Mathematics Department Stanford University
Math 205A 2013—Lecture Supplement #3

Differentiability Theory for Functions and Measures

As a preliminary to the discussion of differentiation of functions and measures, we need the following important covering lemma, which we state and prove in $\mathbb{R}^n$ but which clearly has a natural extension to metric spaces:

Lemma (5-times covering lemma). *Let $\mathcal{B}$ be any collection of closed balls in $\mathbb{R}^n$ with the property that $\cup_{B\in\mathcal{B}}B$ is contained in a bounded set. Then there is a p.w.d. collection $\{B_{\rho_j}(x_j)\}_{j=1,2,\dots} \subset \mathcal{B}$ such that $\cup_{B\in\mathcal{B}}B \subset \cup_{j=1}^\infty B_{5\rho_j}(x_j)$. The subcollection $\{B_{\rho_j}(x_j)\}_{j=1,2,\dots}$ can in fact be chosen so that:*

$$(*)\qquad B\in\mathcal{B} \Longrightarrow \exists\, j \text{ with } B\cap B_{\rho_j}(x_j)\neq\emptyset \text{ and } \rho_j \geq \frac{1}{2}\operatorname{radius} B.$$

Terminology: As in lecture, "p.w.d." is an abbreviation for "pairwise disjoint" and here $B_\rho(y)$ denotes the closed ball with center y and radius $\rho>0$ while $\breve{B}_\rho(y)$ denotes the corresponding open ball.

Proof of the 5-times Lemma: Let $R_0 = \sup\{\operatorname{radius} B : B\in\mathcal{B}\}(<\infty)$, and write $\mathcal{B} = \cup_{k=1}^\infty \mathcal{B}_k$, where $\mathcal{B}_k = \{B\in\mathcal{B} : 2^{-k}R_0 < \operatorname{radius} B \leq 2^{-k+1}R_0\}$. We proceed to inductively select pairwise disjoint subcollections $\mathcal{M}_k\subset\mathcal{B}_k$ as follows:

$\mathcal{M}_1$ is any maximal p.w.d. subcollection of $\mathcal{B}_1$ (meaning contains a maximum number of balls subject to the stated condition of being a p.w.d. collection). Assume now that $k\geq 2$ and that we have already selected $\mathcal{M}_j$ for $j=1,\dots,k-1$. Then select $\mathcal{M}_k$ to be a maximal p.w.d. subcollection of $\{B : B\in\mathcal{B}_k \text{ and } B\cap E=\emptyset\ \forall\, E\in\cup_{j=1}^{k-1}\mathcal{M}_j\}$. Of course we take $\mathcal{M}_k=\emptyset$ in case $\{B\in\mathcal{B}_k : B\cap E=\emptyset\ \forall\, E\in\cup_{j=1}^{k-1}\mathcal{M}_j\}$ is empty. Now we define

$$\mathcal{M} = \cup_{k=1}^\infty \mathcal{M}_k$$

and observe that $\mathcal{M}$ is a countable p.w.d. collection by construction, so the balls in the collection $\mathcal{M}$ can be written as a sequence $\{B_{\rho_j}(x_j)\}_{j=1,2,\dots}$ of p.w.d. balls. We claim that in fact the additional property $(*)$ holds. Indeed if $B\in\mathcal{B}$ then $B\in\mathcal{B}_{k_0}$ for some unique $k_0\geq 1$, and we claim that in fact then $B\cap E\neq\emptyset$ for some $E\in\cup_{j=1}^{k_0}\mathcal{M}_j$. Otherwise for $k_0\geq 2$ we would have both that $B\cap E=\emptyset$ for each ball $E\in\mathcal{M}_{k_0}$ and also $B\cap E=\emptyset$ for each ball $E\in\cup_{j=1}^{k_0-1}\mathcal{M}_j$ which means that $\mathcal{M}_{k_0}\cup\{B\}$ is a p.w.d. collection of balls in $\mathcal{B}_{k_0}$ which do not intersect any ball in the collection $\cup_{j=1}^{k_0-1}\mathcal{M}_j$, thus contradicting the maximality of $\mathcal{M}_{k_0}$. For $k_0=1$ the argument is even simpler: $B\cap E=\emptyset$ for every $E\in\mathcal{M}_1$ implies that $\mathcal{M}_1\cup\{B\}$ is a p.w.d. subcollection of $\mathcal{B}_1$, thus contradicting the maximality of $\mathcal{M}_1$. Thus we have shown that $B\cap B_\rho(x)\neq\emptyset$ for some ball $B_\rho(x)\in\cup_{j=1}^{k_0}\mathcal{M}_j$. But then $\operatorname{radius} B_\rho(x)\geq 2^{-k_0}R_0 = \frac{1}{2}2^{1-k_0}R_0\geq\frac{1}{2}\operatorname{radius} B$. Thus $B\cap B_\rho(x)\neq\emptyset$ and $\rho\geq\frac{1}{2}\operatorname{radius} B$ which is $(*)$. Now, since $(*)$ evidently implies that $B\subset B_{5\rho}(x)$, the proof is complete.

We have now the following important corollary of the 5-times covering lemma:

Corollary 1. *Let $\mathcal{B}$ be any collection of closed balls in $\mathbb{R}^n$ with the property that $\cup_{B\in\mathcal{B}}B$ is contained in a bounded set, and suppose $A\subset\mathbb{R}^n$. If $\mathcal{B}$ covers A finely in the sense that for each $x\in A$ and each $\rho>0$ there is a ball $B\in\mathcal{B}$ such that $x\in B$ and $\operatorname{radius} B<\rho$, then there is a p.w.d. subcollection $\{B_{\rho_j}(x_j)\}_{j=1,2,\dots}\subset\mathcal{B}$ with the properties that $\cup_{B\in\mathcal{B}}B\subset\cup_j B_{5\rho_j}(x_j)$ and*

$$(\ddagger)\qquad A\setminus\left(\cup_{j=1}^N B_{\rho_j}(x_j)\right)\subset\cup_{j=N+1}^\infty B_{5\rho_j}(x_j) \text{ for each } N\geq 1.$$

Proof: The 5-times covering lemma can be applied to $\mathcal{B}$, giving a p.w.d. subcollection of closed balls $\{B_{\rho_j}(x_j)\}_{j=1,2,\ldots} \subset \mathcal{B}_1$ such that

$$B \in \mathcal{B} \Longrightarrow \exists\, j \text{ with } B \cap B_{\rho_j}(x_j) \neq \emptyset \text{ (and hence } B \subset B_{5\rho_j}(x_j)). \tag{1}$$

We claim that this sequence $\{B_{\rho_j}(x_j)\}_{j=1,2,\ldots}$ automatically has the additional property (‡). To see this, take any $N \geq 1$ and let $x \in A \setminus (\cup_{j=1}^N B_{\rho_j}(x_j))$. Since $\mathbb{R}^n \setminus (\cup_{j=1}^N B_{\rho_j}(x_j))$ is an open set and since $\mathcal{B}$ covers A finely, we can certainly find a ball $B \in \mathcal{B}$ with $x \in B \subset \mathbb{R}^n \setminus (\cup_{j=1}^N B_{\rho_j}(x_j))$ and hence for this B the j in (1) must be $\geq N+1$. That is, $x \in B \subset \cup_{j=N+1}^\infty B_{5\rho_j}(x_j)$, which completes the proof.

An important corollary of this is the following Vitali covering lemma.

Lemma (Vitali Covering Lemma). *Let μ be any outer measure on $\mathbb{R}^n$ such that all Borel sets are μ-measurable and such that there is a fixed constant $\beta \in (0,\infty)$ with $\mu(B_{2\rho}(x)) \leq \beta\mu(B_\rho(x)) < \infty$ for each closed ball $B_\rho(x)$ (note that these hypotheses hold with $\mu =$ Lebesgue outer measure λ in case $\beta = 2^n$), let $A \subset \mathbb{R}^n$ be bounded and let $\mathcal{B}$ be any collection of closed balls which cover A finely. Then there is a p.w.d. subcollection $\{B_{\rho_j}(x_j)\}_{j=1,2,\ldots} \subset \mathcal{B}$ such that $\mu(A \setminus (\cup_{j=1}^N B_{\rho_j}(x_j))) \to 0$ as $N \to \infty$.*

Remark 1: Actually the conclusion holds without the hypothesis that $\mu(B_{2\rho}(x)) \leq \beta\mu(B_\rho(x))$, provided that the collection $\mathcal{B}$ not only covers A finely, but actually that for each point $x \in A$ we have balls $B_{\rho_j}(x) \in \mathcal{B}$ (i.e. balls in $\mathcal{B}$ with center at x) with $\rho_j \downarrow 0$. This result (which is important in geometric analysis) requires a more powerful covering lemma (the Besicovich covering lemma) in place of the 5-times covering lemma, and we will not discuss it here.

Proof of the Vitali Lemma: Let U be an open ball containing A and let $\mathcal{B}_1 = \{B \in \mathcal{B} : B \subset U\}$. Evidently $\mathcal{B}_1$ still covers A finely, hence by the corollary above we can choose p.w.d. balls $B_{\rho_1}(x_1), B_{\rho_2}(x_2), \ldots \in \mathcal{B}_1$ such that

$$A \setminus (\cup_{j=1}^N B_{\rho_j}(x_j)) \subset \cup_{j=N+1}^\infty B_{5\rho_j}(x_j) \text{ for each } N \geq 1.$$

Observe that for each j we have $\mu(B_{5\rho_j}(x_j)) \leq \mu(B_{8\rho_j}(x_j)) \leq \beta^3\mu(B_{\rho_j}(x_j))$ by definition of β. So $\mu(\cup_{j=N+1}^\infty B_{5\rho_j}(x_j)) \leq \sum_{j=N+1}^\infty \mu(B_{5\rho_j}(x_j)) \leq \beta^3 \sum_{j=N+1}^\infty \mu(B_{\rho_j}(x_j)) = \beta^3\mu(\cup_{j=N+1}^\infty B_{\rho_j}(x_j)) \leq \beta^3\mu(U) < \infty$, where we used the pairwise disjointness and μ-measurability of the $B_{\rho_j}(x_j)$. Thus $\mu(\cup_{j=N+1}^\infty B_{5\rho_j}(x_j)) \to 0$ as $N \to \infty$, and the proof is complete.

In the following lemmas f is an arbitrary function $: [a,b] \to \mathbb{R}$, and for $x \in (a,b)$ we let

$$\overline{D}f(x) = \limsup_{y\to x} \frac{f(x)-f(y)}{x-y}, \quad \underline{D}f(x) = \liminf_{y\to x} \frac{f(x)-f(y)}{x-y}.$$

Notice that $-\infty \leq \underline{D}f(x) \leq \overline{D}f(x) \leq \infty$, and f is classically differentiable at x if and only if $-\infty < \overline{D}f(x) = \underline{D}f(x) < \infty$. Also, $\underline{D}f(x) \geq 0$ if f is increasing.

Lemma 1. *If $\varepsilon > 0$, $\beta \in \mathbb{R}$, $U \subset (a,b)$ is open, and if $S \subset U$ is an arbitrary set such that $\overline{D}f(x) > \beta$ at each point of S, then there are pairwise disjoint closed intervals $\{[x_j, y_j]\}_{j=1,\ldots,N}$ such that*

$$\cup_j\, [x_j,\, y_j] \subset U, \quad \lambda(S \setminus \cup_{j=1}^N [x_j,\, y_j]) < \varepsilon$$
$$\beta(y_j - x_j) \leq f(y_j) - f(x_j), \quad j = 1, \ldots, N.$$

Proof: We observe that by definition of $\overline{D}f$, for every $x \in S$ we must have $(z_j - x)^{-1}(f(z_j) - f(x)) > \beta$ for some sequence $z_j \to x$ such that $I_{x,j} \subset U$ for each j, where we let $I_{x,j} = [x, z_j]$ if $z_j > x$ and $I_{x,j} = [z_j, x]$ if $z_j < x$. Notice that then the collection $\mathcal{I} = \{I_{x,j} : x \in S,\ j = 1, 2, \ldots\}$ covers S finely and each $I_{x,j} \subset U$. Then by the Vitali covering lemma there are pairwise disjoint intervals $\{[x_j,\, y_j]\}_{j=1,\ldots,N} \subset \mathcal{I}$ such that $\lambda(S \setminus \cup_{j=1}^N [x_j,\, y_j]) < \varepsilon$. Since by definition we have $f(y_j) - f(x_j) > \beta(y_j - x_j)$ for each j, this completes the proof.

Remark 2: Notice that if $\beta > 0$, $a < b$, and if f is increasing (i.e. $a \le x \le y \le b \Rightarrow f(x) \le f(y)$), then we can apply the above lemma with $U = (a,\, b)$ to yield p.w.d. intervals $[x_i, y_i]$ such that $[x_i, y_i] \subset (a,\, b)$, $\beta(y_i - x_i) \le f(y_i) - f(x_i)$ and $\lambda(S \cap (a,\, b) \setminus (\cup_i [x_i, y_i])) < \varepsilon$. Assuming that we order these p.w.d. intervals $[x_i, y_i]$ so that $y_{i-1} < x_i$ for $i \in \{2, \ldots, N\}$, we then have

$$\begin{aligned}\beta\,\lambda(S) &\le \beta\,\lambda(S \setminus \cup_{j=1}^N (x_j,\, y_j)) + \beta \textstyle\sum_{j=1}^N (y_j - x_j) \\ &\le \beta\,\varepsilon + \textstyle\sum_{j=1}^N (f(y_j) - f(x_j)) \\ &\le \beta\,\varepsilon + \textstyle\sum_{j=1}^N (f(y_j) - f(y_{j-1})) \text{ (using notation } y_0 = x_1) \\ &= \beta\,\varepsilon + f(y_N) - f(x_1) \le \beta\,\varepsilon + f(b) - f(a),\end{aligned}$$

which, since $\varepsilon > 0$ is arbitrary, gives

$$\beta\,\lambda(S) \le f(b) - f(a).$$

Notice particularly that if we take $S = \{x \in (a, b) : \overline{D}f(x) = +\infty\}$ then we can apply this for each $\beta > 0$ and hence conclude that $\lambda(S) = 0$, i.e.

$$f : [a, b] \to \mathbb{R} \text{ increasing } \Rightarrow \overline{D}f(x) < \infty, \quad \lambda\text{-a.e. } x \in (a, b).$$

Observe that Lemma 1, with $-f$ in place of f and $\beta = -\alpha$, implies:

Lemma 2. *If $\varepsilon > 0$, $\alpha \in \mathbb{R}$, $U \subset (a, b)$ is open, and if $S \subset U$ is an arbitrary set such that $\underline{D}f(x) < \alpha$ at each point of S, then there are pairwise disjoint closed intervals $\{[x_j, y_j]\}_{j=1,\ldots,N}$ such that*

$$\cup_j\, [x_j,\, y_j] \subset U, \quad \lambda(S \setminus \cup_j [x_j,\, y_j]) < \varepsilon$$
$$f(y_j) - f(x_j) \le \alpha(y_j - x_j), \quad j = 1, \ldots, N.$$

We can now easily prove the following important differentiability theorem for increasing functions:

Theorem 1. *Let $f : [a, b] \to \mathbb{R}$ be an increasing function. Then f is differentiable λ-a.e. in (a, b) (i.e. $\lim_{y\to x} \frac{f(y)-f(x)}{y-x}$ exists and is real for λ-a.e. $x \in (a, b)$). Furthermore the derivative f' (defined to be e.g. zero on the set of measure zero where f is not differentiable) is a non-negative integrable function and*

$$\int_a^b f'(t)\, dt \le f(b) - f(a).$$

Proof: Let $T = \{x \in (a, b) : \overline{D}f(x) > \underline{D}f(x)\}$. Observe that (since $\underline{D}f(x) \ge 0$)

$$T = \cup_{0<\alpha<\beta,\, \alpha,\beta \text{ rational}} S_{\alpha\beta}, \tag{1}$$

where $S_{\alpha\beta} = \{x \in [a, b] : \overline{D}f(x) > \beta > \alpha > \underline{D}f(x)\}$.

Now let $\varepsilon > 0$, $0 < \alpha < \beta$, and choose an open set $U \subset (a, b)$ with $S_{\alpha\beta} \subset U$ and $\lambda(U) < \lambda(S_{\alpha\beta}) + \varepsilon$. Then we can apply Lemma 2 with $S = S_{\alpha\beta}$; this gives p.w.d. intervals $[x_i, y_i]$ with $f(y_i) - f(x_i) \le \alpha(y_i - x_i)$ and $\cup_i [x_i, y_i] \subset U$, so that $\sum_i (y_i - x_i) \le \lambda(U) \le \lambda(S_{\alpha\beta}) + \varepsilon$ and $\lambda(S_{\alpha\beta} \setminus (\cup_j [x_j, y_j])) < \varepsilon$.

Then we apply Remark 2 (following Lemma 1) with $S_{\alpha\beta} \cap (x_i, y_i)$ in place of S and with $(x_j,\, y_j)$ in place of $(a,\, b)$, whence $\beta\lambda(S_{\alpha\beta} \cap (x_j,\, y_j)) \le f(y_j) - f(x_j) \le \alpha(y_j - x_j)$ for each j. Then

$$\beta\lambda(S_{\alpha\beta} \cap [x_j, y_j]) \le f(y_j) - f(x_j) \le \alpha(y_j - x_j), \quad j = 1, \ldots, N,$$

and hence summing on j we have

$$\beta\lambda(S_{\alpha\beta} \cap (\cup_{j=1}^N [x_j, y_j])) \le \alpha \textstyle\sum_{j=1}^N (y_j - x_j) \le \alpha\lambda(U) \le \alpha\lambda(S_{\alpha\beta}) + \alpha\varepsilon,$$

and since $\lambda(S_{\alpha\beta} \setminus (\cup_j [x_j, y_j])) < \varepsilon$ we thus obtain

$$\beta\lambda(S_{\alpha\beta}) \le \alpha\lambda(S_{\alpha\beta}) + (\alpha + \beta)\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary we thus conclude $\beta\lambda(S_{\alpha\beta}) \le \alpha\lambda(S_{\alpha\beta})$, so that $\lambda(S_{\alpha\beta}) = 0$ for each $\alpha < \beta$, whence by (1) we have $\lambda(T) = 0$.

Keeping in mind that $\overline{D}f(x) < \infty$ a.e. $x \in (a, b)$ by Remark 2, we have thus proved that $\overline{D}f(x) = \underline{D}f(x) < \infty$ for a.e. $x \in (a, b)$, which is the same as saying $f\,'$ (the classical derivative of f) exists for a.e. $x \in (a, b)$, as required.

To prove the last part of the theorem, we first extend f to all of $\mathbb{R}$ by defining $g(x) = f(x)$ for $x \in [a, b]$, $g(x) = f(a)$ for $x < a$, and $g(x) = f(b)$ for $x > b$. Then note that $g'(x) = \lim_{n\to\infty} n(f(x + 1/n) - f(x))$ for a.e. $x \in \mathbb{R}$, and hence g' is a non-negative Lebesgue measurable function on $\mathbb{R}$, assuming we define it to e.g. be zero on the set of measure zero where g is not differentiable, and of course $g' = f'$ a.e. on (a, b). Also by Fatou's lemma we have

$$\int_a^b f'(t)\, dt \le \liminf_{n\to\infty} \int_a^b n\big(g(t + 1/n) - g(t)\big)\, dt.$$

But evidently $\int_a^b g(t + 1/n)\, dt = \int_{a+1/n}^{b+1/n} g(t)\, dt$, so $\int_a^b n\big(g(t + 1/n) - g(t)\big)\, dt = n\int_b^{b+1/n} g(t)\, dt - n\int_a^{a+1/n} g(t)\, dt \le f(b) - f(a)$, and hence

$$\int_a^b f'(t)\, dt \le f(b) - f(a)$$

as claimed.

Next we want to discuss Lebesgue's theorem on differentiation of the integral in $\mathbb{R}^n$. As a key preliminary, we need the following lemma.

Lemma 3. *Suppose $f : \mathbb{R}^n \to [0, \infty)$ is locally Lebesgue integrable on $\mathbb{R}^n$ (i.e. λ-measurable and integral over each ball is finite), and suppose $E \subset \mathbb{R}^n$ is λ-measurable. Then*

$$\lim_{\rho\downarrow 0} \rho^{-n} \int_{B_\rho(\xi)\cap E} f(x)\, dx = 0 \text{ for } \lambda\text{-a.e. } \xi \in \mathbb{R}^n \setminus E.$$

Proof: The proof as a simple application of the Vitali covering lemma.

Let $k \in \{1, 2, \ldots\}$, $\alpha > 0$, let K be any compact subset of $E \cap \breve{B}_k(0)$ ($\breve{B}_k(0)$ the open ball of radius k and center 0),

$$S_\alpha = \{\xi \in \breve{B}_k(0) \setminus E : \limsup_{\rho\downarrow 0} \rho^{-n} \int_{B_\rho(\xi)\cap E} f(x)\, dx > \alpha\}.$$

Then for each $\xi \in S_\alpha$ there is a sequence $\rho_j \downarrow 0$ with $\rho_j^{-n} \int_{B_{\rho_j}(\xi)\cap E} f(x)\,dx > \alpha$ for each j, and hence $\mathcal{B} = \{B_\rho(\xi) \subset \breve{B}_k(0) \setminus K : \xi \in S_\alpha \text{ and } \omega_n^{-1}\rho^{-n}\int_{B_\rho(\xi)\cap E} f(x)\,dx > \alpha\}$ covers S_α finely, so by the Vitali covering lemma there are p.w.d. balls $B_{\rho_j}(\xi_j) \in \mathcal{B}$ with

$$\lambda(S_\alpha \setminus (\cup_{j=1}^\infty B_{\rho_j}(\xi_j))) = 0 \text{ and } \int_{B_{\rho_i}(\xi_i)\cap E} f(x)\,dx > \alpha\omega_n\rho_i^n, \quad i = 1, 2, \ldots.$$

Then by subadditivity of λ

$$\begin{aligned}\alpha\lambda(S_\alpha) &\le \alpha\lambda(S_\alpha \setminus (\cup_{j=1}^\infty B_{\rho_j}(x_j))) + \alpha\textstyle\sum_{j=1}^\infty \lambda(B_{\rho_j}(\xi_j)) \\ &\le \textstyle\sum_{j=1}^\infty \displaystyle\int_{B_{\rho_j}(\xi_j)\cap E} f(x)\,dx = \int_{\cup_{j=1}^\infty B_{\rho_j}(\xi_j)\cap E} f(x)\,dx \le \int_{\breve{B}_k(0)\cap E\setminus K} f(x)\,dx,\end{aligned}$$

Now, as proved earlier, we can find an increasing sequence $K_j \subset \breve{B}_k(0) \cap E$ of compact sets with $\lambda(\breve{B}_k(0) \cap E \setminus K_j) \to 0$, so we have actually proved

$$\alpha\lambda(S_\alpha) \le \int_{\mathbb{R}^n} \chi_{\breve{B}_k(0)\cap E\setminus K_j} f(x)\,dx$$

and the right side $\to 0$ as $j \to \infty$ by the dominated convergence theorem, hence $\lambda(S_\alpha) = 0$. Thus $\{\xi \in \breve{B}_k(0) \setminus E : \limsup_{\rho\downarrow 0} \rho^{-n}\int_{B_\rho(\xi)\cap E} f(x)\,dx > 0\} = \cup_{j=1}^\infty S_{1/j}$ is a countable union of sets of measure zero, hence has measure zero, so we have proved

$$\lim_{\rho\downarrow 0} \rho^{-n} \int_{B_\rho(\xi)\cap E} f(x)\,dx = 0 \text{ for } \lambda\text{-a.e. } \xi \in \breve{B}_k(0) \setminus E.$$

Since k is arbitrary this proves the lemma.

The following corollary is important:

Corollary 2. *Let $E \subset \mathbb{R}^n$ be λ-measurable. Then*

$$\lim_{\rho\downarrow 0} \omega_n^{-1}\rho^{-n}\lambda(E \cap B_\rho(\xi)) = \begin{cases} 0 \text{ for } \lambda\text{-a.e. } \xi \in \mathbb{R}^n \setminus E \\ 1 \text{ for } \lambda\text{-a.e. } \xi \in E. \end{cases}$$

Proof: To get the first conclusion simply apply Lemma 3 with $f \equiv 1$. For the second conclusion observe that $1 - \omega_n^{-1}\rho^{-n}\lambda(E \cap B_\rho(\xi)) = \omega_n^{-1}\rho^{-n}\lambda(B_\rho(\xi) \setminus E)$ and so Lemma 3 with $f \equiv 1$ and with $\mathbb{R}^n \setminus E$ in place of E gives the required result.

The Lebesgue differentiation theorem is then as follows:

Theorem 2. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lebesgue integrable (i.e. λ-measurable and integral of $|f|$ over each ball is finite). Then*

(i) $$\lim_{\rho\downarrow 0} \omega_n^{-1}\rho^{-n} \int_{B_\rho(\xi)} f(x)\,dx = f(\xi) \text{ for } \lambda\text{-a.e. } \xi \in \mathbb{R}^n$$

(ii) $$\lim_{\rho\downarrow 0} \rho^{-n} \int_{B_\rho(\xi)} \big|f(x) - f(\xi)\big|\,dx = 0 \text{ for } \lambda\text{-a.e. } \xi \in \mathbb{R}^n.$$

Remarks (a) Notice that of course (ii) $\Rightarrow$ (i) because

$$|\omega_n^{-1}\rho^{-n}\textstyle\int_{B_\rho(\xi)} f(x)\,dx - f(\xi)| = |\omega_n^{-1}\rho^{-n}\int_{B_\rho(\xi)}(f(x) - f(\xi))\,dx| \le \omega_n^{-1}\rho^{-n}\int_{B_\rho(\xi)} |f(x) - f(\xi)|\,dx,$$

but in the proof we first establish (i) and show that (ii) follows directly from it.

(b) The points ξ where the limit in (ii) is valid are called *the Lebesgue points of the function* f.

Proof of Theorem 2: For each $i = 1, 2, \dots$ we have

$$\mathbb{R}^n = \cup_{j=-\infty}^{\infty} A_{ij}, \text{ where } A_{ij} = \{x \in \mathbb{R}^n : (j-1)/i < f(x) \le j/i\}.$$

Notice that then for each $i = 1, 2, \dots$ the sets $A_{ij}, j = 1, 2, \dots$, are p.w.d. λ-measurable, and

$$(1) \qquad \int_{B_\rho(\xi)} f(x)\,dx = \int_{B_\rho(\xi)\cap A_{ij}} f(x)\,dx + \int_{B_\rho(\xi)\setminus A_{ij}} f(x)\,dx,$$

and of course

$$\omega_n^{-1}\rho^{-n}\lambda(B_\rho(\xi)\cap A_{ij})(j-1)/i \le \omega_n^{-1}\rho^{-n}\int_{B_\rho(\xi)\cap A_{ij}} f(x)\,dx \le j/i,$$

hence (1) implies

$$(2) \quad \omega_n^{-1}\rho^{-n}\lambda(B_\rho(\xi)\cap A_{ij})(j-1)/i \le \omega_n^{-1}\rho^{-n}\int_{B_\rho(\xi)} f(x)\,dx - \omega_n^{-1}\rho^{-n}\int_{B_\rho(\xi)\setminus A_{ij}} f(x)\,dx \le j/i.$$

By Lemma 3 (with $E = \mathbb{R}^n \setminus A_{ij}$) and Corollary 2 (with $E = A_{ij}$) we then have

$$(3) \qquad (j-1)/i \le \liminf_{\rho\downarrow 0} \omega_n^{-1}\rho^{-n}\int_{B_\rho(\xi)} f(x)\,dx \le \limsup_{\rho\downarrow 0} \omega_n^{-1}\rho^{-n}\int_{B_\rho(\xi)} f(x)\,dx \le j/i$$

for λ-a.e. $\xi \in A_{ij}$, which means (3) holds for each $\xi \in A_{ij} \setminus E_{ij}$, where $\lambda(E_{ij}) = 0$. Since $(j-1)/i < f(\xi) \le j/i$ for all $\xi \in A_{ij}$, (3) implies

$$(4) \qquad f(\xi) - 1/i \le \liminf_{\rho\downarrow 0} \omega_n^{-1}\rho^{-n}\int_{B_\rho(\xi)} f(x)\,dx \le \limsup_{\rho\downarrow 0} \omega_n^{-1}\rho^{-n}\int_{B_\rho(\xi)} f(x)\,dx \le f(\xi) + 1/i$$

for each $\xi \in A_{ij} \setminus E$ where $E = \cup_{k=1}^{\infty} \cup_{\ell=-\infty}^{\infty} E_{k\ell}$ has measure zero and does not depend on the indices i, j. Since $\cup_{j=-\infty}^{\infty} A_{ij} = \mathbb{R}^n$ we thus have

$$f(\xi) - 1/i \le \liminf_{\rho\downarrow 0} \omega_n^{-1}\rho^{-n}\int_{B_\rho(\xi)} f(x)\,dx \le \limsup_{\rho\downarrow 0} \omega_n^{-1}\rho^{-n}\int_{B_\rho(\xi)} f(x)\,dx \le f(\xi) + 1/i$$

for every $i = 1, 2, \dots$ and every $\xi \in \mathbb{R}^n \setminus E$, and hence

$$\liminf_{\rho\downarrow 0} \omega_n^{-1}\rho^{-n}\int_{B_\rho(\xi)} f(x)\,dx = \limsup_{\rho\downarrow 0} \omega_n^{-1}\rho^{-n}\int_{B_\rho(\xi)} f(x)\,dx = f(\xi), \quad \forall\, \xi \in \mathbb{R}^n \setminus E,$$

so (i) is proved.

To prove (ii), let $q_1, q_2, \dots$ be any countable dense subset of $\mathbb{R}$. Applying (i) to $|f(x) - q_j|$ we have

$$\lim_{\rho\downarrow 0} \omega_n^{-1}\rho^{-n}\int_{B_\rho(\xi)} |f(x) - q_j| = |f(\xi) - q_j|, \ \ \forall \xi \in \mathbb{R}^n \setminus E_j,$$

where $\lambda(E_j) = 0$, hence

$$(5) \qquad \lim_{\rho\downarrow 0} \omega_n^{-1}\rho^{-n}\int_{B_\rho(\xi)} |f(x) - q_j| = |f(\xi) - q_j|, \ \ \forall\, j = 1, 2, \dots \text{ and } \forall\, \xi \in \mathbb{R}^n \setminus E,$$

where $E = \cup_{\ell=1}^{\infty} E_\ell$, so that $\lambda(E) = 0$. If $\varepsilon > 0$ and $\xi \in \mathbb{R}^n \setminus E$, we can select j such that $|f(\xi) - q_j| < \varepsilon$, and hence (5) gives

$$\limsup_{\rho \downarrow 0} \omega_n^{-1} \rho^{-n} \int_{B_\rho(\xi)} |f(x) - f(\xi)| < 2\varepsilon \ \forall \varepsilon > 0,$$

so $\lim_{\rho \downarrow 0} \omega_n^{-1} \rho^{-n} \int_{B_\rho(\xi)} |f(x) - f(\xi)| = 0$ for each $\xi \in \mathbb{R}^n \setminus E$, which is (ii).

The Lebesgue theorem (Theorem 2) has an important corollary in the case $n = 1$:

Corollary 3. *If $a, b \in \mathbb{R}$ with $a < b$ and if $f : [a,b] \to \mathbb{R}$ is Lebesgue integrable, then the function $F(x) = \int_a^x f(t)\,dt$ is differentiable a.e. on (a,b) and $F'(x) = f(x)$ for a.e. $x \in (a,b)$.*

Proof: If $x \in (a,b)$ and $0 < |h| < \min\{b - x, x - a\}$ then

$$\begin{aligned} |h^{-1}(F(x+h) - F(x)) - f(x)| &= \big|h^{-1} \int_x^{x+h} f(t)\,dt - f(x)\big| = \big|h^{-1} \int_x^{x+h} \big(f(t) - f(x)\big)\,dt\big| \\ &\le |h|^{-1} \int_{x-|h|}^{x+|h|} \big|f(t) - f(x)\big|\,dt \end{aligned}$$

which $\to 0$ as $h \to 0$ for a.e. $x \in (a,b)$ by part (ii) of Theorem 2.

The above corollary will play an important role in the theory of absolutely continuous functions on $[a,b]$ which we want to develop below, but first we need to introduce the notion of bounded variation (BV):

Let $\mathcal{P} : x_0 = a < x_1 < x_2 < \cdots < x_N = b$ be any partition of $[a,b]$, $f : [a,b] \to \mathbb{R}$, and define

$$T_{f,\mathcal{P}} = \sum_{j=1}^{N} |f(x_j) - f(x_{j-1})|$$
$$T_f = \sup T_{f,\mathcal{P}},$$

where the sup is over all partitions $\mathcal{P}$ of $[a,b]$. T_f is called the *total variation* of f over the interval $[a,b]$.

Observe that $T_f = T_{f,\mathcal{P}} = f(b) - f(a)$ for each partition $\mathcal{P}$ if f is increasing on $[a,b]$.

Definition: $f : [a,b] \to \mathbb{R}$ has bounded variation (BV) on $[a,b]$ if $T_f < \infty$.

Lemma 4. *$f : [a,b] \to \mathbb{R}$ is BV on $[a,b]$ $\iff$ f can be written as the difference of two increasing functions; i.e. there are increasing $f_1, f_2 : [a,b] \to \mathbb{R}$ such that $f(x) = f_1(x) - f_2(x)$ for all $x \in [a,b]$.*

Proof "⇒": For any partition $\mathcal{P} : a = x_0 < x_1 < x_2 < \cdots < x_N = b$ we define

$$P_{f,\mathcal{P}} = \sum_{j=1}^{N} (f(x_j) - f(x_{j-1}))_+, \quad N_{f,\mathcal{P}} = \sum_{j=1}^{N} (f(x_j) - f(x_{j-1}))_-,$$

where we use the notation $a_+ = \max\{a, 0\}$, $a_- = \max\{-a, 0\}$, so that

$$P_{f,\mathcal{P}} - N_{f,\mathcal{P}} = \sum_{j=1}^{N} (f(x_j) - f(x_{j-1})) = f(b) - f(a)$$
$$P_{f,\mathcal{P}} + N_{f,\mathcal{P}} = \sum_{j=1}^{N} |f(x_j) - f(x_{j-1})| = T_{f,\mathcal{P}}.$$

Observe that then $\sup_{\mathcal{P}} T_{f,\mathcal{P}} < \infty \iff \sup_{\mathcal{P}} P_{f,\mathcal{P}} < \infty \iff \sup_{\mathcal{P}} N_{f,\mathcal{P}} < \infty$ and

$$\sup_{\mathcal{P}} T_{f,\mathcal{P}} < \infty \Rightarrow f(b) - f(a) = \sup_{\mathcal{P}} P_{f,\mathcal{P}} - \sup_{\mathcal{P}} N_{f,\mathcal{P}}.$$

By applying the same argument on the interval $[a, x]$ (where $x \in (a, b]$) we have

$$f(x) = f(a) + f_1(x) - f_2(x), \quad x \in [a, b],$$

where $f_1(x) = \sup_{\text{paritions } \mathcal{P} \text{ of } [a,x]} P_{f|[a,x],\mathcal{P}}$ and $f_2(x) = \sup_{\text{paritions } \mathcal{P} \text{ of } [a,x]} N_{f|[a,x],\mathcal{P}}$ for $x \in (a, b]$ and $f_1(a) = f_2(a) = 0$ are non-negative increasing functions on $[a, b]$, provided $\sup_{\mathcal{P}} T_{f,\mathcal{P}} < \infty$ (i.e. provided f is BV on $[a, b]$).

Proof "$\Leftarrow$": $f = f_1 - f_2$ with $f_1, f_2 : [a, b] \to \mathbb{R}$ increasing $\Rightarrow T_{f,\mathcal{P}} \le T_{f_1,\mathcal{P}} + T_{f_2,\mathcal{P}} = f_1(b) - f_1(a) + f_2(b) - f_2(a)$ for each partition $\mathcal{P}$ of $[a, b]$, so

$$T_f \le f_1(b) - f_1(a) + f_2(b) - f_2(a) < \infty.$$

Next we want to introduce the concept of an absolutely continuous (AC) function:

Definition: $f : [a, b] \to \mathbb{R}$ is AC if for each $\varepsilon > 0$ there is $\delta > 0$ such that $\sum_{i=1}^{N} |f(y_i) - f(x_i)| < \varepsilon$ whenever $[x_1, y_1], \ldots, [x_N, y_N]$ are p.w.d. closed intervals in $[a, b]$ with $\sum_{i=1}^{N}(y_i - x_i) < \delta$.

Remarks: (1) $f : [a, b] \to \mathbb{R}$ is AC $\Rightarrow f$ is uniformly continuous on $[a, b]$, as one sees simply by using the above definition with just one interval ($N = 1$).

(2) For any $f : [a, b] \to \mathbb{R}$, f is AC $\Rightarrow f$ is BV.

To check (2) we let $\delta > 0$ be the δ as in the definition of AC corresponding to $\varepsilon = 1$, and let $\mathcal{Q} : a = y_0 < y_1 < \cdots < y_Q = b$ be any partition of $[a, b]$ with $y_j - y_{j-1} < \delta$ for each $j = 1, \ldots, Q$. Now let $\mathcal{P}$ be *any* partition of $[a, b]$ and let $\widetilde{\mathcal{P}} = \mathcal{P} \cup \mathcal{Q}$. Since refinement evidently does not decrease the value of $T_{f,\mathcal{P}}$ we then have

$$T_{f,\mathcal{P}} \le T_{f,\mathcal{P}\cup\mathcal{Q}} \le T_{f|[y_{j-1},y_j],(\mathcal{P}\cup\mathcal{Q})\cap[y_{j-1},y_j]} \le \sum_{j=1}^{Q} T_{f|[y_{j-1},y_j]} \le Q$$

since $T_{f|[y_{j-1},y_j]} \le 1$ (because $y_j - y_{j-1} < \delta$) for each $j = 1, \ldots, Q$.

We now state a theorem which completely characterizes AC functions, as follows:

Theorem 3. *Let $f : [a, b] \to \mathbb{R}$. Then*
f is AC on $[a, b] \iff \exists$ a Lebesgue integrable g on $[a, b]$ with $f(x) = f(a) + \int_a^x g(t)\,dt \;\; \forall x \in [a, b]$.

Before we begin the proof, we need a simple lemma about non-negative integrable functions on an abstract measure space $(X, \mathcal{A}, \mu)$.

Lemma 5. *Let $(X, \mathcal{A}, \mu)$ be any measure space and $f : X \to [0, \infty)$ any μ-integrable function. Then for each $\varepsilon > 0$ there is a $\delta > 0$ such that $\int_A f\,d\mu < \varepsilon$ for all $A \in \mathcal{A}$ with $\mu(A) < \delta$.*

Proof: For $N = 1, 2, \ldots$, let $f_N = \min\{f, N\}$, so that f_N is an increasing sequence of non-negative $\mathcal{A}$-measurable functions which converges pointwise to f on X, and hence by the monotone convergence theorem we have

$$\int_X \big(f - f_N\big) \to 0 \text{ as } N \to \infty.$$

Thus for given $\varepsilon > 0$ we can select N such that $\int_X \big(f - f_N\big) < \varepsilon/2$, and on the other hand trivially for any set $A \in \mathcal{A}$ we have $\int_A f_N < N\mu(A)$, and so

$$\int_A f = \int_A f_N + \int_A \big(f - f_N\big) \le N\mu(A) + \int_X \big(f - f_N\big) \le N\mu(A) + \varepsilon/2 < \varepsilon$$

provided $\mu(A) < \varepsilon/2N$, and so the lemma is proved with $\delta = \varepsilon/2N$.

Proof of Theorem 3 "⇐": We are given $f(x) = f(a) + \int_a^x g(t)\,dt$ where $g : [a,b] \to \mathbb{R}$ is Lebesgue integrable on $[a,b]$. According to Lemma 5, for a given $\varepsilon > 0$ we can choose $\delta > 0$ such that if A is a λ-measurable subset of $[a,b]$ with $\lambda(A) < \delta$ then $\int_A |g|\,d\lambda < \varepsilon$. So, with this δ, let $[x_i, y_i], i = 1, \ldots, N$, be any p.w.d. intervals in $[a,b]$ with $\sum_{i=1}^N (y_i - x_i) < \delta$. Then $\sum_{i=1}^N |f(y_i) - f(x_i)| = \sum_{i=1}^N |\int_{x_i}^{y_i} g(t)\,dt| \le \sum_{i=1}^N \int_{[x_i,y_i]} |g(t)|\,dt = \int_{\cup_{i=1}^N [x_i,y_i]} |g(t)|\,dt < \varepsilon$, so we have checked the definition of AC.

Proof of Theorem 3 "⇒": Recall from the above discussion that AC ⇒ BV ⇒ $f = f_1 - f_2$ where f_1, f_2 are increasing on $[a,b]$, so by Theorem 1 we have f' is Lebesgue integrable, so to complete the proof we just need to show that $f(x) - \int_a^x f'(t)\,dt$ is constant on $[a,b]$ (then we have the required conclusion with $g = f'$). So let

$$F(x) = f(x) - \int_a^x f'(t)\,dt,$$

and observe that by Corollary 3 we have $F'(x) = 0$ for λ-a.e. $x \in (a,b)$. Thus with

$$S = \{x \in (a,b) : F'(x) \text{ exists and } = 0\}$$

we have $\lambda([a,b] \setminus S) = 0$ and of course, by definition of $F'(x) = 0$, for any given $\varepsilon > 0$ the set S is covered finely by the collection $\mathcal{B}$ of closed intervals $[x,y] \subset (a,b)$ such that $|F(y) - F(x)| \le \varepsilon(y - x)$. Then by the Vitali Covering Lemma, for each $\varepsilon, \delta > 0$ there are p.w.d. closed intervals $[x_1, y_1], \ldots, [x_N, y_N] \subset (a,b)$ with

$$\lambda([a,b] \setminus (\cup_{j=1}^N [x_j, y_j])) = \lambda(S \setminus (\cup_{j=1}^N [x_j, y_j])) < \delta$$
$$|F(y_i) - F(x_i)| \le \varepsilon(y_i - x_i), \quad i = 1, \ldots, N.$$

Without loss of generality we can assume that these intervals $[x_i, y_i]$ are labelled so that $a < x_1 < y_1 < x_2 < y_2 \cdots < x_N < y_N < b$, and then

$$[a,b] \setminus (\cup_{i=1}^N (x_i, y_i)) = \cup_{k=0}^N [y_k, x_{k+1}] \text{ and hence } \sum_{k=0}^N (x_{k+1} - y_k) < \delta,$$

where for convenience of notation we set $y_0 = a$ and $x_{N+1} = b$.

Now f is given to be AC and $\int_a^x f'(t)\,dt$ is AC by the proof of "⇐" above, so F is AC, and hence for any given $\varepsilon > 0$ we can choose the above $\delta > 0$ such that $\sum_{k=0}^N |F(x_{k+1}) - F(y_k)| < \varepsilon$ (notice this inequality holds by definition of AC because $\sum_{k=0}^N (x_{k+1} - y_k) = \lambda([a,b] \setminus \cup_{i=1}^N [x_i, y_i]) < \delta$). Then, with $z_0 = a, z_1 = x_1, z_2 = y_1, \ldots, z_{2N-1} = x_N, z_{2N} = y_N, z_{2N+1} = b$, we have

$$\begin{aligned} |F(b) - F(a)| &= \big|\textstyle\sum_{j=1}^{2N+1} \big(F(z_j) - F(z_{j-1})\big)\big| \\ &= \big|\textstyle\sum_{i=1}^N \big(F(y_i) - F(x_i)\big) + \sum_{k=0}^N \big(F(x_{k+1}) - F(y_k)\big)\big| \\ &\le \varepsilon \textstyle\sum_{i=1}^N (y_i - x_i) + \varepsilon \le (b - a + 1)\varepsilon. \end{aligned}$$

Thus, since $\varepsilon > 0$ is arbitrary, we have proved $F(b) = F(a)$. Since we can repeat the proof on the interval $[a,x]$ for any $x \in (a,b]$, this shows that $F(x)$ is constant (equal to $f(a)$) on $[a,b]$.

We conclude this supplement by showing that the method used to prove Lemma 1 and Lemma 2 above easily modifies to give the following theorem about differentiation of locally finite Borel measures in $\mathbb{R}^n$.

Theorem 4. *Let μ be a Borel measure on $\mathbb{R}^n$ which is finite on bounded Borel sets. Then the density $\Theta_\mu(x) = \lim_{\rho \downarrow 0} \frac{\mu(B_\rho(x))}{\omega_n \rho^n}$ exists and is real for λ-a.e. $x \in \mathbb{R}^n$.*

Proof: We have to show that $\{x : \Theta_{\mu*}(x) < \Theta^*_\mu(x)\}$ has measure zero and also that $\Theta^*_\mu(x) < \infty$ for λ-a.e. $x \in \mathbb{R}^n$, where $\Theta^*_\mu(x) = \limsup_{\rho \downarrow 0} \frac{\mu(B_\rho(x))}{\lambda(B_\rho(x))}$ and $\Theta_{\mu*}(x) = \liminf_{\rho \downarrow 0} \frac{\mu(B_\rho(x))}{\lambda(B_\rho(x))}$.

First observe that if $\beta > 0$, $U \subset \mathbb{R}^n$ is a bounded open set, and if and $S \subset \{x \in U : \Theta^*_\mu(x) > \beta\}$, then (since $x \in S \Rightarrow \frac{\mu(B_{\rho_j}(x))}{\lambda(B_{\rho_j}(x))} > \beta$ for some sequence $\rho_j \downarrow 0$) the set of closed balls $B_\rho(x)$ such that $B_\rho(x) \subset U$ and $\mu(B_\rho(x)) > \beta\lambda(B_\rho(x))$ covers S finely. Hence by Vitali (for Lebesgue measure), there is a p.w.d. collection $B_{\rho_j}(x_j) \subset U$ such that $\mu(B_{\rho_j}(x_j)) > \beta\lambda(B_{\rho_j}(x_j))$ and $\lambda(S \setminus (\cup_{j=1}^N B_{\rho_j}(x_j))) \to 0$ as $N \to \infty$. Thus if $\varepsilon > 0$ there is N such that

$$\begin{aligned}\beta\lambda(S) &\le \beta\lambda(S \cap (\cup_{j=1}^N B_{\rho_j}(x_j))) + \beta\lambda(S \setminus (\cup_{j=1}^N B_{\rho_j}(x_j)))\\ &\le \beta{\textstyle\sum_{j=1}^N}\lambda(S \cap B_{\rho_j}(x_j)) + \beta\lambda(S \setminus (\cup_{j=1}^N B_{\rho_j}(x_j)))\\ &\le {\textstyle\sum_{j=1}^N}\mu(B_{\rho_j}(x_j)) + \beta\varepsilon = \mu(\cup_j B_{\rho_j}(x_j)) + \beta\varepsilon.\end{aligned}$$

Thus since $\varepsilon > 0$ is arbitrary and since $\cup_j B_{\rho_j}(x_j) \subset U$ we thus have

$$\beta\lambda(S) \le \mu(U). \tag{1}$$

Notice that in particular if we take S to be the set of points x in the ball $U = \breve{B}_j(0)$ where $\Theta^*_\mu(x) = \infty$ then we can apply this with each β, thus implying that $\lambda(S) = 0$. Thus (since j is arbitary) we have

$$\Theta^*_\mu(x) < \infty, \quad \lambda \text{ a.e. } x \in \mathbb{R}^n. \tag{2}$$

Next observe that

$$\{x \in \mathbb{R}^n : \Theta_{\mu*}(x) < \Theta^*_\mu(x)\} = \cup_{\alpha,\beta \text{ rational},0<\alpha<\beta,k\in\{1,2,\ldots\}} S_{\alpha,\beta,k}$$

where

$$S_{\alpha,\beta,k} = \{x \in \mathbb{R}^n : |x| < k,\ \Theta_{\mu*}(x) < \alpha < \beta < \Theta^*_\mu(x)\}$$

Now let V be an open set such that $V \supset S_{\alpha,\beta,k}$ and such that $\lambda(V) < \lambda(S_{\alpha,\beta,k}) + \varepsilon$, and let $\mathcal{B}$ be the set of closed balls $B_\rho(x) \subset V$ such that $\mu(B_\rho(x)) < \alpha\lambda(B_\rho(x))$. Then evidently $\mathcal{B}$ covers $S_{\alpha,\beta,k}$ finely, and so by the Vitali lemma there are p.w.d. balls $B_{\rho_j}(x_j)$ in $\mathcal{B}$ with $\lambda(S_{\alpha,\beta,k} \setminus (\cup_{j=1}^N B_{\rho_j}(x_j))) \to 0$ as $N \to \infty$, and for each j

$$\mu(\breve{B}_{\rho_j}(x_j)) \le \mu(B_{\rho_j}(x_j)) \le \alpha\lambda(B_{\rho_j}(x_j)).$$

But then for any given $\varepsilon > 0$ we can select N so that $\lambda(S_{\alpha,\beta,k} \setminus (\cup_{j=1}^N B_{\rho_j}(x_j))) < \varepsilon$ and then for each $j = 1, \ldots, N$ use (1) with $S_{\alpha,\beta,k} \cap \breve{B}_{\rho_j}(x_j)$ in place of S and $U = \breve{B}_{\rho_j}(x_j)$, giving

$$\begin{aligned}\beta\lambda(S_{\alpha,\beta,k} \cap (\cup_{j=1}^N \breve{B}_{\rho_j}(x_j))) &\le {\textstyle\sum_{j=1}^N}\beta\lambda(S_{\alpha,\beta,k} \cap \breve{B}_{\rho_j}(x_j))\\ &\le {\textstyle\sum_{j=1}^N}\mu(\breve{B}_{\rho_j}(x_j)) \le \alpha{\textstyle\sum_{j=1}^N}\lambda(B_{\rho_j}(x_j)) \le \alpha\lambda(\cup_{j=1}^N B_{\rho_j}(x_j))\\ &\le \alpha\lambda(V) \le \alpha\lambda(S_{\alpha,\beta,k}) + \alpha\varepsilon.\end{aligned}$$

Since $\lambda(S_{\alpha,\beta,k} \setminus (\cup_{j=1}^N B_{\rho_j}(x_j))) < \varepsilon$, this gives

$$\beta\lambda(S_{\alpha,\beta,k}) \le \alpha\lambda(S_{\alpha,\beta,k}) + (\alpha + \beta)\varepsilon,$$

and letting $\varepsilon \to 0$ we thus have

$$\beta\lambda(S_{\alpha,\beta,k}) \le \alpha\lambda(S_{\alpha,\beta,k}) < \infty;$$

that is, $\lambda(S_{\alpha,\beta,k}) = 0$ as required.