8. Differentiation of Measures

In the previous sections we discussed at length differentiation of integrals, which can be thought of as a direct generalization of the FTC I. In order to wrap up the subject, we still have to address one topic: differentiation of measures. Noting that \( A \mapsto \int f \, d\mu \) is, for \( f \in L^1(\mu) \) a (signed) measure, the question is what is the result of differentiation theorems when we replace this measure by a general Radon measure.

We begin by some general observations. First we note that differentiation for integrals is not necessarily restricted to balls. Indeed, one can consider more general families of sets:

**Definition 8.1** Let \( \mu \) be a Borel measure on \( \mathbb{R}^d \), let \( \{B_r : r > 0\} \) be a family of Borel subsets of \( \mathbb{R}^d \). We say that \( \{B_r : r > 0\} \) tends nicely to \( x \in \mathbb{R}^d \) if

\[
\exists c > 0 \forall r > 0: \quad B_r \subseteq B(x, r) \land \mu(B_r) \geq c\mu(B(x, r)).
\]  

(8.1)

where \( B(x, r) := \{y \in \mathbb{R}^d : \|x - y\| < r\} \).

Note that we need not have \( x \in B_r \). Indeed, \( B_r := (x + r/2, x + r) \) defines a family \( \{B_r : r > 0\} \) that tends nicely to \( x \) under the Lebesgue measure. Still, we have:

**Theorem 8.2** Suppose \( \mu \) is a Radon measure on \( \mathbb{R}^d \) and let \( f \in L^1(\mu) \). Then there is a Borel set \( C \subseteq \mathbb{R}^d \) with \( \mu(C^c) = 0 \) such that

\[
\lim_{r \downarrow 0} \frac{1}{\mu(B_r)} \int_{B_r} f \, d\mu = f(x)
\]  

(8.2)

holds for all \( x \in C \) and all families \( \{B_r : r > 0\} \) of Borel sets that tend nicely to \( x \).

**Proof.** It is an exercise to show that \( x \mapsto \mu(B(x, r)) \) is Borel measurable. Then

\[
C' := \bigcap_{n \geq 1} \left\{ x \in \mathbb{R}^d : \mu(B(x, 1/n)) > 0 \right\}
\]  

(8.3)

is a Borel set as well. Set

\[
C := \left\{ x \in C' : \lim_{r \downarrow 0} \frac{1}{\mu(B(x, r))} \int_{B(x, r)} |f - f(x)| \, d\mu = 0 \right\}.
\]  

(8.4)

We claim that \( C \) is Borel as well. This is shown directly for \( f := 1_B \), with \( B \) a bounded Borel set, from Borel measurability of \( x \mapsto \int_{B(x, r)} f \, d\mu \) (another exercise). Additivity then implies the same for simple functions and uniform limits extend this to bounded measurable functions with uniformly bounded support. Countable additivity then extends this to all \( f \in L^1(\mu) \).

Let \( x \in A \) and let \( \{B_r : r > 0\} \) be a family that tends nicely to \( x \). Then \( x \in A' \) implies that \( \mu(B_r) \geq c\mu(B(x, r)) > 0 \) for all \( r > 0 \) and \( x \in A \) then shows

\[
\lim_{r \downarrow 0} \frac{1}{\mu(B_r)} \int_{B_r} |f - f(x)| \, d\mu \leq \lim_{r \downarrow 0} \frac{1}{c\mu(B(x, r))} \int_{B(x, r)} |f - f(x)| \, d\mu = 0.
\]  

(8.5)

This yields (8.2). \( \square \)

We formalize the conclusion using the following concept:
**Definition 8.3 (Lebesgue point)** Let \( f \in L^1(\lambda) \). We say that \( x \in \mathbb{R}^d \) is a Lebesgue point of \( f \) if \( (8.2) \) holds (with \( \mu = \lambda \)).

Theorem 7.3 states that the set of the Lebesgue points of \( f \in L^1(\lambda) \) is measurable and that \( \lambda \)-a.e. \( x \in \mathbb{R}^d \) is a Lebesgue point. An interesting corollary arises when the previous statement is specialized to indicators of Borel sets. Indeed from Theorem 7.3 we get:

**Corollary 8.4 (Lebesgue density theorem)** Let \( A \subseteq \mathbb{R}^d \) be a Borel set. Then for \( \lambda \)-a.e. \( x \in A \) we have

\[
\lim_{r \to 0} \frac{\lambda(A \cap B(x, r))}{\lambda(B(x, r))} = 1.
\]

(8.6)

while for \( \lambda \)-a.e. \( x \in A^c \), we have

\[
\lim_{r \to 0} \frac{\lambda(A \cap B(x, r))}{\lambda(B(x, r))} = 0.
\]

(8.7)

The same holds if \( \{B(x, r) : r > 0\} \) is replaced by any family of Borel sets \( \{B_r : r > 0\} \) that tend nicely to \( x \).

**Proof.** By complementation it suffices to prove (8.7). Let \( x \in A^c \). Then \( 1_{A \cap B(x, r)} = 1_{A} 1_{B(x, r)} - 1_{A^c} 1_{B(x, r)} \). Therefore

\[
\lambda(A \cap B(x, r)) = \int_{B(x, r)} |1_{A} - 1_{A^c}| \, d\lambda.
\]

(8.8)

Dividing by \( \lambda(B(x, r)) \) the statement follows from Theorem 7.3; specifically, (7.48), and the observations made in Theorem 8.2.

We now move to the key theorem of this section:

**Theorem 8.5 (Differentiation of measures)** Let \( \mu \) and \( \nu \) be two Radon measures on \( \mathbb{R}^d \). Then

\[
f(x) := \lim_{r \to 0} \frac{\mu(B(\infty, x, r))}{\nu(B(\infty, x, r))}
\]

exists for \( \nu \)-a.e. \( x \in \mathbb{R}^d \).

(8.9)

Moreover, \( \kappa := \mu - f \nu \) is a non-negative (signed) measure with \( \kappa \perp \nu \).

Before we delve into the proof, note the following corollary:

**Corollary 8.6 (Radon-Nikodym derivative is a derivative)** If \( \mu \) and \( \nu \) are Radon measures on \( \mathbb{R}^d \) with \( \mu \ll \nu \), then for \( \nu \)-a.e. \( x \in \mathbb{R}^d \),

\[
\lim_{r \to 0} \frac{\mu(B(\infty, x, r))}{\nu(B(\infty, x, r))} = \frac{d\mu}{d\nu}(x).
\]

(8.10)

**Proof.** Let \( f \) be as in (8.9) and let \( \kappa := \mu - f \nu \). Then \( \mu \ll \nu \) forces \( \kappa = 0 \) and so \( \mu = f \nu \). Hence \( f \) equals the Radon-Nikodym derivative \( \frac{d\mu}{d\nu} \), \( \nu \)-a.e.

In light of the previous corollary, the main ingredient of the proof of Theorem 8.5 is:

**Lemma 8.7** Let \( \nu \) and \( \kappa \) be Radon measures on \( \mathbb{R}^d \) such that \( \kappa \perp \nu \). Then

\[
\lim_{r \to 0} \frac{\kappa(B(\infty, x, r))}{\nu(B(\infty, x, r))} = 0
\]

for \( \nu \)-a.e. \( x \in \mathbb{R}^d \).

(8.11)
Proof. We will use, one last time, the Besicovitch covering lemma. The mutual singularity of the two measures implies
\[ \exists A \in \mathcal{B}(\mathbb{R}^d) : \quad \nu(A^c) = 0 = \kappa(A). \] (8.12)

Fix \( t > 0 \), let \( A' := \bigcap_{n \geq 1} \{ x \in A : \nu(B_\infty(x, 1/n)) > 0 \} \) and set
\[ A_t := \left\{ x \in A' : \limsup_{r \downarrow 0} \frac{\kappa(B_\infty(x,r))}{\nu(B_\infty(x,r))} > t \right\}. \] (8.13)

Then \( A_t \subseteq A \) and so by the outer regularity of \( \kappa \), for each \( \epsilon > 0 \) there is \( O \subseteq \mathbb{R}^d \) open such that
\[ A_t \subseteq O \land \kappa(O) < \epsilon. \] (8.14)

Consider the family
\( \mathcal{B} := \left\{ B_\infty(x,r) : r \in (0,1) \land B_\infty(x,r) \subseteq O \land \kappa(B_\infty(x,r)) > t\nu(B_\infty(x,r)) \right\} \) (8.15)

Then for every \( x \in A_t \), there is \( r \in (0,1) \) such that \( B_\infty(x,r) \in \mathcal{B} \). Proposition 7.6 implies
\[ \exists \{B_i\}_{i \geq 1} \subseteq \mathcal{B} \cup \{ \emptyset \} : \quad A_t \subseteq \bigcup_{i \geq 1} B_i \subseteq O \land \sum_{i \geq 1} 1_{B_i} = c(d). \] (8.16)

We now compute as before
\[ t\nu(A_t) \leq \sum_{i \geq 1} t\nu(B_i) \]
\[ \leq \sum_{i \geq 1} \kappa(B_i) = \int \left( \sum_{i \geq 1} 1_{B_i} \right) d\kappa \]
\[ \leq c(d) \int 1_{\bigcup_{i \geq 1} B_i} d\kappa = c(d) \kappa \left( \bigcup_{i \geq 1} B_i \right) \]
\[ \leq c(d)\kappa(O) < c(d)\epsilon. \] (8.17)

As this holds for all \( \epsilon > 0 \), we have \( \nu(A_t) = 0 \) for all \( t > 0 \). Taking \( t \downarrow 0 \) along a sequence shows
\[ \nu \left( \left\{ x \in A' : \limsup_{r \downarrow 0} \frac{\kappa(B_\infty(x,r))}{\nu(B_\infty(x,r))} > 0 \right\} \right) = 0. \] (8.18)

Since \( \nu(\mathbb{R}^d \setminus A') = 0 \) (prove this!), the claim is proved. \( \square \)

**Proof of Theorem 8.5.** Let \( \mu \) and \( \nu \) be two Radon measures on \( \mathbb{R}^d \). Since both of these are then \( \sigma \)-finite, the Lebesgue-Radon-Nikodym Theorem implies existence of a Radon measure \( \kappa \) such that
\[ \mu = f\nu + \kappa \land \kappa \perp \nu. \] (8.19)

Then for each \( x \in \mathbb{R}^d \),
\[ \frac{\mu(B_\infty(x,r))}{\nu(B_\infty(x,r))} = \frac{1}{\nu(B_\infty(x,r))} \int_{B_\infty(x,r)} f d\nu + \frac{\kappa(B_\infty(x,r))}{\nu(B_\infty(x,r))} \] (8.20)

By Theorem 7.10, the first term converges to \( f(x) \) as \( r \downarrow 0 \) for \( \nu \)-a.e. \( x \in \mathbb{R}^d \). Thanks to Lemma 8.7, the second term in turn converges to zero. \( \square \)