1 Lebesgue differentiation in Euclidean spaces

(First part of Folland Section 3.4)

In this section, \((X, \mathcal{M}) = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})\). Let \(\mu = m\), Lebesgue measure. If \(\nu \ll m\) for a signed measure \(\nu\), then we know that \(d\nu = f \, dm\) for some \(f\). Intuitively, we may think of \(f\) as a ‘density’ which describes the measure \(\nu\) relative to the fixed ‘background’ measure \(m\). In the special setting of Euclidean space, we can try to recover the value \(f(x)\) at a point \(x\) by taking a suitable limit of ‘approximate densities’ using balls around \(x\).

It turns out that this works (except perhaps for a Lebesgue null set of \(x\)) provided that \(\nu\) takes only finite values on balls to begin with. Observe that in this case, for any \(x \in \mathbb{R}^n\) and \(r > 0\), we have

\[
\frac{\nu B(x, r)}{m B(x, r)} = \frac{1}{m B(x, r)} \int_{B(x, r)} f(y) \, dy \quad \text{(Lebesgue integral)}.
\]

**Definition 1.1.** A **locally integrable function** \(\mathbb{R}^n \rightarrow \mathbb{R}\). We denote the space of these by \(L^1_{\text{loc}}\).

[The definition of locally integrable functions \(\mathbb{R}^n \rightarrow \mathbb{C}\) is the obvious analog.]

If \(f \in L^1_{\text{loc}}\) and \(r > 0\), then we define

\[
A_r f(x) = \frac{1}{m B(x, r)} \int_{B(x, r)} f(y) \, dy,
\]

the ‘average’ of \(f\) over the ball \(B(x, r)\).

**Simple observation:**

**Lemma 1.2** (‘Smearing out improves the function’; Folland 3.16). If \(f \in L^1_{\text{loc}}\), then \(A_r f(x)\) is jointly continuous in \(r > 0\) and \(x \in \mathbb{R}^n\).
Here is the main result of this lecture:

**Theorem 1.3** (Modest version of Lebesgue differentiation theorem; Folland 3.18).

If $f \in L^1_{\text{loc}}$, then
$$\lim_{r \to 0} A_r f(x) = f(x)$$
for a.e. $x \in \mathbb{R}^n$.

(In fact we’ll end up with something a bit stronger than this.)

This result really is a special feature of $\mathbb{R}^n$ (and some other spaces that have a lot of ‘geometric structure’ in common with $\mathbb{R}^n$).

**First thoughts:**

1. This is easy if $f$ is continuous at $x$.

2. We may write $\mathbb{R}^n$ as a countable union of boxes $B$ with corners on the integer lattice, and if $x \in B^o$ then it suffices to prove the result for $f_B = f \cdot \chi_B$. This lets us reduce to the case $f \in L^1$ (and even with $\{f \neq 0\}$ bounded).

3. Now recall that we can approximate integrable functions by continuous: if $f \in L^1(m)$ and $\varepsilon > 0$, there is a continuous $g$ such that $\int |f - g| \, dm < \varepsilon$.

**All of the rest of the proof:** We know the result for continuous functions. We will show that the desired result is ‘stable’ under approximations as in item 3. above, so it extends to all integrable functions.

This sounds like it’s just tidying up, but actually this is the main part of the proof. This proof (as we present it here) is a good introduction to an idea that is important much more widely in harmonic analysis: ‘maximal operators’.

Suppose that $f \in L^1(m)$, $\varepsilon > 0$, and $g$ is continuous with $\int |f - g| \, dm < \varepsilon$. We want to compare the limits

$$A_r g(x) \longrightarrow g(x) \quad \text{and} \quad A_r f(x) \longrightarrow f(x).$$

Of course, it would be enough to prove that $A_r (g - f)(x) \longrightarrow 0$ for a.e. $x$. This can’t be quite right, since $g(x)$ and $f(x)$ aren’t necessarily equal. But we can get something like this: we’ll show that, if $\varepsilon$ is small, then for most $x$ the quantity $A_r (f - g)(x)$ stays uniformly small as $r \to 0$.

The next definition captures this kind of control precisely:

**Definition 1.4.** For $f \in L^1_{\text{loc}}$, its **Hardy–Littlewood maximal function** $Hf$.

Clearly $Hf$ is still measurable.

Here’s the ‘smallness’ result that we can prove for the H–L maximal function:
**Theorem 1.5** (The Hardy–Littlewood maximal theorem; Folland 3.17). There is a constant $C > 0$ such that for all $f \in L^1$ and all $\alpha > 0$ we have

$$m\{Hf > \alpha\} \leq \frac{C}{\alpha} \int |f| \, dm.$$  

This theorem, in its turn, is built on a more elementary fact about balls in Euclidean space. This is given in the following lemma, which is the real geometric backbone of this lecture. We’ll use this lemma again later.

**Lemma 1.6** (A version of the Vitali covering lemma; Folland 3.15). Let $C$ be a collection of open balls in $\mathbb{R}^n$, and let $U$ be their union. For any $c < mU$, there are disjoint balls $B_1, \ldots, B_k \in C$ such that $\sum_j mB_j > c/3^n$.

Thus: from any collection of open balls we can choose a finite, disjoint subcollection that cover a substantial fraction of the same set, in terms of Lebesgue measure.

**Proof idea for Lemma 1.6.** Use regularity of $m$ to reduce to the case of a finite collection covering a compact set of measure $> c$. Then choose a disjoint subcollection recursively by at each step taking the biggest ball you can which is disjoint from all those you’ve already chosen. \hfill \Box

**Sketch of [Lemma 1.6 $\Rightarrow$ Theorem 1.5].** By definition, each point in the set $\{Hf > \alpha\}$ is contained in an open ball over which the average value of $f$ exceeds $\alpha$. Now Lemma 1.6 lets us extract a disjoint subfamily of these balls such that (i) $f$ has average $> \alpha$ on each, and (ii) they cover a set whose measure is a large fraction of $m\{Hf > \alpha\}$. Adding up the integrals over these balls gives the correct kind of lower bound on $\int |f| \, dm$. \hfill \Box

**Proof sketch for Theorem 1.3.** Let $f \in L^1(m)$. Let $F_\alpha$ be the set where the convergence $A_r f \rightarrow f$ fails by more than some $\alpha > 0$. Let $\varepsilon > 0$, and choose $g$ continuous so that $\int |f - g| \, dm < \varepsilon$. Since $A_r g \rightarrow g$ everywhere, we must have

$$\limsup_{r \to \infty} \left| A_r(f - g)(x) - (f - g)(x) \right| > \alpha \quad \text{for all } x \in F_\alpha.$$  

But this means $F_\alpha \subseteq \{H(f - g) > \alpha\}$, and now this has measure at most $C\varepsilon/\alpha$ by Theorem 1.5. Since $\varepsilon$ was arbitrary, any such $F_\alpha$ must actually have measure zero.

Important feature of this proof: choose $\alpha$ first, then $\varepsilon$, and then $g$. In the end we learn that $mF_\alpha = 0$ for any $\alpha$, so the choices of $\varepsilon$ and $g$ disappear. \hfill \Box

3