4

Covering Theorems and the Differentiation of Integrals

A number of fundamental problems in geometric analysis—ranging from decompositions of measures to density of sets to approximate continuity of functions—depend on the theory of differentiation of integrals. These results, in turn, depend on a variety of so-called covering theorems for families of balls (and other geometric objects). Thus we come upon the remarkable, and profound, fact that deep analytic facts reduce to rather elementary (but often difficult) facts about Euclidean geometry.

The technique of covering lemmas has become an entire area of mathematical analysis (see, for example, [DGu 75] and [Ste 93]). It is intimately connected with problems of differentiation of integrals, with certain maximal operators (such as the Hardy–Littlewood maximal operator), with the boundedness of multiplier operators in harmonic analysis, and (concomitantly) with questions of summation of Fourier series.

The purpose of the present chapter is to introduce some of these ideas. We do not strive for any sort of comprehensive treatment, but rather to touch upon the key concepts and to introduce some of the most pervasive techniques and applications.

4.1 Wiener's Covering Lemma and Its Variants

Let $S \subseteq \mathbb{R}^N$ be a set. A *covering* of S will be a collection $\mathcal{U} = \{U_\alpha\}_{\alpha \in \mathcal{A}}$ of sets such that $\bigcup_{\alpha \in \mathcal{A}} U_\alpha \supseteq S$. If all the sets of $\mathcal{U}$ are open, then we call $\mathcal{U}$ an *open covering* of S. A *subcovering* of the covering $\mathcal{U}$ is a covering $\mathcal{V} = \{V_\beta\}_{\beta \in \mathcal{B}}$ such that each V_β is one of the U_α. A *refinement* of the covering $\mathcal{U}$ is a collection $\mathcal{W} = \{W_\gamma\}_{\gamma \in \mathcal{G}}$ of sets such that each W_γ is a subset of some U_α. If $\mathcal{U}$ is a covering of a set S, then the *valence* of $\mathcal{U}$ is the least positive integer M such that no point of S lies in more than M of the sets in $\mathcal{U}$.

It is elementary to see that any open covering of a set $S \subseteq \mathbb{R}^N$ has a countable subcover. We also know, thanks to Lebesgue, that any open covering of S has a refinement with valence at most $N + 1$ (see [HW 41, Theorem V 1]).

G.S. Krantz, R.H. Parks (eds.), *Geometric Integration Theory,* doi: 10.1007/978-0-8176-4679-0_4,

Wiener's covering lemma[1] concerns a covering of a set by a collection of balls. When applying the lemma, one must be willing to replace any particular ball by a ball with the same center but triple its radius—see Figure 4.1.

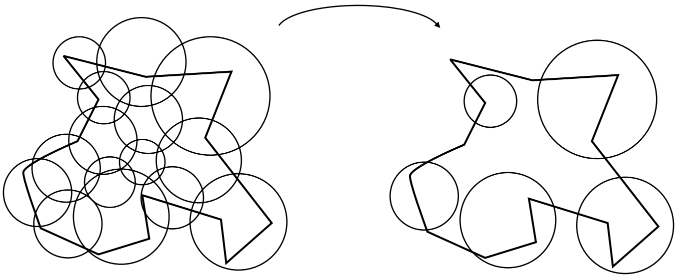

Fig. 4.1. Wiener's covering lemma.

Lemma 4.1.1 (Wiener). *Let $K \subseteq \mathbb{R}^N$ be a compact set with a covering $\mathcal{U} = \{B_\alpha\}_{\alpha \in A}$, $B_\alpha = \mathbb{B}(c_\alpha, r_\alpha)$, by open balls. Then there is a subcollection $B_{\alpha_1}, B_{\alpha_2}, \ldots, B_{\alpha_m}$, consisting of pairwise disjoint balls, such that*

$$\bigcup_{j=1}^{m} \mathbb{B}(c_{\alpha_j}, 3r_{\alpha_j}) \supseteq K.$$

Proof. Since K is compact, we may immediately assume that there are only finitely many B_α. Let B_{α_1} be the ball in this collection that has the greatest radius (this ball may not be unique). Let B_{α_2} be the ball that is disjoint from B_{α_1} and has greatest radius among those balls that are disjoint from B_{α_1} (again, this ball may not be unique). At the jth step choose the (not necessarily unique) ball disjoint from $B_{\alpha_1}, \ldots, B_{\alpha_{j-1}}$ that has greatest radius among those balls that are disjoint from $B_{\alpha_1}, \ldots, B_{\alpha_{j-1}}$. Continue. The process ends in finitely many steps. We claim that the B_{α_j} chosen in this fashion do the job.

For each j, we will write $B_{\alpha_j} = \mathbb{B}(c_{\alpha_j}, r_{\alpha_j})$. It is enough to show that $B_\alpha \subseteq \bigcup_j \mathbb{B}(c_{\alpha_j}, 3r_{\alpha_j})$ for every α. Fix an α. If $\alpha = \alpha_j$ for some j then we are done. If $\alpha \notin \{\alpha_j\}$, let j_0 be the first index j with $B_{\alpha_j} \cap B_\alpha \neq \emptyset$ (there must be one; otherwise, the process would not have stopped). Then $r_{\alpha_{j_0}} \geq r_\alpha$; otherwise, we selected $B_{\alpha_{j_0}}$ incorrectly. But then (by the triangle inequality) $\mathbb{B}(c_{\alpha_{j_0}}, 3r_{\alpha_{j_0}}) \supseteq \mathbb{B}(c_\alpha, r_\alpha)$ as desired. □

For completeness, and because it is such an integral part of the classical theory of measures, we now present the venerable covering theorem of Vitali.[2]

[1] Norbert Wiener (1894–1964).

[2] Giuseppe Vitali (1875–1932).

Proposition 4.1.2. *Let $A \subseteq \mathbb{R}^N$ and let $\mathcal{B}$ be a family of open balls. Suppose that each point of A is contained in arbitrarily small balls belonging to $\mathcal{B}$. Then there exist pairwise disjoint balls $B_j \in \mathcal{B}$ such that*

$$\mathcal{L}^N\left(A \setminus \bigcup_j B_j\right) = 0 .$$

Furthermore, for any $\epsilon > 0$, we may choose the balls B_j in such a way that

$$\sum_j \mathcal{L}^N(B_j) \le \mathcal{L}^N(A) + \epsilon .$$

Proof. The last statement will follow from the substance of the proof. For the first statement, let us begin by making the additional assumption (which we shall remove at the end) that the set $A \equiv A_0$ is bounded. We may select a bounded open set U_0 that contains $\overline{A_0}$ and that is such that $\mathcal{L}^N(U_0)$ exceeds $\mathcal{L}^N(A_0)$ by as small a quantity as we may wish. In fact, we demand that

$$\mathcal{L}^N(U_0) \le (1 + 5^{-N})\mathcal{L}^N(A_0) .$$

Now focus attention on those balls that lie in U_0. By Lemma 4.1.1, we may select a finite, pairwise disjoint collection $B_j = \mathbb{B}(x_j, r_j) \in \mathcal{B}$, $j = 1, \ldots, k_1$, such that $B_j \subseteq U_0$ for each j and

$$\overline{A}_0 \subseteq \bigcup_{j=1}^{k_1} \mathbb{B}(x_j, 3r_j) .$$

Now we may calculate that

$$3^{-N}\mathcal{L}^N(A_0) \le 3^{-N}\sum_j \mathcal{L}^N[\mathbb{B}(x_j, 3r_j)] = 3^{-N}\sum_j 3^N\mathcal{L}^N(B_j) = \sum_j \mathcal{L}^N(B_j) .$$

Let

$$A_1 = A_0 \setminus \bigcup_{j=1}^{k_1} \overline{B_j} .$$

Then

$$\begin{aligned}\mathcal{L}^N(A_1) &\le \mathcal{L}^N\left(U_0 \setminus \bigcup_{j=1}^{k_1} \overline{B_j}\right) \\ &= \mathcal{L}^N\left(U_0 \setminus \bigcup_{j=1}^{k_1} B_j\right) = \mathcal{L}^N(U_0) - \sum_{j=1}^{k_1} \mathcal{L}^N(B_j) \\ &\le (1 + 5^{-N} - 3^{-N})\,\mathcal{L}^N(A_0) \equiv u \cdot \mathcal{L}^N(A_0) ,\end{aligned}$$

where $u \equiv 1 + 5^{-N} - 3^{-N} < 1$. Now A_1 is a bounded subset of $\mathbb{R}^N \setminus \bigcup_{j=1}^{k_1} \overline{B_j}$. Hence we may find a bounded, open set $U_1 \subseteq U_0$ such that

$$A_1 \subseteq U_1 \subseteq \mathbb{R}^N \setminus \bigcup_{j=1}^{k_1} \overline{B_j}$$

and

$$\mathcal{L}^N(U_1) \le (1 + 5^{-N})\, \mathcal{L}^N(A_1)\,.$$

Just as in the first iteration of this construction, we may now find disjoint balls B_j, $j = k_1 + 1, \ldots, k_2$, for which $B_j \subseteq U_1$ and

$$\mathcal{L}^N(A_2) \le u \cdot \mathcal{L}^N(A_1) \le u^2\, \mathcal{L}^N(A_0)\,;$$

here

$$A_2 = A_1 \setminus \bigcup_{j=k_1+1}^{k_2} \overline{B_j} = A_0 \setminus \bigcup_{j=1}^{k_2} \overline{B_j}\,.$$

By our construction, all the balls $B_1, \ldots, B_{k_2}$ are disjoint.

After m repetitions of this procedure, we find that we have balls $B_1, B_2, \ldots, B_{k_m}$ such that

$$\mathcal{L}^N\left(A_0 \setminus \bigcup_{j=1}^{k_m} B_j\right) \le u^m\, \mathcal{L}^N(A_0)\,.$$

Since $u < 1$, the result follows.

For the general case, we simply decompose $\mathbb{R}^N$ into closed unit cubes Q_ℓ with disjoint interiors and sides parallel to the axes and apply the result just proved to each $A_0 \cap Q_\ell$. □

The Maximal Function

A classical construct, due to Hardy and Littlewood,[3] is the so-called maximal function. It is used to control other operators, and also to study questions of differentiation of integrals.

Definition 4.1.3. If f is a locally integrable function on $\mathbb{R}^N$, we let

$$Mf(x) = \sup_{R>0} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} |f(t)|\, d\mathcal{L}^N(t)\,.$$

The operator M is called the *Hardy–Littlewood maximal operator*. The functions to which M is applied may be real-valued or complex-valued. A few facts are immediately obvious about M:

(1) M is not linear, but it is *sublinear* in the sense that

$$M[f + g](x) \le Mf(x) + Mg(x)\,.$$

[3] Godfrey Harold Hardy (1877–1947), John Edensor Littlewood (1885–1977).

(2) Mf is always nonnegative, and it *could* be identically equal to infinity.
(3) Mf makes sense for any locally integrable f.

We will in fact prove that Mf is finite $\mathcal{L}^N$-almost everywhere, for any $f \in L^p$. In order to do so, it is convenient to formulate a weak notion of boundedness for operators. To begin, we say that a measurable function f is *weak type* p, $1 \le p < \infty$, if there exists a $C = C(f)$ with $0 < C < \infty$ such that, for any $\lambda > 0$,

$$\mathcal{L}^N(\{x \in \mathbb{R}^N : |f(x)| > \lambda\}) \le \frac{C}{\lambda^p}.$$

An operator T on L^p taking values in the collection of measurable functions is said to be of *weak type* (p, p) if there exists a $C = C(T)$ with $0 < C < \infty$ such that, for any $f \in L^p$ and for any $\lambda > 0$,

$$\mathcal{L}^N(\{x \in \mathbb{R}^N : |Tf(x)| > \lambda\}) \le C \cdot \left(\frac{\|f\|_{L^p}}{\lambda}\right)^p.$$

A function is defined to be *weak type* ∞ when it is L^∞. For $1 \le p < \infty$, an L^p function is certainly weak type p, but the converse is not true. In fact, we note that the function $f(x) = |x|^{-1/p}$ on $\mathbb{R}^1$ is weak type p, but not in L^p, for $1 \le p < \infty$. The Hilbert transform (see [Kra 99]) is an important operator that is not bounded on L^1 but is in fact weak type (1, 1).

Proposition 4.1.4. *The Hardy–Littlewood maximal operator M is weak type* (1, 1).

Proof. Let $\lambda > 0$. Set $S_\lambda = \{x : |Mf(x)| > \lambda\}$. Because Mf is the supremum of a set of continuous functions, Mf is lower semicontinuous, and consequently, S_λ is open.

Since S_λ is open, we may let $K \subseteq S_\lambda$ be a compact subset with $2\,\mathcal{L}^N(K) \ge \mathcal{L}^N(S_\lambda)$. For each $x \in K$, there is a ball $B_x = \mathbb{B}(x, r_x)$ with

$$\lambda < \frac{1}{\mathcal{L}^N(B_x)} \int_{B_x} |f(t)| d\mathcal{L}^N(t).$$

Then $\{B_x\}_{x \in K}$ is an open cover of K by balls. By Lemma 4.1.1, there is a subcollection $\{B_{x_j}\}_{j=1}^M$ that is pairwise disjoint but such that the threefold dilates of these selected balls still cover K. Then

$$\begin{aligned}
\mathcal{L}^N(S_\lambda) \le 2\,\mathcal{L}^N(K) \le 2\,\mathcal{L}^N\left(\bigcup_{j=1}^M \mathbb{B}(x_j, 3r_j)\right) &\le 2\sum_{j=1}^M \mathcal{L}^N[\mathbb{B}(x_j, 3r_j)] \\
&\le \sum_{j=1}^M 2 \cdot 3^N\, \mathcal{L}^N(B_{x_j}) \\
&\le \sum_{j=1}^M \frac{2 \cdot 3^N}{\lambda} \int_{B_{x_j}} |f(t)| d\mathcal{L}^N(t) \\
&\le \frac{2 \cdot 3^N}{\lambda} \|f\|_{L^1}.
\end{aligned}$$

□

One of the venerable applications of the Hardy–Littlewood operator is the Lebesgue differentiation theorem:

Theorem 4.1.5. *Let f be a locally Lebesgue integrable function on $\mathbb{R}^N$. Then, for $\mathcal{L}^N$-almost every $x \in \mathbb{R}^N$, it holds that*

$$\lim_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} f(t)\, d\mathcal{L}^N(t) = f(x)\,.$$

Proof. Multiplying f by a compactly supported C^∞ function that is identically 1 on a ball, we may as well suppose that $f \in L^1$. We may also assume, by linearity, that f is real-valued. We begin by proving that

$$\lim_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} f(t)\, d\mathcal{L}^N(t)$$

exists.

Let $\epsilon > 0$. Select a function φ, continuous with compact support, and real-valued, such that $\|f - \varphi\|_{L^1} < \epsilon^2$. Then

$$\begin{aligned}
&\mathcal{L}^N\Bigg\{x \in \mathbb{R}^N : \Bigg|\limsup_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} f(t)\, d\mathcal{L}^N(t) \\
&\qquad - \liminf_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} f(t)\, d\mathcal{L}^N(t)\Bigg| > \epsilon\Bigg\} \\
&\le \mathcal{L}^N\Bigg\{x \in \mathbb{R}^N : \limsup_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} |f(t) - \varphi(t)|\, d\mathcal{L}^N(t) \\
&\qquad + \Bigg|\limsup_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} \varphi(t)\, d\mathcal{L}^N(t) \\
&\qquad - \liminf_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} \varphi(t)\, d\mathcal{L}^N(t)\Bigg| \\
&\qquad + \limsup_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} |\varphi(t) - f(t)|\, d\mathcal{L}^N(t) > \epsilon\Bigg\} \\
&\le \mathcal{L}^N\Bigg\{x \in \mathbb{R}^N : \limsup_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} |f(t) - \varphi(t)|\, d\mathcal{L}^N(t) > \frac{\epsilon}{3}\Bigg\} \\
&\quad + \mathcal{L}^N\Bigg\{x \in \mathbb{R}^N : \Bigg|\limsup_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} \varphi(t)\, d\mathcal{L}^N(t) \\
&\qquad - \liminf_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} \varphi(t)\, d\mathcal{L}^N(t)\Bigg| > \frac{\epsilon}{3}\Bigg\} \\
&\quad + \mathcal{L}^N\Bigg\{x \in \mathbb{R}^N : \limsup_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} |\varphi(t) - f(t)|\, d\mathcal{L}^N(t) > \frac{\epsilon}{3}\Bigg\} \\
&\equiv I + II + III\,.
\end{aligned}$$

Now $II = 0$ because the set being measured is empty (since φ is continuous). Each of I and III may be estimated by

$$\mathcal{L}^N \left\{ x \in \mathbb{R}^N : M(f - \varphi)(x) > \epsilon/3 \right\},$$

and this, by Proposition 4.1.4, is majorized by

$$C \cdot \frac{\epsilon^2}{\epsilon/3} = c \cdot \epsilon \,.$$

In sum, we have proved the estimate

$$\mathcal{L}^N \left\{ x \in \mathbb{R}^N : \left| \limsup_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} f(t)\, d\mathcal{L}^N(t) \right.\right.$$
$$\left.\left. - \liminf_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} f(t)\, d\mathcal{L}^N(t) \right| > \epsilon \right\} \le c \cdot \epsilon \,.$$

It follows immediately that

$$\lim_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} f(t)\, d\mathcal{L}^N(t)$$

exists for $\mathcal{L}^N$-almost every $x \in \mathbb{R}^N$.

The proof that the limit actually equals $f(x)$ at $\mathcal{L}^N$-almost every point follows exactly the same lines. We shall omit the details. □

Corollary 4.1.6. *If $A \subseteq \mathbb{R}^N$ is Lebesgue measurable, then for almost every $x \in \mathbb{R}^N$, it holds that*

$$\chi_A(x) = \lim_{r \to 0^+} \frac{\mathcal{L}^N(A \cap \mathbb{B}(x, r))}{\mathcal{L}^N(\mathbb{B}(x, r))}.$$

Proof. Set $f = \chi_A$. Then

$$\int_{\mathbb{B}(x,r)} f(t)\, d\mathcal{L}^N(t) = \mathcal{L}^N(A \cap \mathbb{B}(x, r)),$$

and the corollary follows from Theorem 4.1.5. □

Definition 4.1.7. A function $f : \mathbb{R}^N \to \mathbb{R}$ is said to be *approximately continuous* if, for $\mathcal{L}^N$-almost every $x_0 \in \mathbb{R}^N$ and for each $\epsilon > 0$, the set

$$\{x : |f(x) - f(x_0)| > \epsilon\}$$

has density 0 at x_0, that is,

$$0 = \lim_{r \to 0^+} \frac{\mathcal{L}^N(\{x : |f(x) - f(x_0)| > \epsilon\} \cap \mathbb{B}(x_0, r))}{\mathcal{L}^N(\mathbb{B}(x_0, r))}.$$

Corollary 4.1.8. *If a function* $f : \mathbb{R}^N \to \mathbb{R}$ *is Lebesgue measurable, then it is approximately continuous.*

Proof. Suppose that f is Lebesgue measurable. Let $q_1, q_2, \ldots$ be an enumeration of the rational numbers. For each positive integer i, let E_i be the set of points $x \notin \{z : f(z) < q_i\}$ for which

$$0 < \limsup_{r \to 0^+} \frac{\mathcal{L}^N(\{z : f(z) < q_i\} \cap \mathbb{B}(x, r))}{\mathcal{L}^N(\mathbb{B}(x, r))}$$

and let E^i be the set of points $x \notin \{z : q_i < f(z)\}$ for which

$$0 < \limsup_{r \to 0^+} \frac{\mathcal{L}^N(\{z : q_i < f(z)\} \cap \mathbb{B}(x, r))}{\mathcal{L}^N(\mathbb{B}(x, r))}.$$

By Corollary 4.1.6 and the Lebesgue measurability of f, we know that

$$\mathcal{L}^N(E_i) = 0 \text{ and } \mathcal{L}^N(E^i) = 0 .$$

Thus we see that

$$E \equiv \bigcup_{i=1}^{\infty} (E_i \cup E^i)$$

is also a set of Lebesgue measure zero.

Consider any point $x_0 \notin E$ and any $\epsilon > 0$. There exist rational numbers q_i and q_j such that

$$f(x_0) - \epsilon < q_i < f(x_0) < q_j < f(x_0) + \epsilon.$$

We have $\{x : |f(x) - f(x_0)| > \epsilon\} \subseteq \{z : f(z) < q_i\} \cup \{z : q_j < f(z)\}$. By the definition of E_i and E^j we have

$$0 = \lim_{r \to 0^+} \frac{\mathcal{L}^N(\{z : f(z) < q_i\} \cap \mathbb{B}(x_0, r))}{\mathcal{L}^N(\mathbb{B}(x_0, r))}$$

and

$$0 = \lim_{r \to 0^+} \frac{\mathcal{L}^N(\{z : q_j < f(z)\} \cap \mathbb{B}(x_0, r))}{\mathcal{L}^N(\mathbb{B}(x_0, r))}.$$

It follows that

$$0 = \lim_{r \to 0^+} \frac{\mathcal{L}^N(\{x : |f(x) - f(x_0)| > \epsilon\} \cap \mathbb{B}(x_0, r))}{\mathcal{L}^N(\mathbb{B}(x_0, r))}.$$

Since $x_0 \notin E$ and $\epsilon > 0$ were arbitrary, we conclude that f is approximately continuous. □

4.2 The Besicovitch Covering Theorem

Preliminary Remarks

The Besicovitch covering theorem,[4] which we shall treat in the present section, is of particular interest to geometric analysis because its statement and proof do not depend on a measure. This is a result about the geometry of balls in space.

The Besicovitch Covering Theorem

Theorem 4.2.1. *Let N be a positive integer. There is a constant $K = K(N)$ with the following property. Let $\mathcal{B} = \{B_j\}_{j=1}^M$, where $M \in \mathbb{N} \cup \{\infty\}$, be any finite or countable collection of balls in $\mathbb{R}^N$ with the property that the interior of no ball contains the center of any other. Then we may write*

$$\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_K$$

so that each $\mathcal{B}_j$, $j = 1, \ldots, K$, is a collection of balls with pairwise disjoint closures.

Here by a ball we mean a set B satisfying $\mathbb{B}(x, r) \subseteq B \subseteq \overline{\mathbb{B}}(x, r)$, for some $x \in \mathbb{R}^N$ and some $r > 0$.

It is a matter of some interest to determine what the best possible K is for any given dimension N. Significant progress on this problem has been made in [Sul 94]. See also [Loe 93]. Certainly our proof below will give little indication of the best K.

We shall see that the heart of this theorem is the following lemma about balls.

Lemma 4.2.2. *There is a constant $\widetilde{K} = \widetilde{K}(N)$, depending only on the dimension of the space $\mathbb{R}^N$, with the following property: Let $B_0 = \overline{\mathbb{B}}(x_0, r_0)$ be a ball of fixed positive radius. Let $B_1 = \overline{\mathbb{B}}(x_1, r_1)$, $B_2 = \overline{\mathbb{B}}(x_2, r_2)$, $\ldots$, $B_p = \overline{\mathbb{B}}(x_p, r_p)$ be balls such that*

(1) *Each B_j has nonempty intersection with B_0, $j = 1, \ldots, p$;*
(2) *The radii r_j satisfy $r_j \geq r_0$ for all $j = 1, \ldots, p$;*
(3) *The interior of no ball B_j contains the center of any other B_k for $j, k \in \{0, \ldots, p\}$ with $j \neq k$.*

Then $p \leq \widetilde{K}$.

Here is what the lemma says in simple terms: Fix the ball B_0. Then at most $\widetilde{K}$ pairwise disjoint balls of (at least) the same size can touch B_0. Note here that being "pairwise disjoint" and "intersecting but not containing the center of the other ball" are essentially equivalent: if the second condition holds then shrinking each ball by a factor of one-half makes the balls pairwise disjoint; if the balls are already pairwise disjoint, have equal radii, and are close together, then doubling their size arranges for the first condition to hold.

Our proof of Lemma 4.2.2 is based on the next two lemmas—which in essence rely only on two-dimensional Euclidean geometry (trigonometry)—and on the fact

[4] Abram Samoilovitch Besicovitch (1891–1970).

that we can choose a set of unit vectors in $\mathbb{R}^N$ such that every direction is within a small angle of one of our chosen unit vectors (where the measure of an angle between two vectors is defined to be in the interval $[0, \pi]$).

Lemma 4.2.3. *Suppose the ball $\overline{\mathbb{B}}(q, r)$, with $r \geq 1$, intersects the closed unit ball and does not contain the origin in its interior, i.e., $r \leq |q|$. If u is a unit vector making an angle $\phi \leq \pi/6$ with q, then $\sqrt{3}\, u \in \overline{\mathbb{B}}(q, r)$.*

Proof. Because $\overline{\mathbb{B}}(q, r)$ intersects the closed unit ball and does not contain the origin in its interior, we can write $|q| = x + r$ with $0 \leq x \leq 1 \leq r$. By the law of cosines we have

$$\begin{aligned} |q - \sqrt{3}\, u|^2 &= |q|^2 + 3 - 2\sqrt{3}\, |q| \cos\phi \\ &\leq |q|^2 + 3 - 2\sqrt{3}\, |q| \cos\tfrac{\pi}{6} \\ &= (x + r)^2 + 3 - 3\,(x + r)\,. \end{aligned}$$

Thus it will suffice to show that

$$(x + r)^2 + 3 - 3\,(x + r) \leq r^2$$

or, equivalently,

$$f(x, r) = x^2 + 2xr + 3 - 3x - 3r \leq 0\,.$$

Since for each fixed r, $f(x, r)$ is quadratic in x with positive second derivative and since we are concerned only with the range $0 \leq x \leq 1$, it will suffice to consider only the endpoints $x = 0$ and $x = 1$. But we have

$$f(0, r) = 3 - 3r \leq 0 \quad \text{and} \quad f(1, r) = 1 + 2r + 3 - 3 - 3r = 1 - r \leq 0\,,$$

as required. □

Lemma 4.2.4. *Suppose neither of the balls $\overline{\mathbb{B}}(q_1, r_1)$ and $\overline{\mathbb{B}}(q_2, r_2)$ contains the center of the other ball in its interior. If the point p is in both balls, then the angle between $q_1 - p$ and $q_2 - p$ is at least $\pi/3$.*

Proof. To see this, we denote the angle in question by θ and use the law of cosines to compute

$$|q_1 - q_2|^2 = |q_1 - p|^2 + |q_2 - p|^2 - 2\,|q_1 - p|\,|q_2 - p| \cos\theta\,.$$

So we have

$$\cos\theta \leq \frac{|q_1 - p|^2 + |q_2 - p|^2 - |q_1 - q_2|^2}{2\,|q_1 - p|\,|q_2 - p|}\,.$$

Since neither ball contains the center of the other ball in its interior, we know that $|q_1 - q_2|$ is at least as large as the radius of either ball. So we have both $|q_1 - p| \leq$

$r_1 \leq |q_1 - q_2|$ and $|q_2 - p| \leq r_2 \leq |q_1 - q_2|$. Suppose without loss of generality that $|q_1 - p| \leq |q_2 - p|$. Then we estimate

$$\cos\theta \leq \frac{|q_1 - p|^2 + |q_2 - p|^2 - |q_1 - q_2|^2}{2\,|q_1 - p|\,|q_2 - p|}$$

$$\leq \frac{|q_1 - p|^2}{2\,|q_1 - p|\,|q_2 - p|}$$

$$= \frac{1}{2} \cdot \frac{|q_1 - p|}{|q_2 - p|} \leq \frac{1}{2}\,,$$

as required. □

Proof of Lemma 4.2.2. Suppose for the moment (we confirm this construction later) that we have chosen a set of unit vectors $u_1, u_2, \ldots, u_{\kappa(N)}$ in $\mathbb{R}^N$ with the property that for any unit vector $u \in \mathbb{R}^N$, there is a j such that the angle between u and u_j is strictly less than $\pi/6$ (picture points sufficiently dense on the unit sphere—see the discussion below). The number, $\kappa(N)$, of vectors u_j will be used below.

Consider balls $B_0, B_1, \ldots, B_p$ as in the statement of Lemma 4.2.2 and suppose that $p \geq \kappa(N)^2 + 1$. Without loss of generality, we may assume that $B_0 = \overline{\mathbb{B}}(0, 1)$. The direction to the center of each ball is within an angle strictly less than $\pi/6$ of one of the unit vectors u_j and so, by Lemma 4.2.3, must contain the point $\sqrt{3}\,u_j$. Since there are at least $\kappa(N)^2 + 1$ balls and only $\kappa(N)$ possible u_j's, there must be (at least) one j^* such that $\kappa(N) + 1$ of the balls contain the point $\sqrt{3}\,u_{j^*}$.

Now consider those $\kappa(N) + 1$ balls. The direction from $\sqrt{3}\,u_{j^*}$ to each center is within an angle strictly less than $\pi/6$ of one of the unit vectors u_k. But since there are $\kappa(N) + 1$ balls and only $\kappa(N)$ possible u_k's, there must be two centers within angle less than $\pi/6$ of the same direction and thus within an angle less than $\pi/3$ of each other, contradicting Lemma 4.2.4. We conclude that $p \leq \kappa(N)^2$.

Finally, we show that there exists a set of unit vectors $u_1, u_2, \ldots, u_{\kappa(N)}$ in $\mathbb{R}^N$ with the property that for any unit vector $u \in \mathbb{R}^N$, there is a j such that the angle between u and u_j is strictly less than $\pi/6$. Let

$$\mathcal{F} = \{\,\mathbb{B}(u_j, 1/4) : j = 1, 2, \ldots, \kappa(N)\,\}$$

be a maximal pairwise disjoint family of balls with centers in the unit sphere. All of the balls $\mathbb{B}(u_j, 1/4)$ are contained in $\mathbb{B}(0, 5/4)$, so, by comparing volumes, we see that

$$\kappa(N) \leq \frac{\Omega_N\,(5/4)^N}{\Omega_N\,(1/4)^N} = 5^N\,.$$

[In Remark 4.2.5, we give an alternative construction for the u_j that avoids any use of volume in $\mathbb{R}^N$ or $(N-1)$-dimensional area in the unit sphere.]

To see that the unit vectors $u_1, u_2, \ldots, u_{\kappa(N)}$ have the requisite property, let u be an arbitrary unit vector. There must exist a j such that $|u - u_j| < 1/2$; otherwise, we could add the ball $\mathbb{B}(u, 1/4)$ to the family $\mathcal{F}$, contradicting the maximality of $\mathcal{F}$. Fix such a j and let θ denote the angle between u_j and u. Using the law of cosines we estimate

$$\cos\theta = \frac{|u_j|^2 + |u|^2 - |u_j - u|^2}{2\,|u_j|\,|u|} = 1 - \tfrac{1}{2}|u_j - u|^2$$

$$\geq 7/8 > \sqrt{3}/2 = \cos\tfrac{\pi}{6}\,,$$

so the angle θ is strictly less than $\pi/6$. □

Remark 4.2.5. We now give another, more explicit, construction of a set of unit vectors $\mathcal{U} \subseteq \mathbb{R}^N$ with the property that for any unit vector $u \in \mathbb{R}^N$, there exists $u^* \in \mathcal{U}$ such that the angle between u and u^* is strictly less than $\pi/6$.

The vectors in $\mathcal{U}$ are formed by choosing $\theta_1, \theta_2, \ldots, \theta_{N-1}$ from the set

$$\left\{0,\ \tfrac{\pi}{m},\ \tfrac{2\pi}{m},\ \ldots,\ \tfrac{(m-1)\pi}{m},\ \pi\right\} \tag{4.1}$$

and choosing a sign $\tau \in \{-1, +1\}$. We then set

$$u_{\theta_1,\ldots,\theta_{N-1},\tau}$$
$$= \left(\cos\theta_1,\ \cos\theta_2\ \sin\theta_1,\ \ldots,\ \cos\theta_{N-1}\ \textstyle\prod_{i=1}^{N-2}\sin\theta_i,\ \tau\cdot\prod_{i=1}^{N-1}\sin\theta_i\right).$$

Given a unit vector $u \in \mathbb{R}^N$, there exist $0 \leq \phi_i \leq \pi$, $i = 1, 2, \ldots, N-1$, and $\tau' \in \{-1, +1\}$ such that

$$u = \left(\cos\phi_1,\ \cos\phi_2\ \sin\phi_1,\ \ldots,\ \cos\phi_{N-1}\ \textstyle\prod_{i=1}^{N-2}\sin\phi_i,\ \tau'\cdot\prod_{i=1}^{N-1}\sin\phi_i\right).$$

The sign τ' represents a hemisphere containing u.

The main fact needed to verify that u is within $\pi/6$ of some $u_{\theta_1,\ldots,\theta_{N-1},\tau}$ is that if $\tau = \tau'$, then

$$u\cdot u_{\theta_1,\ldots,\theta_{N-1},\tau} = \cos(\theta_1-\phi_1) - \sum_{k=1}^{N-1}\left(\Big[1-\cos(\theta_k-\phi_k)\Big]\prod_{\ell=1}^{k-1}\sin\theta_\ell\ \sin\phi_\ell\right). \tag{4.2}$$

Equation (4.2) is proved by induction on N.

One completes the construction by choosing a sufficiently large value for m in (4.1). □

H. Federer's concept of a *directionally limited metric space*—see [Fed 69, 2.8.9]—abstracts and formalizes the geometry that goes into the proof of Lemma 4.2.2. More precisely, it generalizes to abstract contexts the notion that a cone in a given direction can contain only a certain number of points with distance $\eta > 0$ from the vertex and distance η from each other. The interested reader is advised to study that source.

Now we can present the proof of Besicovitch's covering theorem.

Proof of Theorem 4.2.1. First consider the case $M < \infty$ (recall that M was the number of balls in $\mathcal{B}$, the given collection of balls).

We have an iterative procedure for selecting balls.

Select B_1^1 to be a ball of maximum radius (this ball may not be unique). Then select B_2^1 to be a ball of maximum radius such that $\overline{B_2^1}$ is disjoint from $\overline{B_1^1}$ (again, this ball may not be unique). Continue until this selection procedure is no longer possible (remember that there are only finitely many balls in total). Set $\mathcal{B}_1 = \left\{B_j^1\right\}$.

Now work with the remaining balls. Let B_1^2 be the ball with greatest radius. Then select B_2^2 to be the remaining ball with greatest radius such that $\overline{B_2^2}$ is disjoint from $\overline{B_1^2}$. Continue in this fashion until no further selection is possible. Set $\mathcal{B}_2 = \left\{B_j^2\right\}$.

Working with the remaining balls, we now produce the family $\mathcal{B}_3$, and so forth (see Figure 4.2). Clearly, since in total there are only finitely many balls, this procedure must stop. We will have produced finitely many—say q—nonempty families of balls, each family consisting of balls having pairwise disjoint closures: $\mathcal{B}_1, \ldots, \mathcal{B}_q$. It remains to say how large q can be.

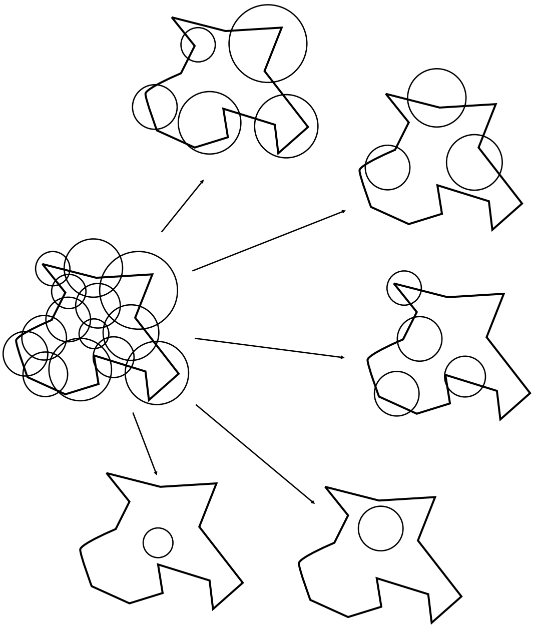

Fig. 4.2. Besicovitch's covering theorem.

Suppose that $q > \widetilde{K}(N) + 1$, where $\widetilde{K}(N)$ is as in the lemma. Let B_1^q be the first ball in the family $\mathcal{B}_q$. The closure of that ball must have intersected the closure of a ball in each of the preceding families (in case there are several such balls in a family, we consider the ball chosen earliest); by our selection procedure, each of those balls must have been at least as large in radius as B_1^q. Thus B_1^q is a ball with at least $\widetilde{K}(N)+1$ "neighbors" as in the lemma. But the lemma says that a ball can have only $\widetilde{K}(N)$ such neighbors. That is a contradiction.

We conclude that $q \le \widetilde{K}(N)+1$. That proves the theorem when M is finite.

When $M=\infty$, recursive application of the above iterative procedure completes the proof of the theorem. We argue as follows:

Suppose that for each $M = 1, 2, \ldots$, the iterative procedure above is carried out for the set of balls $\{B_j\}_{j=1}^M$ resulting in the families of balls $\mathcal{B}_{M,i_1}$, $1 \le i_1 \le \widetilde{K}(N)+1$.

There must be a particular i_1 with $1 \le i_1 \le \widetilde{K}(N)+1$ such that the ball B_1 is assigned to $\mathcal{B}_{M,i_1}$ for infinitely many values of M. We assign B_1 to a family that we label $\mathcal{B}_{i_1}$.

Let $M_{1,1}$ be the smallest value of M for which B_1 is assigned to $\mathcal{B}_{M,i_1}$. Proceeding inductively, we assume that $M_{1,1} < M_{1,2} < \cdots < M_{1,\ell}$ have been defined. Let $M_{1,\ell+1}$ be the smallest value of M that is greater than $M_{1,\ell}$ and is such that B_1 is assigned to $\mathcal{B}_{M,i_1}$. Thus we define the increasing sequence $M_{1,\ell}$, $\ell = 1, 2, \ldots$, with the property that B_1 is assigned to $\mathcal{B}_{M,i_1}$ when our procedure is carried out with $M = M_{1,\ell}$.

There must be a particular i_2 with $1 \le i_2 \le \widetilde{K}(N)+1$ such that the ball B_2 is assigned to $\mathcal{B}_{M,i_2}$ for infinitely many $M \in \{M_{1,1}, M_{1,2}, \ldots\}$. If $i_2 = i_1$ holds, then we assign B_2 to the family $\mathcal{B}_{i_1}$ that already contains B_1. In this case, we see that the closures of B_1 and B_2 do not intersect because there is an $M = M_{1,\ell}$ for which $B_1, B_2 \in \mathcal{B}_{M,i_1} = \mathcal{B}_{M,i_2}$ (in fact, there are infinitely many such M's). On the other hand, if $i_2 \ne i_1$, then we assign B_2 to a new family that we label $\mathcal{B}_{i_2}$.

Let $M_{2,1}$ be the smallest $M \in \{M_{1,1}, M_{1,2}, \ldots\}$ for which B_2 is assigned to $\mathcal{B}_{M,i_2}$. Proceeding inductively, we assume that $M_{2,1} < M_{2,2} < \cdots < M_{2,\ell}$ have been defined. Let $M_{2,\ell+1}$ be the smallest $M \in \{M_{1,1}, M_{1,2}, \ldots\}$ that is greater than $M_{2,\ell}$ and is such that B_2 is assigned to $\mathcal{B}_{M,i_2}$. Thus we define the increasing sequence $M_{2,\ell}$, $\ell = 1, 2, \ldots$, that is a subsequence of $\{M_{1,p}\}_{p=1}^\infty$ and has the property that B_2 is assigned to $\mathcal{B}_{M,i_2}$ when our procedure is carried out with $M = M_{2,\ell}$.

Continuing in this way we assign each ball B_p to one of the families $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_{\widetilde{K}(N)+1}$. □

Remark 4.2.6. Note that there do not exist uncountable families of balls none of which contains the center of any of the other balls. That is because shrinking each ball by a factor of one-half—while keeping the same centers—makes the balls pairwise disjoint.

The next lemma show us one situation in which we can construct a covering of a set by a family of open balls with the property that no ball contains the center of any other ball.

Lemma 4.2.7. *Let $\mathcal{B}$ be a family of open balls centered at points of a compact set A. Suppose $\mathcal{B}$ is such that*

(1) *every point of A is the center of at least one ball in $\mathcal{B}$,*
(2) $\sup\{r : \mathbb{B}(x,r) \in \mathcal{B}\} < \infty$,

(3) $\{\mathbb{B}(x_i, r_i)\}_{i=1}^{\infty} \subseteq \mathcal{B}$ *with* $x_i \to x$ *and* $r_i \to r > 0$ *implies* $\mathbb{B}(x, r) \in \mathcal{B}$.

Then there are finitely many balls $\mathbb{B}(x_i, r_i) \in \mathcal{B}$, $i = 1, 2, \ldots, n$, *such that* $x_i \notin \mathbb{B}(x_j, r_j)$ *whenever* $i \neq j$ *and* $A \subseteq \bigcup_{i=1}^{n} \mathbb{B}(x_i, r_i)$.

Proof. Let $\mathbb{B}(x_1, r_1) \in \mathcal{B}$ be such that r_1 is maximal. Inductively we define $\mathbb{B}(x_{n+1}, r_{n+1})$ to be such that $x_{n+1} \in A \setminus \bigcup_{i=1}^{n} \mathbb{B}(x_i, r_i)$ and r_{n+1} is maximal. If $A \setminus \bigcup_{i=1}^{n} \mathbb{B}(x_i, r_i) = \emptyset$, the construction terminates and we do not define x_{n+1}.

Our construction ensures that we have $x_i \notin \mathbb{B}(x_j, r_j)$ whenever $i \neq j$. We claim that the construction terminates after finitely many steps. To see this fact, we argue by contradiction. Thus we suppose that $\mathbb{B}(x_i, r_i)$ has been defined for $i = 1, 2, \ldots$. Since the balls $\mathbb{B}(x_i, r_i/2)$ are disjoint and all lie in a bounded set, we see that $r_i \downarrow 0$, as $i \to \infty$.

Because A is compact and $\emptyset \neq A \setminus \bigcup_{i=1}^{n} \mathbb{B}(x_i, r_i)$ holds for each n, we see that there is $x \in A \setminus \bigcup_{i=1}^{\infty} \mathbb{B}(x_i, r_i)$. Let $\mathbb{B}(x, r) \in \mathcal{B}$. Since r_i is a nonincreasing sequence with limit 0, there must be an i such that $r_{i+1} < r \leq r_i$, but then we see that $\mathbb{B}(x_{i+1}, r_{i+1})$ was incorrectly chosen. □

Sometimes the requirement that no ball can contain the center of any other ball is too restrictive. In that case the condition we give next may be useful.

Definition 4.2.8. By a *controlled family of balls* we mean a family $\mathcal{B}$ of closed balls with positive radii such that if $\overline{\mathbb{B}}(a, r) \in \mathcal{B}$, $\overline{\mathbb{B}}(b, s) \in \mathcal{B}$, and $\overline{\mathbb{B}}(a, r) \neq \overline{\mathbb{B}}(b, s)$, then

$$\text{either} \qquad |a - b| > r > 4s/5 \qquad \text{or} \qquad |a - b| > s > 4r/5.$$

The next lemma tells us that if we shrink the balls in a controlled family by a factor of one-third, the balls become disjoint. Of course, that also implies that there are no uncountable controlled families.

Lemma 4.2.9. *If* $\overline{\mathbb{B}}(a, r)$ *and* $\overline{\mathbb{B}}(b, s)$ *are members of a controlled family, then* $\overline{\mathbb{B}}(a, r/3) \cap \overline{\mathbb{B}}(b, s/3) = \emptyset$.

Proof. We may assume without loss of generality that

$$|a - b| > r > 4s/5\,.$$

Suppose $p \in \overline{\mathbb{B}}(a, r/3) \cap \overline{\mathbb{B}}(b, s/3)$. Then we have

$$|a - b| \leq |a - p| + |p - b| \leq r/3 + s/3 \leq r/3 + (5/4) \cdot s/3 = 3r/4\,,$$

a contradiction. □

The geometric lemma applicable to balls in a controlled family is given next.

Lemma 4.2.10. *If* $\overline{\mathbb{B}}(a, r)$ *and* $\overline{\mathbb{B}}(b, s)$ *are members of a controlled family and if additionally*

$$4 \leq r \leq |a| \leq r + 1\,,$$

$$4 \leq s \leq |b| \leq s + 1\,,$$

then the angle between $a/|a|$ *and* $b/|b|$ *is at least* $\cos^{-1}(7/8)$.

Proof. Let θ denote the angle between $a/|a|$ and $b/|b|$. Since the balls are members of a controlled family, we may suppose without loss of generality that

$$|a - b| > r > 4s/5\,.$$

Using the law of cosines, we see that

$$\cos\theta = \frac{|a|^2 + |b|^2 - |a-b|^2}{2\,|a|\,|b|} = \frac{|a|}{2\,|b|} + \frac{|b|}{2\,|a|} - \frac{|a-b|^2}{2\,|a|\,|b|}$$

$$\leq \frac{r+1}{2s} + \frac{s+1}{2r} - \frac{r^2}{2rs} = \frac{1}{2s} + \frac{s}{2r} + \frac{1}{2r} \leq \frac{1}{8} + \frac{5}{8} + \frac{1}{8}\,. \qquad \square$$

As before, we have a bound, depending only on the dimension, for how many balls in a controlled family can intersect one particular ball.

Lemma 4.2.11. *There is a constant $K = K(N)$, depending only on the dimension of our space $\mathbb{R}^N$, with the following property: Let $B_0 = \overline{\mathbb{B}}(x_0, r_0)$ be a ball of fixed positive radius. Let $B_1 = \overline{\mathbb{B}}(x_1, r_1)$, $B_2 = \overline{\mathbb{B}}(x_2, r_2)$, $\ldots$, $B_p = \overline{\mathbb{B}}(x_p, r_p)$ be balls such that*

(1) *Each B_j has nonempty intersection with B_0, $j = 1, \ldots, p$;*
(2) *The radii $r_j \geq r_0$ for all $j = 1, \ldots, p$;*
(3) *The balls $\{B_j\}_{j=0}^p$ are members of a controlled family.*

Then $p \leq K$.

Proof. Without loss of generality we may suppose that $x_0 = 0$ and $r_0 = 1$. Divide the balls $B_1, B_2, \ldots, B_p$ into two collections:

$$\mathcal{B}_1 = \{\, B_j : 4 \leq r_j \leq |x_j| \leq r_j + 1 \,\}$$

and

$$\mathcal{B}_2 = \{B_j\}_{j=0}^p \Big\backslash\, \mathcal{B}_1\,.$$

By Lemma 4.2.10, the number of balls in $\mathcal{B}_1$ can be bounded by a number depending only on N. So our task is to bound the number of balls in $\mathcal{B}_2$.

We claim that

$$\bigcup_{B \in \mathcal{B}_2} B \subseteq \overline{\mathbb{B}}(0, 9)\,.$$

Observe that $|x_j| \leq r_j + 1$ holds for every j because $B_0 \cap B_j \neq \emptyset$. Thus

$$\mathcal{B}_2 = \{\, B_j : r_j < 4 \text{ or } |x_j| < r_j \,\}\,.$$

In case $r_j < 4$ holds, we have $|x_j| + r_j \leq 2r_j + 1 < 9$. Also, if $|x_j| < r_j$ and $j \neq 0$, then, because the balls are members of a controlled family, we have $|x_j| > 1 > 4r_j/5$, which yields $|x_j| + r_j < 2r_j < 5/2$.

Since the balls in $\left\{\overline{\mathbb{B}}(x_j, r_j/3)\right\}_{j=0}^p$ are pairwise disjoint (by Lemma 4.2.9) and since $r_j \geq 1$ holds for all the balls in $\mathcal{B}_2$, we see that $\mathcal{B}_2$ contains no more than $9^N/(1/3)^N = 3^{3N}$ balls. $\square$

Theorem 4.2.12. *Let N be a positive integer. There is a constant $K = K(N)$ with the following property. Given a set $A \subseteq \mathbb{R}^N$, a positive finite number R, and a family $\mathcal{B}$ of closed balls of positive radius not exceeding R, if every point of A is the center of at least one ball in $\mathcal{B}$, then there exist $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_K$ such that*

$$A \subseteq \bigcup_{j=1}^{K} \bigcup_{B \in \mathcal{B}_j} B \,,$$

and for each j, the balls in $\mathcal{B}_j$ are pairwise disjoint.

Proof. Enlarge A, if necessary, so that it contains *all* centers of balls in $\mathcal{B}$. It will certainly suffice to prove the result for this possibly larger set, which we will continue to denote by A.

If we construct a controlled family $\mathcal{B}' \subseteq \mathcal{B}$ with

$$A \subseteq \bigcup_{B \in \mathcal{B}'} B \,, \tag{4.3}$$

then we can obtain the desired conclusion by applying the argument used in the proof of Theorem 4.2.1, but with the role of Lemma 4.2.2 filled by Lemma 4.2.11.

We proceed to construct such a controlled family. To this end, we consider the class Ξ of all controlled subfamilies $\mathcal{B}'$ of $\mathcal{B}$ that also satisfy the condition that for any $\overline{\mathbb{B}}(y, s) \in \mathcal{B}$,

$$\left.\begin{array}{lll} \text{either} & |x - y| \le r & \text{holds for some } \overline{\mathbb{B}}(x, r) \in \mathcal{B}', \\ \text{or} & |x - y| > r > 4s/5 & \text{holds for every } \overline{\mathbb{B}}(x, r) \in \mathcal{B}'. \end{array}\right\} \tag{4.4}$$

We note that $\emptyset \in \Xi$, and we partially order Ξ using the relation $\subseteq$. It is easy to see that the union of any subclass of Ξ that is linearly ordered by $\subseteq$ is itself an element of Ξ. Therefore Zorn's lemma[5] tells us that Ξ has a maximal element $\mathcal{B}'$. It remains to verify that $\mathcal{B}'$ satisfies (4.3).

If $\mathcal{B}'$ does not satisfy (4.3), then

$$Y = \{y \in A : |y - x| > r \ \text{ holds for all } \ \overline{\mathbb{B}}(x, r) \in \mathcal{B}'\} \ne \emptyset \,.$$

Select $\overline{\mathbb{B}}(y^*, s^*)$ such that $y^* \in Y$ and

$$s^* > (4/5) \cdot \sup\{s : \exists y \in Y \ \text{ such that } \ \overline{\mathbb{B}}(y, s) \in \mathcal{B}\} \tag{4.5}$$

(this is where we use the fact that the radii of the balls are bounded by $R < \infty$). We will now show that $\mathcal{B}'' = \mathcal{B}' \cup \{\, \overline{\mathbb{B}}(y^*, s^*) \,\}$ is controlled and satisfies the condition (4.4).

To see that $\mathcal{B}''$ is controlled, we need only consider $\overline{\mathbb{B}}(x, r) \in \mathcal{B}'$ and $\overline{\mathbb{B}}(y^*, s^*)$. Since $y^* \in Y$, (4.4) tells us that $|x - y^*| > r > 4s^*/5$, verifying that $\mathcal{B}''$ is controlled.

[5] Max August Zorn (1906–1993).

To check that $\mathcal{B}''$ satisfies (4.4), we consider an arbitrary $\overline{\mathbb{B}}(y, s) \in \mathcal{B}$. If there already exists a $\overline{\mathbb{B}}(x, r) \in \mathcal{B}'$ for which $|x - y| \leq r$ holds, then (4.4) is satisfied. On the other hand, if $|x - y| > r$ holds for every $\overline{\mathbb{B}}(x, r) \in \mathcal{B}'$, then $y \in Y$. We consider $\overline{\mathbb{B}}(y^*, s^*)$. If $|y - y^*| \leq s^*$, then again (4.4) holds. Finally, we have the case in which $|y - y^*| > s^*$ holds. But now we also have $s^* > 4s/5$ by (4.5) and again (4.4) holds.

We have shown that $\mathcal{B}'' \in \Xi$ and we know that $\mathcal{B}'$ is a proper subset of $\mathcal{B}''$. This contradicts the maximality of $\mathcal{B}'$, so we conclude that in fact (4.3) is satisfied. □

Recall the notion of a Radon measure from Definition 1.2.11 in Section 1.2. Using the Besicovitch covering theorem instead of Wiener's covering lemma, we can prove a result like Vitali's (Proposition 4.1.2) for more general Radon measures:

Proposition 4.2.13. *Let μ be a Radon measure on $\mathbb{R}^N$. Let $A \subseteq \mathbb{R}^N$ and let $\mathcal{B}$ be a family of closed balls, with positive radius, such that each point of A is the center of arbitrarily small balls in $\mathcal{B}$. Then there are disjoint balls $B_j \in \mathcal{B}$ such that*

$$\mu\left(A \setminus \bigcup_j B_j\right) = 0 .$$

Proof. We shall follow the same proof strategy as for Proposition 4.1.2. We may as well suppose that $\mu(A) > 0$; otherwise, there is nothing to prove. We also suppose (as we have done in the past) that A is bounded. Let K be as in Theorem 4.2.1.

Let U be a bounded open set with $A \subseteq U$ and choose a compact set C such that $C \subseteq U$ and $\mu(A \cap C) \geq (1/2)\,\mu(A)$. We define $\widetilde{\mathcal{B}}$ to be the family of balls in $\mathcal{B}$ that are centered in $A \cap C$ and contained in U.

By Theorem 4.2.1, we obtain subfamilies $\widetilde{\mathcal{B}}_1, \widetilde{\mathcal{B}}_2, \ldots, \widetilde{\mathcal{B}}_K$ such that each $\widetilde{\mathcal{B}}_j$ is a collection of balls that are pairwise disjoint. We have

$$A \cap C \subseteq \bigcup_{j=1}^{K} \bigcup_{B \in \widetilde{\mathcal{B}}_j} B .$$

Now it is clear that

$$\mu(A \cap C) \leq \sum_{j=1}^{K} \mu\left(\bigcup_{B \in \widetilde{\mathcal{B}}_j} (A \cap B)\right) .$$

Hence there is a particular index j_0 such that

$$\mu(A \cap C) \leq K \cdot \mu\left(\bigcup_{B \in \widetilde{\mathcal{B}}_{j_0}} (A \cap B)\right) .$$

We have

$$\mu(A) \le 2\,\mu(A \cap C) \le 2K \cdot \mu\left(\bigcup_{B \in \widetilde{\mathcal{B}}_{j_0}} (A \cap B)\right).$$

We can choose a finite subfamily $\widehat{\mathcal{B}} \subseteq \widetilde{\mathcal{B}}_{j_0}$ such that

$$\mu(A) \le 3K \cdot \mu\left(\bigcup_{B \in \widehat{\mathcal{B}}} (A \cap B)\right).$$

So setting

$$A_1 = A \setminus \bigcup_{B \in \widehat{\mathcal{B}}} B\,,$$

we conclude that

$$\mu(A_1) \le \mu(A)\,[1 - 1/(3K)]$$

and that A_1 is contained in the bounded open set $U_1 = U \setminus \bigcup_{B \in \widehat{\mathcal{B}}} B$. Now we simply iterate the construction, just as in the proof of Proposition 4.1.2.

We may dispense with the hypothesis that A is bounded just as in the proof of Proposition 4.1.2—making the additional observation that, since the Radon measure μ is σ-finite, it can measure at most countably many hyperplanes parallel to the axes with positive measure (so that we can avoid them when we chop up space into cubes). □

4.3 Decomposition and Differentiation of Measures

Next we turn to differentiation theorems for measures. These are useful in geometric measure theory and also in the theory of singularities for partial differential equations.

Suppose that μ and λ are Radon measures on $\mathbb{R}^N$. We define the *upper derivate* of μ with respect to λ at a point $x \in \mathbb{R}^N$ to be

$$\overline{D}_\lambda(\mu, x) \equiv \limsup_{r \downarrow 0} \frac{\mu[\mathbb{B}(x,r)]}{\lambda[\mathbb{B}(x,r)]}$$

and the *lower derivate* of μ with respect to λ at a point $x \in \mathbb{R}^N$ to be

$$\underline{D}_\lambda(\mu, x) \equiv \liminf_{r \downarrow 0} \frac{\mu[\mathbb{B}(x,r)]}{\lambda[\mathbb{B}(x,r)]}.$$

At a point x where the upper and lower derivates are equal, we define the *derivative of* μ *by* λ to be

$$D_\lambda(\mu, x) = \overline{D}_\lambda(\mu, x) = \underline{D}_\lambda(\mu, x).$$

Remark 4.3.1. It is convenient when calculating these derivates to declare $0/0 = 0$ (this is analogous to other customs in measure theory). The derivates that we have defined are Borel functions. To see this, first observe that $x \mapsto \mu[\mathbb{B}(x,r)]$ is continuous. This is in fact immediate from Lebesgue's dominated convergence theorem. Next notice that our definition of the three derivates does not change if we restrict r to lie in the positive rationals. Since, for each fixed r, the function

$$x \longmapsto \frac{\mu[\mathbb{B}(x,r)]}{\lambda[\mathbb{B}(x,r)]}$$

is continuous, and since the supremum and infimum of a countable family of Borel functions is Borel, we are done.

Definition 4.3.2. Let μ and λ be measures on $\mathbb{R}^N$. We say that μ is *absolutely continuous* with respect to λ if, for $A \subseteq \mathbb{R}^N$,

$$\lambda(A) = 0 \quad \text{implies} \quad \mu(A) = 0\,.$$

It is common to denote this relation by $\mu \ll \lambda$.

Our next result will require the following lemma:

Lemma 4.3.3. *Let μ and λ be Radon measures on $\mathbb{R}^N$. Let $0 < t < \infty$ and suppose that $A \subseteq \mathbb{R}^N$.*

(1) *If $\underline{D}_\lambda(\mu, x) \le t$ for all $x \in A$ then $\mu(A) \le t\lambda(A)$.*
(2) *If $\overline{D}_\lambda(\mu, x) \ge t$ for all $x \in A$ then $\mu(A) \ge t\lambda(A)$.*

Proof. If $\epsilon > 0$ then the Radon property gives us an open set U such that $A \subseteq U$ and $\lambda(U) \le \lambda(A) + \epsilon$. Then the Vitali theorem for Radon measures (Proposition 4.2.13) gives disjoint closed balls $B_j \subseteq U$ such that

$$\mu(B_j) \le (t+\epsilon)\lambda(B_j) \quad \text{(provided the balls are sufficiently small)}$$

and

$$\mu\left(A \setminus \bigcup_j B_j\right) = 0\,.$$

We conclude that

$$\mu(A) \le \sum_j \mu(B_j) \le (t+\epsilon)\sum_j \lambda(B_j)$$

$$\le (t+\epsilon)\lambda(U) \le (t+\epsilon)(\lambda(A)+\epsilon)\,.$$

Letting $\epsilon \to 0$ yields $\mu(A) \le t \cdot \lambda(A)$. This is assertion (1). Assertion (2) may be established in just the same way. □

Theorem 4.3.4. *Suppose that μ and λ are Radon measures on $\mathbb{R}^N$.*

(1) *The derivative* $D_\lambda(\mu, x)$ *exists and is finite* λ*-almost everywhere.*
(2) *For any Borel set* $B \subseteq \mathbb{R}^N$,

$$\int_B D_\lambda(\mu, x)\, d\lambda(x) \le \mu(B)\,,$$

with equality if $\mu \ll \lambda$.
(3) *The relation* $\mu \ll \lambda$ *holds if and only if* $\underline{D}_\lambda(\mu, x) < \infty$ *for* μ*-almost all* $x \in \mathbb{R}^N$.

Proof.
(1) Let $0 < r < \infty$ and $0 < s < t < \infty$. Define

$$A_{s,t}(r) = \{x \in \mathbb{B}(0, r) : \underline{D}_\lambda(\mu, x) \le s < t \le \overline{D}_\lambda(\mu, x)\}$$

and

$$A_t(r) = \{x \in \mathbb{B}(0, r) : \overline{D}_\lambda(\mu, x) \ge t\}\,.$$

Now Lemma 4.3.3 implies that

$$t \cdot \lambda(A_{s,t}(r)) \le \mu(A_{s,t}(r)) \le s \cdot \lambda(A_{s,t}(r)) < \infty$$

and, for $u > 0$,

$$u \cdot \lambda(A_u(r)) \le \mu(A_u(r)) \le \mu[\mathbb{B}(0, r)] < \infty\,.$$

Since $s < t$, these inequalities imply that $\lambda(A_{s,t}(r)) = 0$ and $\lambda(\bigcap_{u>0} A_u(r)) = \lim_{u\to\infty} \lambda(A_u(r)) = 0$. But

$$\begin{aligned} &\mathbb{R}^N \setminus \{x \in \mathbb{R}^N : D_\lambda(\mu, x) \text{ exists and is finite}\} \\ &\quad = \bigcup_{r\in\mathbb{N}} \bigcup_{\substack{0<s<t \\ s,t\in\mathbb{Q}}} A_{s,t}(r) \;\cup\; \bigcup_{r\in\mathbb{N}} \bigcap_{u>0} A_u(r)\,. \end{aligned} \tag{4.6}$$

We see then that the set in (4.6) has λ-measure 0, and this proves assertion (1).

(2) For $1 < t < \infty$ and $p = 0, \pm 1, \pm 2, \ldots,$ we define

$$B_p = \{x \in B : t^p \le D_\lambda(\mu, x) < t^{p+1}\}\,.$$

Then part (1) above and Lemma 4.3.3(2) yield that

$$\begin{aligned} \int_B D_\lambda(\mu, x)\, d\lambda(x) &= \sum_{k=-\infty}^{\infty} \int_{B_k} D_\lambda(\mu, x)\, d\lambda(x) \\ &\le \sum_{k=-\infty}^{\infty} t^{k+1}\, \lambda(B_k) \\ &\le t \cdot \sum_{k=-\infty}^{\infty} \mu(B_k) \\ &\le t \cdot \mu(B)\,. \end{aligned}$$

Letting $t \downarrow 1$ yields then $\int_B D_\lambda(\mu, x)\, d\lambda(x) \le \mu(B)$.

Suppose now that $\mu \ll \lambda$. Then the sets of λ-measure 0 are of course also sets of μ-measure zero. Part (1) tells us that $D_\lambda(\mu, x) = 1/D_\mu(\lambda, x) > 0$ for μ-almost every x. We conclude that $\mu(B) = \sum_{k=-\infty}^{\infty} \mu(B_k)$, and an argument similar to the one just given (using Lemma 4.3.3(2)) gives the inequality $\int_B D_\lambda(\mu, x)\, d\lambda(x) \ge \mu(B)$.

(3) By (1), we know that $\underline{D}_\lambda(\mu, x) < \infty$ at λ-almost every x; if $\mu \ll \lambda$ then this also holds at μ-almost every x.

For the reverse direction in (3), assume that $\underline{D}_\lambda(\mu, x) < \infty$ for μ-almost all $x \in \mathbb{R}^N$. Take $A \subseteq \mathbb{R}^N$ with $\lambda(A) = 0$. For $u = 1, 2, \ldots$, Lemma 4.3.3(2) gives

$$\mu\big(\{x \in A : \underline{D}_\lambda(\mu, x) \le u\} \le u \cdot \lambda(A) = 0\,.$$

We conclude that $\mu(A) = 0$. □

Now we reach our first goal, which is a density theorem and a theorem on the differentiation of integrals for Radon measures.

Theorem 4.3.5. *Let λ be a Radon measure on $\mathbb{R}^N$.*

(1) *If $A \subseteq \mathbb{R}^N$ is λ-measurable then the limit*

$$\lim_{r\downarrow 0} \frac{\lambda(A \cap \mathbb{B}(x, r))}{\lambda[\mathbb{B}(x, r)]}$$

exists and equals 1 *for λ-almost every $x \in A$ and equals* 0 *for λ-almost every $x \in \mathbb{R}^N \setminus A$.*

(2) *If $f : \mathbb{R}^N \to \overline{\mathbb{R}}$ is locally λ-integrable, then*

$$\lim_{r\downarrow 0} \frac{1}{\lambda[\mathbb{B}(x, r)]} \int_{\mathbb{B}(x,r)} f(x)\, d\lambda(x) = f(x)$$

for λ-almost every $x \in \mathbb{R}^N$.

Proof. Part (1) follows from part (2) by setting $f = \chi_A$. To prove (2), we may take $f \ge 0$. Define $\mu(A) = \int_A f(x)\, d\lambda(x)$. Then μ is a Radon measure and $\mu \ll \lambda$. Theorem 4.3.4(2) now yields that

$$\int_E D_\lambda(\mu, x)\, d\lambda(x) = \mu(E) = \int_E f\, d\lambda$$

for all Borel sets E. This clearly entails $f(x) = D_\lambda(\mu, x)$ for λ-almost all $x \in \mathbb{R}^N$. That proves (2). □

We say that two Radon measures μ and λ are *mutually singular* if there is a set $A \subseteq \mathbb{R}^N$ such that $\lambda(A) = 0 = \mu(\mathbb{R}^N \setminus A)$. Now we have a version of the Radon–Nikodym theorem combined with the Lebesgue decomposition.

Theorem 4.3.6. *Suppose that λ and μ are finite Radon measures on $\mathbb{R}^N$. Then there is a Borel function f and a Radon measure ν such that λ and ν are mutually singular and*

$$\mu(E) = \int_E f\, d\lambda + \nu(E)$$

for any Borel set $E \subseteq \mathbb{R}^N$. Furthermore, $\mu \ll \lambda$ if and only if $\nu = 0$.

Proof. Define

$$A = \{x \in \mathbb{R}^N : \underline{D}_\lambda(\mu, x) < \infty\}\,.$$

Recalling that $\lfloor$ denotes the restriction of a measure, we set

$$\mu_1 = \mu \lfloor A \quad \text{and} \quad \nu = \mu \lfloor (\mathbb{R}^N \setminus A)\,.$$

Then obviously $\mu = \mu_1 + \nu$, and λ and ν are mutually singular by Theorem 4.3.4(1). Now Lemma 4.3.3(1) gives $\mu_1 \ll \lambda$; hence μ_1 has the required representation by Theorem 4.3.4(2) with $f(x) = D_\lambda(\mu, x)$. The last statement of the theorem is now obvious. □

We conclude this section with some results concerning densities of measures (see Definition 2.2.1).

Theorem 4.3.7. *Fix $0 < t$. If μ is a Borel regular measure on $\mathbb{R}^N$ and $A \subseteq C \subseteq \mathbb{R}^N$, then*

$$t \le \Theta^{*M}(\mu \lfloor C, x)\,, \textit{ for all } x \in A, \quad \textit{implies} \quad t \cdot \mathcal{S}^M(A) \le \mu(C)\,.$$

Remark 4.3.8. Since spherical measure is always at least as large as Hausdorff measure, we also have the conclusion

$$t \le \Theta^{*M}(\mu \lfloor C, x)\,, \text{ for all } x \in A, \quad \text{implies} \quad t \cdot \mathcal{H}^M(A) \le \mu(C)\,.$$

Proof. Without loss of generality, we may assume that $\mu(C) < \infty$. It will also be sufficient to prove that $t < \Theta^{*M}(\mu \lfloor C, x)$, for all $x \in A$, implies $t \cdot \mathcal{S}^M(A) \le \mu(C)$.

Fix $0 < \delta$. We will estimate the approximating measure $\mathcal{S}^M_{6\delta}(A)$. This estimation will require a special type of covering, which we construct next.

Set

$$\mathcal{B} = \{\, \overline{\mathbb{B}}(x, r) : x \in A, \quad 0 < r \le \delta, \quad t \cdot \Omega_M \cdot r^M \le (\mu \lfloor C)\overline{\mathbb{B}}(x, r)\,\}\,,$$

$$\mathcal{B}_1 = \{\, \overline{\mathbb{B}}(x, r) \in \mathcal{B} : \quad 2^{-1}\delta < r \le \delta\,\}\,,$$

and let $\mathcal{B}_1'$ be a maximal pairwise disjoint subfamily of $\mathcal{B}_1$.

Assuming that $\mathcal{B}_1', \mathcal{B}_2', \ldots, \mathcal{B}_k'$ have already been defined, set

$$\mathcal{B}_{j+1} = \left\{ \overline{\mathbb{B}}(x, r) \in \mathcal{B} :\ 2^{-(j+1)}\delta < r \le 2^{-j}\delta,\ \ \emptyset = \overline{\mathbb{B}}(x, r) \bigcap \bigcup_{i=1}^{j} \bigcup_{B \in \mathcal{B}_i'} B \right\},$$

and let $\mathcal{B}'_{j+1}$ be a maximal pairwise disjoint subfamily of $\mathcal{B}_{j+1}$.

Note that the assumption $\mu(C) < \infty$ ensures that each $\mathcal{B}'_i$ is finite. Also note that, by construction, any two closed balls in the family $\bigcup_{i=1}^{\infty} \mathcal{B}'_i$ are disjoint, so we have

$$\sum_{i=1}^{\infty} \sum_{B \in \mathcal{B}'_i} (\mu \llcorner C)(B) = (\mu \llcorner C) \left(\bigcup_{i=1}^{\infty} \bigcup_{B \in \mathcal{B}'_i} B \right) \leq \mu(C) < \infty . \tag{4.7}$$

Claim. For each n,

$$A \subseteq \left(\bigcup_{i=1}^{n} \bigcup_{B \in \mathcal{B}'_i} B \right) \cup \left(\bigcup_{i=n+1}^{\infty} \bigcup_{B \in \mathcal{B}'_i} \widehat{B} \right) \tag{4.8}$$

holds, where, for each ball $B = \overline{\mathbb{B}}(x, r)$, we set $\widehat{B} = \overline{\mathbb{B}}(x, 3r)$.

To verify the claim, consider $x \notin \bigcup_{i=1}^{n} \bigcup_{B \in \mathcal{B}_i} B$. Since $\bigcup_{i=1}^{n} \bigcup_{B \in \mathcal{B}_i} B$ is closed, there is $\overline{\mathbb{B}}(x, r) \in \mathcal{B}$ such that

$$\emptyset = \overline{\mathbb{B}}(x, r) \bigcap \bigcup_{i=1}^{n} \bigcup_{B \in \mathcal{B}'_i} B .$$

Letting k be such that $2^{-k} < r \leq 2^{-(k-1)}$, we see that if $k > n$ and $\overline{\mathbb{B}}(x, r) \notin \mathcal{B}'_k$, then

$$\emptyset \neq \overline{\mathbb{B}}(x, r) \bigcap \bigcup_{i=n+1}^{k} \bigcup_{B \in \mathcal{B}'_i} B .$$

Thus there is $\overline{\mathbb{B}}(y, t) \in \mathcal{B}'_i$, where $n + 1 \leq i \leq k$, such that $\emptyset \neq \overline{\mathbb{B}}(x, r) \cap \overline{\mathbb{B}}(y, t)$. Since $r \leq 2^{-(k-1)}$ and $2^{-k} < t$, we have $x \in \overline{\mathbb{B}}(y, r + t) \subseteq \overline{\mathbb{B}}(y, 3t)$. The claim is proved.

Let $\epsilon > 0$ be arbitrary. By (4.7) (see also (4.8)), we choose n such that

$$\sum_{i=1}^{\infty} \sum_{B \in \mathcal{B}'_i} (\mu \llcorner C)(B) < \epsilon .$$

Using the claim and letting rad B denote the radius of the ball B, we estimate

$$\begin{aligned}
\mathcal{S}^M_{6\delta}(A) &\leq \left(\sum_{i=1}^{n} \sum_{B \in \mathcal{B}'_i} \Omega_M \, (\operatorname{rad} B)^M \right) + \left(\sum_{i=n+1}^{\infty} \sum_{B \in \mathcal{B}'_i} \Omega_M \, (\operatorname{rad} \widehat{B})^M \right) \\
&= \left(\sum_{i=1}^{n} \sum_{B \in \mathcal{B}'_i} \Omega_M \, (\operatorname{rad} B)^M \right) + 3^M \left(\sum_{i=n+1}^{\infty} \sum_{B \in \mathcal{B}'_i} \Omega_M \, (\operatorname{rad} B)^M \right) \\
&\leq t^{-1} \left(\sum_{i=1}^{n} \sum_{B \in \mathcal{B}'_i} (\mu \llcorner C) B \right) + 3^M \, t^{-1} \left(\sum_{i=n+1}^{\infty} \sum_{B \in \mathcal{B}'_i} (\mu \llcorner C) B \right) \\
&\leq t^{-1} \left[\mu(C) + 3^M \epsilon \right] .
\end{aligned}$$

Since $\epsilon > 0$ was arbitrary, we conclude that $\mathcal{S}^M_{6\delta}(A) \le t^{-1}\,\mu(C)$. The result follows, since $\delta > 0$ was also arbitrary. □

Corollary 4.3.9. *In $\mathbb{R}^N$, the measures $\mathcal{S}^N$, $\mathcal{H}^N$, $\mathcal{T}^N$, $\mathcal{C}^N$, $\mathcal{G}^N$, $\mathcal{Q}^N_t$, and $\mathcal{I}^N_t$ ($1 \le t \le \infty$) all agree with the N-dimensional Lebesgue measure $\mathcal{L}^N$.*

Proof. Noting that $\beta_t(N, N) = 1$, for $1 \le t \le \infty$, and using Proposition 2.1.5, we see that $\mathcal{S}^N$ is the largest of the measures $\mathcal{S}^N$, $\mathcal{H}^N$, $\mathcal{T}^N$, $\mathcal{C}^N$, $\mathcal{G}^N$, $\mathcal{Q}^N_t$, and $\mathcal{I}^N_t$, while $\mathcal{I}^N_1$ is the smallest. Theorem 4.3.7 implies $\mathcal{S}^N \le \mathcal{L}^N$ and (2.9) gives us $\mathcal{I}^N_1 \ge \mathcal{L}^N$, so the result follows. □

Corollary 4.3.10. *If μ is a Borel regular measure on $\mathbb{R}^N$, $A \subseteq \mathbb{R}^N$ is μ-measurable, and $\mu(A) < \infty$, then*

$$\Theta^{*M}(\mu \llcorner A, x) = 0$$

holds for $\mathcal{S}^M$-almost every $x \in \mathbb{R}^N \setminus A$.

Proof. Let j be a positive integer and set

$$C_j = \left\{ x \in (\mathbb{R}^N \setminus A) : j^{-1} \le \Theta^{*M}(\mu \llcorner A, x) \right\}.$$

Arguing by contradiction, suppose that $\mathcal{S}^M(C_j)$ is positive. Then, by the Borel regularity of μ, we can find a closed set $E \subseteq A$ such that

$$\mu(A \setminus E) < j^{-1} \cdot \mathcal{S}^M(C_j).$$

For $x \in C_j$, since E is closed and $x \notin E$, we have

$$j^{-1} \le \Theta^{*M}(\mu \llcorner A, x) = \Theta^{*M}[\,\mu \llcorner (A \setminus E)\,, x]$$

$$= \Theta^{*M}[\,(\mu \llcorner A) \llcorner (\mathbb{R}^N \setminus E)\,, x\,].$$

So we can apply Theorem 4.3.7 (with the roles of μ, A, and B played by $\mu \llcorner A$, $\mathbb{R}^N \setminus E$, and C_j, respectively), to conclude that

$$t \cdot \mathcal{S}^M(C_j) \le (\mu \llcorner A)(\mathbb{R}^N \setminus E) = \mu(A \setminus E)\,,$$

a contradiction.

Thus we have $\mathcal{S}^M(C_j) = 0$ and the result follows. □

4.4 The Riesz Representation Theorem

In this section, we prove a version of the Riesz representation theorem for linear functionals. Anticipating that our main application of this theorem will be to currents with finite mass, we have taken our linear functionals to be defined on the space of real-valued, infinitely differentiable, compactly supported functions on $\mathbb{R}^N$. Standard versions of the theorem apply to linear functionals on the space of *continuous*, compactly supported functions (see, for example, [Fol 84], [Roy 88], or [Rud 87]). In [EG 92], Evans and Gariepy prove a version of the theorem for linear functionals on the space of *vector-valued*, continuous, compactly supported functions.

Theorem 4.4.1 (Riesz Representation Theorem). *Let $\mathcal{D}$ denote the set of real-valued, infinitely differentiable, compactly supported functions on $\mathbb{R}^N$. If $L : \mathcal{D} \to \mathbb{R}$ is a linear functional satisfying*

$$M = \sup \left\{ |L(\phi)| : \phi \in \mathcal{D}, \ \sup_{x \in \mathbb{R}^N} |\phi| \leq 1 \right\} < \infty , \tag{4.9}$$

then there exists a Radon measure λ on $\mathbb{R}^N$ and a λ-measurable function $g : \mathbb{R}^N \to \mathbb{R}$ such that

(1) $\lambda \left(\mathbb{R}^N \right) = M$,

(2) $L(\phi) = \int_{\mathbb{R}^N} \phi \, g \, d\lambda$, *for all $\phi \in \mathcal{D}$.*

Proof. First, we note that it follows immediately from (4.9) that

$$|L(\phi)| \leq M \cdot \sup_x |\phi(x)| , \ \text{for } \phi \in \mathcal{D} . \tag{4.10}$$

Step 1: Definition of the measure λ. We define the function λ on subsets of $\mathbb{R}^N$ by setting $\lambda(\emptyset) = 0$, setting

$$\lambda(U) = \sup \left\{ |L(\phi)| : \phi \in \mathcal{D}, \ \sup_x |\phi(x)| \leq 1, \ \operatorname{supp} \phi \subseteq U \right\} \tag{4.11}$$

when U is a nonempty open set, and setting

$$\lambda(E) = \inf \{ \lambda(U) : U \text{ is open, } E \subseteq U \} \tag{4.12}$$

when E is not an open set.

Ultimately we will show that λ is a measure. It follows immediately that

$$\lambda(\mathbb{R}^N) = M , \tag{4.13}$$

$$A \subseteq B \ \text{ implies } \ \lambda(A) \leq \lambda(B) . \tag{4.14}$$

To show that μ is a measure, we first show that λ is countably subadditive on the family of open sets. To see this, let U_i, $i = 1, 2, \ldots$, be a sequence of open sets. We need to show that

$$\lambda \Big(\bigcup\nolimits_i U_i \Big) \leq \sum_i \lambda(U_i) \tag{4.15}$$

holds. It is no loss of generality to assume that $\sum_i \lambda(U_i) < \infty$.

Suppose that $\phi \in \mathcal{D}$, $\sup_{x \in \mathbb{R}^N} |\phi| \leq 1$, and $\operatorname{supp} \phi \subseteq \cup_i U_i$. Let α_i be a smooth partition of unity for the set $\operatorname{supp} \phi$, subordinate to the cover $\{U_i\}_{i=1}^{\infty}$ (see [KPk 99]).

We estimate

$$\left|L\left(\sum_{i=m}^{n}\phi\cdot\alpha_i\right)\right| = \left|\sum_{i=m}^{n}L\left(\phi\cdot\alpha_i\right)\right| \le \sum_{i=m}^{n}\left|L\left(\phi\cdot\alpha_i\right)\right| \le \sum_{i=m}^{\infty}\lambda(U_i)\,.$$

Thus $L(\sum_i \phi\cdot\alpha_i)$ and $\sum_i |L(\phi\cdot\alpha_i)|$ are convergent. We then have

$$|L(\phi)| = \left|L\left(\phi\sum_i\alpha_i\right)\right| = \left|L\left(\sum_i\phi\cdot\alpha_i\right)\right| \le \sum_i\left|L\left(\phi\cdot\alpha_i\right)\right| \le \sum_i\lambda(U_i)\,,$$

and (4.15) follows.

To complete the proof that λ is a measure, we show that λ is countably subadditive on the family of all subsets of $\mathbb{R}^N$. To see this, we let E_i, $i = 1, 2, \ldots$, be a sequence of sets. We need to show that $\lambda(\bigcup_i E_i) \le \sum_i \lambda(E_i)$. We may suppose without loss of generality that $\sum_i \lambda(E_i) < \infty$.

Let $\epsilon > 0$ be arbitrary. For each i, let U_i be an open set with $\lambda(U_i) \le \lambda(E_i)+2^{-i}\epsilon$. Then, by (4.15), we have

$$\lambda(\textstyle\bigcup_i E_i) \le \lambda(\bigcup_i U_i) \le \displaystyle\sum_i\lambda(U_i) \le \epsilon + \sum_i\lambda(E_i)\,,$$

and the claim follows from the fact that $\epsilon > 0$ was arbitrary.

Step 2: A bound on L. We claim that

$$|L(\phi)| \le \sup_x|\phi(x)|\cdot\lambda\Big(\{x : \phi(x)\ne 0\}\Big)\,,\ \text{for } \phi\in\mathcal{D}\,. \tag{4.16}$$

To see this, fix a nonzero $\phi \in \mathcal{D}$, set $\kappa = \sup_x |\phi(x)|$, and set

$$U = \{x : \phi(x) \ne 0\}\,.$$

Let $\alpha_\ell : \mathbb{R} \to \mathbb{R}$, $\ell = 1, 2, \ldots$, be a sequence of infinitely differentiable functions such that

$$\begin{aligned} \alpha_\ell(t) = 0 \quad &\text{if} \quad |t| \le 1/(2\ell), \\ |\alpha_\ell(t)| \le 1/\ell \ &\text{if } 1/(2\ell) < |t| < 1/\ell, \\ \alpha_\ell(t) = t \quad &\text{if} \quad 1/\ell \le |t|\,. \end{aligned}$$

For ℓ such that $1/\ell \le \sup_x |\phi(x)|$, we have $\kappa = \sup_x \alpha_\ell\circ\phi(x)$ and

$$\operatorname{supp}\alpha_\ell\circ\phi \subseteq U\,,$$

so

$$|L(\alpha_\ell\circ\phi)| \le \kappa\,\lambda(U)\,.$$

Since $\sup_x |\phi - \alpha_\ell\circ\phi| \le 1/\ell$ holds, we conclude from (4.10) that

$$|L(\phi) - L(\alpha_\ell\circ\phi)| = |L(\phi - \alpha_\ell\circ\phi)| \le M/\ell$$

holds. Letting $\ell \to \infty$, we obtain the claim.

Step 3: Showing that λ is a Radon measure. First, we claim that λ is finitely additive on the family of open sets. To see this, let U and V be disjoint open sets. Let $\epsilon > 0$ be arbitrary. Let $\phi_U \in \mathcal{D}$ satisfy

- $\sup_x |\phi_U(x)| \leq 1$,
- $\operatorname{supp} \phi_U \subseteq U$,
- $\lambda(U) \leq |L(\phi_U)| + \epsilon$.

Replacing ϕ_U by $-\phi_U$ if necessary, we may assume that $L(\phi_U) = |L(\phi_U)|$. Choose $\phi_V \in \mathcal{D}$ similarly. Then we have

$$\begin{aligned}\lambda(U) + \lambda(V) &\leq |L(\phi_U)| + |L(\phi_V)| + 2\epsilon \\ &= L(\phi_U) + L(\phi_V) + 2\epsilon \\ &= L(\phi_U + \phi_V) + 2\epsilon \\ &\leq |L(\phi_U + \phi_V)| + 2\epsilon \leq \lambda(U \cup V) + 2\epsilon\,,\end{aligned}$$

and since $\epsilon > 0$ was arbitrary, the claim follows.

Next, we claim that λ satisfies Carathéodory's criterion. To see this, let A and B be sets that are separated by a positive distance.

Let $\epsilon > 0$ be arbitrary. We can find an open set U with $A \cup B \subseteq U$ and $\lambda(U) \leq \lambda(A \cup B) + \epsilon$. Since A and B are at a positive distance from each other, we may assume without loss of generality that $U = U_A \cup U_B$, where U_A and U_B are disjoint open sets containing A and B, respectively. Then we have

$$\lambda(A) + \lambda(B) \leq \lambda(U_A) + \lambda(U_B) = \lambda(U_A \cup U_B) \leq \lambda(A \cup B) + \epsilon\,,$$

and the claim follows from the fact that $\epsilon > 0$ was arbitrary.

Since λ satisfies Carathéodory's criterion, we know that all open sets are λ-measurable. The fact that λ is a Radon measure follows from (4.12) and the fact that $\lambda(\mathbb{R}^N) < \infty$.

Step 4: Extension of *L*. Let $\overline{\mathcal{D}}$ denote the set of functions $f : \mathbb{R}^N \to \mathbb{R}$ such that f is bounded and f is the pointwise limit of a sequence of functions in $\mathcal{D}$. We observe that

- $\overline{\mathcal{D}}$ contains the characteristic function of any open subset of $\mathbb{R}^N$,
- $\overline{\mathcal{D}}$ is a vector space,
- $\overline{\mathcal{D}}$ is closed under multiplication.

We will define the extension of L from $\mathcal{D}$ to $\overline{\mathcal{D}}$.

Let $f \in \overline{\mathcal{D}}$. Let ϕ_i be a sequence of functions in $\mathcal{D}$ with $f = \lim_i \phi_i$. We may assume without loss of generality that the functions ϕ_i are uniformly bounded.

Set

$$\kappa \equiv \sup_i \sup_x \phi_i(x) < \infty\,.$$

Fix $\epsilon > 0$. For each n, set

$$A_n = \{x : \exists\, i, j \geq n \ \text{ such that } \ |\phi_i(x) - \phi_j(x)| \geq \epsilon\,\}\,.$$

Then we have $A_1 \supseteq A_2 \supseteq \cdots$ and $\cap_n A_n = \emptyset$. So $\lambda(A_n) \downarrow 0$ as $n \to \infty$. Fix an n such that $\lambda(A_n) < \epsilon$.

Let $\beta : \mathbb{R} \to \mathbb{R}$ be an infinitely differentiable function satisfying

- β takes its values in [0, 1],
- $\beta(t) = 1$ if $|t| \geq 2\epsilon$,
- $\beta(t) = 0$ if $|t| < \epsilon$.

For $i, j \geq n$, we have

$$|L(\phi_i - \phi_j)| \leq \Big| L[\,\beta \circ (\phi_i - \phi_j) \cdot (\phi_i - \phi_j)\,] \Big| + \Big| L[\,(1 - \beta \circ (\phi_i - \phi_j)) \cdot (\phi_i - \phi_j) \Big| \leq 2\,(\kappa + M)\,\epsilon\,.$$

Thus we see that $L(\phi_i)$ forms a Cauchy sequence. We define

$$L\Big(\lim_{i\to\infty} \phi_i\Big) = \lim_{i\to\infty} L(\phi_i)\,.$$

It is easy to see that the extension of L is well-defined and linear.

The extension of L satisfies an estimate like (4.16); specifically, we claim that if $f \in \overline{\mathcal{D}}$, then it holds that

$$|L(f)| \leq \sup_x |f(x)| \cdot \lambda\Big(\{x : f(x) \neq 0\}\Big)\,. \tag{4.17}$$

To see this, fix the function $f \in \overline{\mathcal{D}}$ and fix a uniformly bounded sequence $\phi_i \in \mathcal{D}$ that converges pointwise to f. It is no loss of generality to assume that

$$\sup_x |f(x)| = \lim_{i\to\infty} \Big(\sup_x |\phi_i(x)| \Big)\,.$$

Set $W = \{x : f(x) \neq 0\}$.

Let $\epsilon > 0$ be arbitrary. Then we can find an open set U with $W \subseteq U$ and $\lambda(U) \leq \lambda(W) + \epsilon$.

Let $\alpha_\ell : \mathbb{R}^N \to \mathbb{R}$ be a sequence of infinitely differentiable functions with values in [0, 1] such that $\{x : \alpha_\ell(x) = 1\}$ increases to χ_U. Then $\phi_i \cdot \alpha_i$ is a uniformly bounded sequence that converges to f.

We have

$$\begin{aligned} |L(\phi_i \cdot \alpha_i)| &\leq \sup_x |(\phi_i \cdot \alpha_i)(x)| \cdot \lambda\Big(\{x : (\phi_i \cdot \alpha_i)(x) \neq 0\}\Big) \\ &\leq \sup_x |f(x)| \cdot \lambda\Big(\{x : \alpha_i(x) \neq 0\}\Big) \\ &\leq \sup_x |f(x)| \cdot \lambda(U) \\ &\leq \sup_x |f(x)| \cdot (\lambda(W) + \epsilon)\,, \end{aligned}$$

and the claim follows.

Step 5: A family of subsets of $\mathbb{R}^N$. Let $\mathcal{O}$ denote the family of subsets A of $\mathbb{R}^N$ for which $\chi_A \in \overline{\mathcal{D}}$. Since

$$\chi_{A\cap B} = \chi_A\,\chi_B\,,$$

$$\chi_{A\cup B} = \chi_A + \chi_B - \chi_A\,\chi_B\,,$$

$$\chi_{A\setminus B} = (1-\chi_B)\,\chi_A\,,$$

we see that $\mathcal{O}$ is closed under finite unions, finite intersections, and complements. Also every element of $\mathcal{O}$ is a Borel set. Note that

$$L(\chi_U) + \lambda(U) \geq 0$$

holds, for any $U \in \mathcal{O}$.

Step 6: Definition of the measure μ. We define the function μ on subsets of $\mathbb{R}^N$ by setting

$$\mu(U) = L(\chi_U) + \lambda(U)\,, \tag{4.18}$$

when U is open, and setting

$$\mu(E) = \inf\{\,\mu(U) : U \text{ is open},\ E \subseteq U\,\}\,, \tag{4.19}$$

when E is not open.

For sets $U, V \in \mathcal{O}$ with $U \subseteq V$, we have

$$\begin{aligned} L(\chi_V) + \lambda(V) &= L(\chi_U + \chi_{V\setminus U}) + \lambda(\,U \cup (V \setminus U)\,) \\ &= L(\chi_U) + L(\chi_{V\setminus U}) + \lambda(U) + \lambda(V \setminus U) \\ &\geq L(\chi_U) + \lambda(U)\,. \end{aligned}$$

If U and V are open with $U \subseteq V$, then we conclude that $\mu(U) \leq \mu(V)$. Then by (4.19), μ is monotone on all sets.

We claim that

$$\mu(E) = L(\chi_E) + \lambda(E)\,, \quad \text{for } E \in \mathcal{O}\,. \tag{4.20}$$

The argument above also shows that if U is open, $E \in \mathcal{O}$, and $E \subseteq U$, then

$$L(\chi_E) + \lambda(E) \leq \mu(U)\,.$$

Let $\epsilon > 0$ be arbitrary. Then we can find an open U with $E \subseteq U$ and

$$\lambda(U) \leq \lambda(E) + \epsilon\,.$$

Since

$$\lambda(U) = \lambda(U \setminus E) + \lambda(E)\,,$$

we have

$$\lambda(U \setminus E) \le \epsilon \,.$$

By (4.17), we have

$$L(\chi_{U\setminus E}) \le \epsilon \,,$$

so

$$L(\chi_U) + \lambda(U) = L(\chi_E) + \lambda(E) + L(\chi_{U\setminus E}) + \lambda(U \setminus E) \le L(\chi_E) + \lambda(E) + 2\epsilon$$

holds. Thus we have

$$\mu(E) \le L(\chi_E) + \lambda(E) + 2\epsilon \,,$$

and the claim follows from the fact that $\epsilon > 0$ was arbitrary.

By (4.20), we see that we obtain the same function μ on subsets of $\mathbb{R}^N$ if we define μ by setting

$$\mu(U) = L(\chi_U) + \lambda(U) \,, \tag{4.21}$$

when $U \in \mathcal{O}$, and setting

$$\mu(E) = \inf \{ \mu(U) : U \in \mathcal{O}, \ E \subseteq U \} \,, \tag{4.22}$$

when $E \notin \mathcal{O}$. We shall use this alternative definition. Ultimately we will show that μ is a measure. We note that the original definition of μ is useful for verifying that μ is a Radon measure.

By (4.17), we see that

$$0 \le \mu(E) \le 2\lambda(E)$$

holds, for every set E. In particular, μ is absolutely continuous with respect to λ. We also note that if $U, V \in \mathcal{O}$, then

$$\begin{aligned}
\mu(V) &= L(\chi_V) + \lambda(V) \\
&= L(\chi_U + \chi_{V\setminus U}) + \lambda(U \cup (V \setminus U)) \\
&= L(\chi_U) + L(\chi_{V\setminus U}) + \lambda(U) + \lambda(V \setminus U) \\
&\ge L(\chi_U) + \lambda(U) \\
&= \mu(U)
\end{aligned}$$

and

$$\begin{aligned}
\mu(U \cup V) &= L(\chi_{U\cup V}) + \lambda(U \cup V) \\
&= L(\chi_U) + L(\chi_V) - L(\chi_{U\cap V}) + \lambda(U \cup V) \\
&= L(\chi_U) + L(\chi_V) - L(\chi_{U\cap V}) + \lambda(U) + \lambda(V) - \lambda(U \cap V) \\
&= \mu(U) + \mu(V) - \mu(U \cap V) \le \mu(U) + \mu(V) \,,
\end{aligned}$$

so λ is finitely additive and finitely subadditive on $\mathcal{O}$.

Step 7: Showing that μ is a Radon measure. First, we claim that μ is countably subadditive on $\mathcal{O}$. To see this, let a sequence $\{U_i\} \subseteq \mathcal{O}$ be given. We need to show that

$$\mu\Big(\bigcup_i U_i\Big) \leq \sum_i \mu(U_i) \tag{4.23}$$

holds.

Let $\epsilon > 0$ be arbitrary. Set

$$A_n = \big(\bigcup_{i=1}^{\infty} U_i\big) \setminus \big(\bigcup_{i=1}^{n} U_i\big) .$$

Then $\lambda(A_n) \to 0$ as $n \to \infty$. Choose n such that $\lambda(A_n) < \epsilon$. We have

$$\begin{aligned}\mu(\bigcup_{i=1}^{\infty} U_i) &= \mu(\bigcup_{i=1}^{n} U_i) + L(\chi_{A_n}) + \lambda(A_n) \\ &\leq \mu(\bigcup_{i=1}^{n} U_i) + 2\epsilon \leq 2\epsilon + \sum_{i=1}^{\infty} \mu(U_i) ,\end{aligned}$$

and the claim follows from the fact that $\epsilon > 0$ was arbitrary.

We see that μ is countably subadditive by using the same argument that showed that λ is subadditive. We can also see that Carathéodory's criterion holds for μ in the same way that we saw that it holds for λ, and we similarly conclude that λ is a Radon measure.

Step 8: Obtaining the function g. By Theorem 4.3.6, there exists a Borel function f such that

$$\mu(E) = \int_E f \, d\lambda$$

holds, for any Borel set E. Set $g = f - 1$. For $U \in \mathcal{O}$, we have

$$L(\chi_U) = \mu(U) - \lambda(U) = \int_U (f-1) \, d\lambda = \int_U g \, d\lambda .$$

For $\phi \in \mathcal{D}$, we obtain

$$L(\phi) = \int \phi \, g \, d\lambda$$

by uniformly approximating ϕ by simple functions of the form $\sum_i \alpha_i \chi_{E_i}$, with $E_i \in \mathcal{O}$, and applying (4.17). □

4.5 Maximal Functions Redux

It is possible to construe the Hardy–Littlewood maximal function in the more general context of measures.

Definition 4.5.1. Let μ be a Radon measure on $\mathbb{R}^N$. If f is a μ-measurable function and $x \in \mathbb{R}^N$ then we define

$$M_\mu f(x) = \sup_{r>0} \frac{1}{\mu[\mathbb{B}(x,r)]} \int_{\mathbb{B}(x,r)} |f(t)| \, d\mu(t) \, .$$

Further, and more generally, if ν is a Radon measure on $\mathbb{R}^N$ then we define

$$M_\mu \nu(x) = \sup_{r>0} \frac{\nu[\mathbb{B}(x,r)]}{\mu[\mathbb{B}(x,r)]} \, .$$

Finally, it is sometimes useful to have the noncentered maximal operator $\widetilde{M}_\mu$ defined by

$$\widetilde{M}_\mu f(x) = \sup_{\mathbb{B}(z,r) \ni x} \frac{1}{\mu[\mathbb{B}(z,r)]} \int_{\mathbb{B}(z,r)} |f(t)| \, d\mu(t) \, .$$

A similar definition may be given for the maximal function of a Radon measure.

The principal result about these maximal functions is the following:

Theorem 4.5.2. *The operator M_μ is weak type $(1, 1)$ in the sense that*

$$\mu \left\{ x \in \mathbb{R}^N : M_\mu \nu(x) > s \right\} \leq C \cdot \frac{\nu(\mathbb{R}^N)}{s} \, .$$

In particular, if $f \in L^1(\mu)$ then

$$\mu \left\{ x \in \mathbb{R}^N : M_\mu f(x) > s \right\} \leq C \cdot \frac{\|f\|_{L^1}}{s} \, .$$

In case the measure μ satisfies the enlargement condition $\mu[\mathbb{B}(x, 3r)] \leq c \cdot \mu[\mathbb{B}(x, r)]$, then we have

$$\mu \left\{ x \in \mathbb{R}^N : \widetilde{M}_\mu \nu(x) > s \right\} \leq c \cdot s^{-1} \cdot \nu \left\{ x \in \mathbb{R}^N : \widetilde{M}_\mu \nu(x) > s \right\} \, .$$

The proof of this result follows the same lines as the development of Proposition 4.1.4, and we omit the details. A full account may be found in [Mat 95].