2. Covering and differentiation

In the first part of this chapter we prove some covering theorems which are among the most fundamental tools of measure theory. They are used to create connections between local and global properties of measures, and they also reflect the geometry of the space. Covering theorems and their applications have been studied much more extensively in Federer [3], Guzmán [1], and Hayes and Pauc [1]. The presentations of Evans and Gariepy [1], Giusti [1], L. Simon [1] and Ziemer [1] are rather close to ours.

We prove two types of covering theorems. The difference between them is that the first ones apply to a larger class of coverings and a narrower class of measures whereas in the second type the coverings are more restricted but the measures can be very general; for example all Radon measures on $\mathbf{R}^n$ are included. In both cases we first prove a geometric result on collections of balls in $\mathbf{R}^n$ and then apply it to get a Vitali-type covering theorem for measures.

At the end of this chapter we apply these covering theorems to prove some basic differentiation theorems for measures.

A $5r$-covering theorem

For $0 < t < \infty$, $x \in \mathbf{R}^n$, $0 < r < \infty$, we shall use the notation

$$tB = B(x, tr) \quad \text{when } B = B(x, r).$$

In a general metric space the centre and radius of a ball need not be unique and for $t = 5$ we use the definition

$$5B = \bigcup \{B' : B' \text{ is a closed ball with } B' \cap B \neq \emptyset \text{ and } d(B') \leq 2d(B)\}.$$

Then $d(5B) \leq 5d(B)$. The special value $t = 5$ appears in covering theorems in a natural way.

A metric space X is called *boundedly compact* if all bounded closed subsets of X are compact. The following theorem holds more generally, for example in separable metric spaces. A similar proof with some technical complications works in that case.

2.1. Theorem. *Let X be a boundedly compact metric space and $\mathcal{B}$ a family of closed balls in X such that*

$$\sup\{d(B) : B \in \mathcal{B}\} < \infty.$$

Then there is a finite or countable sequence $B_i \in \mathcal{B}$ of disjoint balls such that

$$\bigcup_{B \in \mathcal{B}} B \subset \bigcup_i 5B_i.$$

Proof. We simplify slightly by assuming that $\mathcal{B}$ is of the form

$$\mathcal{B} = \{B(x, r(x)) : x \in A\},$$

where A is a bounded subset of X. We comment on the modification required for the general case at the end of the proof. Let

$$M = \sup\{r(x) : x \in A\} \quad \text{and}$$
$$A_1 = \{x \in A : 3M/4 < r(x) \leq M\}.$$

Choose an arbitrary $x_1 \in A_1$ and then inductively

(1)
$$x_{k+1} \in A_1 \setminus \bigcup_{i=1}^{k} B(x_i, 3r(x_i))$$

as long as $A_1 \setminus \bigcup_{i=1}^k B(x_i, 3r(x_i)) \neq \emptyset$. The balls $B(x_i, r(x_i))$ thus chosen are obviously disjoint in view of the definition of A_1 and lie in a compact subset of X. We can only have finitely many of them, say k_1, since we cannot pack infinitely many disjoint balls of radius $3M/4$ into a compact subset of X. Thus we have

$$A_1 \subset \bigcup_{i=1}^{k_1} B(x_i, 3r(x_i)).$$

As $r(x) \leq 2r(x_i)$ for $x \in A_1$, $i = 1, \ldots, k_1$, this gives

$$\bigcup_{x \in A_1} B(x, r(x)) \subset \bigcup_{i=1}^{k_1} B(x_i, 5r(x_i)).$$

Let

$$A_2 = \left\{x \in A : \left(\tfrac{3}{4}\right)^2 M < r(x) \leq \tfrac{3}{4} M\right\},$$

$$A_2' = \left\{x \in A_2 : B(x, r(x)) \cap \bigcup_{i=1}^{k_1} B(x_i, r(x_i)) = \emptyset\right\}.$$

If $x \in A_2 \setminus A_2'$, there is $i \in \{1, \ldots, k_1\}$ such that $B(x, r(x)) \cap B(x_i, r(x_i)) \neq \emptyset$, whence

$$d(x, x_i) \leq r(x) + r(x_i) \leq 3r(x_i).$$

This shows

$$A_2 \setminus A_2' \subset \bigcup_{i=1}^{k_1} B(x_i, 3r(x_i)). \tag{2}$$

Choose $x_{k_1+1} \in A_2'$ arbitrarily and then inductively

$$x_{k+1} \in A_2' \setminus \bigcup_{i=k_1+1}^{k} B(x_i, 3r(x_i)).$$

As above there is k_2 such that the balls $B(x_i, r(x_i))$, $i = 1, \ldots, k_2$, are disjoint and

$$A_2' \subset \bigcup_{i=k_1+1}^{k_2} B(x_i, 3r(x_i)).$$

Combining this with (2) we get as before

$$\bigcup_{x \in A_2} B(x, r(x)) \subset \bigcup_{i=1}^{k_2} B(x_i, 5r(x_i)).$$

Proceeding in this manner we find the required balls.

We made two restrictions on the family $\mathcal{B}$. First we assumed that for each $x \in A$ there is only one ball $B(x, r(x))$. We can reduce to this special case by selecting for each centre x a ball $B(x, r(x)) \in \mathcal{B}$ such that $r(x) > \frac{14}{15} \sup\{r : B(x, r) \in \mathcal{B}\}$ and by observing that in (1) and later the number 3 could be replaced by 8/3. Then we can use the above proof to get the required covering from these balls $B(x, r(x))$.

Secondly we assumed that the centres lie in a bounded set. To avoid this the proof can be modified by choosing the new points x_i not too far from a fixed point $a \in X$; for example if x and y were possible selections and $d(y, a) > 2d(x, a)$ we would make a rule that we cannot pick y. □

Remark. Using the Hausdorff maximality principle one can give a shorter proof and obtain a much more general result; for example families of balls can be replaced by many other families of sets, cf. Federer [3, 2.8.4–6].

Vitali's covering theorem for the Lebesgue measure

We can now easily derive a Vitali-type covering theorem for the Lebesgue measure $\mathcal{L}^n$.

2.2. Theorem. *Let $A \subset \mathbf{R}^n$ and suppose that $\mathcal{B}$ is a family of closed balls in $\mathbf{R}^n$ such that every point of A is contained in an arbitrarily small ball belonging to $\mathcal{B}$, that is,*

$$\text{(1)} \qquad \inf\{d(B) : x \in B \in \mathcal{B}\} = 0 \quad \text{for } x \in A.$$

Then there are disjoint balls $B_i \in \mathcal{B}$ such that

$$\mathcal{L}^n\Big(A \setminus \bigcup_i B_i\Big) = 0.$$

Moreover, given $\varepsilon > 0$ the balls B_i can be chosen so that

$$\sum_i \mathcal{L}^n(B_i) \le \mathcal{L}^n(A) + \varepsilon.$$

Proof. The last statement will be clear from the proof. Assume first that A is bounded. Choose an open set U such that $A \subset U$ and

$$\mathcal{L}^n(U) \le (1 + 7^{-n})\,\mathcal{L}^n(A).$$

Applying Theorem 2.1 to the collection of those balls of $\mathcal{B}$ which are contained in U, we find disjoint balls $B_i = B(x_i, r_i) \in \mathcal{B}$ such that $B_i \subset U$ and

$$A \subset \bigcup_i B(x_i, 5r_i).$$

Then

$$5^{-n}\mathcal{L}^n(A) \le 5^{-n}\sum_i \mathcal{L}^n(B(x_i, 5r_i)) = \sum_i \mathcal{L}^n(B_i),$$

and so there is k_1, such that

$$6^{-n}\mathcal{L}^n(A) \le \sum_{i=1}^{k_1} \mathcal{L}^n(B_i).$$

Letting

$$A_1 = A \setminus \bigcup_{i=1}^{k_1} B_i,$$

we have

$$\mathcal{L}^n(A_1) \leq \mathcal{L}^n\Big(U \setminus \bigcup_{i=1}^{k_1} B_i\Big) = \mathcal{L}^n(U) - \sum_{i=1}^{k_1} \mathcal{L}^n(B_i)$$
$$\leq (1 + 7^{-n} - 6^{-n})\,\mathcal{L}^n(A) = u\mathcal{L}^n(A)$$

where

$$u = 1 + 7^{-n} - 6^{-n} < 1.$$

Now A_1 is contained in the open set $\mathbf{R}^n \setminus \bigcup_{i=1}^{k_1} B_i$, and therefore we can find an open set U_1 such that $A_1 \subset U_1 \subset \mathbf{R}^n \setminus \bigcup_{i=1}^{k_1} B_i$ and

$$\mathcal{L}^n(U_1) \leq (1 + 7^{-n})\,\mathcal{L}^n(A_1).$$

As above there are disjoint balls $B_i \in \mathcal{B}$, $i = k_1 + 1, \dots, k_2$, for which $B_i \subset U_1$ and

$$\mathcal{L}^n(A_2) \leq u\mathcal{L}^n(A_1) \leq u^2\mathcal{L}^n(A),$$

where

$$A_2 = A_1 \setminus \bigcup_{i=k_1+1}^{k_2} B_i = A \setminus \bigcup_{i=1}^{k_2} B_i.$$

Evidently all the balls B_i, $i = 1, \dots, k_2$, are disjoint. After m steps

$$\mathcal{L}^n\Big(A \setminus \bigcup_{i=1}^{k_m} B_i\Big) \leq u^m\mathcal{L}^n(A),$$

and the result follows since $u < 1$.

In the general case we write $\mathbf{R}^n = \bigcup_{i=1}^{\infty} \overline{Q}_i$ where the $\overline{Q}_i$'s are closed cubes such that the corresponding open cubes Q_i are disjoint. Applying the first part of the proof to the sets $A \cap Q_i$ and noting that $\mathcal{L}^n(A \setminus \bigcup_{i=1}^{\infty} Q_i) = 0$, we complete the proof. □

2.3. Remarks. (1) For families $\mathcal{B}$ satisfying condition (1) of Theorem 2.2 the conclusion of Theorem 2.1 can be strengthened: the disjoint sequence (B_i) can be found in such a way that for every $m = 1, 2, \dots$

$$\bigcup \mathcal{B} \subset \bigcup_{i=1}^{m} B_i \cup \bigcup_{i=m+1}^{\infty} 5B_i.$$

Essentially the same argument as that of 2.1 applies, see e.g. Federer [3, 2.8.6] or L. Simon [1, 3.4].

(2) All that we really used of the Lebesgue measure in the proof of Theorem 2.2 was the equality $\mathcal{L}^n(B(x,5r)) = 5^n\mathcal{L}^n(B(x,r))$, in fact only the inequality "$\leq$". It is rather straightforward to modify the above proof to see that the theorem remains valid if $\mathcal{L}^n$ is replaced by any Radon measure μ on $\mathbf{R}^n$ such that for some τ, $1 < \tau < \infty$,

$$\limsup_{r\downarrow 0} \{\mu(B(y,\tau r))/\mu(B(y,r)) : x \in B(y,r)\} < \infty$$

$$\text{for } \mu \text{ almost all } x \in \mathbf{R}^n.$$

Moreover, the balls can be replaced by more general families of closed sets and $\mathbf{R}^n$ by more general spaces, see Federer [3, 2.8] for example. However, the above theorem is not valid even for all very nice Radon measures on $\mathbf{R}^n$, as the following example shows.

2.4. Example. Let μ be the Radon measure on $\mathbf{R}^2$ defined by

$$\mu(A) = \mathcal{L}^1(\{x \in \mathbf{R} : (x,0) \in A\}),$$

that is, μ is the length measure on the x-axis. The family

$$\mathcal{B} = \{B((x,y),y) : x \in \mathbf{R}, 0 < y < \infty\}$$

covers $A = \{(x,0) : x \in \mathbf{R}\}$ in the sense of Theorem 2.2 but for any countable subcollection $B_1, B_2, \ldots$ we have

$$\mu\Big(A \cap \bigcup_{i=1}^{\infty} B_i\Big) = 0.$$

Here A touches only the boundaries of the balls of $\mathcal{B}$. By a slight modification we could find a family $\mathcal{B}$ such that each point of A is an interior point of arbitrarily small balls of $\mathcal{B}$ and yet the conclusion of Theorem 2.2 fails. However, if we should require that each point of A is the centre (in fact, not too far from the centre would be enough) of arbitrarily small balls of $\mathcal{B}$, we would get the conclusion of Theorem 2.2. Next we shall develop a covering theorem of this type.

Besicovitch's covering theorem

Again we shall first prove a theorem on families of balls in $\mathbf{R}^n$. This is called Besicovitch's covering theorem, which originates from Besicovitch [6] and [7]. More general covering theory was developed simultaneously by Morse [1]. For some recent developments concerning the best constants in the Besicovitch covering theorem, see Loeb [1], J. M. Sullivan [1] and Füredi and Loeb [1].

We shall begin with a simple lemma from plane geometry. Instead of the following elementary geometric considerations one can also easily deduce it from the cosine formula for the angle of a triangle in terms of the side-lengths.

2.5. Lemma. *Suppose that $a, b \in \mathbf{R}^2$, $0 < |a| < |a-b|$ and $0 < |b| < |a-b|$. Then the angle between the vectors a and b is at least $60°$, that is,*

$$\big|\, a/|a| - b/|b| \,\big| \geq 1.$$

Proof. We have $a \notin B(b, |b|)$ and $b \notin B(a, |a|)$. Let L be the mid-normal to the segment $[0, a]$ with the end-points 0 and a, and let H be the closed half-plane with boundary L such that $O = 0 \in H$. Let T be the triangle OAB as in Figure 2.1.

Then $b \in H \setminus T$, which yields that the angle between a and b is at least $60°$. □

2.6. Lemma. *There is a positive integer $N(n)$ depending only on n with the following property. Suppose there exist k points $a_1, \dots, a_k$ in $\mathbf{R}^n$ and k positive numbers $r_1, \dots, r_k$ such that*

$$a_i \notin B(a_j, r_j) \quad \text{for } j \neq i, \quad \text{and} \quad \bigcap_{i=1}^{k} B(a_i, r_i) \neq \emptyset.$$

Then $k \leq N(n)$.

Proof. We may assume $a_i \neq 0$ for all $i = 1, \dots, k$ and

$$0 \in \bigcap_{i=1}^{k} B(a_i, r_i).$$

Then

$$|a_i| \leq r_i < |a_i - a_j| \quad \text{for } i \neq j.$$

Applying Lemma 2.5 with $a = a_i$ and $b = a_j$ for $i \neq j$ in the two-dimensional plane containing 0, a_i and a_j, we obtain

$$\big|\, a_i/|a_i| - a_j/|a_j| \,\big| \geq 1 \quad \text{for } i \neq j. \tag{1}$$

Since the unit sphere S^{n-1} is compact there is an integer $N(n)$ with the following property: if $y_1, \dots, y_k \in S^{n-1}$ with $|y_i - y_j| \geq 1$ for $i \neq j$, then $k \leq N(n)$. By (1), $N(n)$ is what we want. □

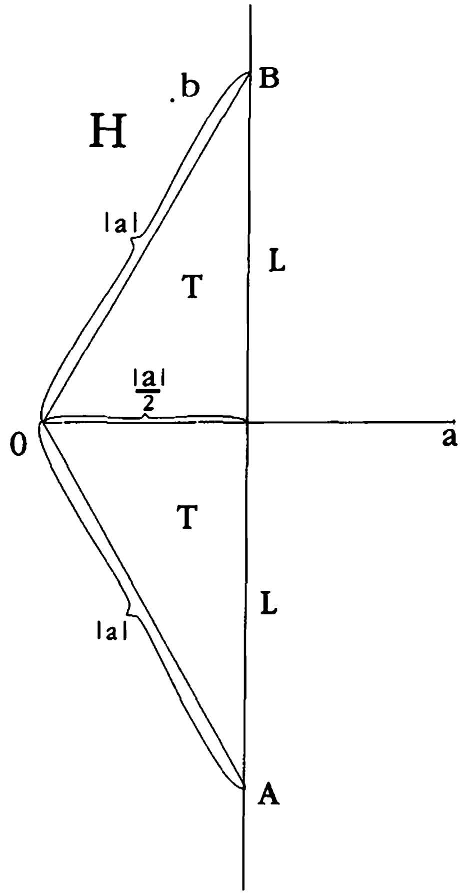

Figure 2.1.

2.7. Besicovitch's covering theorem. *There are integers $P(n)$ and $Q(n)$ depending only on n with the following properties. Let A be a bounded subset of $\mathbf{R}^n$, and let $\mathcal{B}$ be a family of closed balls such that each point of A is the centre of some ball of $\mathcal{B}$.*

(1) *There is a finite or countable collection of balls $B_i \in \mathcal{B}$ such that they cover A and every point of $\mathbf{R}^n$ belongs to at most $P(n)$ balls B_i, that is,*

$$\chi_A \leq \sum_i \chi_{B_i} \leq P(n).$$

(2) *There are families $\mathcal{B}_1, \ldots, \mathcal{B}_{Q(n)} \subset \mathcal{B}$ covering A such that each $\mathcal{B}_i$ is disjoint, that is,*

$$A \subset \bigcup_{i=1}^{Q(n)} \bigcup \mathcal{B}_i$$

and

$$B \cap B' = \emptyset \quad \text{for } B, B' \in \mathcal{B}_i \text{ with } B \neq B'.$$

Proof. (1) For each $x \in A$ pick one ball $B(x, r(x)) \in \mathcal{B}$. As A is bounded, we may assume that

$$M_1 = \sup_{x \in A} r(x) < \infty.$$

Choose

$$x_1 \in A \quad \text{with } r(x_1) \geq M_1/2$$

and then inductively

$$x_{j+1} \in A \setminus \bigcup_{i=1}^{j} B(x_i, r(x_i)) \quad \text{with } r(x_{j+1}) \geq M_1/2$$

as long as possible. Since A is bounded, the process terminates, and we get a finite sequence $x_1, \ldots, x_{k_1}$.

Next let

$$M_2 = \sup \left\{ r(x) : x \in A \setminus \bigcup_{i=1}^{k_1} B(x_i, r(x_i)) \right\}.$$

Choose

$$x_{k_1+1} \in A \setminus \bigcup_{i=1}^{k_1} B(x_i, r(x_i)) \quad \text{with } r(x_{k_1+1}) \geq M_2/2,$$

and again inductively

$$x_{j+1} \in A \setminus \bigcup_{i=1}^{j} B(x_i, r(x_i)) \quad \text{with } r(x_{j+1}) \geq M_2/2.$$

Continuing this process we obtain an increasing sequence of integers $0 = k_0 < k_1 < k_2 < \ldots$, a decreasing sequence of positive numbers M_i with $2M_{i+1} \leq M_i$, and a sequence of balls $B_i = B(x_i, r(x_i)) \in \mathcal{B}$ with the following properties. Let

$$I_j = \{k_{j-1} + 1, \ldots, k_j\} \quad \text{for } j = 1, 2, \ldots.$$

Then

$$(3) \qquad M_j/2 \le r(x_i) \le M_j \quad \text{for } i \in I_j,$$

$$(4) \qquad x_{j+1} \in A \setminus \bigcup_{i=1}^{j} B_i \quad \text{for } j = 1, 2, \ldots,$$

$$(5) \qquad x_i \in A \setminus \bigcup_{m \ne k} \bigcup_{j \in I_m} B_j \quad \text{for } i \in I_k.$$

The first two properties follow immediately from the construction. To verify the third property, let $m \ne k$, $j \in I_m$ and $i \in I_k$. If $m < k$, $x_i \notin B_j$ by (4). If $k < m$, then $r(x_j) < r(x_i)$, $x_j \notin B_i$ by (4), and so $x_i \notin B_j$.

Since $M_i \to 0$, (3) implies $r(x_i) \to 0$, and it follows from the construction that

$$A \subset \bigcup_{i=1}^{\infty} B_i.$$

To establish also the second statement of (1), suppose a point x belongs to p balls B_i, say

$$x \in \bigcap_{i=1}^{p} B_{m_i}.$$

We shall show that $p \le P(n) = 16^n N(n)$ with $N(n)$ as in Lemma 2.6.

Using (5) and Lemma 2.6 we see that the indices m_i can belong to at most $N(n)$ different blocks I_j, that is,

$$\operatorname{card}\left\{j : I_j \cap \{m_i : i = 1, \ldots, p\} \ne \emptyset\right\} \le N(n).$$

Consequently it suffices to show that

$$(6) \qquad \operatorname{card}\left(I_j \cap \{m_i : i = 1, \ldots, p\}\right) \le 16^n \quad \text{for } j = 1, 2, \ldots.$$

Fix j and write

$$I_j \cap \{m_i : i = 1, \ldots, p\} = \{\ell_1, \ldots, \ell_q\}.$$

By (3) and (4) the balls $B(x_{\ell_i}, \frac{1}{4} r(x_{\ell_i}))$, $i = 1, \ldots, q$, are disjoint and they are contained in $B(x, 2M_j)$. Hence, with $\alpha(n) = \mathcal{L}^n(B(0,1))$,

$$\begin{aligned} q\alpha(n)(M_j/8)^n &\le \sum_{i=1}^{q} \mathcal{L}^n\left(B(x_{\ell_i}, \tfrac{1}{4} r(x_{\ell_i}))\right) \\ &\le \mathcal{L}^n\left(B(x, 2M_j)\right) = \alpha(n)(2M_j)^n, \end{aligned}$$

and so $q \le 16^n$ as desired. This proves (6), and thus also (1).

(2) Let $B_1, B_2, \dots$ be the balls found in (1). Letting $B_i = B(x_i, r_i)$, there are for each $\varepsilon > 0$ only finitely many balls B_i with $r_i \ge \varepsilon$ because of (1) and the boundedness of A. Thus we may assume $r_1 \ge r_2 \ge \dots$. Let $B_{1,1} = B_1$ and then inductively if $B_{1,1}, \dots, B_{1,j}$ have been chosen, $B_{1,j+1} = B_k$ where k is the smallest integer with

$$B_k \cap \bigcup_{i=1}^{j} B_{1,i} = \emptyset.$$

We continue this as long as possible getting a finite or countable disjoint subfamily

$$\mathcal{B}_1 = \{B_{1,1}, B_{1,2}, \dots\}$$

of $\{B_1, B_2, \dots\}$.

If A is not covered by $\bigcup \mathcal{B}_1$, we define first $B_{2,1} = B_k$ where k is the smallest integer for which $B_k \notin \mathcal{B}_1$. Again we define inductively $B_{2,j+1} = B_k$ with the smallest k such that

$$B_k \cap \bigcup_{i=1}^{j} B_{2,i} = \emptyset.$$

With this process we find subfamilies $\mathcal{B}_1, \mathcal{B}_2, \dots$ of $\{B_1, B_2, \dots\}$, each $\mathcal{B}_i$ being disjoint. We claim that

$$A \subset \bigcup_{k=1}^{m} \bigcup \mathcal{B}_k \quad \text{for some } m \le 4^n P(n) + 1.$$

Suppose m is such that there is $x \in A \setminus \bigcup_{k=1}^{m} \bigcup \mathcal{B}_k$. We then have to show that $m \le 4^n P(n)$. Since the balls B_i cover A we can find i with $x \in B_i$. Then for each $k = 1, \dots, m$, $B_i \notin \mathcal{B}_k$, which means by the construction of $\mathcal{B}_k$ that $B_i \cap B_{k,i_k} \neq \emptyset$ for some i_k for which $r_i \le r_{k,i_k}$, r_i and r_{k,i_k} being the radii of B_i and B_{k,i_k}, respectively. Hence there are balls B'_k of radius $r_i/2$ contained in $(2B_i) \cap B_{k,i_k}$ for all $k = 1, \dots, m$. Since each point of $\mathbf{R}^n$ is contained in at most $P(n)$ balls B_{k,i_k}, $k = 1, \dots, m$, this is also true for the smaller balls B'_k, that is

$$\sum_{k=1}^{m} \chi_{B'_k} \le P(n)\, \chi_{\bigcup_{k=1}^{m} B'_k}.$$

Using the fact $B'_k \subset 2B_i$, we then have

$$\begin{aligned} 2^n\alpha(n)\, r_i^n &= \mathcal{L}^n(2B_i) \geq \mathcal{L}^n\Big(\bigcup_{k=1}^m B'_k\Big) \\ &= \int \chi_{\bigcup_{k=1}^m B'_k}\, d\mathcal{L}^n \geq P(n)^{-1} \int \sum_{k=1}^m \chi_{B'_k}\, d\mathcal{L}^n \\ &= P(n)^{-1} \sum_{k=1}^m \mathcal{L}^n(B'_k) = mP(n)^{-1}\, 2^{-n}\, \alpha(n)\, r_i^n. \end{aligned}$$

Hence $m \leq 4^n P(n)$ as required. □

Vitali's covering theorem for Radon measures

We can now easily establish a Vitali-type covering theorem for arbitrary Radon measures on $\mathbf{R}^n$.

2.8. Theorem. *Let μ be a Radon measure on $\mathbf{R}^n$, $A \subset \mathbf{R}^n$ and $\mathcal{B}$ a family of closed balls such that each point of A is the centre of arbitrarily small balls of $\mathcal{B}$, that is,*

$$\inf\{r : B(x,r) \in \mathcal{B}\} = 0 \quad \text{for } x \in A.$$

Then there are disjoint balls $B_i \in \mathcal{B}$ such that

$$\mu\Big(A \setminus \bigcup_i B_i\Big) = 0.$$

Proof. We may assume $\mu(A) > 0$. Suppose first A is bounded. By Definition 1.5 (4) there is an open set U such that $A \subset U$ and

$$\mu(U) \leq \big(1 + (4Q(n))^{-1}\big)\, \mu(A),$$

where $Q(n)$ is as in Besicovitch's covering theorem 2.7. By that theorem we can find $\mathcal{B}_1, \ldots, \mathcal{B}_{Q(n)} \subset \mathcal{B}$ such that each $\mathcal{B}_i$ is disjoint and

$$A \subset \bigcup_{i=1}^{Q(n)} \bigcup \mathcal{B}_i \subset U.$$

Then

$$\mu(A) \leq \sum_{i=1}^{Q(n)} \mu\Big(\bigcup \mathcal{B}_i\Big),$$

and consequently there is an i with

$$\mu(A) \leq Q(n)\,\mu\Big(\bigcup \mathcal{B}_i\Big).$$

Further, for some finite subfamily $\mathcal{B}'_i$ of $\mathcal{B}_i$ we have

$$\mu(A) \leq 2Q(n)\,\mu\Big(\bigcup \mathcal{B}'_i\Big).$$

Letting

$$A_1 = A \setminus \bigcup \mathcal{B}'_i,$$

we get

$$\begin{aligned}\mu(A_1) &\leq \mu\Big(U \setminus \bigcup \mathcal{B}'_i\Big) = \mu(U) - \mu\Big(\bigcup \mathcal{B}'_i\Big)\\ &\leq \big(1 + \tfrac{1}{4}Q(n)^{-1} - \tfrac{1}{2}Q(n)^{-1}\big)\,\mu(A) = u\mu(A)\end{aligned}$$

with $u = 1 - \frac{1}{4}Q(n)^{-1} < 1$. We can now continue by the same principle as in the proof of Theorem 2.2.

In order to get rid of the assumption that A is bounded, we may modify the last step of the proof of Theorem 2.2 making use of the fact that $\mu(V)$ can be positive for at most countably many parallel hyperplanes V. □

Differentiation of measures

We shall now turn to the differentiation theory of measures.

2.9. Definition. *Let μ and λ be locally finite Borel measures on $\mathbf{R}^n$. The upper and lower derivatives of μ with respect to λ at a point $x \in \mathbf{R}^n$ are defined by*

$$\overline{D}(\mu,\lambda,x) = \limsup_{r\downarrow 0} \frac{\mu(B(x,r))}{\lambda(B(x,r))},$$

$$\underline{D}(\mu,\lambda,x) = \liminf_{r\downarrow 0} \frac{\mu(B(x,r))}{\lambda(B(x,r))}.$$

At the points x where the limit exists we define the derivative of μ by

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

2.10. Remarks. Here we interpret $0/0 = 0$. The above derivatives are Borel functions. Let us consider the proof only in the case $\lambda = \mathcal{L}^n$, which is essentially all we shall need. More generally, see for example Federer [3, 2.9.6]. Show first that the function $x \mapsto \mu(B(x,r))$ is upper semicontinuous (that is, $x_i \to x$ implies $\limsup_{i\to\infty} \mu(B(x_i,r)) \le \mu(B(x,r))$). Then using the facts that $\mu(B(x,r))$ is monotonic and $\mathcal{L}^n(B(x,r))$ continuous in r, prove that the upper and lower limits do not change if r is restricted to positive rationals. Thus the Borel measurability of the upper and lower derivatives reduces to the fact that the suprema and infima of countable families of Borel functions are Borel functions.

Later on we shall encounter other functions of the same kind which can be shown to be Borel functions by similar reasoning.

2.11. Definition. *Let μ and λ be measures on $\mathbf{R}^n$. We say that μ is absolutely continuous with respect to λ if*

$$\lambda(A) = 0 \quad \textit{implies } \mu(A) = 0 \quad \textit{for all } A \subset \mathbf{R}^n.$$

In this case we write

$$\mu \ll \lambda.$$

The following theorem contains the basic ingredients of the differentiation of μ with respect to λ.

2.12. Theorem. *Let μ and λ be Radon measures on $\mathbf{R}^n$.*

(1) *The derivative $D(\mu, \lambda, x)$ exists and is finite for λ almost all $x \in \mathbf{R}^n$.*
(2) *For all Borel sets $B \subset \mathbf{R}^n$,*

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

with equality if $\mu \ll \lambda$.
(3) *$\mu \ll \lambda$ if and only if $\underline{D}(\mu, \lambda, x) < \infty$ for μ almost all $x \in \mathbf{R}^n$.*

For the proof we will need the following lemma.

2.13. Lemma. *Let μ and λ be Radon measures on $\mathbf{R}^n$, $0 < t < \infty$ and $A \subset \mathbf{R}^n$.*

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

Proof. (1) Let $\varepsilon > 0$. Using Definition 1.5 (4) we find an open set U such that $A \subset U$ and $\lambda(U) \le \lambda(A) + \varepsilon$. An application of Theorem 2.8 gives disjoint closed balls $B_i \subset U$ such that

$$\mu(B_i) \le (t+\varepsilon)\,\lambda(B_i) \quad \text{and} \quad \mu\Big(A \setminus \bigcup_i B_i\Big) = 0.$$

Then

$$\begin{aligned}\mu(A) &\le \sum_i \mu(B_i) \le (t+\varepsilon)\sum_i \lambda(B_i)\\ &\le (t+\varepsilon)\,\lambda(U) \le (t+\varepsilon)(\lambda(A)+\varepsilon).\end{aligned}$$

Letting $\varepsilon \downarrow 0$, we get $\mu(A) \le t\lambda(A)$, which proves (1). (2) can be proven in the same way. □

Proof of Theorem 2.12. For $0 < r < \infty$, $0 < s < t < \infty$, let

$$\begin{aligned}A_{s,t,r} &= \{x \in B(r) : \underline{D}(\mu,\lambda,x) \le s < t \le \overline{D}(\mu,\lambda,x)\},\\ A_{t,r} &= \{x \in B(r) : \overline{D}(\mu,\lambda,x) \ge t\}.\end{aligned}$$

By Lemma 2.13

$$\begin{aligned}&t\lambda(A_{s,t,r}) \le \mu(A_{s,t,r}) \le s\lambda(A_{s,t,r}) < \infty,\\ &u\lambda(A_{u,r}) \le \mu(A_{u,r}) \le \mu(B(r)) < \infty.\end{aligned}$$

These inequalities yield $\lambda(A_{s,t,r}) = 0$ since $s < t$, and $\lambda(\bigcap_{u>0} A_{u,r}) = \lim_{u\to\infty} \lambda(A_{u,r}) = 0$. But the complement of the set $\{x : \exists D(\mu,\lambda,x) < \infty\}$ is the union of the sets $A_{s,t,r}$ and $\bigcap_{u>0} A_{u,r}$ where s and t run through the positive rationals with $s < t$ and r runs through the positive integers. Hence it is of λ measure zero, which settles (1).

To prove (2) choose $1 < t < \infty$ and let

$$B_p = \{x \in B : t^p \le D(\mu,\lambda,x) < t^{p+1}\}, \quad p = 0, \pm 1, \pm 2, \ldots.$$

Then by part (1) of this theorem already proved and by part (2) of Lemma 2.13,

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

Letting $t \downarrow 1$, we get $\int_B D(\mu, \lambda, x)\, d\lambda x \le \mu(B)$.

If $\mu \ll \lambda$, the sets of λ measure zero also have μ measure zero. Hence, noting also that by (1) $D(\mu, \lambda, x) = D(\lambda, \mu, x)^{-1} > 0$ for μ almost all x, we have $\mu(B) = \sum_{p=-\infty}^{\infty} \mu(B_p)$, and a similar argument as above making use of part (1) of Lemma 2.13 gives the opposite inequality.

By (1), $\underline{D}(\mu, \lambda, x) < \infty$ λ almost everywhere, and hence if $\mu \ll \lambda$ this also holds μ almost everywhere.

Finally, to prove the other half of (3), suppose $\underline{D}(\mu, \lambda, x) < \infty$ for μ almost all $x \in \mathbf{R}^n$. Let $A \subset \mathbf{R}^n$ with $\lambda(A) = 0$. For $u = 1, 2, \ldots$ Lemma 2.13 (1) gives

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

and so $\mu(A) = 0$. □

As a corollary we obtain immediately a density theorem and a theorem on differentiation of integrals.

2.14. Corollary. *Let λ be a Radon measure on $\mathbf{R}^n$.*

(1) *If $A \subset \mathbf{R}^n$ is λ measurable, then the limit*

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

exists and equals 1 for λ almost all $x \in A$ and equals 0 for λ almost all $x \in \mathbf{R}^n \setminus A$.

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

$$\lim_{r \downarrow 0} \frac{1}{\lambda(B(x, r))} \int_{B(x,r)} f\, d\lambda = f(x) \quad \text{for } \lambda \text{ almost all } x \in \mathbf{R}^n.$$

Proof. (1) follows from (2) with $f = \chi_A$. To prove (2) we may assume $f \ge 0$. Define the Radon measure μ by $\mu(A) = \int_A f\, d\lambda$. Then $\mu \ll \lambda$ and Theorem 2.12 (2) gives

$$\int_B D(\mu, \lambda, x)\, d\lambda x = \mu(B) = \int_B f\, d\lambda$$

for all Borel sets B. Obviously this means that $f(x) = D(\mu, \lambda, x)$ for λ almost all $x \in \mathbf{R}^n$, which proves (2). □