Math 202B — UCB, Spring 2014 — M. Christ

See §3.4 of our text for this material.

1 Differentiation Theorem

Arbitrary Lebesgue measurable sets, and arbitrary (but locally integrable) Lebesgue measurable functions, possess a degree of regularity. One manifestation of this is Lusin's Theorem. If $f : [a,b] \to \mathbb{C}$ is measurable, then for any $\varepsilon > 0$, there exists a continuous function $\varphi : [a,b] \to \mathbb{C}$ such that the measure of $\{x : f(x) \neq \varphi(x)\}$ is $< \varepsilon$. This lecture is about a different form of regularity.

Denote by $B(x,r)$ the ball in $\mathbb{R}^d$ centered at x with radius $r \in (0,\infty)$. The quantity

$$m(B(x,r))^{-1} \int_{B(x,r)} f(y)\, dm(y)$$

represents an average of f over $B(x,r)$. This quantity is defined for all locally integrable Lebesgue measurable functions; recall that local integrability means that $\int_{|x|\le R} |f(x)|\, dm(x) < \infty$ for all finite radii R. Recall that $m(B(x,r)) = \omega_d r^d$ where ω_d is a dimensional constant.

Theorem.

- For any Lebesgue measurable set $E \subset \mathbb{R}^d$,

$$\lim_{r\to 0^+} \frac{m(E\cap B(x,r))}{m(B(x,r))} \text{ exists and equals } \begin{cases} 1 & \text{for almost every } x \in E \\ 0 & \text{for almost every } x \notin E. \end{cases}$$

- For any locally integrable Lebesgue measurable function f,

$$\lim_{r\to 0^+} m(B(x,r))^{-1} \int_{B(x,r)} f\, dm \text{ exists and equals } f(x) \quad \text{for almost every } x \in \mathbb{R}^d.$$

□

The first conclusion says that for any E, if x is almost any point of E, then in any sufficiently small ball centered at x, the overwhelming majority of all points belong to E. Moreover, if x is almost any point of the complement of E, then the reverse holds.

The conclusion obviously fails at some points: If $E = [a,b] \subset \mathbb{R}^1$, and if $x = a$ or $x = b$, then the ratio $m(E \cap B(x,r))/2r$ equals $\frac{1}{2}$ for all $r < (b-a)$.

This theorem has lots of applications. A typical one is this: If $f \in L^1(\mathbb{R}^d)$ or $f \in L^\infty(\mathbb{R}^d)$, if $u(t,x) = f * h_t(x)$ is the solution of the heat equation that we constructed last week, then $\lim_{t\to 0^+} u(t,x) = f(x)$ for almost every $x \in \mathbb{R}^d$. (Recall that we had proved that $u(t,\cdot) \to f$ in the sense that $\|u(t,\cdot) - f\|_{L^1} \to 0$ as $t \to 0+$ if $f \in L^1$, but we had said nothing about pointwise convergence for discontinuous functions.)

The theorem has a stronger form: If f is locally integrable then

$$\lim_{t\to 0^+} m(B(x,r))^{-1} \int_{B(x,r)} |f(y) - f(x)|\, dm(y) = 0$$

for almost every $x \in \mathbb{R}^d$. $\square$

I believe that in the Fall, you learned a version of this theorem for $d = 1$: $F(x) = \int_{[0,x]} f\,dm$ is differentiable almost everywhere (in the Math 1A sense), and its derivative equals $f(x)$. The connection is that (for $h > 0$)

$$h^{-1}(F(x+h) - F(x)) = h^{-1} \int_{[x,x+h]} f$$

is almost, though not quite, one of the averages in the above Theorem. A closer connection can be found by noting that

$$\big(h^{-1} \int_{[x,x+h]} f\,dm\big) - f(x) = h^{-1} \int_{[x,x+h]} (f(y) - f(x))\,dm(y)$$

and consequently

$$\begin{aligned} \big|\big(h^{-1} \int_{[x,x+h]} f\,dm\big) - f(x)\big| &\le h^{-1} \int_{[x,x+h]} |f(y) - f(x)|\,dm(y) \\ &\le 2(2h)^{-1} \int_{[x-h,x+h]} |f(y) - f(x)|\,dm(y) \\ &= 2 \cdot m(B(x,h))^{-1} \int_{B(x,h)} |f(y) - f(x)|\,dm(y); \end{aligned}$$

the stronger third form of the theorem asserts that this quantity tends to zero for almost every x. $\square$

2 Hardy-Littlewood maximal function

The key to an approach to this theorem that works in higher dimensions is the *Hardy-Littlewood maximal function*, which is defined by replacing the limit $r \to 0^+$ by a supremum over all r.

Definition. For $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ and each $x \in \mathbb{R}^d$,

$$M(f)(x) = \sup_{r>0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f|\,dm.$$

$\square$

The first thing to notice is that $f \mapsto M(f)$ is not a linear operator. It is *sublinear*:

$$\begin{aligned} M(f+g) &\le M(f) + M(g) \\ M(tf) &= |t| M(f) \end{aligned}$$

for any locally integrable (Lebesgue measurable) functions f, g and any $t \in \mathbb{C}$. $\square$

Obviously $M(f)(x) \le \|f\|_{L^\infty}$ for all $f \in L^\infty$ and all points x. The most important property of M is:

Maximal Theorem. Let $d \geq 1$. There exists $A < \infty$ such that for all $f \in L^1(\mathbb{R}^d)$ and all $\alpha > 0$,

$$m(\{x : M(f)(x) > \alpha\}) \leq A\alpha^{-1}\|f\|_{L^1}.$$

□

Recall the inequality of Markov and/or Chebyshev: If $g \in L^1$ then

$$m(\{x : |g(x)| > \alpha\}) \leq A\alpha^{-1}\|g\|_{L^1}.$$

Thus the conclusion of the theorem is exactly what one would conclude from Markov's inequality, *if* one knew that $\|M(f)\|_{L^1} \leq A\|f\|_{L^1}$ for all $f \in L^1$. Thus the form of this conclusion is rather natural.

It would be nice to know that $\|M(f)\|_{L^1} \leq A\|f\|_{L^1}$ for all $f \in L^1$, but this is wholly and irretrievably false; in fact, if $M(f) \in L^1$ then $f = 0$ almost everywhere. So the weaker conclusion represents the strongest statement that is actually true. □

3 Proof of differentiation theorem

Proof of the first theorem, using the second: I'll prove the third form of the conclusion.

One may restrict attention to radii $r \leq 1$ (or any other positive quantity), since only the limit as $r \to 0$ does not depend on larger values. It suffices to show that for any $R < \infty$, the conclusion holds for almost all $x \in B(0, R)$. For such x, $\int_{B(x,r)} |f(y) - f(x)|\, dm(y)$ depends only on the restriction of f to $B(0, R + r)$ and therefore for $0 < r \leq 1$, depends only on the restriction of f to $B(0, R + 1)$. Therefore it suffices to consider $\tilde{f} = f\mathbf{1}_{B(0,R+1)}$. If f is locally integrable, $\tilde{f} \in L^1$.

If φ is continuous then obviously $m(B(x, r))^{-1} \int_{B(x,r)} \varphi \to \varphi(x)$ as $r \to 0$, for every x, for as we have already seen,

$$\begin{aligned}\left|\varphi(x) - \frac{1}{m(B(x,r))}\int_{B(x,r)} \varphi(y)\, dm(y)\right| &\leq \frac{1}{m(B(x,r))}\int_{B(x,r)} \big|\varphi(x) - \varphi(y)\big|\, dm(y) \\ &\leq \max_{|y-x|\leq r} |\varphi(y) - \varphi(x)|,\end{aligned}$$

which tends to zero if φ is continuous at x.

To treat a general $f \in L^1$, let $\varepsilon > 0$ and choose $\varphi \in C_c(\mathbb{R}^d)$ such that $\|f - \varphi\|_{L^1} < \varepsilon$. We have proved that such a function φ exists. Set $g = f - \varphi$. Then $f = \varphi + g$, and each summand has its own advantage; one is continuous, while the other is small. To exploit these advantages, consider

$$\limsup_{r\to 0^+} \left| f(x) - \frac{1}{m(B(x,r))}\int_{B(x,r)} f(y)\, dm(y)\right|$$

which is

$$\le \limsup_{r\to 0^+} \left| g(x) - \frac{1}{m(B(x,r))} \int_{B(x,r)} g(y)\, dm(y) \right|$$
$$+ \limsup_{r\to 0^+} \left| \varphi(x) - \frac{1}{m(B(x,r))} \int_{B(x,r)} \varphi(y)\, dm(y) \right|$$
$$\le |g(x)| + \limsup_{r\to 0^+} \frac{1}{m(B(x,r))} \int_{B(x,r)} |g(y)|\, dm(y)$$
$$\le |g(x)| + \sup_{r>0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |g(y)|\, dm(y)$$
$$= |g(x)| + M(g)(x),$$

using first the result shown above for the continuous function φ, then the definition of M.

Note that we have sacrificed information by replacing a lim sup by a sup. The idea is that because g is small, even this sup will be manageably small.

Let $\alpha > 0$ be arbitrary, and set

$$E_\alpha = \{x : \limsup_{r\to 0^+} \left| f(x) - \frac{1}{m(B(x,r))} \int_{B(x,r)} f(y)\, dm(y) \right| > \alpha\}.$$

We aim to prove that

$$m(E_\alpha) = 0$$

for every $\alpha > 0$. Therefore

$$m\big(\{x : \limsup_{r\to 0^+} \left| f(x) - \frac{1}{m(B(x,r))} \int_{B(x,r)} f(y)\, dm(y) \right| > 0\}\big) = 0,$$

which is the conclusion of the Theorem.

If $x \in E_\alpha$ then $|g(x)| + M(g)(x) > \alpha$, so $|g(x)| > \frac{1}{2}\alpha$ or $M(g)(x) > \frac{1}{2}\alpha$ (or both, of course). Therefore

$$E_\alpha \subset \{x : |g(x)| > \tfrac{1}{2}\alpha\} \bigcup \{x : M(g)(x) > \tfrac{1}{2}\alpha\}.$$

Now

$$m(\{x : |g(x)| > \tfrac{1}{2}\alpha\}) \le 2\alpha^{-1}\|g\|_{L^1}$$

by Markov's inequality, while

$$m(\{x : |M(g)(x)| > \tfrac{1}{2}\alpha\}) \le 2A\alpha^{-1}\|g\|_{L^1}$$

by the Hardy-Littlewood maximal theorem. Therefore in all,

$$m(E_\alpha) \le 2\alpha^{-1}\|g\|_1 + 2A\alpha^{-1}\|g\|_1 \le 2(A+1)\alpha^{-1}\varepsilon.$$

This holds for every $\varepsilon > 0$. Therefore $m(E_\alpha) = 0$. □

Note how essential it is that the constant A does not depend on the function g; this ensures that the product $(A+1)\varepsilon$ can be made arbitrarily small by choosing ε to be arbitratily small; this could break down if A were to depend on g and thereby on ε.

On Wednesday 4/30 we finished here.

---

4 Proof of the Maximal Theorem

4.1 The one-dimensional case

Assume for now that the dimension d equals 1.

Covering Lemma. Let $K \subset \mathbb{R}$ be a compact set. Let $\{I_\gamma : \gamma \in \mathcal{C}\}$ be an open cover of K by intervals I_γ. Then there exists a finite subcover $\{I_{\gamma_j}\}$ for K such that no point of $\mathbb{R}$ is contained in more than two of the intervals I_{γ_j}. □

Assuming the Lemma for now, here is a proof of the Maximal Theorem: Let $f \in L^1$ and let $\alpha > 0$. Let $E_\alpha = \{x \in \mathbb{R} : M(f)(x) > \alpha\}$. Let K be any compact subset of E_α.

For any $x \in E_\alpha$ there exists $r > 0$ such that $m(B(x,r))^{-1} \int_{B(x,r))} |f|\, dm > \alpha$. Choose such an $r = r_x$, and set $I_x = B(x, r_x)$ (where $B(x,r)$ denotes an *open* ball).

Now $\mathcal{C} = \{I_x : x \in K\}$ is an open cover of K. Let $\{J_j : 1 \le j \le N\}$ be a finite subcover such that no point of K is contained in more than 2 intervals J_j.

By the criterion applied in choosing the intervals I_x,

$$m(J_j)^{-1} \int_{J_j} |f|\, dm > \alpha \text{ for every index } j.$$

This condition can be equivalently written as an upper bound for $m(J_j)$:

$$m(J_j) \le \alpha^{-1} \int_{J_j} |f|\, dm.$$

Now

$$\begin{aligned}
m(K) &\le \sum_j m(J_j) \\
&\le \sum_j \alpha^{-1} \int_{J_j} |f|\, dm \\
&= \sum_j \alpha^{-1} \int |f| \mathbf{1}_{J_j}\, dm \\
&= \alpha^{-1} \int |f| \cdot \sum_j \mathbf{1}_{J_j}\, dm \\
&= \alpha^{-1} \int |f| \cdot 2\, dm \\
&= 2\alpha^{-1} \|f\|_1.
\end{aligned}$$

Since Lebesgue measure is inner regular, $m(E_\alpha)$ equals the supremum of $m(K)$ over all compact subsets K. Thus we have proved that

$$m(E_\alpha) \le 2\alpha^{-1} \|f\|_{L^1}.$$

□

4.2 Proof of the Covering Lemma (1D)

Suppose first that K is a closed bounded interval $[a, b]$. We may assume that $\mathcal{C}$ is a finite collection; for if not, just replace it by a finite subcover. We will prove by induction on the number of elements of $\mathcal{C}$ that there exists a subcover $\mathcal{C}'$ of $[a, b]$ such that no point is contained in more than two elements of $\mathcal{C}'$, *and* each endpoint a, b is contained in only one element of $\mathcal{C}'$.

First, discard from $\mathcal{C}$ all intervals that are disjoint from $[a, b]$. Then choose an interval $I^* \in \mathcal{C}$ that contains a, whose right endpoint c is maximal among all right endpoints of intervals in $\mathcal{C}$ that contain a.

If $c > b$ then this single interval constitutes a cover with the required properties. Otherwise consider the interval $[c, b]$, for which the collection $\tilde{\mathcal{C}} = \mathcal{C} \setminus \{I^*\}$ is a finite open cover. Discard all intervals in $\tilde{\mathcal{C}}$ that are disjoint from $[c, b]$, and still call the resulting subcollection $\tilde{\mathcal{C}}$. Since $\tilde{\mathcal{C}}$ has fewer elements than $\mathcal{C}$, by the induction hypothesis there exists a subcover $\tilde{\mathcal{C}}'$ of $[c, b]$ such that no point of $\mathbb{R}$ belongs to more than 2 elements of $\tilde{\mathcal{C}}'$, and c, b each belong to only one.

I claim that $\mathcal{C}' = \tilde{\mathcal{C}}' \cup \{I^*\}$ is a cover for $[a, b]$ with the required properties. It covers $[a, c)$ and $[c, b]$, so it is a cover. Any interval $J \in \tilde{\mathcal{C}}$ must be disjoint from a, for J intersects $[c, b]$, so the right endpoint of J is $> c$; if $a \in J$ then J would have been chosen in place of I^*.

b lies in only one of the intervals in $\tilde{\mathcal{C}}'$, and does not lie in I^*, so lies in only one element of $\mathcal{C}'$.

If $x \in [c, b]$ then $x \notin I^*$, and x belongs to at most two elements of $\tilde{\mathcal{C}}'$, so x belongs to at most two elements of $\mathcal{C}'$.

If $x \in [a, c)$ then $x \in I^*$, so we need to show that x belongs to at most *one* interval $J \in \tilde{\mathcal{C}}'$. Any interval $J \in \tilde{\mathcal{C}}'$ containing x, must intersect $[c, b]$ and therefore must contain c. There can be only one such interval. □

If K is a general compact set then let $\tilde{K}$ be the smallest closed bounded interval that contains K. Create a cover of $\tilde{K}$ by open intervals, by augmenting the given cover $\mathcal{C}$ with all open intervals contained in $\mathbb{R} \setminus K$. Apply the special case proved above to $\tilde{K}$, to produce a cover of $\tilde{K}$ by open intervals, with the required overlap property. Some of the intervals in this cover may not belong to the original cover; those are contained entirely in $\mathbb{R} \setminus K$, so may be discarded. The resulting subcollection is a cover of K with the required properties. □

4.3 Higher dimensions

Now consider any dimension $d \geq 1$. The statement of the Covering Lemma generalizes to $\mathbb{R}^d$ (with 2 replaced by a larger number that depends on the dimension), but the proof is more intricate. A more common approach uses a different type of covering lemma.

To any open ball $B = B(x, r)$, associate the enlarged ball $B^\star = B(x, 3r)$.

Vitali Covering Lemma. Let $d \geq 1$. Let $\{B_j\}$ be a finite collection of open balls in $\mathbb{R}^d$. There exists a pairwise disjoint subcollection $\{\tilde{B}_i\}$ such that for each j there exists i such that $B_j \subset \tilde{B}_i^\star$. □

The Maximal Theorem is easily deduced from this lemma, using the fact that

$$\cup_j B_j \subset \cup_i \tilde{B}_i^\star$$

and consequently

$$m(\cup_j B_j) \le m(\cup_i \tilde{B}_i^\star) \le \sum_i m(\tilde{B}_i^\star) \le \sum_i 3^d m(\tilde{B}_i).$$

Now with the notation of the argument given above for $d = 1$,

$$\begin{aligned} m(K) &\le \sum_i m(\tilde{B}_i^\star) \\ &= 3^d \sum_i m(\tilde{B}_i) \\ &\le 3^d \alpha^{-1} \sum_i \int_{\tilde{B}_i} |f| \\ &= 3^d \alpha^{-1} \int_{\cup_i \tilde{B}_i} |f| \\ &\le 3^d \alpha^{-1} \|f\|_1. \end{aligned}$$

This sum can be controlled as above since in the proof of the maximal theorem one will have $m(\tilde{B}_i) \le \alpha^{-1} \int_{\tilde{B}_i} |f| \ldots$ □

Proof of the Vitali lemma: Let $\{B_j\}$ be given. Reorder these so that $|B_j| \ge |B_{j+1}|$ for all j.

Apply the following recursive selection procedure, in which $B_1, B_2, B_3, \ldots$ are examined in sequence. Each is either selected, or rejected, upon being examined. The procedure then moves on from B_j to B_{j+1}. $\{B_i^*\}$ is the set of all balls that are selected.

Select B_1. Let $n \ge 2$ and suppose that $B_1, B_2, \ldots, B_{n-1}$ have been examined. If B_n intersects some ball $\tilde{B}_i$ that was selected at an earlier step, then B_n is rejected. Otherwise B_n is selected.

The procedure halts after finitely many steps, since the list of balls is finite. The construction certainly ensures that the selected balls are pairwise disjoint.

If B_n is a rejected ball then there exists a ball $\tilde{B}_i$ that intersects B_n and was selected at an earlier step. Therefore the radius of $\tilde{B}_i$ is $\ge$ the radius of B_n. *It follows from the triangle inequality that*

$$B_n \subset \tilde{B}_i^\star.$$

Therefore

$$\cup_i \left\{ \tilde{B}_i^\star \right\} \supset \cup_j B_j \supset K;$$

$\left\{ \tilde{B}_i \star \right\}$ is a cover for K. □