is the Sobolev–space generalization of the divergence theorem.

Theorem 2.16 (Generalized divergence theorem)**.** *Let $U \subset \mathbb{R}^n$ be bounded with Lipschitz boundary, and let $1 \leq p < \infty$. For every $\varphi \in C^1(\mathbb{R}^n;\mathbb{R}^n)$ and every $f \in W^{1,p}(U)$ there holds*

$$\int_U f \operatorname{div} \varphi \, dx = -\int_U Df \cdot \varphi \, dx + \int_{\partial U} (\varphi \cdot \nu)\, Tf \, d\mathcal{H}^{n-1},$$

where ν denotes the outer unit normal on ∂U and Tf is the trace of f on ∂U given by Theorem 2.14.

Proof. Let $U \subset \mathbb{R}^n$ be bounded with Lipschitz boundary, fix $1 \leq p < \infty$, and let $\varphi \in C^1(\mathbb{R}^n;\mathbb{R}^n)$ and $f \in W^{1,p}(U)$ be given. By Theorem 2.12 there exists a sequence $\{f_k\} \subset W^{1,p}(U) \cap C^\infty(\overline{U})$ with

$$f_k \longrightarrow f \quad \text{in } W^{1,p}(U).$$

For each k the classical divergence theorem on the Lipschitz domain U applied to the smooth function f_k yields

$$\int_U f_k \operatorname{div} \varphi \, dx = -\int_U \nabla f_k \cdot \varphi \, dx + \int_{\partial U} (\varphi \cdot \nu)\, f_k \, d\mathcal{H}^{n-1}. \tag{$*$}$$

Passing to the limit $k \to \infty$ in the interior terms is immediate: since $f_k \to f$ in $L^p(U)$ and $\nabla f_k \to \nabla f$ in $L^p(U;\mathbb{R}^n)$, and since φ and $\operatorname{div} \varphi$ are bounded on $\overline{U}$, we have by DCT

$$\int_U f_k \operatorname{div} \varphi \, dx \longrightarrow \int_U f \operatorname{div} \varphi \, dx, \qquad \int_U \nabla f_k \cdot \varphi \, dx \longrightarrow \int_U \nabla f \cdot \varphi \, dx.$$

By Theorem 2.14 the trace operator $T : W^{1,p}(U) \to L^p(\partial U; \mathcal{H}^{n-1})$ is bounded and linear, and $Tf_k = f_k|_{\partial U}$ for each k. Thus

$$\int_{\partial U} (\varphi \cdot \nu)\, Tf_k \, d\mathcal{H}^{n-1} \longrightarrow \int_{\partial U} (\varphi \cdot \nu)\, Tf \, d\mathcal{H}^{n-1},$$

Passing to the limit in $(*)$ now yields

$$\int_U f \operatorname{div} \varphi \, dx = -\int_U \nabla f \cdot \varphi \, dx + \int_{\partial U} (\varphi \cdot \nu)\, Tf \, d\mathcal{H}^{n-1},$$

which is exactly the claimed identity. $\square$

3 Co-Area Formula for BV functions

3.1 BV Functions $BV_{\mathrm{loc}}(U)$

This section is devoted to developing the notion of *variation* for functions $f : U \subset \mathbb{R}^n \to \mathbb{R}$, and to introducing the broader class of functions of *bounded variation*, denoted $BV_{\mathrm{loc}}(U)$. These functions play a central role in geometric measure theory: among their many applications, they provide a natural framework for defining and studying the *perimeter* of highly irregular sets $E \subset \mathbb{R}^n$.

We begin by recalling the classical definition of the variation of a function $f : U \subset \mathbb{R} \to \mathbb{R}$ via partitions of an interval, together with an equivalent expression when f is C^1. This one–dimensional viewpoint motivates the general definition of bounded variation for functions $f : U \subset \mathbb{R}^n \to \mathbb{R}$. We then establish a fundamental property of BV functions: every such function admits a distributional derivative represented by an $\mathbb{R}^n$-valued signed Radon measure. Finally, we discuss the relationship between Sobolev functions and functions of bounded variation.

Recall 3.1 (Variation on an interval)**.** Let $f : [a, b] \to \mathbb{R}$. For any finite partition $\mathcal{P} = \{\, a = x_0 < x_1 < \cdots < x_N = b \,\}$, define

$$V(f, \mathcal{P}) := \sum_{i=1}^{N} |f(x_i) - f(x_{i-1})|.$$

The *total variation* of f on $[a, b]$ is

$$\mathrm{Var}_{[a,b]}(f) := \sup_{\mathcal{P}} V(f, \mathcal{P}),$$

and we say f has bounded variation on $[a, b]$ if $\mathrm{Var}_{[a,b]}(f) < \infty$.
If in addition $f \in C^1([a, b])$, then the classical identity

$$\mathrm{Var}_{[a,b]}(f) = \int_a^b |f'(x)|\, dx \tag{1}$$

holds, giving the familiar representation of the variation in terms of the derivative.

Although the next definition of bounded variation in higher dimensions may at first appear unrelated to the one recalled above, we will later see that it recovers the classical one-dimensional formula expressed through the integral of the gradient.

Definition 3.2 (Functions of bounded variation)**.** Let $U \subset \mathbb{R}^n$ be open and let $f \in L^1(U)$.

(a) The *variation* of f on U is defined by

$$V(f, U) := \sup\Big\{ \int_U f \operatorname{div} \varphi \, dx \;\Big|\; \varphi \in C_c^1(U; \mathbb{R}^n),\ |\varphi(x)| \le 1 \text{ for all } x \in U \Big\}.$$

If $V(f,U) < \infty$, we say that f has *bounded variation* on U.

(b) The space of functions of bounded variation on U is

$$BV(U) := \{ f \in L^1(U) : V(f,U) < \infty \}.$$

We equip $BV(U)$ with the norm

$$\|f\|_{BV(U)} := \|f\|_{L^1(U)} + V(f,U),$$

with respect to which $BV(U)$ is a Banach space.

(c) We write

$$BV_{\mathrm{loc}}(U) := \Big\{ f \in L^1_{\mathrm{loc}}(U) : V(f,W) < \infty \text{ for every open } W \subset\subset U \Big\},$$

and refer to such functions as being of locally bounded variation.

Remark 3.3. If $f \in C^1(U) \cap BV(U)$, then the above definition of variation agrees with the classical formula. Indeed, for any $\varphi \in C^1_c(U;\mathbb{R}^n)$ with $\|\varphi\|_{L^\infty} \le 1$, we get:

$$\left| \int_U f \operatorname{div} \varphi \, dx \right| = \left| - \int_U Df \cdot \varphi \, dx \right| \le \int_U |Df \cdot \varphi| \, dx \le \int_U |Df| \, dx.$$

To see that equality is attained, choose

$$\varphi(x) := \begin{cases} \dfrac{Df(x)}{|Df(x)|}, & \text{if } Df(x) \neq 0, \\ 0, & \text{otherwise,} \end{cases}$$

Note that φ need not be compactly supported. This is easily remedied by choosing an exhaustion of U by compact sets $\{K_j\}_{j=1}^\infty$ with $K_j \subset \operatorname{int}(K_{j+1})$ and $\bigcup_j K_j = U$. For each j pick a cutoff $\chi_j \in C^1_c(U)$ satisfying

$$0 \le \chi_j \le 1, \qquad \chi_j \equiv 1 \text{ on } K_j, \qquad \operatorname{spt}(\chi_j) \subset \operatorname{int}(K_{j+1}).$$

Define $\varphi_j := \chi_j \varphi$. Then $\varphi_j \in C^1_c(U;\mathbb{R}^n)$, $\|\varphi_j\|_{L^\infty} \le 1$, and by DCT,

$$- \int_U f \operatorname{div} \varphi_j \, dx \to \int_U |Df| \, dx,$$

which leads to $V(f,U) = \int_U |Df| \, dx$
Thus the general notion of variation recovers the classical one-dimensional identity recalled in (1).

To show the existence of a derivative for $f \in BV(U)$ as a vector-valued signed Radon measure, we first recall a slight variant of the Riesz–Markov–Kakutani theorem where the functional is defined on $C_c(\mathbb{R}^n;\mathbb{R}^m)$.

Theorem 3.4 (Riesz–Markov-Kakutani Representation). *Let $T : C_c(\mathbb{R}^n;\mathbb{R}^m) \to \mathbb{R}$ be a linear functional such that, for every compact set $K \subset \mathbb{R}^n$, the restriction of T to $\{\varphi \in C_c(\mathbb{R}^n;\mathbb{R}^m) : \operatorname{spt}(\varphi) \subset K\}$ is bounded. Then there exists a unique Radon measure μ on $\mathbb{R}^n$ and a unique μ-measurable function*

$$\sigma : \mathbb{R}^n \to \mathbb{R}^m, \qquad |\sigma(x)| = 1 \textit{ for } \mu\textit{-a.e. } x,$$

such that

$$T(\varphi) = \int_{\mathbb{R}^n} \varphi \cdot \sigma \, d\mu \qquad \textit{for all } \varphi \in C_c(\mathbb{R}^n;\mathbb{R}^m).$$

Remark 3.5. In the proof of Theorem 3.4, the representing measure μ is constructed such that for any open $V \subset \mathbb{R}^n$,

$$\mu(V) = \sup\Big\{ T(\varphi) : \varphi \in C_c(\mathbb{R}^n;\mathbb{R}^m),\ |\varphi| \leq 1,\ \operatorname{spt}(\varphi) \subset V \Big\}.$$

From this, we may be recoverer $\mu(A)$ for any set A by the Radon property.

Notation 3.6. By a *vector-valued signed Radon measure* μ on $U \subset \mathbb{R}^n$, we mean a measure of the form

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

where each μ_i is a signed Radon measure on U. For any measurable function $f : U \to \mathbb{R}$, the integral of f with respect to μ is understood componentwise:

$$\int_U f \, d\mu := \left(\int_U f \, d\mu_1, \ldots, \int_U f \, d\mu_n \right).$$

Theorem 3.7. *(Existence of Vector-Valued Signed Radon Measure) Let $f \in BV_{\mathrm{loc}}(U)$. Then there exists a vector-valued signed Radon measure*

$$Df = (D_1 f, \ldots, D_n f)$$

such that for all $\varphi \in C_c^1(U;\mathbb{R}^n)$,

$$\int_U f \operatorname{div} \varphi \, dx = -\int_U \varphi \cdot d(Df).$$

Moreover, Df admits the following structure: there exists a Radon measure $\|Df\|$ on U and a $\|Df\|$-measurable function

$$\sigma : U \to \mathbb{R}^n, \quad |\sigma(x)| = 1 \quad \|Df\|\textit{-a.e.},$$

such that each component $D_i f$ is represented as

$$D_i f(A) = \int_A \sigma_i \, d\|Df\|, \quad i = 1, \dots, n,$$

for any set $A \subset U$.

Proof. We first define the linear functional

$$T : C_c^1(U; \mathbb{R}^n) \longrightarrow \mathbb{R}, \quad T(\varphi) := -\int_U f \operatorname{div} \varphi \, dx.$$

To apply Theorem 3.4, we need to show T restricted to functions suported on a compact set is bounded, as well as extend this linear functional to $C_c(U; \mathbb{R}^n)$.
We first note that the restriction of T to functions supported in a open set $W \subset\subset U$ is bounded as for all $\varphi \in C_c^1(U; \mathbb{R}^n)$ with $\operatorname{supp}(\varphi) \subset W$, we have

$$|T(\varphi)| \leq V(f, W) \, \|\varphi\|_{L^\infty} \quad \text{for all } \varphi \in C_c^1(U; \mathbb{R}^n) \text{ with } \operatorname{supp}(\varphi) \subset V. \tag{2}$$

Now, fix an arbitrary compact set $K \subset U$, and let $\psi \in C_c(U; \mathbb{R}^n)$ with $\operatorname{supp}(\psi) \subset K$. To extend T to such a ψ, we construct a sequence $(\varphi_n) \subset C_c^1(U; \mathbb{R}^n)$ converging uniformly to ψ on K as follows:

1. Choose an open set W s.t $K \subset W \subset \overline{W} \subset U$. Then $\operatorname{dist}(K, U \setminus W) > 0$.
2. Let $\{\eta_\epsilon\}_{\epsilon > 0}$ be the standard mollifiers from 2.4. By 2.2, we know that

$$\psi * \eta_\epsilon \longrightarrow \psi \quad \text{uniformly on } V \text{ as } \epsilon \to 0.$$

3. Choose a sequence $\epsilon_n \to 0$ sufficiently small so that $\operatorname{supp}(\psi * \eta_{\epsilon_n}) \subset V$. Define

$$\varphi_n := \psi * \eta_{\epsilon_n} \in C_c^1(U; \mathbb{R}^n).$$

By (2), the sequence $T(\varphi_n)$ is Cauchy and therefore convergent. Define

$$T(\psi) := \lim_{n \to \infty} T(\varphi_n),$$

which provides a well-defined extension of T to $C_c(U; \mathbb{R}^n)$ which satisfies hypothesis of Theorem 3.4.
By Theorem 3.4, there exists a Radon measure μ and a μ-measurable function $\sigma : U \to \mathbb{R}^n$ such that $|\sigma(x)| = 1$ for μ-a.e. x and

$$T(\varphi) = \int_U \varphi \cdot \sigma \, d\mu \quad \text{for all } \varphi \in C_c^1(U; \mathbb{R}^n).$$

Taking $\|Df\| := \mu$ and defining

$$Df_i(A) := \int_A \sigma_i \, d\|Df\| \quad \text{for Borel sets } A \subset U,$$

we obtain the desired representation, concluding the proof of the theorem.

$\square$

Remark 3.8 ($||Df||$ is the Variational measure)**.** Let $f \in BV_{\text{loc}}(U)$. For any $W \subset\subset U$, by Remark 3.5 we have

$$\|Df\|(W) = \sup\left\{ \int_W f \operatorname{div} \varphi \, dx : \varphi \in C_c^1(W; \mathbb{R}^n), \|\varphi\|_{L^\infty} \leq 1 \right\} = V(f, W).$$

The measure $\|Df\|$ is called the *variational measure* of f.

Proposition 3.9. $W^{1,1}_{\text{loc}}(U) \subset BV_{\text{loc}}(U)$

Proof. Let $f \in W^{1,1}_{\text{loc}}(U)$. For any $\varphi \in C_c^1(U; \mathbb{R}^n)$, we have

$$\int_U f \operatorname{div} \varphi \, dx = -\int_U \nabla f \cdot \varphi \, dx,$$

where here ∇f denotes the weak gradient, to distinguish it from the vector-valued Radon measure Df introduced above.

Taking the supremum over all φ with $\|\varphi\|_{L^\infty} \leq 1$ and $\operatorname{supp}(\varphi) \subset W \subset\subset U$, we obtain

$$V(f, W) \leq \int_W |\nabla f| \, dx < \infty,$$

showing that $f \in BV_{\text{loc}}(U)$.
Moreover, choosing

$$\varphi = \frac{\nabla f}{|\nabla f|} \chi_{\{|\nabla f| > 0\}}$$

for each compact $V \subset\subset U$, we achieve equality in the supremum. It follows in particular that for any $W \subset\subset U$,

$$V(f, W) = \int_W |\nabla f| \, dx.$$

By the uniqueness in the Riesz–Markov–Kakutani representation, it follows that the total variation measure satisfies

$$\|Df\| = |\nabla f| \, dx, \qquad \sigma = \frac{\nabla f}{|\nabla f|}, \quad \text{and} \quad Df(A) = \int_A \nabla f \, dx$$

for sets $A \subset U$. $\square$

Proposition 3.10. *Let $f \in BV_{\text{loc}}(U)$. Then its distributional derivative Df admits a Lebesgue decomposition with respect to $\mathcal{L}^n$ of the form*

$$Df = \nabla f\, dx + Df_{\text{s}},$$

where $\nabla f \in L^1_{\text{loc}}(U;\mathbb{R}^n)$ is the density of the absolutely continuous part of Df, and Df_{s} is the singular part of Df.

Proof. Let $f \in BV_{\text{loc}}(U)$ and let Df denote its vector-valued signed Radon measure derivative. By Lebesgue decomposition theorem, we can decompose

$$Df = Df_{\text{ac}} + Df_{\text{s}}$$

with respect to the Lebesgue measure $\mathcal{L}^n$. In particular, $Df_{\text{ac}} = \nabla f\, dx$ represents the absolutely continuous part, while Df_{s} is singular. This shows that, in contrast to Sobolev functions which only have an absolutely continuous derivative, BV functions may possess an additional singular component in their derivative. □

3.2 Density Theorems for BV Functions

Theorem 3.11 (Lower semicontinuity of the variation measure)**.** *Suppose $f_k \in BV(U)$ for $k = 1, 2, \dots$ and $f_k \to f$ in $L^1_{\text{loc}}(U)$. Then*

$$\|Df\|(U) \le \liminf_{k\to\infty} \|Df_k\|(U).$$

Proof. Let $\varphi \in C^1_c(U;\mathbb{R}^n)$ satisfy $\|\varphi\|_{L^\infty} \le 1$. Since $\operatorname{supp}\varphi$ is compact, by DCT

$$\int_U f_k \operatorname{div}\varphi\, dx \longrightarrow \int_U f \operatorname{div}\varphi\, dx, \qquad (k\to\infty),$$

For each fixed φ we therefore have

$$\int_U f \operatorname{div}\varphi\, dx = \lim_{k\to\infty}\int_U f_k \operatorname{div}\varphi\, dx \le \liminf_{k\to\infty} \sup_{\substack{\psi\in C^1_c(U;\mathbb{R}^n)\\ \|\psi\|_\infty\le 1}} \int_U f_k \operatorname{div}\psi\, dx = \liminf_{k\to\infty}\|Df_k\|(U),$$

Taking the supremum over all such φ on the left-hand side gives

$$\|Df\|(U) = \sup_{\substack{\varphi\in C^1_c(U;\mathbb{R}^n)\\ \|\varphi\|_\infty\le 1}} \int_U f \operatorname{div}\varphi\, dx \le \liminf_{k\to\infty}\|Df_k\|(U),$$

completing the proof. □

Theorem 3.12 (Density of smooth functions in BV)**.** *Assume $f \in BV(U)$. Then there exists a sequence*

$$\{f_k\}_{k=1}^\infty \subset BV(U)\cap C^\infty(U)$$

such that $f_k \to f$ in $L^1(U)$, and $\|Df_k\|(U) \longrightarrow \|Df\|(U)$ as $k \to \infty$.

Proof. The proof is largely very similar to the density proof for Sobolov functions given in Theorem 2.10, and we omit it for brevity. □

3.3 Perimeter of a Set

Definition 3.13 (Sets of locally finite perimeter)**.** Let $E \subset \mathbb{R}^n$ be a measurable set and let $U \subset \mathbb{R}^n$ be open. We say that E *has locally finite perimeter in U* if the characteristic function χ_E belongs to $BV_{\mathrm{loc}}(U)$.

In this case, χ_E admits a vector-valued Radon measure $D\chi_E$ as its distributional derivative, and possesses an associated variational measure $\|D\chi_E\|$ (see Remark 3.8). In this case, we shall henceforth denote $\|D\chi_E\|$ simply by $\|\partial E\|$.

Note that it satistifes the following: if $W \subset U$ is compactly contained, then

$$\|\partial E\|(W) = \sup\left\{\int_{E\cap W} \operatorname{div}\varphi\, dx : \varphi \in C_c^1(W;\mathbb{R}^n),\ \|\varphi\|_{L^\infty} \le 1\right\}.$$

This quantity is a generalization of the notion of perimeter and can be viewed as the "perimeter of E in W." The next proposition makes this precise.

Proposition 3.14. *Let $E \subset \mathbb{R}^n$ be a bounded open set with Lipschitz boundary. Let $U \subset \mathbb{R}^n$ be open and assume $\mathcal{H}^{n-1}(U \cap \partial E) < \infty$. Then $\chi_E \in BV(U)$; moreover the distributional derivative of χ_E is the measure*

$$D\chi_E(A) = \int_A -\nu_E\, d\Big(\mathcal{H}^{n-1}\lfloor_{\partial E}\Big), \qquad A \subset U$$

where ν_E denotes the outer unit normal to E (defined $\mathcal{H}^{n-1}$–a.e. on ∂E). In particular

$$\|\partial E\| = \|D\chi_E\| = \mathcal{H}^{n-1}\lfloor_{\partial E} \quad \text{on } U,$$

and consequently

$$\|\partial E\|(U) = \mathcal{H}^{n-1}(U \cap \partial E).$$

Proof. Let $\varphi \in C_c^1(U;\mathbb{R}^n)$. Since E has Lipschitz boundary and is bounded, the generalized divergence theorem (Theorem 2.16) gives

$$\int_E \operatorname{div}\varphi\, dx = \int_{\partial E} \varphi \cdot \nu_E\, d\mathcal{H}^{n-1}.$$

Because $\operatorname{supp}\varphi \subset U$, this can be written as

$$\int_U \chi_E\ \operatorname{div}\varphi\, dx = \int_{U\cap\partial E} \varphi \cdot \nu_E\, d\mathcal{H}^{n-1}.$$

Taking absolute values and using $|\varphi \cdot \nu_E| \le \|\varphi\|_{L^\infty}$, we obtain

$$\left| \int_U \chi_E \operatorname{div} \varphi \, dx \right| \le \|\varphi\|_{L^\infty} \mathcal{H}^{n-1}(U \cap \partial E).$$

Taking the supremum over all $\varphi \in C_c^1(U; \mathbb{R}^n)$ with $\|\varphi\|_{L^\infty} \le 1$ yields

$$V(\chi_E, U) \le \mathcal{H}^{n-1}(U \cap \partial E) < \infty,$$

so $\chi_E \in BV(U)$.
From the identity above we have, for all $\varphi \in C_c^1(U; \mathbb{R}^n)$,

$$\int_U \chi_E \operatorname{div} \varphi \, dx = \int_U \varphi \cdot (\nu_E) \, d\Big(\mathcal{H}^{n-1} \lfloor_{\partial E}\Big).$$

But the vector-valued Radon derivative satisfies

$$\int_U \chi_E \operatorname{div} \varphi \, dx = - \int_U \varphi \cdot \sigma \, d\Big(\|\partial E\|\Big).$$

and so by uniqueness,

$$\sigma = -\nu \;\text{ and }\; \|D\chi_E\| = \mathcal{H}^{n-1} \lfloor_{\partial E} \quad \text{on } U,$$

Therefore

$$\|\partial E\|(U) = \mathcal{H}^{n-1}(U \cap \partial E),$$

completing the proof. □

3.4 The Co-Area Formula for BV Functions

The preceding proposition shows that, for an open set E with Lipschitz boundary, our generalized notion of perimeter coincides with the classical geometric quantity $\mathcal{H}^{n-1}(\partial E)$. This confirms that the BV–theoretic definition is a genuine extension of the classical one.

We now turn to the coarea formula for BV functions, a fundamental result that provides further intuition for why the $||Df||$ indeed captures variation of f. Indeed, the coarea formula expresses the total variation of a BV function as an integral of the perimeters of its level sets. Before stating and proving it, we first establish the following preliminary lemma:

Lemma 3.15. *Let $f \in BV(U)$ and define*

$$E_t := \{ \, x \in U : f(x) > t \, \}, \qquad t \in \mathbb{R}.$$

Then the map

$$t \longmapsto \|\partial E_t\|(U)$$

is Lebesgue measurable.

Proof. Consider the function

$$F : \mathbb{R} \times U \to \{0, 1\}, \quad F(t, x) := \chi_{\{f(x)>t\}} = \chi_{E_t}(x).$$

This is measurable because f is measurable, and for each $x \in U$, the map $t \mapsto \chi_{E_t}(x)$ is measurable.
Fix $\varphi \in C_c^1(U; \mathbb{R}^n)$. Then the function

$$x \longmapsto \chi_{E_t}(x) \operatorname{div} \varphi(x)$$

belongs to $L^1(U)$ for each fixed t, since $\operatorname{div} \varphi$ is continuous and compactly supported. The linear functional

$$L^1(U) \to \mathbb{R}, \quad g \mapsto \int_U g(x)\, dx$$

is continuous, so the composition

$$t \longmapsto \int_U \chi_{E_t}(x) \operatorname{div} \varphi(x)\, dx$$

is measurable.
Finally, by the definition of the variation measure, there exists a sequence $\{\varphi_n\} \subset C_c^1(U; \mathbb{R}^n)$ with $\|\varphi_n\|_\infty \leq 1$ such that

$$\int_U \chi_{E_t}(x) \operatorname{div} \varphi_n(x)\, dx \longrightarrow \|\partial E_t\|(U) \qquad \text{as } n \to \infty.$$

Since the pointwise limit of measurable functions is measurable, it follows that

$$t \longmapsto \|\partial E_t\|(U)$$

is measurable. $\square$

Theorem 3.16 (Coarea formula for BV functions)**.** *Let $U \subset \mathbb{R}^n$ be open.*

(i) If $f \in BV(U)$, then for Lebesgue–a.e. $t \in \mathbb{R}$ the level set

$$E_t := \{x \in U : f(x) > t\}$$

has finite perimeter in U, and moreover

$$||Df||(U) \;=\; \int_{-\infty}^{\infty} \|\partial E_t\|(U)\, dt.$$

(ii) Conversely, if $f \in L^1(U)$ satisfies

$$\int_{-\infty}^{\infty} \|\partial E_t\|(U)\, dt < \infty,$$

then $f \in BV(U)$ and the identity above holds.

Remark 3.17. Before proving Theorem 3.16, it is helpful to recall the classical coarea formula for Lipschitz mappings and compare it with the BV version.
The coarea formula states that if $f : \mathbb{R}^n \to \mathbb{R}^m$ is locally Lipschitz, with $n \geq m$, then for every $\mathcal{L}^n$–measurable set $A \subset \mathbb{R}^n$,

$$\int_A Jf(x)\,dx = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}\Big(A \cap f^{-1}\{y\}\Big)\,dy,$$

In the special case $m = 1$, this becomes the scalar coarea identity: for any locally Lipschitz function $f : U \to \mathbb{R}$,

$$\int_U |Df(x)|\,dx = \int_{\mathbb{R}} \mathcal{H}^{n-1}\Big(U \cap f^{-1}(t)\Big)\,dt.$$

Note that by Rademacher's theorem any locally Lipschitz function $f : U \to \mathbb{R}$ is differentiable almost everywhere, with $\nabla f \in L^1_{\text{loc}}(U)$. Hence $f \in W^{1,1}_{\text{loc}}(U)$, and in particular $f \in BV_{\text{loc}}(U)$. Thus, the BV coarea formula in Theorem 3.16 may be seen as a natural extension of this identity to functions $f \in BV(U)$, where the role of the the $(n-1)$–dimensional measure of $f^{-1}(t) \cap U$ is replaced by the perimeter of E_t.

For brevity, We omit trying to compare $\mathcal{H}^{n-1}\Big(U \cap f^{-1}(t)\Big)$ and $\|\partial E_t\|(U)$ when f is locally Lipschitz as this requires establishing further theory such as the notion of the reduced boundary and its relation to the perimeter measure.
We now culminate this report with a proof of the co-area formula for BV functions.

Proof of the Co-Area Formula for BV Functions. **Step 1.** We first show that

$$|Df|(U) \;\leq\; \int_{-\infty}^{\infty} \|\partial E_t\|(U)\,dt.$$

Fix $\varphi \in C^1_c(U;\mathbb{R}^n)$ with $\|\varphi\|_\infty \leq 1$. We first assume $f \geq 0$. By the layer–cake representation,

$$f(x) = \int_0^\infty \chi_{E_t}(x)\,dt, \qquad \text{with } \; E_t = \{x \in U : \; f(x) > t\}.$$

Since $\operatorname{div}\varphi$ is integrable and compactly supported, Tonelli's theorem yields

$$\int_U f(x)\ \operatorname{div}\varphi(x)\,dx = \int_0^\infty \left(\int_U \chi_{E_t}(x)\ \operatorname{div}\varphi(x)\,dx\right) dt.$$

If $f \leq 0$, apply the same argument to $-f$, and combining both cases gives, for an arbitrary $f \in BV(U)$,

$$\int_U f\ \operatorname{div}\varphi = \int_{-\infty}^{\infty} \left(\int_U \chi_{E_t}\ \operatorname{div}\varphi\right) dt. \tag{$*$}$$

For each t, the inner integral satisfies

$$\left|\int_U \chi_{E_t}\ \operatorname{div}\varphi\right| \leq \|\partial E_t\|(U),$$

Thus, integrating in t and using $(*)$,

$$\left| \int_U f \ \operatorname{div} \varphi \right| \leq \int_{-\infty}^{\infty} \|\partial E_t\|(U)\, dt.$$

Finally, take the supremum over all $\varphi \in C_c^1(U; \mathbb{R}^n)$ with $\|\varphi\|_\infty \leq 1$; we get:

$$|Df|(U) = \sup_{\|\varphi\|_\infty \leq 1} \left| \int_U f \ \operatorname{div} \varphi \right| \leq \int_{-\infty}^{\infty} \|\partial E_t\|(U)\, dt.$$

This completes Step 1.

Step 2. Assume first that $f \in C^\infty(U) \cap BV(U)$ and define, for $t \in \mathbb{R}$,

$$m(t) := \int_{\{f \leq t\}} |Df|.$$

Since m is nondecreasing, $m'(t)$ exists for Lebesgue a.e. t. We shall prove that

$$m'(t) \geq \|\partial E_t\|(U) \qquad \text{for a.e. } t,$$

Fix $t \in \mathbb{R}$ and let $r > 0$. Define the Lipschitz cutoff

$$\eta_r(s) := \begin{cases} 1, & s \leq t, \\ 1 - \dfrac{s-t}{r}, & t < s < t+r, \\ 0, & s \geq t+r. \end{cases}$$

Note that $\eta_r \circ f \in C_c^\infty(U)$ when $\varphi \in C_c^1(U; \mathbb{R}^n)$ is fixed (or at least $\eta_r \circ f$ is Lipschitz and compactly supported on the support of φ), and that

$$\eta_r'(s) = -\frac{1}{r} \mathbf{1}_{(t,t+r)}(s), \qquad \text{hence} \qquad \nabla(\eta_r \circ f)(x) = \eta_r'(f(x))\, \nabla f(x) = -\frac{1}{r} \mathbf{1}_{\{t < f(x) < t+r\}}\, \nabla f(x).$$

Let $\varphi \in C_c^1(U; \mathbb{R}^n)$ with $\|\varphi\|_{L^\infty} \leq 1$. Using the chain rule and integration by parts for smooth functions we get

$$\begin{aligned} \int_{E_t} \operatorname{div} \varphi\, dx \ &= \ \int_U \chi_{E_t} \ \operatorname{div} \varphi\, dx \ \leq \ \int_U \eta_r(f) \ \operatorname{div} \varphi\, dx \\ &= \ -\int_U \varphi \cdot \nabla(\eta_r \circ f)\, dx \ = \ -\int_U \varphi \cdot \Big(\eta_r'(f) \nabla f \Big)\, dx \\ &= \ \frac{1}{r} \int_{\{t < f < t+r\}} \varphi \cdot \nabla f\, dx \ \leq \ \frac{1}{r} \int_{\{t < f < t+r\}} |\nabla f|\, dx. \end{aligned}$$

Taking the supremum over all $\varphi \in C_c^1(U; \mathbb{R}^n)$ with $\|\varphi\|_\infty \leq 1$ yields

$$\|\partial E_t\|(U) \leq \frac{1}{r} \int_{\{t < f < t+r\}} |\nabla f|\, dx.$$

But for smooth f the measure $|Df|$ coincides with $|\nabla f|\,dx$, hence

$$\int_{\{t<f<t+r\}} |\nabla f|\,dx = m(t+r) - m(t).$$

Therefore

$$\|\partial E_t\|(U) \leq \frac{m(t+r) - m(t)}{r}.$$

Letting $r \downarrow 0$ and using existence of $m'(t)$ for a.e. t gives

$$\|\partial E_t\|(U) \leq m'(t),$$

which is the claimed inequality.

Step 3. Assume $f \in C^\infty(U) \cap BV(U)$. We'll show, using step 1 and step 2, that the co=area formula indeed holds for this special class of f.

From Step 2 we have $m'(t) \geq \|\partial E_t\|(U)$ for a.e. t, and thus

$$\int_{-\infty}^{\infty} \|\partial E_t\|(U)\,dt \leq \int_{-\infty}^{\infty} m'(t)\,dt \leq m(\infty) - m(-\infty) = |Df|(U).$$

Combining this with the inequality from Step 1 produces the chain

$$|Df|(U) \leq \int_{-\infty}^{\infty} \|\partial E_t\|(U)\,dt \leq \int_{-\infty}^{\infty} m'(t)\,dt \leq |Df|(U).$$

Hence,

$$|Df|(U) = \int_{-\infty}^{\infty} \|\partial E_t\|(U)\,dt,$$

and moreover $m'(t) = \|\partial E_t\|(U)$ for almost every t. This proves the coarea formula for $f \in C^\infty(U) \cap BV(U)$.

Step 4. Let $f \in BV(U)$ and let $\{\varphi_k\} \subset C^\infty(U) \cap BV(U)$ be the sequence provided by Theorem 3.12, so that

$$\varphi_k \longrightarrow f \quad \text{in } L^1(U), \qquad \text{and} \qquad \|D\varphi_k\|(U) \longrightarrow \|Df\|(U).$$

For each k and $t \in \mathbb{R}$ set

$$E_t^k := \{x \in U : \varphi_k(x) > t\}, \qquad E_t := \{x \in U : f(x) > t\}.$$

We claim that there exists a subsequence $\{k_j\}$ such that for Lebesgue–a.e. $t \in \mathbb{R}$,

$$\chi_{E_t^{k_j}} \longrightarrow \chi_{E_t} \quad \text{in } L^1(U) \qquad (j \to \infty).$$

Proof of the claim. Note that for any $a, b \in \mathbb{R}$

$$|a - b| \;=\; \int_{-\infty}^{\infty} \left|\mathbf{1}_{\{a>t\}} - \mathbf{1}_{\{b>t\}}\right| dt.$$

Applying this identity with $a = \varphi_k(x)$ and $b = f(x)$ and then integrating in x gives, by Tonelli/Fubini,

$$\int_U |\varphi_k(x) - f(x)|\, dx \;=\; \int_U \int_{-\infty}^{\infty} \left|\chi_{E_t^k}(x) - \chi_{E_t}(x)\right| dt\, dx \;=\; \int_{-\infty}^{\infty} \Big(\int_U \left|\chi_{E_t^k} - \chi_{E_t}\right| dx \Big) dt.$$

Define the nonnegative functions

$$a_k(t) := \int_U \left|\chi_{E_t^k} - \chi_{E_t}\right| dx = \|\chi_{E_t^k} - \chi_{E_t}\|_{L^1(U)}.$$

As $\|\varphi_k - f\|_{L^1(U)} \longrightarrow 0$, it follows that a_k converges to 0 in $L^1(\mathbb{R})$; in particular $\{a_k\}$ has a subsequence $\{a_{k_j}\}$ for which $a_{k_j}(t) \to 0$ for almost every $t \in \mathbb{R}$.
Thus, after passing to this subsequence $\{k_j\}$, we have

$$\|\chi_{E_t^{k_j}} - \chi_{E_t}\|_{L^1(U)} = a_{k_j}(t) \longrightarrow 0 \qquad \text{for a.e. } t,$$

□

Step 5. To finish the proof we pass to the limit along the subsequence $\{\varphi_{k_j}\}$ and use lower semicontinuity of the variation measure.

Fix t for which $\chi_{E_t^{k_j}} \to \chi_{E_t}$ in $L^1(U)$ (this holds for a.e. t by Step 4). Since each $\chi_{E_t^{k_j}} \in BV(U)$, Theorem 3.11 applied to the sequence $\chi_{E_t^{k_j}}$ yields

$$\|\partial E_t\|(U) \le \liminf_{j\to\infty} \|\partial E_t^{k_j}\|(U).$$

This inequality holds for almost every $t \in \mathbb{R}$. Integrating in t and applying Fatou's lemma gives

$$\int_{-\infty}^{\infty} \|\partial E_t\|(U)\, dt \le \int_{-\infty}^{\infty} \liminf_{j\to\infty} \|\partial E_t^{k_j}\|(U)\, dt \le \liminf_{j\to\infty} \int_{-\infty}^{\infty} \|\partial E_t^{k_j}\|(U)\, dt.$$

By Step 3 (the coarea identity for the smooth approximants) we have, for each k,

$$\int_{-\infty}^{\infty} \|\partial E_t^k\|(U)\, dt = \|D\varphi_k\|(U).$$

Therefore

$$\int_{-\infty}^{\infty} \|\partial E_t\|(U)\, dt \le \liminf_{j\to\infty} \|D\varphi_{k_j}\|(U) = \|Df\|(U),$$

where the last equality uses the convergence $\|D\varphi_k\|(U) \to \|Df\|(U)$ from the density theorem. Combining this with the inequality from Step 1, we obtain equality.

$$\|Df\|(U) = \int_{-\infty}^{\infty} \|\partial E_t\|(U)\, dt,$$

which completes the proof of the coarea formula for $f \in BV(U)$. □

References

[1] Giovanni Alberti. A brief introduction to geometric measure theory. https://www.math.stonybrook.edu/~bishop/classes/math638.F20/Alberti_GMT_brief_intro.pdf, 2020. Lecture notes for Math 638, Stony Brook University.

[2] Lawrence C. Evans and Ronald F. Gariepy. *Measure Theory and Fine Properties of Functions*. CRC Press, revised edition edition, 1992.

[3] Herbert Federer. *Geometric Measure Theory*. Berlin, Heidelberg, New York, Springer, 1969.

[4] Leon Simon. *Introduction to Geometric Measure Theory*. Stanford University, 2014. Available online: https://web.stanford.edu/class/math285/ts-gmt.pdf.