

March 4, 2021

We'll finish our discussion of the Hahn-Banach theorem today – recall that this theorem tells us that a bounded linear functional on a subspace of a normed space satisfying $|u(t)| \leq C\|t\|$ (on the subspace) can be extended to a bounded linear functional on the whole space with the same bound. The proof is important to see, but what's more important is how we can use it as a tool. We mentioned that we can show that the dual of ℓ^∞ is not ℓ^1 , and here's something else we can do:

Theorem 50

Let V be a normed space. Then for all $v \in V \setminus \{0\}$, there exists an $f \in V'$ (a bounded linear functional) with $\|f\| = 1$ and $f(v) = \|v\|$.

Proof. First, define the linear map $u : \mathbb{C}v \rightarrow \mathbb{C}$ (here, $\mathbb{C}v$ denotes the span of v) by defining $u(\lambda v) = \lambda\|v\|$ (this is well-defined because every element in the span of v can be uniquely represented this way, and it's also clearly linear because only λ is varying). Then it is indeed true that $|u(t)| \leq \|t\|$ for all $t \in \mathbb{C}v$, and also $u(v) = \|v\|$. Therefore, by Hahn-Banach, there exists an element of the dual space f extending u , such that $\|f(t)\| \leq \|t\|$ for all t . So we've found a linear functional so that $f(v) = u(v) = \|v\|$, and also with operator norm 1 (we know it is exactly 1 because we have equality when applying f to $\frac{v}{\|v\|}$), and we're done. \square

Definition 51

The **double dual** of a normed space V , denoted V'' , is the dual of V' .

In other words, V'' is the set of bounded linear functionals on the set of bounded linear functionals on V .

Example 52

Fix an element $v \in V$, and define the element $T_v : V' \rightarrow \mathbb{C}$ by setting

$$T_v(v') = v'(v)$$

for all linear functionals $v' \in V'$. Then T_v is an element of the double dual.

To check this, we should make sure T_v is linear in the argument v' , and this is true because we're applying functionals to a fixed v :

$$T_v(v'_1 + v'_2) = (v'_1 + v'_2)(v) = v'_1(v) + v'_2(v).$$

We should also check that T_v is bounded: indeed,

$$|T_v(v')| = |v'(v)| \leq \|v\| \cdot \|v'\|$$

(because v' is some bounded linear functional with norm $\|v'\|$). And since $\|v\|$ is a constant, we've found that the norm of T_v is at most $\|v\|$, and thus T_v is indeed in the double dual of V .

Definition 53

Let V and W be normed spaces. A bounded linear operator $T \in \mathcal{B}(V, W)$ is **isometric** if for all $v \in V$, $\|Tv\| = \|v\|$.

Theorem 54

Let $v \in V$, and define the element $T_v : V' \rightarrow \mathbb{C}$ of the double dual via $T_v(v') = v'(v)$. Then $T : V \rightarrow V''$ sending v to T_v is isometric.

Proof. We've already done a lot of the work here: we showed already that $v \mapsto T_v$ is a bounded linear operator (noting that $T_v(v')$ is linear in both v and in v'). So the map T sending $v \mapsto T_v$ is in $\mathcal{B}(V, V'')$, and we just need to show that it is isometric.

Since $\|T_v\| \leq \|v\|$ from our work above, we know that $\|T\| \leq 1$, and it suffices to show equality for all v . It's clear that $\|T_0\| = \|0\|$, and now if $v \in V \setminus \{0\}$ is a nonzero vector, then there exists some $f \in V'$ such that $\|f\| = 1$ and $f(v) = \|v\|$ (by Theorem 50). So now

$$\|v\| = f(v) = |f(v)| = |T_v(f)| \leq \|T_v\| \cdot \|f\|,$$

and thus $\|v\| \leq \|T_v\|$. Putting this together with the reverse inequality above yields the result – $\|T_v\| = \|v\|$, and thus T is isometric. \square

Notice that isometric bounded operators are one-to-one, because the only thing that can be sent to the zero vector is the zero vector if lengths are preserved. It's natural to ask whether operators are also onto (surjective), and there is a special categorization for that:

Definition 55

A Banach space V is **reflexive** if $V = V''$, in the sense that the map $v \mapsto T_v$ is onto.

Example 56

For all $1 < p < \infty$, we know that ℓ^p is reflexive (since the dual of ℓ^p is ℓ^q , whose dual is ℓ^p again). But ℓ^1 is not reflexive, because the dual of its dual ℓ^∞ is not ℓ^1 . And the space c_0 of sequences converging to 0 is also not reflexive – we can identify $(c_0)'$ with ℓ^1 , whose dual is ℓ^∞ .

With this, we've concluded our general discussion about Banach spaces, and we are now moving to **Lebesgue measure and integration**. We've been talking about ℓ^p spaces so far on sequences, and it makes sense to try to define L^p spaces on functions in a similar way. But using Riemann integration isn't quite good enough – Lebesgue integration has better convergence theorems, in the sense that they're more widely useful. And for a concrete example, consider the space of Riemann integrable functions on $[0, 1]$

$$L^1_{\mathbb{R}}([0, 1]) = \{f : [0, 1] \rightarrow \mathbb{C} : f \text{ Riemann integrable on } [0, 1]\}.$$

(We integrate a complex-valued function by integrating the real and imaginary parts separately here.) We may try to define a norm via

$$\|f\|_1 = \int_0^1 |f(x)| dx$$

(it's not quite a norm because we can have a function which is nonzero at only a single point, but let's pretend it's a norm), and the problem we'll run into is that we don't have a Banach space! So more general integration will help us get completeness, which is important for applications like differential equations.

To get the Lebesgue L^p spaces, we can take the **completion** of the $L^1_{\mathbb{R}}$ space that we defined above, much like the real numbers can be defined as the completion of the rational numbers. But we can do things from the ground

up instead, and we'll indeed see along the way that the Riemann integrable functions are dense in the set of Lebesgue integrable functions.

Fact 57

Our goal is to make a new definition of integration that is more general than Riemann integration: it will still be a method of calculating area under a curve, but we'll build it up in a way that allows for more powerful formalism.

And the way we'll define this is to start with functions 1_E that are 1 on some set E and 0 otherwise, which we call **indicator functions**. We'll get to definitions and theorems in a second, but we know what we want those functions to integrate to in some special cases: if $E = [a, b]$, then the integral $\int 1_E(x)dx = \int 1_{[a,b]}(x)dx$ should be the area under the curve, which is $b - a$. So the way we'll define integrals over more complicated functions E to look a lot like the "length" of E , and that's more generally going to be called the **measure** $m(E)$ of E .

In other words, the first step of getting an integral defined is to get a measure defined on subsets of \mathbb{R} , and this is what will be called the **Lebesgue measure**. From our discussion above, there are a **few properties** of this Lebesgue measure that we already know we want to have:

1. We want to be able to measure any subset of the real numbers (because Riemann integration can't deal with functions like $1_{\mathbb{Q}}$). In other words, we want to define the function m on $\mathcal{P}(\mathbb{R})$, the powerset of \mathbb{R} .
2. As a sanity check, if I is an interval, $m(I)$ should be the length of I (and the measure shouldn't care about whether we have open or closed intervals).
3. The measure of a whole set should be the sum of the measures of its chunks: more formally, if $\{E_n\}$ is a countable collection of disjoint sets and $E = \cup_n E_n$, then we want $m(E) = \sum_n m(E_n)$.
4. Translation invariance should hold: if E is a subset of \mathbb{R} , and $x \in \mathbb{R}$ is some constant, then $m(x + E) = m(E)$.

But unfortunately, even these four properties are impossible to satisfy at the same time – it turns out that there is **no function** $m : \mathcal{P}(\mathbb{R}) \rightarrow [0, \infty]$ that satisfies these conditions! (We can search up the **Vitali construction** for more details.) So what we'll do is to drop the first assumption – we'll try to define a function m on only some of the subsets of \mathbb{R} , while still satisfying properties (2), (3), (4), and we'll show that the set of such **Lebesgue measurable sets** is indeed pretty large.

The strategy for doing this comes from Caratheodory: we'll first define a function $m^* : \mathcal{P}(\mathbb{R}) \rightarrow [0, \infty)$ called the **outer measure**, which satisfies conditions (2), (4), and "almost (3)," and then we'll restrict m^* to appropriately well-behaved subsets of \mathbb{R} to get our actual construction.

Definition 58

For any interval $I \subset \mathbb{R}$, let $\ell(I)$ denote its length (regardless of whether it is open, closed, or half-closed). For any subset $A \subset \mathbb{R}$, we define the **outer measure** $m^*(A)$ via

$$m^*(A) = \inf \left\{ \sum_n \ell(I_n) : \{I_n\} \text{ countable collection of open intervals with } A \subset \bigcup_n I_n \right\}.$$

(Through this definition, we can see that $m^*(A) \geq 0$ for all A .)

Basically, we can cover any subset of the real numbers with a union of open intervals, and we take the minimum possible length over all coverings. (The idea is that as we make the intervals smaller, we can get more information about the subset A , and the infimum gives us the best possible information about "how much length" is in A .)

Example 59

Consider the set $A = \{0\}$ containing just a single point.

Then $m^*(\{0\}) = 0$, because we can cover $\{0\}$ with the interval $(-\frac{\epsilon}{2}, \frac{\epsilon}{2})$ for any $\epsilon > 0$, and this interval has measure ϵ . So $0 \leq m^*(\{0\}) \leq \epsilon$ for all ϵ , and taking $\epsilon \rightarrow 0$ gives us $m^*(\{0\}) = 0$. A similar argument showing that any finite set of points has measure zero, and in fact the measure of a countable set is always zero:

Theorem 60

If $A \subset \mathbb{R}$ is countable, then $m^*(A) = 0$.

For example, even though there are a lot of rational numbers and they are dense in \mathbb{R} , we're saying that they don't actually fill up a lot of space – the measure of \mathbb{Q} is zero.

Proof. (We can check the case where A is finite ourselves.) If A is countably infinite, then there is a bijection from A to \mathbb{N} , so we can enumerate the elements as $\{a_1, a_2, a_3, \dots\} = \{a_n : n \in \mathbb{N}\}$. Pick some $\epsilon > 0$ – we'll show that $m^*(A) \leq \epsilon$.

For each $n \in \mathbb{N}$, let I_n be the interval $(a_n - \frac{\epsilon}{2^{n+1}}, a_n + \frac{\epsilon}{2^{n+1}})$. Because I_n is an interval containing a_n , the set A must be contained in the (countable) union of intervals I_n , and then the outer measure is an infimum over all possible unions, so

$$m^*(A) \leq \sum_n \ell(I_n) = \sum_n \frac{\epsilon}{2^n} = \epsilon.$$

Finally, taking $\epsilon \rightarrow 0$ yields the result. □

We can now talk about what it means for the outer measure to “almost satisfy (3)” in the set of properties above, and the argument is pretty similar to what we did just now. But first, we establish a quick fact:

Lemma 61

If $A \subset B$, then $m^*(A) \leq m^*(B)$.

Proof. Any covering of B is a covering of A , so the infimum (of interval length sums) over all coverings of A should be at most the infimum over all coverings of B . □

Theorem 62

Let $\{A_n\}$ be a countable collection of subsets of \mathbb{R} , not necessarily disjoint. Then

$$m^*\left(\bigcup_n A_n\right) \leq \sum_n m^*(A_n).$$

(This is basically “half” of the additivity condition that we wanted.)

Proof. First of all, if there is some n such that $m^*(A_n) = \infty$ (meaning we can't cover the set by a collection of intervals whose sum of lengths is finite), or if $\sum_n m^*(A_n) = \infty$, then the inequality is true (because the right-hand side is already ∞). So we can just consider the case where all of the outer measures of A_n are finite, and the sum of those outer measures also converges.

The strategy here is going to come up frequently: instead of proving an inequality of the form $X \leq Y$ (for two quantities X and Y), we can equivalently prove that $X \leq Y + \varepsilon$ for any $\varepsilon > 0$. We'll do that here: fix some $\varepsilon > 0$, and now define the collection $\{I_{nk}\}_{k \in \mathbb{N}}$ of intervals to be a covering of A_n with total length $\sum_{k=1}^{\infty} \ell(I_{nk}) < m^*(A_n) + \frac{\varepsilon}{2^n}$ (we can't always achieve the infimum given by the outer measure, but we can always achieve a slightly larger number). Now because A_n is covered by $\{I_{nk}\}_k$, the union of the A_n s must be contained in the union $\cup_{n,k \in \mathbb{N}} I_{nk}$ (which is indeed a countable union of intervals as well). Thus,

$$m^* \left(\bigcup_n A_n \right) \leq \sum_{n,k} \ell(I_{nk}) = \sum_n \sum_k \ell(I_{nk})$$

and now we can sum over k to find that this is

$$< \sum_n \left(m^*(A_n) + \frac{\varepsilon}{2^n} \right) = \sum_n m^*(A_n) + \varepsilon.$$

Taking $\varepsilon \rightarrow 0$ gives the desired result. □

In particular, we should notice the similarities in this proof with the one in Theorem 60 – the previous proof we did was basically a special case where each A_n was a single point.

In our homework, we'll be able to check that the outer measure is indeed translation-invariant (so it satisfies (4)), and it seems like the next step is to show that $m^*(I) = \ell(I)$ for an interval I (so (2) is also satisfied). This may be intuitive, but it'll take a bit of work to show! So that'll be the first thing we do next lecture, and it'll complete our construction of the outer measure and allow us to define the Lebesgue measure.

MIT OpenCourseWare
<https://ocw.mit.edu>

18.102 / 18.1021 Introduction to Functional Analysis
Spring 2021

For information about citing these materials or our Terms of Use, visit: <https://ocw.mit.edu/terms>.