Lecture Notes of Functional Analysis - Part 1

Degree Course: Master's Program in Mathematics
Teacher: Sisto Baldo

These notes are just a fairly detailed summary of what went on in class. In no way they are meant as a replacement for actual classes, human interaction with the teacher, and/or the reading of reference texts, You are of course strongly encouraged to take advantage of ALL these different learning opportunities.

Contents

1 Lecture of october 1, 2015 (2 hours)

The lecture began with a very brief presentation of the course: syllabus, exams, learning material...

We will begin by recalling (or introducing) the basics of Lebesgue measure and integration theory. Meanwhile, and with very little additional effort, we will learn about abstract measures and integrals. For this part of the course, most proofs were omitted in class because most of the students already covered the subject in bachelor classes.

In your previous calculus courses, you probably saw the definition of Peano-Jordan measure, which is one of the simplest and most natural methods of defining (in a rigorous way) the area of a subset of the plane, or the volume of a subset of $\mathbf{R}^3$...

Let us recall the main definitions:

DEFINITION: An *interval* or *rectangle* in $\mathbf{R}^n$ is a subset $I \subset \mathbf{R}^n$ which is the cartesian product of 1-dimensional intervals: $I = (a_1, b_1) \times (a_2, b_2) \times \ldots \times (a_n, b_n)$. We allow one or both the endpoints of these 1-dimensional intervals to be included, and we also allow empty or degenerate intervals. The *measure* of the interval I above is, by definition, the number

$$|I| = \prod_{i=1}^{n} (b_i - a_i).$$

One readily checks that for $n = 2$ our interval is a rectangle with edges parallel to the axes (and its measure coincides with its area), while for $n = 3$ I will be a rectangular prism, and the measure is simply the volume.

A subset A of $\mathbf{R}^n$ is called Peano-Jordan measurable if its "n-dimensional volume" can be approximated, both from within and from without, by means of *finite unions* of intervals. More precisely, we have the following

DEFINITION (Measurable set in the sense of Peano-Jordan): A subset $A \subset \mathbf{R}^n$ is measurable in the sense of Peano-Jordan if it is bounded and for every $\varepsilon > 0$ there are a finite number of intervals $I_1, \ldots, I_N, J_1, \ldots, J_K \subset \mathbf{R}^n$ such that I_i have pairwise disjoint interiors, J_i have pairwise disjoint interiors,

$$\bigcup_{i=1}^{N} I_i \subset A \subset \bigcup_{i=1}^{K} J_i$$

and finally

$$\sum_{i=1}^{K} |J_i| - \sum_{i=1}^{N} |I_i| \leq \varepsilon.$$

If this is the case, we define the *Peano-Jordan measure* of A as

$$\begin{aligned}|A| = & \quad \sup\{\sum_{i=1}^{N} |I_i| : I_i \textit{ with pairwise disjoint interiors}, \ \bigcup_{i=1}^{N} I_i \subset A\} \\ = & \quad \inf\{\sum_{i=1}^{K} |J_i| : J_i \textit{ with pairwise disjoint interiors}, \ \bigcup_{i=1}^{K} J_i \supset A\}.\end{aligned}$$

In the last expression, we can drop the requirement that the intervals J_i have pairwise disjoint interiors: the infimum takes care of that!

We immediately check that rectangles are Peano-Jordan measurable, while the set of points with rational coordinate within a rectangle is not.

Likewise, given a Riemann-integrable function $f : [a, b] \to \mathbf{R}$, $f \geq 0$, the region bounded by the graph of f, the x axis and the lines $x = a$, $x = b$ is Peano-Jordan measurable, and its measure coincides with the Riemann integral of f. More generally:

EXERCISE: Let $g, h : [a, b] \to \mathbf{R}$ be two Riemann-integrable functions of one variable with $g(x) \leq h(x)$ for all $x \in [a, b]$. Consider the set$A = \{(x, y) \in \mathbf{R}^2 : \ x \in [a, b], \ g(x) \leq y \leq h(x)$. Show that A is Peano-Jordan measurable and

$$|A| = \int_a^b (h(x) - g(x)) \, dx.$$

A set of this kind is called a *simple set* with respect to the x-axis... Simple sets with respect to the y-axis are defined in a similar way, and there are natural extensions in higher dimension.

Peano-Jordan measure is a nice object, but it behaves badly with respect to *countable* operations: it is certainly true that a finite union of P.J.-measurable sets is again P.J-measurable, but this is false for countable unions: for instance, the set of points with rational coordinates within a rectangle is a countable union of *points*, which are of course measurable.

For this and other reasons, it is convenient to introduce a more general notion of measure, which will be Lebesgue measure. The definition is very similar to that of P.J. measure, but we will allow *countable unions* of intervals.

DEFINITION (Outer Lebesgue Measure): If $A \subset \mathbf{R}^n$, its *outer Lebesgue measure* is defined as

$$m(A) = \inf\{\sum_{i=1}^{\infty} |I_i| : \ I_i \textit{ intervals}, \ \bigcup_{i=1}^{\infty} I_i \supset A\}.$$

Notice that we do not require that the intervals have disjoint interiors. Moreover, since we allow degenerate or empty intervals, finite coverings are possible.

Lebesgue measure enjoys of the following elementary properties:

THEOREM (Elementary properties of Lebesgue outer measure): Let $m : \mathcal{P}(\mathbf{R}^n) \to [0, +\infty]$ *denote (outer) Lebesgue Measure*[1]*. The following holds:*

(i) $m(\emptyset) = 0$, $m(\{x\}) = 0$ *for every* $x \in \mathbf{R}^n$.

(ii) If $A \subset \bigcup_{i=1}^{\infty} A_i$, *with* $A, A_1, A_2, \ldots \subset \mathbf{R}^n$, *then*

$$m(A) \leq \sum_{i=1}^{\infty} m(A_i)$$

(countable subadditivity of Lebesgue measure). In particular, $A \subset B$ *implies* $m(A) \leq m(B)$ *(monotonicity of Lebesgue measure).*

(iii) In the definition of Lebesgue measure, we may ask without loss of generality that the intervals I_i *are open.*

(iv) $m(I) = |I|$ *for every interval* $I \subset \mathbf{R}^n$*. Moreover,* $m(\mathbf{R}^n) = +\infty$.

DIM.: (i) is a simple exercise. To prove (ii), we begin by recalling that the sum of a series of non-negative numbers *does not depend* on the order of summation.

Fix $\varepsilon > 0$ and an index i: by definition of infimum, we find a sequence of intervals $\{I_j^i\}_j$ such that $\bigcup_{j=1}^{\infty} I_j^i \supset A_i$ and

$$\sum_{j=1}^{\infty} |I_j^i| < m(A_i) + \frac{\varepsilon}{2^i}.$$

Then $\{I_j^i\}_{i,j}$ is a countable covering of A whose members are intervals, and by definition of Lebesgue measure we get

$$m(A) \leq \sum_{i,j=1}^{\infty} |I_j^i| \leq \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |I_j^i| \leq \sum_{i=1}^{\infty} (m(A_i) + \frac{\varepsilon}{2^i}) = \sum_{i=1}^{\infty} m(A_i) + \varepsilon,$$

and (ii) follows because ε can be taken arbitrarily small.

Monotonicity is an immediate consequence of (ii). Let us show (iii): if $A \subset \mathbf{R}^n$, for every $\varepsilon > 0$ we can find intervals I_j such that $\bigcup_{j=1}^{\infty} I_j \supset A$ and

$$\sum_{j=1}^{\infty} |I_j| < m(A) + \frac{\varepsilon}{2}.$$

[1] $\mathcal{P}(\mathbf{R}^n)$ denotes the set of all subsets of $\mathbf{R}^n$.

For every $j = 1, 2, \ldots$ let $I'_j \supset I_j$ be a slightly larger *open* interval, chosen in such a way that $|I'_j| < |I_j| + \frac{\varepsilon}{2^{j+1}}$. We then get

$$\sum_{j=1}^{\infty} |I'_j| < \sum_{j=1}^{\infty} (|I_j| + \frac{\varepsilon}{2^{j+1}}) < m(A) + \frac{\varepsilon}{2} + \frac{\varepsilon}{2}$$

and (iii) is proved.

Surprisingly enough, (iv) is slightly harder to prove. By (iii), it immediately follows from the following

CLAIM: If I is an interval, then for every sequence I_j of open intervals with $\bigcup_{j=1}^{\infty} I_j \supset I$ one has

$$(*) \quad |I| \leq \sum_{j=1}^{\infty} |I_j|.$$

Indeed, (*) is easy enough if there is only a finite number of intervals I_j, a bit less in the general case of a countable covering. On the other hand, if $J \subset I$ is *closed and bounded*, we can use compactness to choose a finite number of intervals $I_1, I_2, \ldots, I_N$ in our covering in such a way that we still have $J \subset \bigcup_{j=1}^{N} I_j$. Since (*) holds for finite coverings, we deduce that

$$|J| \leq \sum_{j=1}^{N} |I_j| \leq \sum_{j=1}^{\infty} |I_j|.$$

But of course J can be chosen in such a way that its measure is arbitrarily close to that of I, and (*) is proved. Q.E.D.

As a trivial consequence of this theorem, every *countable* subset of $\mathbf{R}^n$ has measure 0, because points have measure 0 and Lebesgue measure is countably subadditive.

Lebesgue measure in $\mathbf{R}^n$ is only a particular case of a more general object, called an *outer measure*:

DEFINITION (Outer measure): An *outer measure* on a set X is a function $\mu : \mathcal{P}(X) \to [0, +\infty]$ such that $\mu(\emptyset) = 0$, and which is countably subadditive: if $A, A_1, A_2, A_3, \ldots \subset X$ and $A \subset \bigcup_{j=1}^{\infty} A_j$, then

$$\mu(A) \leq \sum_{j=1}^{\infty} \mu(A_j).$$

Of course, monotonicity of μ follows from countable subadditivity: if $A \subset B$ then $\mu(A) \leq \mu(B)$.

A new example of outer measure is obtained by *restricting* Lebesgue measure to a subset $A_0 \subset \mathbf{R}^n$: this is the measure $\tilde{m}$ defined as

$$\tilde{m}(A) := m(A \cap A_0).$$

Another example is the measure δ_0 (*Dirac's delta at* 0), which is the measure on $\mathbf{R}^n$ defined by

$$\delta_0(A) = \begin{cases} 1 & if\ 0 \in A, \\ 0 & otherwise. \end{cases}$$

Yet another examples is *counting measure* defined by

$$\#(A) = \begin{cases} number\ of\ elements\ in\ A & if\ A\ is\ finite, \\ +\infty & otherwise. \end{cases}$$

In general, Lebesgue mesaure does not enjoy of good properties on *every* subset of $\mathbf{R}^n$: it behaves much better on a particular class of sets, the *measurable sets*:

DEFINITION (Lebesgue-measurable sets in the sense of Caratheodory): A set $A \subset \mathbf{R}^n$ is *Lebesgue-measurable* or *m-measurable* if the following equality

$$m(T) = m(T \cap A) + m(T \setminus A)$$

holds for every subset $T \subset \mathbf{R}^n$. Roughly speaking, we are requiring that A "split well" the measure of every subset of $\mathbf{R}^n$.

Notice that due to countable subadditivity we always have $m(T) \leq m(T \cap A) + m(T \setminus A)$: so, to prove measurability it is enough to prove the opposite inequality

$$m(T) \geq m(T \cap A) + m(T \setminus A) \qquad \forall T \subset \mathbf{R}^n.$$

In the same way, given an outer measure μ, A is said to be μ-measurable if $\mu(T) = \mu(T \cap A) + \mu(T \setminus A)$ for every $T \subset X$.

REMARK: In the near future we will need the following fact: if $A \subset \mathbf{R}^n$ is Lebesgue-measurable and $\tilde{m}$ denotes the restriction of Lebesgue measure to *any* subset $A_0 \subset \mathbf{R}^n$, then A is also $\tilde{m}$-measurable. Indeed, if $T \subset \mathbf{R}^n$ we have

$$\tilde{m}(T) = m(T \cap A_0) = m((T \cap A_0) \cap A) + m((T \cap A_0) \setminus A) =$$
$$m((T \cap A) \cap A_0) + m((T \setminus A) \cap A_0) = \tilde{m}(T \cap A) + \tilde{m}(T \setminus A).$$

This is clearly still true, with the same proof, if m and $\tilde{m}$ are replaced with an arbitrary outer measure μ and its restriction $\tilde{\mu}$ to a set A_0.

The following theorem shows two main things: first, if we start with measurable sets and make countable unions, complements, contable intersections, we don't leave the category of measurable sets. Moreover, Lebesgue measure or an abstract outer measure show some very good properties when restricted to the measurable sets. The main of these is *countable additivity*: the measure of the union of a contable family of pairwise disjoint measurable sets is simply the *sum* of their measures.

THEOREM (Properties of measurable sets and of the measure restricted to measurable sets): Let μ be an outer measure on a set X. The following facts hold true:

(i) If A is μ-measurable, then $A^C = X \setminus A$ is μ-measurable. Moreover, if $\mu(A) = 0$ then A is μ-measurable.

(ii) Contable union or intersection of μ-measurable sets is μ-measurable.

(iii) If $\{A_i\}_i$ is a countable family of pairwise disjoint μ-measurable sets and $A = \bigcup_{i=1}^{\infty} A_i$, then

$$\mu(A) = \sum_{i=1}^{\infty} \mu(A_i)$$

(countable additivity of μ on measurable sets).

(iv) If $\{A_i\}$ is an increasing sequence of μ-measurable sets, i.e. if $A_1 \subset A_2 \subset A_3 \subset \ldots$, and $A = \bigcup_{i=1}^{\infty} A_i$ then

$$\mu(A) = \lim_{i \to +\infty} \mu(A_i).$$

(v) If $\{A_i\}$ is a decreasing sequence of μ-measurable sets, i.e. if $A_1 \supset A_2 \supset A_3 \supset \ldots$, if $\mu(A_1) < +\infty$ and if we denote $A = \bigcap_{i=1}^{\infty} A_i$, then

$$\mu(A) = \lim_{i \to +\infty} \mu(A_i).$$

PROOF: (i) is obvious if we remark that the measurability condition can be rewritten as

$$\mu(T) \geq \mu(T \cap A) + \mu(T \cap A^C) \qquad \forall T \subset X.$$

Also, it is trivial that a set with measure 0 is measurable. In particular, we deduce that $\emptyset$ and X are μ-measurable.

Let us show for the moment a weaker version of (ii): if A and B are measurable, then $A \cup B$ is measurable. Indeed, if $T \subset X$ we have

$$\mu(T) = \mu(T \cap A) + \mu(T \setminus A) =$$
$$\mu((T \cap A) \cap B) + \mu((T \cap A) \setminus B) + \mu((T \setminus A) \cap B) + \mu((T \setminus A) \setminus B).$$

Look at the last row: the union of the sets within the first 3 terms is exactly $T \cap (A \cup B)$: by the subadditivity of μ we then infer that the sum of those terms is $\geq \mu(T \cap (A \cup B))$. Since the set in the last term is simply $T \setminus (A \cup B)$, we then get:

$$\mu(T) \geq \mu(T \cap (A \cup B)) + \mu(T \setminus (A \cup B)),$$

and $A \cup B$ is measurable.

From this and (i) we also get the measurability of $A \cap B$, since $A \cap B = (A^C \cup B^C)^C$. By induction, if follows that *finite* union and intersections of measurable sets are measurable. We will complete the proof of (ii) (i.e., for *countable* union and intersections) only at the end.

Let us prove (iii): the claim is easy to show for the union of *two* disjoint measurable sets A and B, because $\mu(A \cup B) = \mu((A \cup B) \cap A) + \mu((A \cup B) \setminus A) = \mu(A) + \mu(B)$. By induction, (iii) holds for the union of a finite family of pairwise disjoint measurable sets.

In the general case of a countable family, countable subadditivity already gives $\mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i)$, while monotonicity ensures that for any $N \in \mathbf{N}$

$$\mu(A) \geq \mu\left(\bigcup_{i=1}^{N} (A_i)\right) = \sum_{i=1}^{N} \mu(A_i),$$

where the last equality holds because we proved (iii) for finite unions... By taking the supremum over all N we get

$$\mu(A) \geq \sum_{i=1}^{\infty} \mu(A_i),$$

and (iii) is proved.

Let us show (iv): it suffices to apply (iii) on the sequence of pairwise disjoint measurable sets defined by $B_1 = A_1$, $B_i = A_i \setminus A_{i-1}$ $(i \geq 2)$. We get

$$\mu(A) = \sum_{i=1}^{\infty} \mu(B_i) = \lim_{N \to +\infty} \sum_{i=1}^{N} \mu(B_i) = \lim_{N \to +\infty} \mu(A_N).$$

We now prove (v): we define the increasing sequence of measurable sets $B_i = A_1 \setminus A_i$, $i = 2, 3, \ldots$. It follows that

$$A_1 = A \cup \bigcup_{i=2}^{\infty} B_i$$

and by (iv) we get

$$\mu(A_1) \leq \mu(A) + \lim_{i \to +\infty} [\mu(A_1) - \mu(A_i)],$$

whence $\lim_{i \to +\infty} \mu(A_i) \leq \mu(A)$. The opposite inequality holds by monotonicity, and (v) is proved.

To conclude the proof of the theorem, we only need to show (ii).

Let $A = \bigcup_{i=1}^{\infty} A_i$, where the sets A_i are measurable. We must show that A is measurable.

Let $T \subset \mathbf{R}^n$ and consider the increasing sequence of measurable sets $B_N := \bigcup_{i=1}^{N} A_i$: these are measurable because of the *finite* version of (ii) we already proved. But they are also measurable for the outer measure $\tilde{\mu}$ obtained by taking the restriction of μ to the set T (i.e. the measure defined by $\tilde{\mu}(A) := \mu(T \cap A)$ for all $A \subset \mathbf{R}^n$). By monotonicity we get:

$$(***) \quad \mu(T) = \mu(T \cap B_N) + \mu(T \setminus B_N) \geq \mu(T \cap B_N) + \mu(T \setminus A)$$

On the other hand, by applying (iv) to the measure $\tilde{\mu}$ we obtain

$$\lim_{N \to +\infty} \mu(T \cap B_N) = \lim_{N \to +\infty} \tilde{\mu}(B_N) = \tilde{\mu}(A) = \mu(T \cap A)$$

and measurability of A follows by passing to the limit for $N \to +\infty$ in (***).
Measurability of $\bigcap_{i=1}^{\infty} A_i$ follows as usual by writing

$$\bigcap_{i=1}^{\infty} A_i = \left(\bigcup_{i=1}^{\infty} A_i^C \right)^C .$$

Q.E.D.

The family of measurable sets of an outer measure form what is called a σ-*algebra*. Moreover, an outer measure restricted to the family of its measurable sets is called a *measure*:

DEFINITION (σ-algebra, measure): Given a set X, a family $\mathcal{A} \subset \mathcal{P}(X)$ of subsets of X is called a σ-algebra if $X \in \mathcal{A}$, $A^C \in \mathcal{A}$ whenever $A \in \mathcal{A}$, and if $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ whenever $A_i \in \mathcal{A}$ for all $i = 1, 2, \ldots$.

Given X and a σ-algebra $\mathcal{A}$ on X, a measure is a function $\mu : \mathcal{A} \to [0, +\infty]$ such that $\mu(\emptyset) = 0$ and which is countably additive: if $A_1, A_2, A_3, \ldots$ belong to $\mathcal{A}$ and are pairwise disjoint, then

$$\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i).$$

After quite general results, which hold for all outer measures, les us go back for the moment to Lebesgue measure: the following theorem shows that there is plenty of Lebesgue measurable sets.

THEOREM (Regularity of Lebesgue measure): Open and closed subsets of $\mathbf{R}^N$ are Lebesgue-measurable. Moreover, if $A \subset \mathbf{R}^N$ is Lebesgue-measurable, then for every $\varepsilon > 0$ there exist B open, C closed with $C \subset A \subset B$ and $m(B \setminus C) < \varepsilon$.

To prove the theorem, we need an easy topological fact: the following proposition shows that every open subset of $\mathbf{R}^N$ is a countable union of intervals.

PROPOSITION: Every open set $A \subset \mathbf{R}^n$ is a countable union of open intervals.

DIM.: Consider the family $\mathcal{F}$ of all cubes in $\mathbf{R}^n$ of the type $(q_1 - r, q_1 + r) \times (q_2 - r, q_2 + r) \times \ldots \times (q_n - r, q_n + r)$, where all q_i ed r are rational numbers. This is clearly a countable family of intervals.

Let us show that A is the union of the following subfamily:

$$\mathcal{F}' = \{I \in \mathcal{F} : I \subset A\}.$$

Indeed, since A is open, for every $x \in A$ there exists an open ball $B_{r(x)}(x) \subset A$. Whitin this ball there is a cube with center in x, whithin which we can find an element $I_x \in \mathcal{F}$ with $x \in I_x$: this is true because rational are dense in $\mathbf{R}$. But then $A = \bigcup\limits_{I \in \mathcal{F}'} I$. Q.E.D.

Proof of the theorem on regularity of Lebesgue measure: It is an easy exercise to show that intervals are Lebesgue-Measurable: an interval is indeed a finite intersection of *half-spaces*, which in turn are Lebesgue-measurable: if $T \subset \mathbf{R}^n$, fix $\varepsilon > 0$ an let $\{I_i\}$ be a countable family of intervals covering T and such that $\sum_{i=1}^{\infty} |I_i| < m(T) + \varepsilon$. Define $I'_i = I_i \cap S$, $I''_i = I_i \cap (\mathbf{R}^n \setminus S)$: these are still (possibly empty) intervals, the sum of which measures is exactly $|I_i|$. Moreover, the family $\{I_i'\}$ covers $T \cap S$, and $\{I_i''\}$ covers $T \cap S^C$: we then get

$$m(T) + \varepsilon > \sum_{i=1}^{\infty} |I_i'| + \sum_{i=1}^{\infty} |I_i''| \geq m(T \cap S) + m(T \cap S^C)$$

and measurability of S follows because ε is arbitrary.

Intervals are then measurable, and so are open sets (because they are obtained as a countable union of intervals). Closed sets are measurable because their complements are open.

Let now A be measurable, $\varepsilon > 0$: we show that there exists an open set $B \supset A$ with $m(B \setminus A) < \varepsilon/2$. Suppose for now that $m(A) < +\infty$. By definition of Lebsegue measure, there are intervals $I_1, I_2, \ldots$ with $\bigcup_{i=1}^{\infty} I_i \supset A$ and $\sum_{i=1}^{\infty} |I_i| \leq m(A) + \varepsilon/2$. We know that we may suppose without loss of generality that the intervals I_i be open. If $B = \bigcup_{i=1}^{\infty} I_i$, then B is open and by subadditivity

$$m(B) \leq \sum_{i=1}^{\infty} m(I_i) \leq m(A) + \varepsilon/2,$$

whence $m(B \setminus A) = m(B) - m(A) \leq \varepsilon/2$.

If $m(A) = +\infty$ the claim still true: we take $\varepsilon > 0$ and we show that there exists $B \supset A$, B open, such that $m(B \setminus A) < \varepsilon$.

To this end, consider the measurable sets $A_N = A \cap B_N(0)$, $N = 1, 2, \ldots$: their measure if finite, and their union is A. For each of these we can find an open set $B_N \supset A_N$, B_N such that $m(B_N \setminus A_N) < \frac{\varepsilon}{2^{N+1}}$: define $B = \bigcup_{N=1}^{\infty} B_N$.

Now, B is open and contains A, moreover $B \setminus A \subset \bigcup_{N=1}^{\infty} (B_N \setminus A_N)$: by subadditivity we get $m(B \setminus A) \leq \sum_{N=1}^{\infty} m(B_N \setminus A_N) < \frac{\varepsilon}{2}$.

To conclude the proof, we show that given a measurable A and $\varepsilon > 0$, there exists a closed set $C \subset A$ with $m(A \setminus C) < \varepsilon/2$. It is enough to choose an open set $F \supset A^C$ such that $m(F \setminus A^C) < \varepsilon/2$. Then $C = F^C$ is closed, $C \subset A$, and $m(A \setminus C) = m(F \setminus A^C) < \varepsilon/2$. Q.E.D.

The following exercise builds on the regularity of Lebesgue measure:

EXERCISE (Borel hull): Recall that *Borel sigma-algebra* in $\mathbf{R}^n$ (o in a topological space) is, by definition, the smallest σ-algebra which cointains all open sets. Its elements are called *Borel sets*. Show that for any Lebesgue measurable set A there are Borel sets B, C such that $C \subset A \subset B$ and $m(B \setminus C) = 0$. As a consequence, every Lebesgue measurable set is the union of a Borel set and a set of measure 0.

Altough there are lots of Lebesgue measurable sets, not every subset of $\mathbf{R}^n$ is measurable:

EXAMPLE (Vitali non measurable set): Let $n = 1$, and consider the interval $(0, 1) \subset \mathbf{R}$. Define the following equivalence relation on $(0, 1)$: we will say that $x \sim y$ if and only if $x - y \in \mathbf{Q}$. This equivalence relation gives us a partition of $(0, 1)$ in infinitely many equivalence classes: we choose a set A which contains *exactly one element* of each class (of course, to do this we need the axiom of choice). We will show that A is *not* Lebesgue-measurable.

For each $q \in \mathbf{Q} \cap [0, 1)$ define the sets $A_q = \{x + q : x \in A\}$. Since Lebesgue measure is obviously translation-invariant, we have $m(A_q) = m(A)$. Moreover, we have $m(A) = m(A_q) = m(A_q \cap (0, 1)) + m(A_q \setminus (0, 1))$ because intervals are Lebesgue-measurable. Let now $B_q = A_q \setminus (0, 1)$ and define $\tilde{B}_q = \{x : x + 1 \in B_q\}$: we have $m(\tilde{B}_q) = m(B_q)$ by translation invariance.

We next denote $\tilde{A}_q = (A_q \cap (0, 1)) \cup \tilde{B}_q$: we clearly have $m(\tilde{A}_q) = m(A)$. Now, it is easy to check that the sets A_q are pairwise disjoint for $q \in \mathbf{Q} \cap [0, 1)$ and that $\bigcup_q \tilde{A}_q = (0, 1)$.

If we suppose by contradiction that A is measurable, then so are also the sets $\tilde{A}_q$ and by countable additivity we get

$$1 = m([0, 1)) = \sum_{n=1}^{\infty} m(\tilde{A}_q) = \sum_{n=1}^{\infty} m(A).$$

This is impossible: the mesaure of A is either zero or positive. If we had $m(A) = 0$, the r.h.s. would be 0, while if $m(A) > 0$ it would be $+\infty$: in neither case it can be equal to 1. Thus A cannot be Lebesgue-measurable.

2 Lecture of october 5, 2015 (2 hours)

Before we can define Lebesgue integral, we need to define the important class of the *measurable* functions.

DEFINITION (measurable functions): Let $A \subset X$ be a measurable set (w.r.t. a fixed outer measure, for instance Lebesgue measure on $\mathbf{R}^n$), $f : A \to \overline{\mathbf{R}}$. With $\overline{\mathbf{R}}$ we denote the set $\mathbf{R} \cup \{+\infty\} \cup \{-\infty\}$: in this constest, we will adopt the "funny" convenction that $0 \cdot \pm\infty = 0$, while the expression $+\infty - \infty$ has to be considered meaningless as usual.

The function f is said to be *measurable* if for all $a \in \mathbf{R}$ the sets $f^{-1}((a,+\infty]) = \{x \in A : f(x) > a\}$ are measurable.

The following result gives an equivalent (and more "topological") characterization of measurability:

PROPOSITION (Characterization of measurable functions): A function $f : A_0 \to \mathbf{R} \cup \{+\infty\}$ *(with* $A_0 \subset X$*) is measurable iff* $f^{-1}(\{+\infty\})$, $f^{-1}(\{-\infty\})$ *are measurable and* $f^{-1}(U)$ *is measurable for every open set* $U \subset \mathbf{R}$.

PROOF.: If we know that $f^{-1}(\{+\infty\})$, $f^{-1}(\{-\infty\})$ are measurable and $f^{-1}(U)$ is measurable for every open $U \subset \mathbf{R}$, then f is measurable since $f^{-1}((a,+\infty]) = f^{-1}((a,+\infty)) \cup f^{-1}(\{+\infty\})$.

To see the other implication, suppose f is measurable. We can write

$$f^{-1}(\{+\infty\}) = \bigcap_{N=1}^{\infty} f^{-1}((N,+\infty]),$$

whence $f^{-1}(\{+\infty\})$ is measurable because it is a countable intersection of measurable sets.

From the measurability of f it follows then that $f^{-1}((a,+\infty))$ is measurable for every $a \in \mathbf{R}$. Also the sets $f^{-1}([a,+\infty))$, $f^{-1}((-\infty,a))$ and $f^{-1}((-\infty,a])$ are measurable for every $a \in \mathbf{R}$. Indeed, $f^{-1}([a,+\infty)) = \bigcap_{N=1}^{\infty} f^{-1}((a-\frac{1}{N},+\infty))$ is measurable being the intersection of a countable family of measurable sets. The counterimages of left halflines like $f^{-1}([-\infty,a))$ and $f^{-1}([-\infty,a])$ are measurable because their complements are measurable. Then also $f^{-1}(\{-\infty\})$ is measurable: $f^{-1}(\{-\infty\}) = \bigcap_{N=1}^{\infty} f^{-1}([-\infty,-N])$...and then also the sets $f^{-1}((-\infty,a))$ are measurable.

But then the counterimages of open intervals are measurable, because $f^{-1}((a,b)) = f^{-1}((-\infty,b)) \cap f^{-1}(a,+\infty))$.

If finally $U \subset \mathbf{R}$ is open, we write $U = \bigcup_{i=1}^{\infty} I_i$, where $I_i \subset \mathbf{R}$ are open intervals. Then $f^{-1}(U) = \bigcup_{i=1}^{\infty} f^{-1}(I_i)$ is measurable. Q.E.D.

As a consequence, the domain A of a measurable function is necessarily measurable (because $A = f^{-1}(\mathbf{R}) \cup f^{-1}(\{+\infty\}) \cup f^{-1}(\{-\infty\})$). Moreover, a real valued continuous function defined on an open set of $\mathbf{R}^n$ is certainly Lebesgue-measurable (Why?).

Measurable functions are "stable" under a whole lot of algebraic and limit operations:

PROPOSITION (Stability of measurable functions): Suppose f, g *are measurable functions,* $\lambda \in \mathbf{R}$ *and* $\{f_n\}$ *is a sequence of measurable functions. Then* *(i) the set* $\{x : \ f(x) > g(x)\}$ *is measurable;*

(ii) if $\phi : \overline{\mathbf{R}} \to \overline{\mathbf{R}}$ *is continuous, then* $\phi \circ f$ *is measurable (in its domain);*

(iii) the functions $f+g$*,* λf*,* $|f|$*,* $\max\{f,g\}$*,* $\min\{f,g\}$ *and* fg *are all measurable within their domains;*

(iv) the functions $\sup f_n$*,* $\inf f_n$*,* $\limsup f_n$*,* $\liminf f_n$ *and* $\lim f_n$ *are all measurable in their domains.*

PROOF: To show (i), observe that if $f(x) > g(x)$, there is a rational number $q \in (g(x), f(x))$. We can then write

$$\{x : \ f(x) > g(x)\} = \bigcup_{q \in \mathbf{Q}} \left(f^{-1}((q, +\infty]) \cap g^{-1}([-\infty, q)) \right),$$

and we have obtained our set as a countable union of measurable sets.

(ii) is obvious for real valued functions, thanks to our characterization of measurable functions: indeed, the preimage of open sets under ϕ is open! In the general case of functions taking values in $\overline{\mathbf{R}}$, we must only specify the topology of $\overline{\mathbf{R}}$: the open sets are simply unions of open sets in $\mathbf{R}$ and of open half lines. Once this is understood, the proof is as before.

To see (iii), let f, g be measurable and consider their sum $f+g$ (which is defined on the intersection of the domains, minus the points where the sum is of the form $+\infty - \infty$ or $-\infty + \infty$). This is measurable because

$$(f+g)^{-1}((a, +\infty]) = \{x : f(x) > a - g(x)\}$$

is measurable by (i). From (ii) we then infer the measurability of λf, $|f|$ and f^2 (compositions with continuous functions). If f, g are real valued, we can write $\max\{f(x), g(x)\} = \frac{1}{2}(f(x) + g(x) + |f(x) - g(x)|)$, $\min\{f(x), g(x)\} = \frac{1}{2}(f(x) + g(x) - |f(x) - g(x)|)$, $f(x)g(x) = \frac{1}{2}((f(x)+g(x))^2 - f^2(x) - g^2(x))$, whence the measurability of $\max\{f,g\}$, $\min\{f,g\}$ and fg. In the general case, the above identities give measurability on the restrictions to the measurable set where f and g are finite. What is left is easily decomposed in a small number of measurable pieces, where the functions are constant: for instance, $f(x)g(x)$ is identically $+\infty$ on the set $\{x \in X : \ f(x) = +\infty, \ g(x) > 0\}$, it is zero on the set $\{x \in X : \ f(x) = +\infty, \ g(x) = 0\}$, etc. In conclusion, we easily deduce the measurability of fg.

Consider now (iv) and let $f(x) = \sup\{f_n(x) : n = 1, 2, \ldots\}$. We have $f^{-1}((a, +\infty)) = \bigcup_n f_n^{-1}((a, +\infty))$, so that f is measurable, and $\inf_n f_n(x)$ is likewise measurable.

The function $\liminf_{n\to+\infty} f_n(x)$ is measurable because $\liminf_{n\to+\infty} f_n(x) = \sup_n \inf\{f_m(x) : m \geq n\}$, and so is $\limsup_{n\to+\infty} f_n(x)$. The set where $\liminf_n f_n$ and $\limsup_n f_n$ coincide is measurable: this is precisely the domain of $\lim_n f_n$, which is thus measurable. Q.E.D.

An importan sublcass of measurable functions is that of *simple functions*: in the definition of Lebesgue integral they play the crucial role step functions have in the theory of Riemann integral.

Recall that given $A \subset X$, its *characteristic function* is

$$\mathbf{1}_A(x) = \begin{cases} 1 & if\ x \in A, \\ 0 & if\ x \notin A. \end{cases}$$

DEFINITION: A *simple function* $\phi : \mathbf{R}^n \to \mathbf{R}$ is a finite linear combination of characteristic functions of measurable sets. In other words, ϕ is simple if there are measurable sets $A_1, A_2, \ldots, A_N$ and real numbers $c_1, c_2, \ldots, c_N$ such that $\phi(x) = \sum_{i=1}^{N} c_i \mathbf{1}_{A_i}(x)$. With no loss of generality, we may suppose that the sets A_i are pairwise disjoint.

An equivalent definition is the following: simple functions are measurable functions whose image is a *finite subset of* $\mathbf{R}$.

If ϕ is simple and $\phi(x) \geq 0$ for all x, we define in a natural way its (Lebesgue) integral w.r.t. the measure μ:

$$\int_X \phi(x)\, d\mu(x) = \sum_{i=1}^{N} c_i\ \mu(A_i).$$

Observe that step function in $\mathbf{R}^n$ are just simple functions for which the sets A_i are intervals: for this kind of functions (and Lebesgue measure), the new definition of integral coincides with Riemann's. Moreover, the integral of simple functions enjoys the usual properties of monotonicity, homogeneity and additivity w.r.t. the integrand functions.

As we will see, the Lebesgue integral of a non-negative measurable function is defined in a way very similar to the (lower) Riemann integral, just by replacing step functions with simple functions:

$$\int f(x)\, d\mu(x) := \sup\{\int \phi(x)\, d\mu(x) :\ \phi\ simple,\ \phi \leq f\}.$$

However, to prove that this object has all the usual properties we expect from the integral, we will need an approximation result: the following, fundamental theorem guarantees that every non-negative measurable function can be approximated from below with an increasing sequence of simple functions.

THEOREM (Approximation of measurable functions with simple functions): Let $f : X \to [0, +\infty]$ be a measurable function. Then there exists a sequence $\phi_k : X \to [0, +\infty)$ of simple functions such that $f \geq \phi_{k+1} \geq \phi_k$ ($k = 1, 2, 3, \ldots$) and such that

$$\lim_{k \to +\infty} \phi_k(x) = f(x) \quad \forall x \in X.$$

PROOF: For each fixed $k = 1, 2, \ldots$ and $j = 0, 2, \ldots, k2^k - 1$ we define the measurable sets $E_{k,j} = f^{-1}([\frac{j}{2^k}, \frac{j+1}{2^k}))$, while we define $E_{k,k2^k} = f^{-1}([k2^k, +\infty])$.

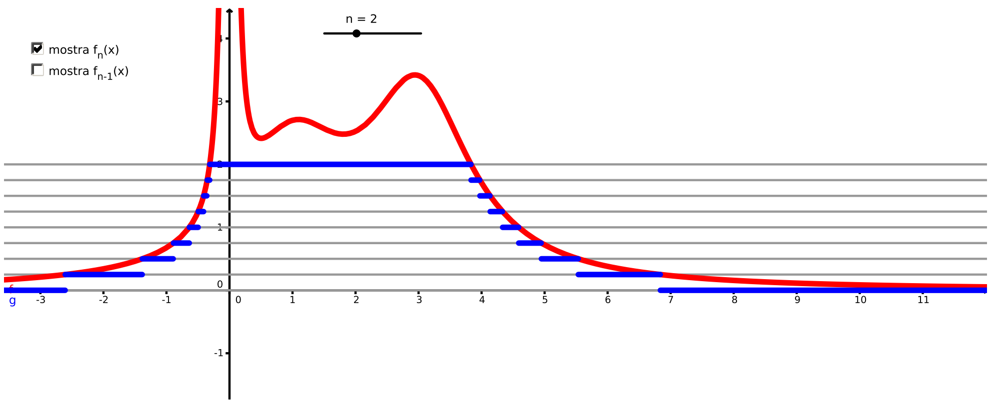
Consider the simple functions[2]

$$\phi_k(x) = \sum_{j=0}^{k2^k} \frac{j}{2^k} \mathbf{1}_{E_{k,j}}(x).$$

By constructions, these functions are measurable and are less or equal than f. Moreover, they form an increasing sequence: indeed, for every k and $j = 1, \ldots, k2^k - 1$ we have $E_{k,j} = E_{k+1,2j} \cup E_{k+1,2j+1}$. Only $E_{k,k2^k}$ is slightly different: this set in the following step is subdivided into $2^k + 1$ parts.

It is also easy to check that $\phi_k(x) \to f(x)$ for every x: in case $f(x) < +\infty$, for large enough k we have $f(x) - \phi_k(x) \leq 2^{-k}$, while if $f(x) = +\infty$ it holds $x \in E_{k,2^k}$ for every k and $\phi_k(x) = k \to +\infty$.

The following is my attempt to visualize the construction of the functions ϕ_k with a GeoGebra worksheet[3].



[2]These functions can be expressed in the following, synthetic way: $\phi_k(x) = \min\{k, 2^{-k}[2^k f(x)]\}$, where $[\cdot]$ denotes the integer part (floor) function.

[3]`http://www.geogebratube.org/material/show/id/51513`

Q.E.D.

We finally reached the point where the Lebesgue integral of a non-negative measurable function has to be introduced: as a matter of fact, I already told you the definition

DEFINITION: The *Lebsegue integral* of a measurable function $f : X \to [0, +\infty]$ is defined by

$$\int_X f(x)\, d\mu(x) = \sup\{\int_X \phi(x)\, d\mu(x) : \ \phi \ simple, \ \phi \leq f\}.$$

More generally, if $f : A \to [0, +\infty]$ is measurable, we define $\int_A f(x)\, d\mu(x)$ as $\int_X \tilde{f}(x)\, d\mu(x)$, where $\tilde{f} : X \to [0, +\infty]$ is obtained by extending f to 0 outside A.

The following integral convergence result will be extremely important in the theory of Lebesgue integral.

THEOREM (Beppo Levi's or of the monotone convergence): Let $\{f_k\}$ be a sequence of non-negative measurable functions, $f_k : X \to [0, +\infty]$, and suppose the sequence is increasing: $f_{k+1}(x) \geq f_k(x)$ for every $x \in X$ and for every $k = 1, 2, 3, \ldots$. Then, if we denote $f(x) = \lim_{k\to+\infty} f_k(x)$, we have

$$\int_X f(x)\, d\mu(x) = \lim_{k\to+\infty} \int_X f_k(x)\, d\mu(x).$$

PROOF: Notice that f is measurable, being the pointwise limit of measurable functions. Morevore, the sequence $k \mapsto \int_X f_k(x)\, d\mu(x)$ is increasing: denote by α its limit. As $f \geq f_k$ for each k, we obviously have $\int_X f(x)\, d\mu(x) \geq \alpha$: in particular, if $\alpha = +\infty$ the theorem is trivially true. If $\alpha \in \mathbf{R}$, we need to prove the opposite inequality

$$\int_X f(x)\, d\mu(x) \leq \alpha.$$

To this end, let us fix $c \in (0, 1)$ and a simple function $s : X \to [0, +\infty)$ with $s \leq f$. The function s can be expressed as $s(x) = \sum_{j=1}^N s_j \mathbf{1}_{A_j}(x)$, where A_j are pairwise disjoint measurable sets. Define $E_k = \{x \in X : f_k(x) \geq$

$cs(x)\}$. Thanks to the fact that $f_k \to f$ and $c < 1$, we have $\bigcup\limits_{k=1}^{\infty} E_k = X$, and the sequence E_k is increasing because so are the functions f_k. We define next $A_{j,k} = A_j \cap E_k$: thanks to the continuity of measure on increasing sequences, we have $\mu(A_{j,k}) \to \mu(A_j)$ as $k \to +\infty$. We then infer

$$\alpha = \lim_{k\to+\infty} \int\limits_X f_k(x)\, d\mu(x) \geq$$

$$\lim_{k\to+\infty} \int\limits_{E_k} f_k(x)\, d\mu(x) \geq \lim_{k\to+\infty} \int\limits_{E_k} c\, s(x)\, d\mu(x) =$$

$$\lim_{k\to+\infty} c \sum_{j=1}^{N} s_j\, \mu(A_{j,k}) = c \sum_{j=1}^{\infty} s_j\, \mu(A_j) = c \int\limits_X s(x)\, d\mu(x).$$

By taking the supremum over all simple functions $s \leq f$ and all $c < 1$, we get the inequality we need. Q.E.D.

Let us see some important consequences of the monotone convergence theorem:

(i) *Additivity of integral w.r.t. the integrand:* Let $f, g : X \to [0, +\infty]$ be measurable functions. Then

$$\int_X (f(x) + g(x))\, d\mu(x) = \int_X f(x)\, d\mu(x) + \int_X g(x)\, d\mu(x).$$

Indeed, we can choose two increasing sequences of simple functions, $\{s_k\}$, $\{u_k\}$ with $s_k \to f$, $u_k \to g$. Lebesgue integral is clearly additive on the set of simple functions: by the monotone convergence theorem we can pass to the limit and we get the desired equality.

(ii) *Countable additivity of the integral w.r.t. the integration set:* If $\{A_i\}$ is a sequence of pairwise disjoint measurable sets and f is a non-negative measurable function defined on $A = \bigcup\limits_{i=1}^{\infty} A_i$, then

$$\int\limits_A f(x)\, d\mu(x) = \sum_{i=1}^{\infty} \int\limits_{A_i} f(x)\, d\mu(x).$$

It is enough to consider the increasing seguence of measurable functions $g_k(x) = \sum\limits_{i=1}^{k} f(x)\, \mathbf{1}_{A_i}(x)$, whose limit is $g(x) = f(x)\, \mathbf{1}_A(x)$. Again, the thesis follows from Beppo Levi's theorem.

(iii) *Integration by series:* If $\{f_k\}$ is a sequence of non-negative measurable functions defined on a set A, then

$$\int_A \sum_{i=1}^{\infty} f_k(x)\, d\mu(x) = \sum_{i=1}^{\infty} \int_A f_k(x)\, d\mu(x).$$

To prove this, it is enough to apply the monotone convergence theorem and the additivity of the Lebesgue integral to the partial sums of our series.

DEFINITION (Integral of arbitrarily signed functions): How can we proceed if we wish to integrate an arbitrarily signed, measurable function $f : A \to \overline{\mathbf{R}}$? We define the positive and negative part of f as follows:

$$f^+(x) := \max\{0, f(x)\}, \qquad f^-(x) := -\min\{0, f(x)\}.$$

We clearly have $f(x) = f^+(x) - f^-(x)$ and $|f(x)| = f^+(x) + f^-(x)$. Moreover, both f^+ and f^- are *non-negative*: if their integrals are not *both* $+\infty$, f is called *Lebesgue-integrable* and we define

$$\int_A f(x)\, d\mu(x) := \int_A f^+(x)\, d\mu(x) - \int_A f^-(x)\, d\mu(x).$$

If *both* the integrals of f^+ and f^- are *finite*, f is said to be *summable* and its integral is finite. Obviously, a measurable function is summable if and only if the integral of $|f|$ is finite.

Notice that additivity w.r.t. the integrand and the integration set still hold for summable functions. In particular, as the integral is clearly 1-homogeneous by its definition, Lebesgue integral is *linear* on the vector space of summable functions.

The following are two celebrated integral convergence theorems. Don't be fooled by the fact that the first one is called Fatou's *lemma*: it is a fundamental result!

THEOREM (Fatou's Lemma): Let $f_k : X \to [0, +\infty]$ be a sequence of non-negative measurable functions, $f(x) = \liminf_{k\to+\infty} f_k(x)$. Then

$$\int_X f(x)\, d\mu(x) \le \liminf_{k\to+\infty} \int_X f_k(x)\, d\mu(x).$$

PROOF: We already know that f is a non-negative measurable funtion. We have $f(x) = \lim_{k\to+\infty} g_k(x)$, with $g_k(x) = \inf\{f_h(x) : \ h \geq k\}$. Now, g_k form an increasing sequence of non-negative measurable functions: by the monotone convergence theorem we have

$$\int_X f(x)\ d\mu(x) = \lim_{k\to+\infty} \int_X g_k(x)\ d\mu(x).$$

Our thesis then follows from the monotonicity of Lebesgue integral, as $g_k(x) \leq f_k(x)$. Q.E.D.

Let me mention a couple of things: Fatou's Lemma is in general false for functions with arbitrary sign. Take for instance $A = \mathbf{R}$, $\mu = m$ (Lebesgue measure), $f_k(x) = -1/k$ (constant functions). Then $f_k(x) \to 0$, but

$$\int_{\mathbf{R}} f_k(x)\ dx = -\infty, \qquad \int_{\mathbf{R}} 0\ dx = 0^4.$$

Since in this example the sequence f_k is increasing, this also shows that the monotone convergence theorem fails unless the functions are non-negative. The same functions, with the opposite sign, also show that in Fatou's Lemma we may well have strict inequality.

The following is probably the most famous result in Lebesgue's theory:

THEOREM (Lebesgue's or of the dominated convergence): Let $f_k : X \to \overline{\mathbf{R}}$ be a sequence of measurable functions, and suppose there exists a summable function $\phi : X \to [0, +\infty]$ such that $|f_k(x)| \leq \phi(x)$ for every k and for every x. If the limit $f(x) = \lim\limits_{k\to+\infty} f_k(x)$ exists, then

$$\lim_{k\to+\infty} \int_X |f_k(x) - f(x)|\ d\mu(x) = 0,$$
$$\int_X f(x)\ d\mu(x) = \lim_{k\to+\infty} \int_X f_k(x)\ d\mu(x).$$

PROOF: The limit function f is measurable, and it is also sommable because its absolute value is dominated by ϕ. Moreover, $|f_k(x) - f(x)| \leq |f_k(x)| + |f(x)| \leq 2\phi(x)$. It follows that the sequence $2\phi(x) - |f_k(x) - f(x)|$ is non-negative, and converges to the pointwise limit $2\phi(x)$. From Fatou's Lemma it then follows that

$$\liminf_{k\to+\infty} \int_X (2\phi(x) - |f_k(x) - f(x)|)\ d\mu(x) \geq \int_X 2\phi(x)\ d\mu(x),$$

[4] For Lebesgue measure we write dx instead of $dm(x)$

whence, eliminating the integral of 2ϕ:

$$\limsup_{k\to+\infty}\int_X |f_k(x)-f(x)|\,d\mu(x)\le 0,$$

which is the first assert in the thesis. The second assert follows from

$$\left|\int_X f_k(x)\,d\mu(x)-\int_X f(x)\,d\mu(x)\right|\le\int_X |f_k(x)-f(x)|\,d\mu(x).$$

Q.E.D.

To be totally convinced that Lebesgue's theory is reasonable, it is important to compare Lebesgue's integral with Riemann's.

Before we do that, let us introduce an useful notation: we say that a certain property is true *for almost every* $x\in A$ (shortly: a.e in A, *almost everywhere in* A), if the set of points where the property fails has measure 0.

Of course, this definition makes sense for Lebesgue measure, or for any fixed outer measure μ: to avoid confusion, we will often say that something is true "μ-a.e.".

For instance, given two functions $f,g:\mathbf{R}^n\to\overline{\mathbf{R}}$, they are *$\mu$-a.e. equal* if $\mu(\{x:\ f(x)\neq g(x)\})=0$.

It is a very simple exercise to check that a function which is a.e. equal to a measurable function is also measurable (because sets with measure 0 are measurable). Moreover, if we change a function on a set with measure 0, its integral does not change.

The following result also holds:

PROPOSITION: Let $f:A\to[0,+\infty]$ *be a measurable function such that* $\int\limits_A f(x)\,d\mu(x)=0$. *Then* $f=0$ *μ-a.e. in* A.

PROOF: We can write

$$\{x\in A:\ f(x)>0\}=\bigcup_{n=1}^{\infty}\{x\in A:\ f(x)>\frac{1}{n}\}.$$

All the sets in the r.h.s. have measure 0: if we had $\mu(E_{\overline{n}})>0$, where $E_{\overline{n}}=\{x\in A:\ f(x)>\frac{1}{\overline{n}}\}$, we would get

$$\int\limits_A f(x)\,d\mu(x)\ge\int\limits_{E_{\overline{n}}} f(x)\,d\mu(x)\ge\mu(E_{\overline{n}})/\overline{n}>0,$$

which contradicts the hypothesis. Q.E.D.

The following theorem shows that Riemann integral coincides with Lebesgue integral (w.r.t the Lebesgue Measure) on the set of Riemann integrable functions (bounded and on a bounded interval: for generalized Riemann integrals things are slightly more complicated[5].

We state and prove the result in dimension 1: the generalization to higher dimensions is proved in the same way.

THEOREM: Let $f : [a, b] \to \mathbf{R}$ be a bounded Riemann-integrable function. Then f is Lebesgue-measurable and its Lebesgue and Riemann integrals coincide.

PROOF of the comparison between Riemann and Lebesgue integrals: Until the proof is finished, we need to distinguish between Riemann and Lebesgue integrals: given $f : [a, b] \to \mathbf{R}$, we denote by $\int_a^b f(x)\, dx$ its Lebesgue integral, by $\mathcal{R} \int_a^b f(x)\, dx$ its Riemann integral (provided they exist). But we need to stress that at least for step functions we already know that Lebesgue and Riemann integrals coincide.

By the definition of Riemann integral, we find two sequences of step functions $\{\psi_n\}$ and $\{\phi_n\}$, with $\psi_n \geq f \geq \phi_n$ and

$$\lim_{n\to+\infty} \int_a^b \psi_n\, dx = \lim_{n\to+\infty} \int_a^b \phi_n\, dx = \mathcal{R} \int_a^b f\, dx.$$

Let now $\overline{\psi}(x) = \inf\{\psi_n(x) : \ n = 1, 2, \ldots\}$, $\underline{\phi}(x) = \sup\{\phi_n(x) : \ n = 1, 2, \ldots\}$. These two functions are measurable and $\underline{\phi} \leq f \leq \overline{\psi}$. By monotonicity of integral we have

$$\int_a^b \psi_n(x)\, dx \geq \int_a^b \overline{\psi}(x)\, dx,$$

whence passing to the limit

$$\mathcal{R} \int_a^b f(x)\, dx \geq \int_a^b \overline{\psi}(x)\, dx,$$

and likewise

$$\mathcal{R} \int_a^b f(x)\, dx \leq \int_a^b \underline{\phi}(x)\, dx.$$

---

[5] As an exercise, show that the generalized Riemann integral of a *non negative* function coincides with its Lebesgue integral whenever it exists. It is enough to use the monotone convergence theorem and the result for ordinary Riemann integrals. The same holds when the function is absolutely integrable (use the dominated convergence theorem). If instead f is integrable in the generalized Riemann sense, but the integral of $|f|$ is $+\infty$, then f is not Lebesgue integrable because both the integrals of f^+ and f^- are $+\infty$.

As $\overline{\psi} \geq \underline{\phi}$, we deduce that $\int_a^b (\overline{\psi} - \underline{\phi})\, dx = 0$, whence $\overline{\psi} - \underline{\phi} = 0$ a.e., that is $\overline{\psi} = \underline{\phi} = f$ a.e. in $[a, b]$. It follows that f is measurable and its Lebesgue integral is equal to its Riemann Integral. Q.E.D.

Actually, one could prove that a bounded function is Riemann integrable if and only if it is *continuous almost everywhere* (Vitali Theorem). We will not prove this result because of time constraints.

3 Lecture of october 6, 2015 (2 hours)

The next, very important theorem is a big improvement over the results we had for the Riemann integral:

THEOREM (Fubini and Tonelli): Let $f : \mathbf{R}^2 \to \overline{\mathbf{R}}$ be a measurable function. Then

(i) If $f \geq 0$, then for a.e. $y \in \mathbf{R}$ the function $x \mapsto f(x, y)$ is measurable on $\mathbf{R}$. Moreover, the function $y \mapsto \int_{\mathbf{R}} f(x, y)\, dx$ is measurable and one has

$$(*) \quad \int_{\mathbf{R}^2} f(x, y)\, dx\, dy = \int_{\mathbf{R}} \left(\int_{\mathbf{R}} f(x, y)\, dx \right) dy.$$

Obviously, the same holds also if we interchange the role of x and y.

(ii) If f is $\overline{\mathbf{R}}$-valued and $\int_{\mathbf{R}} \left(\int_{\mathbf{R}} |f(x, y)|\, dx \right) dy < +\infty$, then f is summable. The same holds if we reverse the order of integration.

(iii) If f is $\overline{\mathbf{R}}$-valued and summable, (i) still holds.

We will give a more general statement for a larger class of measures: this will require the introduction of the product measure.

Notice for the moment that when f is not summable, the statement is no longer true and the two iterated integral can well be different[6].

The theorem is easy to generalize to higher dimension: the ambient space could be $\mathbf{R}^n \times \mathbf{R}^k$, while $x \in \mathbf{R}^n$, $y \in \mathbf{R}^k$...

While we are at it, let us state a (non-optimal) version of the change of variables theorem for multiple integrals:

THEOREM (Change of variables in multiple integrals): Let $\Phi : A \to B$ be a diffeomorphism, where A and B are open subsets of $\mathbf{R}^n$, $f : B \to \overline{\mathbf{R}}$ an

[6] Consider for instance the function $f(x, y) = (x-y)/(x+y)^3$: the two iterated integrals on the square $[0, 1] \times [0, 1]$ are finite and different.

integrable function. Then we have

$$\int_B f(y)\, dy = \int_A f(\Phi(x))|\det(\nabla\Phi(x))|\, dx.$$

Let us go back to Fubini's Theorem: we begin with the definition of *product measure*. In the following, I will give more details of those we saw in class.

DEFINITION: Let μ be a measure on R^n, ν a measure on $\mathbf{R}^m$[7]. The *product measure* $\mu \times \nu$ is an application $\mu \times \nu : \mathcal{P}(\mathbf{R}^{n+m}) \to [0, +\infty]$ defined for every $S \subset \mathbf{R}^{n+m}$ by

$$\begin{aligned}(\mu \times \nu)(S) \quad &= \inf\{\sum_{i=1}^{\infty} \mu(A_i)\nu(B_i) : \\ &S \subset \bigcup_{i=1}^{\infty}(A_i \times B_i),\ A_i\ \mu - measurable,\ B_i\ \nu - measurable\}.\end{aligned}$$

It is a simple exercise to check that this is indeed an outer measure (whence the name): proceed as for Lebesgue measure.

Moreover, the product of Lebesgue measures on euclidean spaces give Lebesgue measure on the product space:

REMARK: Denote m_n, m_m, m_{n+m} Lebesgue measure on $\mathbf{R}^n$, $\mathbf{R}^m$, $\mathbf{R}^{n+m}$ respectively. Then $m_{n+m} = m_n \times m_m$. Indeed, if in the coverings which define the measure we put the extra condition that A_i and B_i are intervals, we obtain exactly Lebesgue Measure on the product space: we deduce that $m_n \times m_m \leq m_{n+m}$.

To get the opposite inequality, take a covering of $S \subset \mathbf{R}^{n+m}$ as in the definition of product measure. Notice that we may suppose w.l.o.g. that $m(A_i) \leq 1$, $m(B_i) \leq 1$ for every i (otherwise, subdivide the sets in the covering into smaller, pairwise disjoint measurable sets). Fix $\varepsilon > 0$: by definition of Lebesgue measure, for every i we find a covering of A_i with intervals $I_{i,j}$ such that $\sum_j m_n(I_{i,j}) < m_n(A_i) + \varepsilon/2^i$, and likewise a covering of B_i with intervals $J_{i,k}$ such that $\sum_k m_m(J_{i,k}) < m_m(B_i) + \varepsilon/2^i$. The intervals $I_{i,j} \times I_{i,k}$ are still a countable covering of S, and we easily check that the sum of their measures is less than $\sum_i m_n(A_i)m_m(B_i) + 3\varepsilon$: the infimum made over

[7]More generally, μ can be a measure on a set X, ν a measure on a set Y. The product measure is then defined on $X \times Y$. All the following results hold in this more general setting, with the obvious exception of those referring to Lebesgue measure.

coverings with products of measurable sets or with intervals is the same and $m_{n+m} = m_n \times m_m$.

The following is the general version of Fubini's Theorem. To understand the statement, we need the notion of σ-finiteness:

DEFINITION: A measure μ on $\mathbf{R}^n$ is *σ-finite* if we can split $\mathbf{R}^n$ in a countable partition of measurable sets with finite measure. More generally, a set is σ-finite w.r.t. a measure μ if it is measurable and it can be covered with a countable family of measurable sets with finite measure: it is easy to see that the sets of this family can be chosen to be pairwise disjoint.

A function $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ is σ-finite if $\mathbf{R}^n$ can be split into a countable union of measurable sets, over each of which f is summable.

EXERCISE: Show that Lebesgue measure is σ-finite. Show next that a Lebesgue-measurable function $f : \mathbf{R}^n \to [0, +\infty]$ is σ-finite if and only if $m(\{x \in \mathbf{R}^n : f(x) = +\infty\}) = 0$[8].

THEOREM(Fubini, general version): Let μ be a measure on X, ν a measure on Y. Then the product measure $\mu \times \nu$ has the following property: if $S \subset X \times Y$ there exists a $(\mu \times \nu)$-measurable set S' such that $S \subset S'$ and $(\mu \times \nu)(S) = (\mu \times \nu)(S')$ (S' is called a measurable envelope of S). If A is μ-measurable and B is ν-measurable then $A \times B$ is $(\mu \times \nu)$-measurable and $(\mu \times \nu)(A \times B) = \mu(A)\nu(B)$.

If $S \subset X \times Y$ is σ-finite with respect to $\mu \times \nu$, then the slices

$$S_y = \{x \in X : \ (x, y) \in S\}, \qquad S_x = \{y \in Y : \ (x, y) \in S\}$$

are μ-measurable for ν-almost every fixed y, and ν-measurable for μ-almost every fixed x. Moreover

$$(\mu \times \nu)(S) = \int_Y \mu(S_y)\, d\nu(y) = \int_X \nu(S_x)\, d\mu(x).$$

If $f : X \times Y \to [0, +\infty]$ is σ-finite or if $f : X \times Y \to \overline{\mathbf{R}}$ is summable w.r.t. the measure $\mu \times \nu$, then the maps

$$\phi : x \mapsto \int_Y f(x, y)\, d\nu(y), \quad \psi : y \mapsto \int_X f(x, y)\, d\mu(x)$$

[8] Consider the sets $A_i = \{x \in \mathbf{R}^n : i \leq f(x) < i + 1\}$: together with the set where $f = +\infty$, they form a partition of $\mathbf{R}^n$. If $m(A_i) < +\infty$, it is easy to check that $\int_{A_i} f(x)\, dx < +\infty$. Otherwise, thanks to σ-finiteness, A_i can be decomposed into a sequence of sets with finite measure...

are well-defined for μ-a.e. x, μ-a.e. y, and are μ-integrable andν-integrable respectively. Moreover the following holds:

$$\int_{X\times Y} f(x,y)\, d(\mu\times\nu)(x,y) = \int_X \phi(x)\, d\mu(x) = \int_Y \psi(y)\, d\nu(y).$$

REMARK: The statement we wrote for Lebesgue measure follows immediately from this: the only thing which is not immediately obvius is what happens if we computed the iterated integral of a function $f : \mathbf{R}^2 \to [0, +\infty]$ which is measurable but *not σ-finite* w.r.t. Lebesgue measure in the plane. We have to check that the iterated integral of this function is $+\infty$, and so it "predicts" correctly that the function is not summable.

But our hypothesis implies that the set $S = \{(x,y) : \ f(x,y) = +\infty\}$ has positive Lebesgue measure (otherwise f would be σ-finite). Now, Lebesgue measure of S is obtained by integrating the measure of the slices S_y. Then the set $\{y : \ m(S_y) > 0\}$ must have positive measure. For every y in that set we obviously have $\int_{\mathbf{R}} f(x,y)\, dy = +\infty$ (because $f \equiv +\infty$ on S).

In conclusion, when we compute the iterated integral of f we integrate the constant $+\infty$ on a set with positive measure: the result is $+\infty$.

We omit the proof of Fubini's Theorem: the interested readers can ask me for references!

Having finished (for the moment) the part on measure theory, we will begin to study the basics of linear functional analysis: in particular, we will concentrate on Banach and Hilbert spaces. We will see many examples and also - I fear - some complements on measaure theory.

At the beginning, our interest will focus on *linear algebra in infinite dimensional spaces*! As you probably already know, from a purely algebraic viewpoint, there are few dissimilarities from the finite dimensional case : if we accept the axiom of choice (as we will do), every vector space over $\mathbf{R}$ or over $\mathbf{C}$ has a basis, i.e. a maximal set of linearly independent vectors. Moreover, each element of the space can be written in a unique way as a *finite* linear combinantion of basis vectors.

In analysis, however, this is not enough: for instance, as a bare minimum we need a notion of continuity (and thus a topology or, better yet, a metric), which is compatible with vector space operations. In other words, we at least need a topology for which *sum of vectors* and *product by a scalar* are continuous.

DEFINITION: Let X be a vector space over $\mathbf{R}$ or $\mathbf{C}$, equipped with a topology τ. The space (X, τ) is a *topological vector space* if and only if vector space operations (sum and produts by a scalar) are continuous w.r.t. the obvious product topologies induced by τ.

The simplest examples of topological vector spaces (and -almost- the only examples we will touch in this course) are *normed spaces*:

DEFINITION (Norm, normed space): Let X be a vector space over $\mathbf{R}$ (or over $\mathbf{C}$, with the obvious changes...). A *norm* on X is a map $\|\cdot\| : X \mapsto \mathbf{R}$ such that

(i) $\|x\| \geq 0 \ \forall x \in X$, $\|x\| = 0$ iff $x = 0$;

(ii) $\|\lambda x\| = |\lambda|\|x\|$ for every $x \in X$, $\lambda \in \mathbf{R}$ (homogeneity);

(iii) $\|x + y\| \leq \|x\| + \|y\|$ for every $x, y \in X$ (triangle inequality).

A vector space equipped with a norm is a *normed space*: it is a metric space with the induced distance

$$d(x, y) := \|x - y\|, \quad x, y \in X.$$

It is very simple to verify that a normed space with the induced metric is a topological vector space.

As an aside, let me notice that since a normed space is a metric space, it is possible to test continuity on sequences: for instance a function $f : X \to \mathbf{R}$ is continuous iff for every $\overline{x}$ and every sequence $x_k \to \overline{x}$ we have $f(x_k) \to f(\overline{x})$.

Exactly as with euclidean space, a property we will often need is *completeness*:

DEFINITION (Banach) space: A normed space X is a *Banach space* if it is complete, i.e. if every Cauchy sequence in X converges.

The following are examples of normed spaces:

1. $\mathbf{R}^n$ is a normed space with the usual euclidean norm: if $x = (x_1, \ldots, x_n)$, then $|x| = \sqrt{x_1^2 + \ldots + x_n^2}$. Norms on $\mathbf{R}^n$ are also $|x|_1 = \sum_{i=1}^n |x_i|$ and $|x|_\infty = \max\{|x_i| : \ i = 1, \ldots, n\}$: it is a simple exercise to check that, and also useful is to draw the balls w.r.t. these metrics.

   More general norms on $\mathbf{R}^n$ are the following:

   $$|x|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}, \quad 1 \leq p < +\infty.$$

In this case, checking the triangle inequality is not so simple: we will see it later.

It is well known that $\mathbf{R}^n$ with the euclidean norm is complete. It is complete also with all the other norms we mentioned: we will see in a few days that on $\mathbf{R}^n$ *all norms are equivalent* in the sense that they induce the same topology and have the same Cauchy sequences.

2. $(C^0([a,b]), \|\cdot\|_\infty)$ is also an example of a Banach space. Indeed, let $\{f_n\}$ be a Cauchy sequence in our space: for every $\varepsilon > 0$ there exists $\nu \in \mathbf{N}$ such that $\|f_n - f_m\|_\infty < \varepsilon$ for every $m, n \geq \nu$. Then, for every $x \in [a,b]$ we have

$$(**)|f_m(x) - f_n(x)| < \varepsilon$$

for $m, n \geq \nu$, and the real sequence $\{f_n(x)\}$ is a Cauchy sequence. By the completeness of $\mathbf{R}$, there is a real number $f(x)$ such that $f_n(x) \to f(x)$. Passing to the limit as $m \to +\infty$ in (**) we get

$$|f_n(x) - f(x)| \leq \varepsilon \quad \forall n \geq \nu,$$

and taking the supremum over $x \in [a,b]$ $\|f_n - f\|_\infty \leq \varepsilon$ for $n \geq \nu$. Then $f_n \to f$ uniformly. The function f is continuous being the uniform limit of continuous functions.

In infinite dimension, there are always linear maps which are discontinuous:

PROPOSITION: Let $(X, \|\cdot\|)$ *be a infinitely dimensional vector space over* $\mathbf{R}$.*Then there are discontinuous linear functionals* $T : X \to \mathbf{R}$.

PROOF: We will exhibit a member T of the *algebraic dual space* of X which is discontinuous. Take an algebraic basis $\mathcal{B} = \{x_\alpha\}_{\alpha \in I}$ of X: this is by assumption an infinite set. By normalizing the basis vectors, we may assume as well that $\|x_\alpha\| = 1$ for every $\alpha \in I$.

To specify T, we only need to decide its values on the basis vectors: by linearity, this fixes the linear functional. Now choose a countable subset $\mathcal{B}' = \{\tilde{x}_n\} \subset \mathcal{B}$ and define $T(\tilde{x}_n) = n$, $n = 1, 2, \ldots$, $T(x_\alpha) = 0$ if $x_\alpha \in \mathcal{B} \setminus \mathcal{B}'$. The linear functional defined in this way is discontinuous: the sequence $y_n = x_n/\sqrt{n}$ converges to 0 in norm, but $T(y_n) = \sqrt{n} \to +\infty$. Q.E.D.

4 Lecture of october 7, 2015 (2 hours)

In our discussion of the completeness of $\mathbf{R}^n$ with various norms, we briefly touched the notion of *equivalent norms*: two norms on the same vector space are said to be equivalent, it they induce the same topology, i.e. if the open sets for the two norms are the same.

Let now X be a vector space with two norms $\|\cdot\|_1$, $\|\cdot\|_2$. Let us check that the two norms are equivalent *if and only if* there are two constants $c, C > 0$ such that

$$(*)\ c\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1 \quad \forall x \in X.$$

Indeed, in a metric space the open sets are defined starting from open balls: the topology is the same if and only if, given a ball for any one of the two norms, it is possible to find a ball for the other norm (with a suitably small radius) which is contained in the first.

By homogeneity of the norm, balls in a normed metric spaces are obtained from the unit ball by scaling and translating. But the two inequalities (*) imply that $B^{(1)}_{1/C}(0) \subset B^{(2)}_1(0)$and $B^{(2)}_c(0) \subset B^{(1)}_1(0)$.

Conversely, if the two norms induce the same topology, then (*) holds. Indeed, the ball $B^{(1)}_1(0)$ is an open set for both norms: there exists $r > 0$ such that $B^{(2)}_r(0) \subset B^{(1)}_1(0)$. Given $x \in X$, $x \neq 0$ we obviously have $y = r/2\frac{x}{\|x\|_2} \in B^{(2)}_r(0)$, whence $\|y\|_1 < 1$ and $r/2\|x\|_1 \leq \|x\|_2$. The other inequality is proved by means of a similar argument.

Thanks to (*), we see that two equivalent norms have the same Cauchy sequences: completeness is preserved when passing to equivalent norms. Instead, the norms $\|\cdot\|_\infty$ and $\|\cdot\|_1$ on the space $C^0([a,b])$ of continuous functions are not equivalent (the first is complete, the second is not). In fact, the topology induced by the norm $\|\cdot\|_\infty$ is *stronger* (it has more open sets) than the topology induced by $\|\cdot\|_1$... Open balls for the norm $\|\cdot\|_\infty$ are not open w.r.t. the norm $\|\cdot\|_1$, while the opposite is true.

Let us look at another example:

- The space $(C^0([a,b]), \|\cdot\|_1)$ is not a Banach space, because it is not complete. Take for instance $[a,b] = [-1,1]$ and consider the sequence

  $$u_n(x) = \begin{cases} -1 & if x \leq -1/n, \\ 1 & if\ x \geq 1/n, \\ nx & if\ -1/n < x < 1/n. \end{cases}$$

  This is a Cauchy sequence w.r.t. the norm $\|\cdot\|_1$, but it does not converge to any continuous function: indeed, it converges in the norm to the discontinuous function $\operatorname{sgn}(x)$. In this course, we will study the completion of this space: it is the Banach space $L^1([a,b])$ of Lebesgue summable functions, with the norm $\|\cdot\|_1$.

We will see next time that *any* two norms over $\mathbf{R}^n$ are equivalent. This is no longer the case in infinite dimension: it is possible to define non equivalent norms on the same space. For instance, $(C^0([a,b]), \|\cdot\|_\infty)$and $(C^0([a,b]), \|\cdot\|_1)$ have different topologies. In particular, it is easy to construct a sequence of continuous functions which converges to 0 in the norm $\|\cdot\|_1$ but not in the norm $\|\cdot\|_\infty$. This has a consequence which is surprising at a first glance:the *identity map* $Id : (C^0([a,b]), \|\cdot\|_1) \to (C^0([a,b]), \|\cdot\|_\infty)$ is discontinuous at 0. So here is an example of a *linear and invertible map* between two infinitely dimensional normed spaces, which are not even continuous!

Let us go back to our examples:

1. The space $L^p(\Omega)$ is defined as the set of measurable functions such that $\|u\|_{L^p} < +\infty$, where

   $$\|u\|_{L^p(\Omega)} := \left(\int_\Omega |u(x)|^p \, dx \right)^{1/p},$$

   quotiented w.r.t. the equivalence relation

   $$u \sim v \Leftrightarrow u(x) = v(x) \ for \ a.e. \ x \in \Omega.$$

   This is a norm on L^p, as we will see later in the course.

   More generally, if Ω is a set and μ is an outer measure on Ω, for a μ-measurable function $u : \Omega \to \overline{\mathbf{R}}$ we define

   $$\|u\|_{L^p(\mu)} := \left(\int_\Omega |u(x)|^p \, d\mu(x) \right)^{1/p}.$$

   The space $L^p(\mu)$ is then defined as above.

   The spaces $L^\infty(\Omega)$, $L^\infty(\mu)$ are slightly harder to define: we postpone the definition a little bit.

2. The space of sequences ℓ^p is defined as follows: given a real sequence $\{x_n\}$ and $p \in [1, +\infty)$, let

   $$\|\{x_n\}\|_{\ell^p} := \left(\sum_{n=1}^\infty |x_n|^p \right)^{1/p}.$$

   The space ℓ^p is then the space of those sequences which have finite norm.

We will now see that this space is a particular case of $L^p(\mu)$, obtained when μ is the counting measure over $\mathbf{N}$.

Indeed, one can check that given a sequence $\{a_n\}_{n\in\mathbf{N}}$ of nonnegative real numbers, one has

$$\int_{\mathbf{N}} a_n \, d\mu(n) = \sum_{n=1}^{\infty} a_n.$$

Indeed, let $\{a_n^N\}_n$ be the sequence truncated to the first N terms, i.e. $a_n^N = a_n$ if $1 \le n \le N$, $a_n^N = 0$ otherwise. As $N \to +\infty$, these sequences increase and converge pointwise to the original sequence. By the monotone convergence theorem we thus get that the integral of $\{a_n\}$ is the limit of the integrals of the truncated sequences: but those are simply the partial sums of the series (because the truncated sequences are "simple functions" for the counting measure!).

By applying the dominated convergence theorem, one verifies that the statement still holds for absolutely convergent sequences (with arbitrary sign).

As an execise, we also checked that *summability* with respect to the counting measure of an arbitrary family of numbers $\{a_\alpha\}_{\alpha\in A}$ (where A is a possibly uncountable set of indices), implies that $\{\alpha \in A : \ a_\alpha \neq 0\}$ is at most countable. Indeed, whe can write the latter set of indices as the union of the following sequence of *finite* sets: $\{\alpha \in A : \ |a_\alpha| > 1/k\}$, $k = 1, 2, 3, \ldots$

3. All the spaces ℓ^p, $L^p(\Omega)$, $L^p(\mu)$ are honest Banach spaces: details will come later!

4. The space $\mathcal{C}^1([a, b])$ with the norm $\|f\|_{\mathcal{C}^1} := \|f\|_\infty + \|f'\|_\infty$ is a Banach space.

   Indeed, let us check that this obvious norm is complete. Let $\{f_n\}$ be a Cauchy sequence in $\mathcal{C}^1$. Then both sequences $\{f_n\}$ and $\{f_n'\}$ are Cauchy sequences for the norm $\|\cdot\|_\infty$, which is a complete norm on the space $\mathcal{C}^0$: it follows that there are continuous functions f, g such that $f_n \to f$ and $f_n' \to g$ uniformly. Now, for every $x \in [a, b]$ we have

$$f_n(x) = f_n(a) + \int_a^x f_n'(t) \, dt.$$

   Passing to the limit as $n \to +\infty$ we get

$$f(x) = f(a) + \int_a^x g(t) \, dt,$$

whence $f \in C^1$ and $f' = g$.

5. The space $\mathcal{C}^1([a,b])$ with the norm $\|\cdot\|_\infty$ is not complete: there are sequence of $\mathcal{C}^1$ functions which converge uniformly to functions which are not differentiable. For instance, take $[a,b] = [-1,1]$ and $f_n(x) = \sqrt{x^2 + 1/n}$: this sequence converges uniformly to $|x|$.

The proposition about the existence of discontinuous linear functionals has a... sad but inevitable consequence: when an analyst mentions the *dual* of a normed vector space X (or more generally of a topological vector space), he *never* refers to the algebraic dual spaces, i.e. to the space of *all* linear functionals over X, but rather to the vector subspace of the *continuous* linear functionals, which is also called the *topological dual space*.

DEFINITION (topological dual space): Given a normed vector space (a topological vector space) X, its *(topological) dual space* is the vector space X' of all *continuous* linear functionals $T : X \to \mathbf{R}$.

The following characterization holds:

THEOREM (Characterization of the continuous linear functionals on a normed space): let $(X, \|\cdot\|)$ be a normed vector space, $T : X \to \mathbf{R}$ a linear functional. The following facts are equivalent:

(i) T is continuous;

(ii) T is continuous at the origin;

(iii) T is a bounded functional: there exists a constant $C > 0$ such that

$$|T(x)| \leq C\|x\| \quad \forall x \in X;$$

(iv) The kernel of T, $\ker(T)$ is a closed subspace of X.

PROOF: Obviously $(i) \Rightarrow (ii)$, while $(ii) \Rightarrow (i)$ comes from the fact that $|T(x) - T(\overline{x})| = |T(x - \overline{x})|$ by linearity. $(iii) \Rightarrow (ii)$ is also obvious.

Conversely, let us show that $(ii) \Rightarrow (iii)$: by definition of continuity at 0 (with $\varepsilon = 1$), there exists $\delta > 0$ such that $|T(x)| \leq 1$ whenever $\|x\| \leq \delta$. Then for every $y \in X$, $y \neq 0$, we have

$$|T(y)| = |T(\frac{\|y\|}{\delta} \cdot y\frac{\delta}{\|y\|})| = \frac{\|y\|}{\delta}|T(y\frac{\delta}{\|y\|})| \leq \frac{1}{\delta}\|y\|,$$

and T is bounded.

It it obvious that $(i) \Rightarrow (iv)$, because the preimage of the closed set $\{0\}$ under the continuous function T is closed.

Next time, we will finish the proof by showing that $(iv) \Rightarrow (ii)$.

5 Lecture of october 12, 2015 (2 hours)

To finish the proof of last result, we show that $(iv) \Rightarrow (ii)$. If $T \equiv 0$ we have nothing to prove. Otherwise, suppose $\ker(T)$ is closed and assume by contradiction that T is discontinuous at 0. Then there is a sequence $\{x_n\} \subset X$ with $\|x_n\| \to 0$ and such that $T(x_n) \not\to 0$. Up to subsequences, this implies that there is a constant $c > 0$ such that $|T(x_n)| \geq c$ for every n.

Take an arbitrary point $y \in X$ and consider the sequence $y_n = y - \frac{x_n}{T(x_n)}T(y)$: one immediately checks that $y_n \in \ker(T)$ and $y_n \to y$ in the norm. It follows that $y \in \overline{\ker(T)} = \ker(T)$, whence $\ker(T) = X$ and $T \equiv 0$, which contradicts our hypothesis that $T \not\equiv 0$. Q.E.D.

We discovered that a linear functional $T : X \to \mathbf{R}$ is continuous if and only if it is bounded, i.e. if there is a constant $C > 0$ such that

$$|T(x)| \leq C\|x\| \quad \forall x \in X.$$

The *norm* of $T \in X'$ is the *smallest* constant C for which the inequality holds. Precisely, we define

$$(*)\ \|T\|_{X'} := \sup\{\frac{|T(x)|}{\|x\|} :\ x \in X,\ x \neq 0\} = \sup\{|T(x)| :\ x \in X,\ \|x\| \leq 1\}.$$

We begin with a couple of examples I didn't have the time to discuss yesterday:

- The space $\mathcal{C}^1([a,b])$ with the norm $\|f\|_{\mathcal{C}^1} := \|f\|_\infty + \|f'\|_\infty$ is a Banach space.

  Indeed, let us check that this obvious norm is complete. Let $\{f_n\}$ be a Cauchy sequence in $\mathcal{C}^1$. Then both sequences $\{f_n\}$ and $\{f_n'\}$ are Cauchy sequences for the norm $\|\cdot\|_\infty$, which is a complete norm on the space $\mathcal{C}^0$: it follows that there are continuous functions f, g such that $f_n \to f$ and $f_n' \to g$ uniformly. Now, for every $x \in [a,b]$ we have

  $$f_n(x) = f_n(a) + \int_a^x f_n'(t)\ dt.$$

  Passing to the limit as $n \to +\infty$ we get

  $$f(x) = f(a) + \int_a^x g(t)\ dt,$$

  whence $f \in C^1$ and $f' = g$.

- The space $\mathcal{C}^1([a,b])$ with the norm $\|\cdot\|_\infty$ is not complete: there are sequence of $\mathcal{C}^1$ functions which converge uniformly to functions which are not differentiable. For instance, take $[a,b]=[-1,1]$ and $f_n(x)=\sqrt{x^2+1/n}$: this sequence converges uniformly to $|x|$.

We already mentioned the fact that all norms on $\mathbf{R}^n$ are equivalent. Here is the proof:

THEOREM: All norms on $\mathbf{R}^n$ *are equivalent.*

PROOF: Because of transitivity of the equivalence between norms, it suffices to show that the euclidean norm $|\cdot|$ is equivalent to any other norm $\|\cdot\|$. We have to check that there are constants $c, C>0$ such that $c|x|\le\|x\|\le C|x|$ for every $x\in\mathbf{R}^n$. If $x\neq 0$, dividing by $|x|$ and using the homogeneity of norms, we see that our inequalities are true if and only if

$$(**)\ c\le\|x\|\le C \qquad \forall x\in\mathbf{R}^n, |x|=1.$$

Now, the function $x\mapsto\|x\|$ is continuous w.r.t. the euclidean topology. As a matter of fact, it is lipschitz continuous because of the following inequalities:

$$\begin{aligned}|\|x\|-\|y\|| &\le \|x-y\| = \|\sum_{i=1}^n (x_i-y_i)e_i\| \le \sum_{i=1}^n |x_i-y_i|\|e_i\| \le \\ & |x-y|(\sum_{i=1}^n \|e_i\|),\end{aligned}$$

where e_i are the vectors of the canonical basis of $\mathbf{R}^n$ (and we used the homogeneity of the norm and the triangle inequality).

On the other hand, the unit sphere $S=\{x\in\mathbf{R}^n:\ |x|=1\}$ is compact for the euclidean topology, so (**) is fulfilled with $c=\min\{\|x\|:\ x\in S\}$, $C=\max\{\|x\|:\ x\in S\}$, which exist by Weierstrass theorem (and $c\neq 0$ because the norm $\|\cdot\|$ is non degenerate). Q.E.D.

From this result, it follows that all norms on a real, finite dimensional vector space are equivalent, and that every linear isomorphism between such a space and $\mathbf{R}^n$ with the euclidean norm is a homeomorphism (exercise!). So all normed, finite dimensional vector spaces are isomorphic to $\mathbf{R}^n$ as normed spaces.

Let us go back to the topological dual of a normed space: this is always complete!

THEOREM (dual of a normed space): Let X be a (not necessarily complete) normed vector space over $\mathbf{R}$. Consider the topological dual X' of X, equipped with the dual norm (*). Then $(X', \|\cdot\|_{X'})$ is a Banach space.

PROOF: Checking that $\|\cdot\|_{X'}$ is a norm is an easy exercise we leave to the reader.

We need to show completeness: let $\{T_n\} \subset X'$ be a Cauchy sequence in the dual norm $\|\cdot\|_{X'}$. For a fixed $\varepsilon > 0$, there exists $\nu \in \mathbf{N}$ such that $\|T_n - T_m\|_{X'} \leq \varepsilon$ for $m, n \geq \nu$. Then, for every $x \in X$ and $m, n \geq \varepsilon$ we have

$$(**)\ |T_n(x) - T_m(x)| \leq \varepsilon \|x\|,$$

and the real sequence $\{T_n(x)\}$ is Cauchy and converges to a real number we denote $T(x)$. This pointwise limit $T : X \to \mathbf{R}$ is clearly linear. To conclude, we only need to verify that T is continuous and $T_n \to T$ in the norm of X'.

Passing to the limit as $m \to +\infty$ in $(**)$ we get

$$|T_n(x) - T(x)| \leq \varepsilon \|x\| \quad \forall x \in X,\ \forall n \geq \nu,$$

i.e. $\|T_n - T\|_{X'} \leq \varepsilon$ for $n \geq \nu$. T is also bounded because, by the previous inequality,

$$\|T\|_{X'} \leq \|T_\nu\|_{X'} + \|T_\nu - T\|_{X'} < +\infty.$$

Q.E.D.

REMARK/EXERCISE: With a similar argument, you can prove the following important fact. Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ be normed spaces, with Y a Banach space. Then the vector space $\mathcal{L}(X;Y)$ of continuous linear maps between X and Y is a Banach space with the norm

$$\|T\|_{\mathcal{L}(X;Y)} = \sup\{\|T(x)\|_Y :\ x \in X,\ \|x\|_X \leq 1\}.$$

EXAMPLE: We show that the supremum in the definition of the dual norm is not always a maximum. Consider the vector space $\ell^1 = \{\{x_n\}_{n\in\mathbf{N}} : \sum_{n=1}^{\infty} |x_n| < +\infty\}$, with the norm

$$\|\{x_n\}\|_{\ell^1} = \sum_{n=1}^{\infty} |x_n|.$$

We will show later on that this is a Banach space.

Consider next the linear functional $T : \ell^1 \to \mathbf{R}$ defined by

$$T(\{x_n\}) = \sum_{n=1}^{\infty} (1 - 1/n) x_n.$$

One easily checks that this is well defined on ℓ^1 (the series is absolutely convergent) and it is linear. It is also bounded:

$$(***)|T(\{x_n\})| \leq \sum_{n=1}^{\infty}(1-1/n)|x_n| \leq \|\{x_n\}\|_{\ell^1},$$

and notice that the *last inequality is strict if* $\{x_n\}$ *is different from the zero sequence.*

From (***) we infer $\|T\|_{(\ell^1)'} \leq 1$. On the other hand, the dual norm is 1: there is a sequence e_k in ℓ^1 (a sequence of sequences...) such that $\|e_k\|_{\ell^1} = 1$ and $T(e_k) \to 1$. It suffices to choose as e_k the "k-th canonical basis vector", i.e. the sequence whose k-th element is 1, while all other elements are 0: we thus have $T(e_k) = 1 - 1/k$ and

$$\|T\|_{(\ell^1)'} = 1.$$

The maximum in the definition of dual norm is not attained because (***) is strict for all non zero sequences.

6 Lecture of october 14, 2015 (2 hours)

One of the most important and useful results for the study of the dual space of a Banach space is the following:

THEOREM (Hahn-Banach): Let X be a real vector space, $p : X \to [0, +\infty)$ a map such that

(i) $p(\lambda x) = \lambda p(x)$ for all $x \in X$ and for all $\lambda > 0$ (positive homogeneity);

(ii) $p(x + y) \leq p(x) + p(y)$ for all $x, y \in X$ (subadditiviy).

Let Y be a proper subspace of X, $T : Y \to \mathbf{R}$ a linear functional such that $T(x) \leq p(x)$ for every $x \in Y$. Then there exists a linear functional $\tilde{T} : X \to \mathbf{R}$ wich extends T (i.e. $\tilde{T}(x) = T(x)$ for every $x \in Y$) and such that $\tilde{T}(x) \leq p(x)$ for every $x \in X$.

As a particular case, if X is a normed space, every bounded linear functional on Y extends to a linear functional defined on the whole space X and whith the same dual norm.

Before we see the proof, let us derive some important consequences of the Hahn-Banach theorem. One of these is the following: *if $x \in X$ there exists $T \in X'$ such that $\|T\| = 1$ and $T(x) = \|x\|$.*

Indeed, this functional can be defined first on the straight line $\mathbf{R}\{x\}$ by letting $T(tx) = t\|x\|$ (its dual norm is obviously 1), and then extended to the whole space X with the Hahn-Banach theorem.

In particular, this also implies that for every $x \in X$ one has

$$\|x\| = \max\{T(x) : \ T \in X', \ \|T\| \leq 1\} :$$

one inequality is obvious, the equal sign follows by using the functional constructed above!

Another consequence of the theorem is the fact that *the dual space of a normed spaces separates points*: given any two vectors $x, y \in X$, we can always find an element of the dual spaces which takes *different values* on those two points. Indeed, it is enough to choose a functional which takes the value $\|x-y\|$ on the vector $x-y$: it follows that $T(x)-T(y) = T(x-y) \neq 0$. Notice that this property does not hold for a general topological vector space: there are examples of quite honest (metrizable and complete) topological vector spaces, whose topological dual contains only the zero functional!

PROOF of the Hahn-Banach Theorem: The main part of the proof consists in proving the following assert: *if Z is a proper subspace of X and $T : Z \to \mathbf{R}$ is linear with $T(x) \leq p(x)$ for every $x \in Z$, then T extends to a strictly larger subspace in such a way that the inequality still holds for the extension.*

To this aim, choose $x_0 \in X \setminus Z$: we extend T to a functional $\tilde{T}$ defined on the space $Z \oplus \mathbf{R}\{x_0\}$. By linearity we have, for every $x \in Z$ and all $t \in \mathbf{R}$:

$$\tilde{T}(x + tx_0) = T(x) + t\alpha,$$

where $\alpha = T(x_0)$ is a real number we must choose in such a way that

$$T(x) + t\alpha \leq p(x + tx_0) \quad \forall x \in Z, \ \forall t \in \mathbf{R}.$$

By using again the linearity of T and the positive homogeneity of p, and by distinguishing the cases $t > 0$ and $t < 0$, we easily check that the last inequality is equivalent to the following two:

$$\begin{aligned} T(x) + \alpha &\leq p(x + x_0) \ \forall x \in Z, \\ T(y) - \alpha &\leq p(y - x_0) \ \forall y \in Z, \end{aligned}$$

that is, α must satisfy

$$T(y) - p(y - x_0) \leq \alpha \leq p(x + x_0) - T(x) \quad \forall x, y \in Z.$$

This is certainly possible, provided the left hand side is always smaller or equal than the right hand side, for every choice of $x, y \in Z$. But this is true because $T(x) + T(y) = T(x+y) \leq p(x+y) \leq p(x+x_0) + p(y-x_0)$ for all $x, y \in Z$. Our claim is proved.

To conclude the proof we need one of the forms of the choice axiom, for instance Hausdorff's maximality principle. Consider indeed the family $\mathcal{F}$ of all linear functionals defined on subspaces Z of X which contain Y, which are also extensions of the original functional T and are dominated by the function p. We next put an order relation on $\mathcal{F}$: if $R : Z_1 \to \mathbf{R}$ and $S : Z_2 \to \mathbf{R}$ are two elements of $\mathcal{F}$, then $R \preceq S$ iff $Z_1 \subset Z_2$ and S extends R. By Hausdorff's maximality principle, we can find a *maximal* totally ordered subset $\mathcal{G} \subset \mathcal{F}$,

$$\mathcal{G} = \{S_\sigma : Z_\sigma \to \mathbf{R}, \sigma \in I\}.$$

This set has an upper bound: this is the functional S defined on the subspace

$$Z = \bigcup_{\sigma \in I} Z_\sigma$$

(why is it a subspace?) by $S(x) = S_\sigma(x)$ whenever $x \in Z_\sigma$ (this is a good definition because $\mathcal{G}$ is totally ordered).

But then $Z = X$ and S is the required extension: otherwise, the claim we proved at the beginning of this discussion would allow us to extend S to a strictly larger subspace, thus contradicting the maximality of the totally ordered subset $\mathcal{G}$: this concludes the proof of our main statement.

The "particular case" for normed spaces follows immediately by choosing $p(x) = \|T\|_{Y'} \, \|x\|$. Q.E.D.

EXERCISE: In general, given a linear functional T which is bounded on a subspace Y of $(X, \|\cdot\|)$, its extension $\tilde{T} \in X'$ given by the Hahn-Banach theorem is not unique: there are easy examples also in finitely dimensional spaces. Show that if Y is a *dense* subspace of X, that is if $\overline{Y} = X$, then the extension is unique.

Our discussion of the consequences of the Hahn-Banach theorem is far from concluded. But before we proceed further, we will make a more "concrete" digression, to build a good number of important examples of infinitely dimensional Banach spaces. In particular, some time ago we introduced - without proofs - the spaces ℓ^p: let us verify that the norm $\|\cdot\|_{\ell^p}$ is indeed a norm, and those spaces are complete. The same will be done for the spaces $L^p(\mu)$, with μ an arbitrary outer measure.

To this aim, we begin with a simple inequality in $\mathbf{R}$.

If $p \in (1, +\infty)$, its conjugate exponent q is defined through the equation

$$\frac{1}{p} + \frac{1}{q} = 1.$$

As a natural extension, the conjugate exponent of $p = 1$ is $q = +\infty$, and vice versa.

The following inequality holds (Young's inequality):

LEMMA: If p, q are conjugate exponent, $1 < p < +\infty$, then

$$ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q \qquad \forall a > 0, b > 0.$$

PROOF: By the concavity of the logarithm we get:

$$\log(ab) = \frac{1}{p}\log a^p + \frac{1}{q}\log b^q \leq \log\left(\frac{1}{p}a^p + \frac{1}{q}b^q\right),$$

whence the thesis.

A very important tool in the study of the spaces ℓ^p, L^p is Hölder inequality. We will need it, in particular, to show that these are normed spaces. The case $p = 1$ and $q = +\infty$ will be analyzed tomorrow, because we need to define the L^∞ norm!

PROPOSITION (Hölder inequality): Let $1 \leq p \leq +\infty$, q its conjugate exponent. Then for μ-measurable functions $u, v : X \to \overline{\mathbf{R}}$ the following holds:

$$\int_X |u(x)v(x)|\, d\mu(x) \leq \|u\|_{L^p(\mu)}\, \|v\|_{L^q(\mu)}.$$

PROOF: We first consider the case $1 < p < +\infty$: $p = 1$, $q = \infty$ is postponed till later.

Hölder's inequality is a simple consequence of Young's inequality. Notice first that the L^p norms are homogeneous. We may then suppose without loss of generality that $\|u\|_{L^p(\mu)} = \|v\|_{L^q(\mu)} = 1$. By Young's inequality we get

$$|u(x)v(x)| \leq \frac{1}{p}|u(x)|^p + \frac{1}{q}|v(x)|^q,$$

whence, integrating over X

$$\int_X |uv|\, d\mu \leq \frac{1}{p} + \frac{1}{q} = 1,$$

which is the inequality we had to prove[9].

We will see tomorrow the case $p = 1$, $q = +\infty$, *after* we define the L^∞ norm!

7 Lecture of october 15, 2015 (2 hours)

To prove Hölder's inequality with $p = 1$, we still need to define $L^\infty(\mu)$ and its norm, with μ a general outer measure.

DEFINITION ($L^\infty(\Omega)$): If μ is an outer measure on X and $u : X \to \overline{\mathbf{R}}$ is measurable, define

$$\|u\|_{L^\infty(\mu)} = \text{esssup}\{|u(x)| : x \in X\} := \inf\{t \in \mathbf{R} : \mu(\{x \in X : |u(x)| > t\}) = 0\}.$$

Loosely speaking, this is the sup of $|u(x)|$ *up to sets of measure zero.* Of course, $L^\infty(\mu)$ will be defined as the space of the functions with finite norm, quotiented by the usual equivalence relation.

REMARK: One has

$$\mu(\{x \in X : |u(x)| > \|u\|_{L^\infty(\mu)}\}) = 0.$$

Indeed, if $\|u\|_{L^\infty(\mu)} = +\infty$ there is nothing to prove, otherwise we can write this set as a countable union of sets with zero measure:

$$\{x \in X : |u(x)| > \|u\|_{L^\infty(\mu)}\} = \bigcup_{n=1}^{\infty}\{x \in X : |u(x)| > \|u\|_{L^\infty(\mu)} + \frac{1}{n}\}.$$

Let us check that Hölder's inequality is true in the limit casee $p = 1$, $q = +\infty$: we may suppose $\|v\|_{L^\infty} < +\infty$ (otherwise the inequality is obvious). We know that the set $A = \{x \in X : |v(x)| > \|v\|_\infty\}$ has measure 0. It follows that the integral of any non negative measurable function on X and on $X \setminus A$ are the same. Then

$$\int_X |u(x)v(x)|\, d\mu(x) \le \|v\|_\infty \int_X |u(x)|\, d\mu(x) = \|v\|_\infty \|u\|_{L^1},$$

as we wanted. Q.E.D.

With Hölder's inequality we finally prove that L^p norms are norms:

[9] In case one of the norms is zero, or if one or both are $+\infty$, the inequality is obvious!

THEOREM: If $1 \le p \le +\infty$, the spaces $(L^p(\mu), \|\cdot\|_{L^p(\mu)})$ are normed vector spaces.

PROOF: L^p norms are obviously non negative and homogeneous, they take finite values on the L^p spaces and are not degenerate. Indeed we saw that

$$\int_X |u(x)|^p \, d\mu(x) = 0 \Rightarrow u(x) = 0 \; per \; \mu - a.e. \; x \in X,$$

and so L^p norm is non degenerate thanks to the equivalence relation we have put on measurable functions.

We are left to prove the triangle inequality, which in this case is called *Minkowski's inequality.* Again, this is obvious in the limit cases $p = 1$ and $p = +\infty$.

If $1 < p < +\infty$, take two measurable functions u, v and apply Hölder's inequality:

$$\int_X |u+v|^p \le \int_X |u||u+v|^{p-1} + \int_X |v||u+v|^{p-1} \le (\|u\|_p + \|v\|_p) \left(\int_X |u+v|^p \right)^{(p-1)/p}.$$

Minkowski's inequality is then obtained by dividing both sides by $\left(\int_\Omega |u+v|^p \right)^{(p-1)/p}$.
But to do this, we must know that this quantity is finite as soon as $\|u\|_p$ and $\|v\|_q$ are finite: this follows from the convexity of the function $s \mapsto s^p$ because

$$\left| \frac{|u(x)| + |v(x)|}{2} \right|^p \le \frac{1}{2}|u(x)|^p + \frac{1}{2}|v(x)|^p.$$

Q.E.D.

REMARK: What we prove for the spaces $L^P(\mu)$ also applies to the spaces ℓ^p: we already proved that the latter are just particular case of the first, with $X = \mathbf{N}$ and μ the counting measure.

We are now in position to prove that the spaces $L^p(\mu)$ are indeed Banach spaces:

THEOREM (Riesz-Fischer): Let μ be an outer measure on a set X, $1 \le p \le +\infty$. Then the spaces $L^p(\mu)$ are complete. Moreover, given a sequence $\{f_n\}$ which converges to some function f in the norm L^p, we can extract a subsequence $\{f_{n_k}\}$ such that $f_{n_k}(x) \to f(x)$ for μ-a.e. $x \in X$.

PROOF:

We begin with the case $1 \leq p < +\infty$: $p = +\infty$ is different (and also easier!) and will be shown at the end.

Let $\{f_n\}$ be a Cauchy sequence in $L^p(\mu)$: it is easy to construct an increasing sequence of indices n_k in such a way that

$$\|f_{n_{k+1}} - f_{n_k}\|_{L^p} \leq \frac{1}{2^k}, \quad k = 1, 2, 3, \dots$$

Next, consider the functions

$$g_K(x) = \sum_{k=1}^{K} |f_{n_{k+1}}(x) - f_{n_k}(x)|, \quad g(x) = \sum_{k=1}^{\infty} |f_{n_{k+1}}(x) - f_{n_k}(x)|$$

(where the last series makes perfect sense, because its terms are non negative) . By the triangle inequality and our choice of n_k, we immediately check that $\|g_K\|_{L^p} \leq 1$ for every K. Moreover, the monotone convergence theorem ensures that $\|g_K\|_{L^p} \to \|g\|_{L^p}$, whence $g \in L^p(\mu)$, so that $|g(x)| < +\infty$ for almost every $x \in X$.

Consider now the telescopic sums

$$\sum_{k=1}^{K} (f_{n_{k+1}}(x) - f_{n_k}(x)) = f_{n_K}(x) - f_{n_1}(x)$$

and the corresponding series $\sum_{k=1}^{\infty} (f_{n_{k+1}}(x) - f_{n_k}(x))$: the latter is absolutely convergent for almost all x (the series of absolute values converges to $g(x)$), so the limit

$$\lim_{k \to +\infty} f_{n_k}(x) =: f(x)$$

exists for almost every $x \in X$. On the zero measure set where we do not have convergence, we can define the pointwise limit in an arbitrary way, for instance by putting $f(x) = 0$.

From the previous inequalities we also get

$$|f_{n_k}(x)| \leq (f_{n_1}(x) + g(x)) \in L^p(\mu),$$

and the pointwise limit satisfies the same inequality, so that $f \in L^p(\mu)$. Finally, the dominated convergence theorem ensures that $f_{n_k} \to f$ in $L^p(\Omega)$ (the sequence $f_{n_k} - f$ is dominated by $2g + |f_{n_1}| \in L^p$).

It is a very easy exercise to show that when a Cauchy sequence has a convergent subsequence, the *whole* sequence converges to the same limit: this fact holds in any metric space.

Let us consider now the case $p = +\infty$. Let then $\{f_n\}$ be a Cauchy sequence in L^∞. For every $k \in \mathbf{N}$ there exists an index n_k such that $\|f_m - f_n\|_\infty < \frac{1}{k}$ for every $m, n \geq n_k$. Then the sets

$$A_k = \{x \in X : |f_n(x) - f_m(x)| > 1/k \ for\ some\ m, n \geq n_k\}, \quad A = \bigcup_{k=1}^{\infty} A_k$$

have all measure zero (by definition of the L^∞ norm). Then for every $x \in X \setminus A$ the sequence $f_n(x)$ is a Cauchy sequence in $\mathbf{R}$, and it converges to a pointwise limit $f(x)$ (which we extend as above by putting for instance $f(x) = 0$ for $x \in A$). Passing to the limit as $m \to +\infty$ in the inequality

$$|f_n(x) - f_m(x)| \leq 1/k \quad \forall x \in X \setminus A,\ \forall m, n \geq n_k$$

we get

$$|f_n(x) - f_(x)| \leq 1/k \quad \forall x \in X \setminus A,\ \forall n \geq n_k,$$

whence $\|f_n - f\|_\infty \leq 1/k$ for all $n > n_k$ and $f_n \to f$ in L^∞.

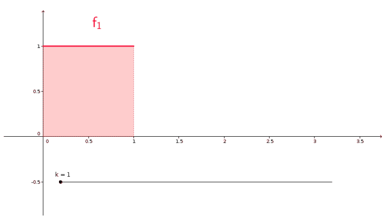
To finish the proof, we only have to check the last claim that whenever $f_n \to f$ in L^p we have a subsequence for which we have pointwise convergence a.e. to f. Of course, if f_n converges in L^p, it is a Cauchy sequence. The theorem we just proved gives us a subsequence which converges a.e. and in L^p to some functions: this is necessarily equal a.e. to f by the uniqueness of the limit in L^p. Q.E.D.

REMARK: In general, if $1 \leq p < +\infty$, convergence in $L^p(\mu)$ does not imply convergence a.e. of the *whole* sequence. This is true in the space L^∞ (see the proof of the theorem!).

We now give an example of a sequence $\{u_i\}$ which converges to 0 in $L^p([0,1])$ (for every $1 \leq p < +\infty$), but which does not converge to 0 in any point of the interval $[0,1]$. The construction goes as follows: if $i \in [2^k, 2^{k+1} - 1] \cap \mathbf{N}$, we put

$$u_i(x) = \mathbf{1}_{[0,2^{-k}]}\left(x - \frac{i - 2^k}{2^k}\right).$$

The following is an animation of the sequence:



We will next study the dual space of $L^p(\mu)$: we will see that in "most cases", it can be identified with $L^q(\mu)$, where q is the conjugate exponent to p.

More precisely, we will consider the linear map

$$\begin{array}{rcl}\Phi : L^q(\mu) & \to & (L^p(\mu))' \\ v & \mapsto & T_v\end{array}$$

where $T_v(u) := \int_X v(x)u(x)\,d\mu(x)$.

First of all, $T_v(u)$ is well defined: indeed, by Hölder's inequality the function uv is summable. Moreover, T_v is linear and, again by Hölder,

$$|T_v(u) \le \|v\|_{L^q} \cdot \|u\|_{L^p} \quad \forall u \in L^p,$$

so that T_v is continuous and $\|T_v\|_{(L^p)'} \le \|v\|_{L^q}$.

Next time, we will see that for $1 < p \le +\infty$ we actually have $\|T_v\|_{(L^p)'} = \|v\|_{L^q}$: the map Φ is a *linear isometry* between the two spaces. The same holds also for $p = 1$ if we make some mild assumption on the measure.

Moreover, for $1 \le p < +\infty$ the map Φ is surjective. So, for $1 \le p < +\infty$ (and μ "good enough" in case $p = 1$), the map Φ is actually a linear *isomestric isomprphism* between L^q and $(L^p)'$.

8 Lecture of october 19, 2015 (2 hours)

EXERCISE: Show that if the measure of Ω is finite, then a function in $L^p(\Omega)$ belongs to $L^r(\Omega)$ for every $r \in [1, p]$ (in other words, the spaces L^p become smaller and smaller when the exponent grows). Find an example which shows that this is no longer true if the measure of Ω is $+\infty$ *(Hint: Apply Hölder's inequality to the product $|u(x)|^r \cdot 1$, using as exponents p/r and its conjugate...]*

An explicit study of the dual spaces of ℓ^p would not be overly difficult:

PROPOSITION: Let $1 \le p \le +\infty$, q coniugate to p, $\{y_k\} \in \ell^q$. Define a linear map $T_{\{y_k\}} : \ell^p \to \mathbf{R}$ as follows:

$$(*)\ T_{\{y_k\}}(\{x_k\}) = \sum_{k=1}^{\infty} y_k x_k.$$

Then the map $\Phi : \{y_k\} \mapsto T_{\{y_k\}}$ is a well-defined linear isometry from ℓ^q into $(\ell^p)'$. If $1 \le p < +\infty$, then the map Φ is surjective, and provides an isometri isomorphism between ℓ^q and the dual of ℓ^p.

REMARK: Let μ be an outer measure on a set X We can construct in the same way an isometric injection of $L^q(\mu)$ in the dual space of $L^p(\mu)$: if $1 < p \le +\infty$ and q is the conjugate exponent, to each $v \in L^q(\Omega)$ we associate a functional $T_v \in (L^p)'$ defined by

$$T_v(u) = \int_X u(x)v(x)\, dx \quad \forall u \in L^p(X).$$

We show that $\Phi : v \mapsto T_v$ is a linear isometry between $L^q(\mu)$ and $(L^p(\mu))'$.

By Hölder's inequality it follows that $\|T_v\|_{(L^p)'} \le \|v\|_{L^q}$. But choose $u(x) = sgn(v(x))|v(x)|^{q-1}/\|v\|_{L^q}^{q-1}$: one immediately checks that the L^p norm of this function is 1 and $T_v(u) = \|v\|_{L^q}$, so we get equality of the norms.

Notice that this works also for $p = +\infty$ and $q = 1$: the map above is an isometry, with identical proof!

We will discuss the case $p = 1$, $1 = +\infty$ next time: in this case, *if the measure μ is sufficiently "nice"*, then the map Φ is again a linear ismometry. For instance, this is true for the Lebesgue measure or for the counting measure on $\mathbf{N}$. But there are "patological" measures for which the result is false.

This whole discussion shows the first part of the proposition on the dual spaces of ℓ^p. To show surjectivity in the case $1 \le p < +\infty$ would be a relatively easy exercise for the spaces ℓ^p, but we will prove later the result later for a large class of measures μ (the class of the σ-finite meausres).

As a matter of fact, for $1 < p < +\infty$ the linear isometry Φ is *always* an isomorphism between the Banach spaces $L^q(\mu)$ and $(L^p(\mu))'$, for every measure μ.

The linear isometry $\Phi : L^q(\mu) \to (L^p(\mu))'$ in general is *not surjective* if $q = 1$, $p = +\infty$: we will show with an example that the *dual space of ℓ^∞ is strictly larger than ℓ^1* or, more precisely, its image under Φ.

On the contrary, for $p = 1$, $q = \infty$ and for "nice measures", it is both an isometry and surjective.

In order to be sure that the map $\Phi : L^\infty(\mu) \to (L^1(\mu))'$ defined above is an isometry, we need an hypothesis on the measure: we need to know that *every measurable set with infinite measure has a measurable subset with finite and strictly positive measure.* Suppose that this is true, and notice that for every $\varepsilon > 0$ there is a subset $A \subset X$ such that $0 < \mu(A) < +\infty$ and $|v(x)| \geq \|v\|_{L^\infty} - \varepsilon$ for all $x \in A$[10] Consider the function

$$u(x) = \begin{cases} \operatorname{sgn}(v(x))/m(A) & if\ x \in A, \\ 0 & otherwise \end{cases}$$

Then $\|u\|_{L^1} = 1$ and $\int\limits_\Omega u(x)v(x)\,dx \geq \|v\|_{L^\infty} - \varepsilon$. But $\varepsilon > 0$ is arbitrary, and we deduce equality of the norms.

Again, in this case one can show that the map Φ is surjective: for "nice enough" measures, the dual of L^1 can be identified with L^∞.

By summarizing, the following holds:

THEOREM(Duals of the L^p spaces): Let μ be an outer measure on a set X. Consider the map

$$\begin{array}{llll} \Phi: & L^q(\mu) & \to & (L^p(\mu))' \\ & v & \mapsto & T_v \end{array}$$

where $T_v(u) := \int\limits_X u(x)v(x)\ d\mu(x)$ for every $u \in L^p(\mu)$. If $1 < p < +\infty$, then the map Φ is a isometric isomorphism. If $p = 1$, $q = +\infty$ and the measure μ has the property that every measurable set with infinite measure has a measurable subset with finite and strictly positive measure, then Φ is again an isometric isomorphism.

Finally, if $p = +\infty$, $q = 1$, then the map Φ is a linear isometry, but in general it is not surjective (see the following two examples).

We will prove later the missing part of the theorem, i.e. the surjectivity of Φ for finite p, and we will only do this for σ-finite measures.

EXAMPLE: Consider the linear subspace c of ℓ^∞ of those sequences $\{a_k\}_k$ which have a *finite limit* as $k \to +\infty$. Define a linear functional $T : c \to \mathbf{R}$ as follows:

$$T(\{a_k\}_k) = \lim_{k\to+\infty} a_k.$$

This is an element of the dual space of $(c, \|\cdot\|_{\ell^\infty})$: one immediately checks its norm is 1.

By the Hahn-Banach theorem, T can be extended to an element of norm 1 of $(\ell^\infty)'$ we still denote T: such a functional is known as a *Banach limit.*

I claim that $T \notin \Phi(\ell^1)$, i.e. there exists no sequence $\{y_k\}_k \in \ell^1$ such that

$$(*)\ T(\{a_k\}) = \sum_{k=1}^\infty y_k a_k \quad \forall \{a_k\} \in \ell^\infty.$$

[10] If $\mu(\{x \in X :\ |v(x)| \geq \|v\|_{L^\infty} - \varepsilon\}) = +\infty$, use the hypothesis on the measure to replace this set with a smaller one!

Suppose by contradiction such a $\{y_k\}$ exists, and apply T to the elements e^h of the "canonical basis" (i.e. $e^h_k = \delta_{hk}$)): obviously for each fixed h we have $T(e^h) = 0$, and plugging this into (*) we get $y_h = 0$. This contradicts the fact that T is not the zero functional.

EXAMPLE: The dual of $L^\infty(\Omega)$ is also strictly larger than $L^1(\Omega)$. We show an example of a bounded linear functional $T \in (L^\infty([-1,1]))'$ for which there is no function $v(x) \in L^1([-1,1])$ such that

$$T(u) = \int_{-1}^{1} uv \, dx \quad \forall u \in L^\infty.$$

We first define T on the subspace $C^0([-1,1])$ by putting $T(u) = u(0)$ for every continuous function u. This is a linear functional of norm 1 (check it!), which we extend to the whole space L^∞ thanks to the Hahn-Banach theorem. Suppose by contradiction there is a function v as above: show that one gets $v(x) = 0$ for a.e. $x \in [-1,1]$, a contradiction because $T \not\equiv 0$ (use the following fact: given the function $\mathrm{sgn}(v(x))$, it is possible to construct a sequence u_h of continuous functions with $u_h(x) \to \mathrm{sgn}(v(x))$ for a.e. $x \in [-1,1]$, $u_h(0) = 0$ and $-1 \le u_h(x) \le 1$ for every $x \in [-1,1]$. The existence of such a sequence follows from some approximation results for L^p functions we will prove later on. Use then the dominated convergence theorem.).

In the last two examples we used the Hahn-Banach theorem. We will now see a couple of important "geometrical"consequences of this theorem, which tell us that sometimes a pair of convex sets can be separated by means of a closed hyperplane.

To this aim, we introduce the fundamental notion of the *Minkowski functional* of a convex set:

PROPOSITION (Minkowski functional): Let $(X, \|\cdot\|)$ be a normed space, C a convex open subset of X containing 0. We define the Minkowski functional of C in the following way:

$$p(x) = \inf\{t > 0 : \ \frac{x}{t} \in C\}.$$

Then $p(x)$ is a well-defined, real, positively homogeneous and subadditive function[11]*. Moreover, there is a constant $K > 0$ such that*

$$(*) \ p(x) \le K\|x\| \quad \forall x \in X,$$

and finally

$$(**) \ C = \{x \in X : \ p(x) < 1\}.$$

[11] Compare with the statement of Hahn-Banach theorem

PROOF: Let $r > 0$ be such that $B_r(0) \subset C$ (possible because C is open): for every $x \in X$ we then have $r\frac{x}{2\|x\|} \in C$, whence $p(x) \leq \frac{2}{r}\|x\|$ and (*) is proved. In particular, $p(x)$ is everywhere finite. Positive homogeneity is immediate. Let us show (**): if $x \in C$, we use the fact that C is open to find $r > 0$ such that $(1+r)x \in C$. It follows that $p(x) \leq \frac{1}{1+r} < 1$. If conversely $p(x) < 1$ we can find $0 \leq t < 1$ such that $\frac{x}{t} \in C$. But then $x = t(\frac{x}{t}) + (1-t)0 \in C$ thanks to convexity of C.

We are left to prove subadditivity (the triangle inequality): let $x, y \in X$. By definition of $p(x)$, for every $\varepsilon > 0$ we have

$$x_0 = \frac{x}{p(x)+\varepsilon} \in C, \quad y_0 = \frac{y}{p(y)+\varepsilon} \in C.$$

Take a convex combination $tx_0 + (1-t)y_0$ with

$$t = \frac{p(x)+\varepsilon}{p(x)+p(y)+2\varepsilon}, \quad (1-t) = \frac{p(y)+\varepsilon}{p(x)+p(y)+2\varepsilon}.$$

We thus deduce that

$$\frac{x+y}{p(x)+p(y)+2\varepsilon} \in C,$$

whence

$$p(x+y) \leq p(x)+p(y)+2\varepsilon.$$

Since ε is arbitrary, we have subadditivity. Q.E.D.

EXERCISE: Show that if $C = B_1(0)$ is the unit open ball of our normed space, the corresponding Minkowski functional is $p(x) = \|x\|$.

9 Lecture of october 20, 2015 (2 hours)

To show the geometric consequences of the Hahn-Banach theorem we will need the following lemma, which is of independent interest:

LEMMA: Let C be a nonempty, convex open set in X, $x_0 \in X \setminus C$. Then there exists $T \in X'$ such that

$$T(x) < T(x_0) \quad \forall x \in C.$$

PROOF: Choose $\overline{y} \in C$ and define $\tilde{C} = C - \{\overline{y}\}$, $\overline{x} = x_0 - \overline{y}$. $\tilde{C}$ is a convex open set containing the origin and $\overline{x} \notin C$. If p denotes the Minkowski functional of $\tilde{C}$, we have $p(\overline{x}) \geq 1$ while $\tilde{C} = \{x : \ p(x) < 1\}$.

On the 1-dimensional subspace $Y = \mathbf{R}\{\overline{x}\}$ consider the linear functional $T \in Y'$ such that $T(t\overline{x}) = tp(\overline{x})$. By positive homogeneity of the Minkowski functional we immediately check that $T(t\overline{x}) \leq p(t\overline{x})$ for all $t \in \mathbf{R}$.

By Hahn-Banach, we can extend T to a linear functional defined on the whole of X, in such a way that $T(x) \leq p(x)$ for every $x \in X$. We also have $T \in X'$ thanks to (*) in previous proposition.

Thus $T(x) < 1$ for all $x \in \tilde{C}$, while $T(\overline{x}) = p(\overline{x}) \geq 1$. By adding $\overline{y}$ and using linearity of T we deduce $T(x) < 1 + T(\overline{y})$ for all $x \in C$, while $T(x_0) \geq 1 + T(\overline{y})$. Q.E.D.

Here is the "geometric version" of Hahn-Banach theorem:

THEOREM (Hahn-Banach, geometric version): Let $(X, \|\cdot\|)$ be a normed space, A and B nonempty, disjoint convex subsets of X. Then

(i) If A is open, there are $T \in X'$, $T \neq 0$, $\alpha \in \mathbf{R}$ such that

$$T(x) \leq \alpha \leq T(y) \quad \forall x \in A, \ \forall y \in B.$$

Geometrically, we can say that the (affine) closed hyperplane $T(x) = \alpha$ separates the convex sets A and B.

(ii) If A and B are closed and A is also compact, then there are $T \in X'$, $T \neq 0$, $\alpha \in \mathbf{R}$ and $\varepsilon > 0$ such that

$$T(x) \leq \alpha - \varepsilon \ \forall x \in A, \quad \alpha + \varepsilon \leq T(y) \ \forall y \in B.$$

Geometrically, the (affine) closed hyperplane $T(x) = \alpha$ strictly separates the convex sets A and B.

.

PROOF: We first show (i): put $C = A - B = \{x - y : \ x \in A, \ y \in B\}$ (beware, this is an *algebraic* difference of sets!).

One immediately checks that C is convex. It is also open, because we can write $A - B = \bigcup_{y \in B} (A - \{y\})$.

Apply the lemma with $x_0 = 0$: notice indeed that $0 \notin C$ because A and B are disjoint. We find $T \in X'$ such that

$$T(x) < 0 = T(0) \quad \forall x \in C,$$

whence (linearity of T)

$$T(x) < T(y) \quad \forall x \in A, \ \forall y \in B.$$

(i) then follows by putting $\alpha = \sup\{T(x) : \ x \in A\}$.

To prove (ii), we set $A_\varepsilon = A + B_\varepsilon(0)$, $B_\varepsilon = B + B_\varepsilon(0)$. These are obviously open convex sets. I claim that $A_\varepsilon \cap B_\varepsilon = \emptyset$ for small enouch ε: otherwise, we could find a sequence $z_n \in A_{\frac{1}{n}} \cap B_{\frac{1}{n}}$. Now, we could write $z_n = a_n + w_n = b_n + w_n'$, with $a_n \in A$, $b_n \in B$ and $\|w_n\| < 1/n$, $\|w_n'\| < 1/n$. Using the compactness of A and up to subsequences, we would find then $a_n \to \overline{a} \in A$. But then (since $w_n \to 0$, $w_n' \to 0$) we would also have $b_n \to \overline{a}$, whence $\overline{a} \in B$ (since B is closed). This is impossible since $A \cap B = \emptyset$.

Take then a small enough ε and apply (i) to the convex sets A_ε, B_ε: we find $T \in X'$ and $\alpha \in \mathbf{R}$ such that

$$\begin{aligned} T(a+w) &\leq \alpha \quad \forall a \in A, \ w \in B_\varepsilon(0) \\ T(b+w') &\geq \alpha \quad \forall b \in B, \ w' \in B_\varepsilon(0). \end{aligned}$$

Passing to the sup on w and to the inf on w' we get

$$\begin{aligned} T(a) + \varepsilon\|T\| &\leq \alpha \quad \forall a \in A \\ T(b) - \varepsilon\|T\| &\geq \alpha \quad \forall b \in B. \end{aligned}$$

Q.E.D.

The following is an important consequence of the "geometric" Hahn-Banach theorem:

COROLLARY: Let $(X, \|\cdot\|)$ be a normed space, Y a vector subspace. Then Y is dense iff every linear functional $T \in X'$ which vanishes identically on Y is the zero functional.

PROOF: If Y is dense, then obviously every continuous linear functional which vanishes on Y is identically zero.

Conversely, suppose Y is not dense, i.e. $\overline{Y}$ is a proper subspace of X. We will find a non-zero bounded linear functional which vanishes identically on Y. To this end, take $x_0 \in X \setminus \overline{Y}$ and apply (ii) of the previous theorem to the convex sets $\overline{Y}$ and $\{x_0\}$ (the second of which is compact). There exists $T \in X'$ such that $T(x) < T(x_0)$ for every $x \in \overline{Y}$. By linearity of T we immediately deduce $T \equiv 0$ on Y (a non-zero linear functional has *never* a bounded image!). So, $T(x) = 0$ for every $x \in Y$, while $T(x_0) \neq 0$. Q.E.D.

REMARK: Incidentally, the proof of the corollary also suggests that is in not always possible to separate disjoint convex sets with an hyperplane: there is no hyperplane separating a *proper dense subspace* of X and a point outside the subspace!

Another important result for the study of dual spaces is the following

THEOREM (Banach-Steinhaus): Let $(X, \|\cdot\|)$ be a Banach space, $\{T_k\} \subset X'$ be a sequence of continuous linear functionals on X which is pointwise bounded, i.e. such that

$$\sup\{|T_k(x)| : k \in \mathbf{N}\} < +\infty \quad \forall x \in X.$$

Then the sequence is uniformly bounded: there exists $C > 0$ such that $\|T_k\|_{X'} \leq C$ for every $k \in \mathbf{N}$.

To prove the theorem we need the following important result:

THEOREM (Baire): Let (X, d) be a complete metric space. If $\{F_k\}$ is a sequence of closed sets in X with empty interiors, then $\bigcup_{k\in\mathbf{N}} F_k$ has empty interior.

Passing to the complements, this means that in a complete metric space a countable intersection of open and dense subsets is still dense.

PROOF: Let $\Omega \subset X$ be a fixed nonempty open set: we need to show that $\Omega \setminus \bigcup_{k\in\mathbf{N}} F_k$ is nonempty.

Choose $x_1 \in X$, $1 > r_1 > 0$ such thate $B_{r_1}(x_1) \subset \Omega$ and $B_{r_1}(x_1) \cap F_1 = \emptyset$: this is possible because the complement of F_1 is open and dense. Now, there are points of the complement of F_2 whitin $B_{r_1}(x_1)$ (because F_2 has empty interior): we can choose x_2, r_2 such that $B_{r_2}(x_2) \cap F_2 = \emptyset$, $\overline{B_{r_2}(x_2)} \subset B_{r_1}(x_1)$ and $r_2 < 1/2$.

Proceeding in the same way, we build a sequence $\{x_k\} \subset X$ and positive real numbers $\{r_k\}$ such that

$$B_{r_k}(x_k) \cap F_k = \emptyset, \ \overline{B_{r_k}(x_k)} \subset B_{r_{k-1}}(x_{k-1}), \ r_k < 1/k.$$

Thanks to the fact that the radius of the balls goes to zero, one immediately sees that $\{x_k\}$ is a Cauchy sequence. Its limit $\overline{x}$ has the property that

$$\overline{x} \in \bigcap_{k\in\mathbf{N}} B_{r_k}(x_k).$$

On the other hand,

$$\bigcap_{k\in\mathbf{N}} B_{r_k}(x_k) \subset \Omega \setminus \bigcup_{k\in\mathbf{N}} F_k,$$

so the latter set is nonempty. Q.E.D.

10 Lecture of october 21, 2015 (2 hours)

We can now prove the theorem of Banach-Steinhaus:

PROOF of the Banach-Steinhaus theorem: Put

$$F_k = \{x \in X : \; |T_h(x)| \le k \; \forall h \in \mathbf{N}\}.$$

These are closed sets whose union is X by the pointwise boundedness hypothesis. By the Baire theorem, there exists an index $\overline{k} \in \mathbf{N}$ such that $F_{\overline{k}}$ has a nonempty interior.

Choose $\overline{x} \in X$, $r > 0$ in such a way that $B_r(\overline{x}) \subset F_{\overline{k}}$: we get

$$|T_h(\overline{x} + y)| \le \overline{k} \quad \forall y \in X, \|y\| < r, \; \forall h \in \mathbf{N}.$$

If $x \in X$, $\|x\| \le 1$, we get for every h:

$$|T_h(x)| = \frac{2}{r}|T_h\left(\frac{r}{2}x\right)| = \frac{2}{r}(|T_h\left(\overline{x} + \frac{r}{2}x - \overline{x}\right)| \le$$
$$\frac{2}{r}\left(|T_h\left(\overline{x} + \frac{r}{2}x\right)| + |T_h(\overline{x})|\right) \le \frac{4}{r}\overline{k}.$$

Passing to the sup over x we obtain our thesis. Q.E.D.

REMARK/EXERCISE: The following is an easy but important consequence of the Banach-Steinhaus Theorem: if $\{T_k\} \subset X'$ is a sequence of linear functionals such that $T_k(x) \to T(x) \in \mathbf{R}$ for every $x \in X$ (i.e. T_n converges to some real function T), then $T \in X'$ and

$$\|T\| \le \liminf_{k \to +\infty} \|T_k\|.$$

Prove this result and show with an example that in general we do not have the convergence of T_k to T in the norm of X'. *(This statement is a fairly immediate consequence of the theorem: linearity of T is trivial, and by Banach-Steinhaus the functionals T_k are equibounded in norm...hence their pointwise limit is bounded. The inequality on the norm is an easy consequence. Finally, to construct the required counterexample consider the space ℓ^2 and the sequence e^k of the "dual basis" elements of $(\ell^2)'$ (i.e. $e_k(\{x_n\}) = x_k$).*

We introduce a couple of fundamental concepts.

DEFINITION (bidual, reflexive Banach space): Let $(X, \|\cdot\|)$ be a normed space. Let X'' be its *bidual*, i.e. the vector space of all continuous linear functionals on X'.

As we well know from our first-year linear algebra course, there is a canonical way to associate to each element $x \in X$ an element of the bidual space. Precisely, we associate to x the functional $S_x : X' \to \mathbf{R}$ defined by

$$S_x(T) = T(x) \quad \forall T \in X'.$$

The map $J : X \to X''$ sending x into S_x is a *linear isometry*: indeed, by the Hahn-Banach theorem we can write

$$\|x\| = \max\{T(x) : \ T \in X',\ \|T\|_{X'} \leq 1\} \quad \forall x \in X,$$

whence $\|x\| = \|S_x\|_{X''}$.

In finite dimension, the map J is a (canonical) isomorphism between X and X''. This is no longer true if the dimension is infinite: in some cases $J(X)$ is a proper subset of X''. A normed space is called *reflexive* precisely when $J(X) = X''$ (i.e. if the space "coincides with its bidual").

Thanks to our detailed study of the spaces ℓ^p, it is easy to see that ℓ^1 and ℓ^∞ are *not* reflexive, while the spaces ℓ^p are for every $1 < p < +\infty$.

Reflexivity, like *separability*[12] plays a crucial role in the theory of weak topologies.

The injection of X in its bidual allows us to obtain another important consequence of the Banach-Steinhaus Theorem:

REMARK/EXERCISE: Let $(X, \|\cdot\|)$ be a Banach space, A a subset. Then A is bounded if and only if for every $T \in X'$, the image $T(A)$ is a bounded subset of $\mathbf{R}$. (*HINT: Obviously, if A is bounded, then $T(A)$ is bounded for every $T \in X'$. To prove the converse, it suffices to show that every sequence $\{x_n\} \subset A$ such that $\{T(x_n)\}$ is bounded for every $T \in X'$, is actually bounded in norm. To this end, apply the Banach-Steinhaus Theorem to the sequence $S_{x_n} \in X''$: it is bounded pointwise by our hypothesis, hence it is bounded in norm. We can then conclude because $\|x_n\| = \|S_{x_n}\|_{X''}$.*)

REMARK: We will now show a very concrete consequence of an abstract result like the Banach-Steinhaus theorem: we prove that *there are continuous and 2π-periodic functions, whose Fourier series does not converge at some point.*.

If $f : \mathbf{R} \to \mathbf{R}$ is a continuous, 2π-periodic function, we recall that its Fourier series is

$$a_0/2 + \sum_{n=1}^{\infty} [a_n \cos nx + b_n \sin nx],$$

[12] *A space is separable if it has a countable dense subset.*

where

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos nt\, dt, \quad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin nt\, dt.$$

There is a well-known convergence theorem, ensuring that if f is C^1, then its Fourier series converges at every point to $f(x)$ (and convergence is also uniform). However, the regularity required on f may seem "excessive": to compute Fourier coefficients, we certainly don't need differentiability... But we will show that if f is only C^0, we cannot be sure there is convergence at every point!

To prove this, we need the following well-known formula, expressing the N-th partial sum of the Fourier series:

$$f_N(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{\sin((N+\frac{1}{2})(y-x))}{2\sin((y-x)/2)} f(y)\, dy.$$

We show that there are continuous function for which $f_N(0)$ does not converge to $f(0)$: to this aim, consider the Banach space $X = C^0(2\pi)$ of continuous, 2π-periodic functions with the norm $\|\cdot\|_\infty$. If we define

$$T_N : f \mapsto \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{\sin((N+\frac{1}{2})y)}{2\sin(y/2)} f(y)\, dy,$$

the functionals T_N are well defined elements of $(L^\infty(2\pi))'$, and then also of X' (because X is a closed subspace of L^∞). Moreover, if we put

$$g_N(y) = \frac{\sin((N+\frac{1}{2})y)}{2\sin(y/2)},$$

one easily checks that $\|T_N\|_{X'} = \|g_N\|_{L^1([-\pi,\pi])}$.[13]

Now, from our construction of the functionals we have $f_N(0) = T_N(f)$. If we had $f_N(0) \to f(0)$ for every $f \in X$, in particular we would have

$$\sup_N |T_N(f)| < +\infty \quad \forall f \in X.$$

By the Banach-Steinhaus theorem, we could conclude that the norms of the functionals T_N are equibounded: but this is false, because $\|g_N\|_{L^1} \to +\infty$ as $N \to +\infty$.[14]

[13] Obviously, T has this norm as an elemento of L^∞, whence $\|T_n\|_{X'} \leq \|g_N\|_{L^1}$. On the other hand, it is easy to construct a sequence σ_k of periodic continuous functions such that $\lim_{k\to+\infty} \sigma_k(x) \to \mathrm{sgn}(g_N(x))$ for a.e.. $x \in [-\pi,\pi]$ anch such that $\|\sigma_k\|_\infty \leq 1$. By the dominated convergence theorem we have $T_N(\sigma_k) \to \|g_N\|_{L^1}$, whence the thesis.

[14] With a change of variables, and recalling that $|\sin t| \leq |t|$, this comes from the non-summability of the function $\frac{\sin t}{t}$ on the half line $[1,+\infty)$.

A very important result is the following

THEOREM (of the open mapping): Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ be Banach spaces, $T : X \to Y$ a continuous and surjective linear map. Then T is open: there exists $r > 0$ such that $T(B_1(0)) \supset B_r(0)$.

PROOF: To begin with, we show that there exists $r > 0$ such that

$$(*) \qquad \overline{T(B_1(0))} \supset B_{2r}(0).$$

Indeed, by the surjectivity of T we have $Y = \bigcup_{n=1}^{\infty} \overline{T(B_n(0))}$. By Baire lemma we know that at least one of these closed sets has nonempty interior. But by the homogeneity of the norm and the linearity of T, all these sets are homotetic, so $\overline{T(B_1(0))}$ has nonempty interior.

This interior is a symmetric, convex open sets (the closure of a symmetric open set is symmetric and convex, the same holds for the interior...): we deduce that 0 belongs to the interior of $\overline{T(B_1(0))}$ and $(*)$ is proved.

We will finish the proof next time, by showing that we actually have

$$T(B_1(0)) \supset B_r(0).$$

11 Lecture of october 27, 2015 (2 hours)

To conclude the proof of the open mapping theorem, we have to show that

$$T(B_1(0)) \supset B_r(0),$$

which is our thesis.

Indeed, let $\|y\| < r$: we look for a point $x \in X$ such that $\|x\| < 1$ and $T(x) = y$.

Since by $(*)$ we have $B_r(0) \subset \overline{T(B_{1/2}(0))}$, for every $\varepsilon > 0$ we can find $z \in X$ such that $\|z\| < 1/2$ and $\|y - T(z)\| < \varepsilon$. Choosing $\epsilon = r/2$ we find $z_1 \in X$ such that $\|z_1\| < 1/2$ and $\|y - T(z_1)\| < r/2$.

As $B_{r/2} \subset \overline{T(B_{1/4}(0))}$, by repeating the same argument with $y - T(z_1)$ at the place of y and $\varepsilon = r/4$, we find $z_2 \in X$ such that $\|z_2\| < 1/4$ e $\|y - T(z_1) - T(z_2)\| < r/4$.

Proceeding in the same way, we construct a sequence $\{z_n\} \subset X$ such that $\|z_n\| < 1/2^n$ and $\|y - T(z_1 + z_2 + \ldots + z_n)\| < r/2^n$. The sequence $x_n = z_1 + z_2 + \ldots + z_n$ is clearly a Cauchy sequence, whence $x_n \to \overline{x}$ in X. We obviously have $\|\overline{x}\| < 1$ and $y = T(\overline{x})$ thanks to the continuity of T. Q.E.D.

REMARKS/COROLLARIES: An important consequence of the theorem is the following: if $T : X \to Y$ is an *algebraic isomorphism* between Banach spaces and T *is continuous*, then the inverse map $T^{-1} : Y \to X$ is continuous and T is a Banach spaces isomorphims. Indeed, the fact that T is open implies the boundedness of T^{-1}.

Another important consequence: if $\|\cdot\|$ and $\|\cdot\|'$ are two Banach norms on X, and there exists $C > 0$ such that $\|x\| \leq C\|x\|'$ for every $x \in X$, then the two norms are equivalent. It is enough to apply the previous remark to the identity map between the two Banach spaces.

Another important consequence of the open mapping theorem is the

THEOREM (Of the closed graph): Let X, Y be Banach spaces, $T : X \to Y$ a linear map. Then T is continuous if and only if its graph $G_T = \{(x, T(x)) : x \in X\}$ is closed in $X \times Y$.

PROOF: Suppose T is continuous and $(x_n, T(x_n)) \to (\overline{x}, \overline{y})$. By the continuity of T we get $\overline{y} = T(\overline{x})$ whence $(\overline{x}, \overline{y}) \in G_T$: the graph is closed.

Suppose conversely that G_T is closed. Then G_T is a closed linear subspace of the Banach space $X \times Y$, and so is itself a Banach space.

The map $\Phi : x \mapsto (x, T(x))$ is clearly an algebraic isomorphism between X and G_T, with inverse map p_1 (the projection on the first factor of $X \times Y$). As p_1 is continuous, by the open mapping theorem we get that Φ is also continuous. But then $T = p_2 \circ \Phi$ is continuous (where p_2 is projection on the second factor). Q.E.D.

To conclude this first discussion of Banach spaces, let us examine compactness. In finite dimension, we have an abundance of compact sets (all closed bounded sets are compact!), and this has been useful to prove many theorems. In infinite dimension, however, not all bounded and closed sets are compact:

THEOREM: Let $(X, \|\cdot\|)$ be a normed space. Then the dimension of X is finite if and only if the unit closed ball

$$\overline{B} = \{x \in X : \ \|x\| \leq 1\}$$

is compact.

To prove the theorem, we need the following result:

LEMMA (Riesz): Let $(X, \|\cdot\|)$ be a normed space, Y a proper closed subspace. Then there exists $\overline{x} \in X$ such that $\|\overline{x}\| = 1$ and $\operatorname{dist}(\overline{x}, Y) \geq 1/2$.

PROOF: Choose $x_0 \in X \setminus Y$. As Y is closed, the distance δ between x_0 and Y is positive. Moreover, by definition of distance there exists $y_0 \in Y$ such that $\|x_0 - y_0\| < 2\delta$. Put

$$\overline{x} = \frac{x_0 - y_0}{\|x_0 - y_0\|}.$$

Clearly $\|\overline{x}\| = 1$, and if $y \in Y$ we have:

$$\|\overline{x} - y\| = \frac{1}{\|x_0 - y_0\|}\|x_0 - (y_0 + y\|x_0 - y_0\|)\| > 1/2,$$

where we used the fact that $y_0 + y\|x_0 - y_0\| \in Y$. Q.E.D.

PROOF of the theorem on the (non) compactness of the unit closed ball: If the dimension of X is finite, then the closed unit ball is compact: we can assume w.l.o.g. that we are in the case $X = \mathbf{R}^n$, where all norms are equivalent to the euclidean norm. It follows that our ball is an euclidean bounded, closed set, hence compact.

Conversely, suppose the dimension of X is infinite. We construct an increasing sequence of subspaces $Y_1 \subset Y_2 \subset Y_3 \dots$ in such a way that $\dim(Y_k) = k$, $k = 1, 2, \dots$

Now fix $x_1 \in Y_1$, $\|x_1\| = 1$. By Riesz lemma with $X = Y_2$ and $Y = Y_1$, we can find $x_2 \in Y_2$ such that $\|x_2\| = 1$ and $\mathrm{dist}(x_2, Y_1) > 1/2$.

Proceeding in the same way we find a sequence $\{x_k\}$ such that $\|x_k\| = 1$, $x_k \in Y_k$ and $\mathrm{dist}(x_k, Y_{k-1}) > 1/2$. This sequence contains only norm-one vectors, and the distance between any two of its elements is larger than $1/2$. Such a sequence has obviously no Cauchy subsequences: the unit closed ball is not compact. Q.E.D.

In a metric space (and thus in a Banach space), compact subsets are characterized as follows. We need a definition:

DEFINITION (Total boundedness): Let (X, d) be a metric space. A subset $K \subset X$ is *totally bounded* if, for every $\varepsilon > 0$, it is possible to cover K with a *finite number* of balls with radius ε.

THEOREM: Let (X, d) be a metric space, $K \subset X$. Then K is compact if and only if it is complete and totally bounded.

Moreover, a totally bounded subset K of a complete metric space is relatively compact: from any sequence with values in K, we can extract a subsequence converging in X.

This characterization gives a good "geometric" idea of how compact sets look like in an infinite dimensional space:

REMARK: Given a compact subset K of a normed space and any $\varepsilon > 0$, there exists a finite dimensional subspace Y_ε whose distance from every point of the set K is less than ε. Indeed, by the total boundedness we can cover K by a finite number of balls $B_\varepsilon(x_1)$,...,$B_\varepsilon(x_N)$. We can then define $Y_\varepsilon = \text{span}\{x_1, \ldots, x_N\}$.

We can summarize this by saying that in infinite dimension compact sets are rather "skinny"...and so balls are not compact: in particular, it is false that from a norm-bounded sequence we can extract a convergent subsequence!

PROOF of the characterization of compact sets: We know that in a metric space, sequential and topological compactness are the same.

Suppose now K is compact: we show that K is complete and totally bounded. Let $\{x_k\} \subset K$ be a Cauchy sequence. By compactness, it has a subsequence converging to some $\overline{x} \in K$. But it is easy to check that if a Cauchy sequence has a converging subsequence, the whole sequence converges to the same limit!

Choose then $\varepsilon > 0$ and consider the family of open balls $\{B_\varepsilon(x)\}_{x \in K}$. This is an open covering of K: by compactness, we can extract a finite subcovering...which gives us a finite number of balls of radius ε covering K.

We will show the opposite implication tomorrow.

12 Lecture of october 28, 2015 (2 hours)

To conclude the proof of the characterization of compact sets in a metric space, suppose K is complete and totally bounded. Let $\{x_k\} \subset K$ be a sequence in K: we show that it is possible to extract a subsequence which converges to some point of K.

By the total boundedness, we can cover K with a finite number of open balls of radius 1. Of necessity, infinitely many terms of the sequence will fall within one of these balls, which we will call B_1. Let $\{x_k^{(1)}\}$ be the subsequence of $\{x_k\}$ formed by those elements which belong to B_1.

Cover now K with a finite number of balls of radius 1/2: within one of those, which we call B_2, we will have infinitely many terms of $\{x_k^{(1)}\}$. Let $\{x_k^{(2)}\}$ be the subsequence of those elements of $\{x_k^{(1)}\}$ which belong to B_2. We proceed in the same way, covering K with balls having radius 1/3, 1/4...

By recurrence, we construct a sequence of subsequences such that $\{x_k^{(n)}\}$ is a subsequence of $\{x_k^{(n-1)}\}$ and all its elements are contained within a ball of radius $1/n$.

We then take the *diagonal subsequence*, defined by $\tilde{x}_k = x_k^{(k)}$ (the k-th element of the diagonal subsequence is the k-th element of the k-th subsequence.

The sequence $\{\tilde{x}_k\}$ is a subsequence of $\{x_k^{(n)}\}$ for $k \geq n$: in particular, it is a subsequence of $\{x_k\}$, and is obviously a Cauchy sequence (because, from the n-th term on, it is contained within a ball of radius $1/n$, and this is true for every fixed n). Thus $\tilde{x}_k \to \overline{x} \in K$ by the completeness assumption.

The same proof works also when X is complete, while K is only totally bounded: in that case, we only say that $\overline{x} \in X$. Q.E.D.

A possible cure to the lack of compactness of the closed ball is obtained through the concept of *weak convergence*: if we are in a "good enough" Banach space, from every bounded sequence we can extract a weakly convergent subsequence. We begin by giving the relevant definition:

DEFINITION: Let $(X, \|\cdot\|)$ be a normed space. We say that a sequence $\{x_k\} \subset X$ *weakly converges* to $\overline{x} \in X$, and we write $x_k \rightharpoonup \overline{x}$, if and only if

$$T(x_k) \to T(\overline{x}) \quad \forall T \in X'.$$

EXAMPLES/REMARKS/EXERCISES: Observe first that a norm-convergent sequence converges also weakly (by the continuity of each $T \in X'$): *strong convergence* implies *weak convergence*

In finite dimension, strong and weak convergence coincide: indeed, a sequence in $\mathbf{R}^n$ converges if and only if its components converge.

This is no longer true in infinite dimension: we will see in a moment that in a reflexive Banach space there are always weakly convergent sequences which do *not* converge in norm.

Explicit examples are easily obtained in the spaces ℓ^p and $L^p(\Omega)$, because we know wery well the dual spaces!

If $1 \leq p < +\infty$, consider a sequence $x^n = \{x_k^n\}_k \in \ell^p$. By definition, $x^n \rightharpoonup \overline{x} = \{\overline{x}_k\}_k$ in ℓ^p iff for every $T \in (\ell^p)'$ we have $T(x^n) \to T(\overline{x})$, i.e. iff

$$\lim_{n \to +\infty} \sum_{k=1}^{\infty} x_k^n y_k = \sum_{k=1}^{\infty} \overline{x}_k y_k \quad \forall \{y_k\} \in \ell^q.$$

For instance, the sequence e^n of the "basis vectors" in ℓ^p converges weakly to 0 for $1 < p < +\infty$ (not in ℓ^1), but it does not converge strongly.

Weak convergence in $L^p(\Omega)$ is similarly characterized: given $\{u_k\} \subset L^p(\Omega)$ (with $1 \leq p < +\infty$), we have $u_k \rightharpoonup u$ in L^p if and only if

$$\lim_{n \to +\infty} \int_\Omega u_k(x) v(x)\, dx = \int_\Omega u(x) v(x)\, dx \quad \forall v \in L^q(\Omega).$$

Weak convergence is mainly useful because of the following

THEOREM (Banach-Alaoglu): If X is a reflexive Banach space, its closed unit ball is (sequentially) weakly compact: from every norm-bounded sequence it is possible to extract a weakly convergence subsequence.

We will not prove this theorem in its full generality, but later we will give a proof in the particular case where X is a Hilbert space.

REMARK: If $\{x_n\} \subset X$ (with X a Banach space) is a sequence such that $x_n \rightharpoonup x$, then $\sup\{\|x_n\| : n \in \mathbf{N}\} < +\infty$: every weakly convergent sequence is norm-bounded. Indeed, from the definition of weak convergence we infer that $\{T(x_n)\}$ is a bounded subset of $\mathbf{R}$ for every $T \in X'$: we already observed that this implies the boundedness of $\{x_n\}$ (as a consequence of the Banach-Steinhaus theorem).

REMARK: In an infinite dimensional, reflexive Banach space we always have sequences which converge weakly but not strongly. Indeed, the closed unit ball is not strongly compact: we constructed a sequence of vectors of norm 1, which does not have strongly convergent subsequences. By Banach-Alaoglu, the same sequence has a weakly convergent subsequence!

REMARK (we saw no details in class...): In the space ℓ^1 (which is not reflexive!) every weakly convergent sequence converges strongly. This is really a patological example, and the proof is not so easy!

It is easy to see that it suffices to prove that a sequence $\{x^n\} \subset \ell^1$ with $x^n \rightharpoonup 0$ converges also strongly. By one of the previous remarks, we have $\|x^n\|_{\ell^1} \leq C$ for every n.

We have to show that $\|x^n\|_{\ell^1} \to 0$. Suppose by contradiction this is false: up to subsequences, this implies that $\|x^n\|_{\ell^1} \geq c > 0$ for every n. We show that we can extract a further subsequence which does not converge weakly to 0, thut contradicting our hypothesis.

Now, if $x^n = \{x^n_k\}_k$, then

$$\lim_{n\to+\infty} x^n_{\overline{k}} = 0 \quad \forall \overline{k} \in \mathbf{N}:$$

weak convergence to 0 implies convergence to 0 of all components (apply the definition of weak convergence with $\{y_k\} = e^{\overline{k}}$). So, for every fixed $N \in \mathbf{N}$ we have

$$\lim_{n\to+\infty} \sum_{k=1}^{N} |y^n_k| = 0.$$

From this we see that we can choose a strictly increasing sequence of natural numbers $k_1 < k_2 < k_3 < \ldots$ and a subsequence of x^n (which we still denote

x^n) in such a way that

$$\sum_{k=k_n+1}^{k_{n+1}} |x_k^n| \geq \frac{3}{4}\|x^n\|_{\ell^1},\ n = 1, 2, 3, \ldots$$

Define now a sequence $\{y_k\} \in \ell^\infty$ as follows: if $k_n + 1 \leq k \leq k_{n+1}$, then $y_k = \mathrm{sgn}(x_k^n)$. Then, for every fixed n, we have

$$T_{\{y_k\}}(x^n) = \sum_{k=k_n+1}^{k_{n+1}} |x_k^n| + \sum_{other\ k's} x_k^n y_k \geq \frac{3}{4}\|x^n\|_{\ell^1} - \frac{1}{4}\|x^n\|_{\ell^1} \geq \frac{1}{2}c.$$

Obviously this sequence does not converge to zero as $n \to +\infty$, thus contradicting the fact that $x^n \rightharpoonup 0$.

EXERCISE: If X is a Banach space, $C \subset X$ is a nonempty, convex and strongly closed set (i.e., it is closed in the topology induced by the norm), then C is sequentially weakly closed: if $\overline{x} \in X$ is the weak limit of a sequence in C, then $\overline{x} \in C$.

Let indeed $x_n \rightharpoonup \overline{x}$, $\{x_n\} \subset C$. Suppose by contradiction $\overline{x} \notin C$. We can apply the geometric form of the Hahn-Banach theorem to the convex sets $\{\overline{x}\}$ and C, the first of which is compact and the second closed. We find $T \in X'$, $T \neq 0$ and $\varepsilon > 0$ such that $T(x) + \varepsilon < T(\overline{x})$ for every $x \in C$, and in particular $T(x_n) + \varepsilon < T(\overline{x})$ for every n. This is impossible because $T(x_n) \to T(\overline{x})$ by definition of weak convergence!

EXERCISE: Let X be a Banach space, $F : X \to \mathbf{R}$ be a convex continuous function. Then F is (sequentially) weakly lower semicontinuous: for every sequence $x_n \rightharpoonup \overline{x}$ we have $F(\overline{x}) \leq \liminf\limits_{n \to +\infty} F(x_n)$.

In particular, as the norm is a convex function, we have

$$\|\overline{x}\| \leq \liminf_{n \to +\infty} \|x_n\| \quad \forall x_n \rightharpoonup \overline{x}.$$

Let $\ell = \liminf_{n \to +\infty} F(x_n)$. If $\ell = +\infty$, there is nothing to prove. Let then $\ell < +\infty$: choose $s \in \mathbf{R}$, $s > \ell$ and consider the sublevel set

$$C_s = \{x \in X : \ F(x) \leq s\}.$$

This is a closed convex set by the convexity and strong continuity of F. By our previous exercise C_s is sequentially weakly closed. Moreover, up to subsequences we may assume $F(x_n) \to \ell$: for n large enough we have

$\{x_n\} \subset C_s$. But then, by the weak closure, $\overline{x} \in C_s$ whence $F(\overline{x}) \leq s$. The thesis follows because $s > \ell$ is arbitrary.

EXERCISE: If X is a reflexive Banach space, C a nonempty convex closed subset, $x_0 \in X$, show that there exist a point of C whose distance from x_0 is minimal.

Up to translations, we may assume $x_0 = 0$: we show that C has a point of minimal norm.

Let indeed $\{y_n\} \subset C$ be a sequence such that $\|y_n\| \to \inf\{\|y\| : y \in C\} = \delta$. Obviously, this sequence is norm bounded: by Banach-Alaoglu we have a subsequence which converge weakly to some point $\overline{y}$. By weak closure, $\overline{y} \in C$, and by the lower semicontinuity of the norm we conclude that $\overline{y}$ is the minimum point: $\|\overline{y}\| \leq \liminf \|y_n\| = \delta$.

EXERCISE (not seen in class): If the Banach space X is not reflexive, the result of our previous exercise may fail. Consider indeed the space $C^0([0,1])$ with the sup norm and the set

$$C = \{u \in C^0([0,1]) : \int_0^{1/2} u(x)\,dx - \int_{1/2}^1 u(x)\,dx = 1\}.$$

This is a closed convex set, and the infimum of the norms of its elements is 1. On the other hand, C ha no element of norm 1: the minimum is not attained!

Closure and convexity of C is obvious: observe that $\Phi : u \mapsto \int_0^{1/2} u(x)\,dx - \int_{1/2}^1 u(x)\,dx$ is a bounded linear functional, so C is a closed hyperplane.

We first check that $\operatorname{dist}(0, C) \geq 1$: indeed, no function with norm strictly less than 1 belongs to C (our difference of integrals is less or equal than

$$1/2\left(\operatorname{essup}\{u(x) : x \in [0,1/2]\} - \operatorname{essinf}\{u(x) : x \in [1/2,1]\}\right),$$

which is below $\|u\|_\infty$).

This same inequality tells us that if $\|u\|_\infty = 1$, to belong the C the function u should be 1 a.e. on $[0,1/2]$, -1 a.e. in $[1/2,1]$: no continuous function has this property.

But the distance is exactly 1, as we see by considering the sequence of continuous functions $u_n(x)$ which is $1 + 1/n$ on $[0, 1/2 - 1/(n+1)]$, $-1 - 1/n$ on $[1/2 + 1/(n+1), 1]$ and is linear on $[1/2 - 1/(n+1), 1/2 + 1/(n+1)]$: those functions belong to C, and $\|u_n\|_\infty = 1 + 1/n \to 1$.

13 Lecture of october 29, 2015 (2 hours)

In some cases, weak compactness is not enough and we need strong compactness theorems. The study of strong compactness in C^0 is of course very important: indeed, this is probably the most natural space we can think of!

We have the following result:

THEOREM (Ascoli-Arzelà): Let $u_n : A \to B$ be a sequence of continuous functions, where A and B are compact metric spaces. If the sequence u_n is equicontinuous, i.e. if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $x, y \in A$, $d_A(x, y) < \delta$ imply $d_B(u_n(x), u_n(y)) < \varepsilon$ for every n, then u_n has a subsequence which converge uniformly to some continuous function $u : A \to B$.

PROOF: Remark first that the set $C^0(A; B)$ of continuous functions from A to B is a complete metric space with the uniform distance $d(u, v) = \sup\{d_b(u(x), v(x)) : \ x \in A\}$.

To prove the theorem it suffices to show that $\mathcal{F} = \{u_n : \ n \in \mathbf{N}\}$ is a totally bounded subset of $C^0(A, B)$. Fix $\varepsilon > 0$, and use the total boundedness of B to write $B = B_1 \cup B_2 \cup \ldots \cup B_N$, where B_j are balls of radius ε. Use then equicontinuity to find δ such that $d_A(x, y) < \delta$ implies $d_B(u_n(x), u_n(y)) < \varepsilon$, and then the total boundedness of A to write$A = A_1 \cup \ldots \cup A_M$, where A_i are balls of radius δ and center a_i.

For each multiindex $(j_1, j_2, \ldots j_M) \in \{1, 2, \ldots, N\}^M$ (there is a finite number of those) consider the set of function

$$\mathcal{W}_{(j_1, j_2, \ldots j_M)} = \{u \in \mathcal{F} : \ u(a_i) \in B_{j_i}, \ i = 1, \ldots, M\}.$$

Each element of the original sequence belongs to one of these sets. Moreover, each set of function is either empty, or is diameter is less than 4ε, and is thus contained in a ball of radius 4ε: indeed, if $u, v \in \mathcal{W}_{(j_1, j_2, \ldots j_M)}$ and $x \in A$, choose i such that $x \in A_i$. By the equicontinuity we get $d_B(u(x), v(x)) \leq d_B(u(x), u(a_i)) + d_B(u(a_i), v(a_i)) + d_B(v(a_i), v(x)) < 4\varepsilon$.

We thus covered $\mathcal{F}$ with a finite number of balls of radius 4ε. Q.E.D.

REMARK: In the most common case of real valued functions, the Ascoli-Arzelà theorem is usually stated as follows: each sequence of functions in $C^0(A; \mathbf{R})$ (with A a compact metric space) which is *equicontinuous* and *equibounded*, has a subsequence which converges uniformly to some continuous function.

Indeed, equiboundedness ensures that the functions in the sequence take values in the compact interval $[-M, M]$ for M large enough.

The Ascoli-Arzelà theorem is used to prove a lot of important theorems in analysis: for instance, one can use it to prove the Peano theorem, ensuring the existence of local solutions of the Cauchy problem for non-linear O.D.E.s

We will now begin to discuss the theory of Hilbert spaces. Before we give the definition, we recall the definition of scalar product and of the norm induced by a scalar product.

DEFINITION: Let X be a real vector space. A *scalar product* over X is a map

$$\begin{array}{rcl} <\cdot,\cdot>: X \times X & \to & \mathbf{R} \\ (x, y) & \mapsto & <x, y> \end{array}$$

which is bilinear (i.e. linear in each of its arguments x and y), symmetric (i.e. $<x, y>=<y, x>$ for every x, y) and positively defined (i.e. $<x, x>\geq 0$, with equality iff $x = 0$).

From a scalar product we get a norm on X as follows:

$$\|x\| :=<x, x>^{1/2} .$$

Of course, we have to verify that this is a norm. This, and other simple facts, are summarized in the following proposition:

PROPOSITION: Let $<\cdot,\cdot>$ be a scalar product on X, $\|\cdot\|$ the induced norm. Then the following hold

(i) For every $x, y \in X$ we have the Cauchy-Schwarz inequality

$$<x, y>\leq \|x\| \, \|y\|;$$

(ii) The induced norm... is a norm;

(iii) The parallelogram law holds:

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) \quad \forall x, y \in X;$$

(iv) The polarization law holds:

$$<x, y>= \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2) \quad \forall x, y \in X.$$

PROOF: If $x, y \in X$ and $t \in \mathbf{R}$ we have:

$$0 \leq \|ty + x\|^2 =< ty + x, ty + x >= t^2\|y\|^2 + 2t < x, y > +\|x\|^2.$$

The discriminant of this quadratic polynomial is thus less or equal than 0: this is exactly (i).

We show (ii): the norm is obviously homogeneous and non degenerate. We have to prove the triangle inequality. For every x, y we have, by the Cauchy-Schwarz inequality:

$$\|x + y\|^2 =< x + y, x + y >= \|x\|^2 + 2 < x, y > +\|y\|^2 \leq$$
$$\|x\|^2 + 2\|x\| \ \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$

(iii) and (iv) are easily proved by expanding the scalar products. (iii) is called the *parallelogram law* because, if we interpret the vectors x and y as the edges of a parallelogram, then $x + y$ and $x - y$ represent the diagonals. The identity is then the expression of a well known result in euclidean geometry. Q.E.D.

The parallelogram identity allows to characterize which norms are induced by a scalar product:

PROPOSITION: Let $(X, \|\cdot\|)$ be a normed space. Then the map

$$a(x, y) := \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2), \quad x, y \in X$$

is a scalar product which induces the given norm if and only if the norm satisfies the parallelogram law.

PROOF: If the norm is induced by a scalar product, we already know the parallelogram law is satisfied, and the scalar product is recovered thanks to the polarization identity.

Conversely, suppose the norm satisfies the parallelogram law and define $a(x, y)$ as in the statement. This function is symmetric and $a(x, x) = \|x\|^2 \geq 0$ with equality iff $x = 0$. Moreover, $a(x, 0) = a(0, y) = 0$ and $a(-x, y) = -a(x, y)$. The function $a(x, y)$ is also continuous.

Let now $x_1, x_2, y \in X$: from the parallelogram law we get

$$(*) \qquad a(x_1, y) + a(x_2, y) = \frac{1}{4}(\|x_1 + y\|^2 - \|x_1 - y\|^2 + \|x_2 + y\|^2 - \|x_2 - y\|^2) =$$
$$\frac{1}{8}(\|x_1 + x_2 + 2y\|^2 + \|x_1 - x_2\|^2 - \|x_1 + x_2 - 2y\|^2 - \|x_1 - x_2\|^2) =$$
$$\frac{1}{2}a(x_1 + x_2, 2y).$$

In particular, by letting $x_1 = x$, $x_2 = 0$ the last identity becomes

$$(**)\ a(x,y) = \frac{1}{2}a(x,2y) \quad \forall x,y.$$

By replacing (**) within (*) we get

$$a(x_1,y) + a(x_2,y) = a(x_1+x_2,y) \quad \forall x_1,x_2,y.$$

By applying this formula repeatedly we easily see that $a(mx,y) = m\,a(x,y)$ for every $m \in \mathbf{Z}$. Then, by using again (**) and the symmetry:

$$a(\frac{m}{2^n}x,y) = \frac{m}{2^n}a(x,y) \quad \forall x,y \in X,\ \forall m \in \mathbf{Z}, n \in \mathbf{N}.$$

Now, the set of numbers of the form $m/2^n$ is dense in $\mathbf{R}$: by the continuity of a we conclude that $a(tx,y) = ta(x,y)$ for every $x,y \in X$ and every $t \in \mathbf{R}$: $a(x,y)$ is thus a scalar product. Q.E.D.

DEFINITION (Hilbert space): A real vector space X, equipped with a scalar product $< \cdot,\cdot >$ is a *Hilbert space* if it is a Banach space with the norm induced by the scalar product.

EXAMPLES: Typical prototypes of Hilbert spaces are the spaces ℓ^2 with the scalar product $< \{x_k\},\{y_k\} >:= \sum_{k=1}^{\infty} x_k y_k$ and $L^2(\Omega)$ with the scalar prduct $< u,v >:= \int_\Omega u(x)v(x)\,dx$.

The following theorem is a stronger version of something we already know is valid in a reflexive Banach space. But the proof will be independent from the Banach Alaoglu theore, which we did not prove!

THEOREM (projection on a closed convex set): Let X be a Hilbert space, C a nonempty, closed, convex subset of X, $x_0 \in X$. Then there exists a unique $\overline{y} \in C$ such that $\|x_0 - \overline{y}\| = \mathrm{dist}(x_0, C)$.

PROOF: After a translation, we may suppose that $x_0 = 0$: we must now prove that in C there is a unique element of minimal norm. Let now $\delta = \inf\{\|y\| :\ y \in C\}$, and let $\{y_n\} \subset C$ be a sequence such that $\|y_n\| \to \delta$ (such a sequence exists by the definition of infimum!).

We prove that $\{y_n\}$ is a Cauchy sequence in X: to this aim, consider the parallelogram law with $x/2$, $y/2$ at the place of x,y... We easily get the identity

$$\|x-y\|^2 = 2(\|x\|^2 + \|y\|^2) - 4\left\|\frac{x+y}{2}\right\|^2,$$

which holds for every $x, y \in X$. Notice also that, if x and y are in C, then by convexity $\frac{x+y}{2} \in C$: by the identity just obtained and the definition of δ we get

$$(***) \qquad \|y_n - y_m\|^2 = 2(\|y_n\|^2 + \|y_m\|^2) - 4\left\|\frac{y_n + y_m}{2}\right\|^2 \leq 2(\|y_n\|^2 + \|y_m\|^2) - 4\delta^2.$$

The last quantity vanishes as $m, n \to +\infty$, so $\{y_n\}$ is a Cauchy sequence and it converges to some point $\overline{y} \in X$. Since C is closed, $\overline{y} \in C$. Moreover $\|\overline{y}\| = \delta$ by the continuity of the norm: $\overline{y}$ is our element of minimal norm in C.

Let us prove uniqueness: if we have also $\|\tilde{y}\| = \delta$ with $\tilde{y} \in C$, we can apply (***) with $y_n = \overline{y}$, $y_m = \tilde{y}$ and we get

$$\|\overline{y} - \tilde{y}\| \leq 0,$$

whence $\overline{y} = \tilde{y}$. Q.E.D.

14 Lecture of november 2, 2015 (2 hours)

COROLLARY: In the hypotheses of previous theorem, the point $\overline{y} \in C$ having minimal distance from x_0 is characterized by the inequality

$$(*) \; < x_0 - \overline{y}, y - \overline{y} > \leq 0 \qquad \forall y \in C.$$

In particular, if Y is a closed vector subspace of the Hilbert space X, $x_0 \in X$, then there exists a unique point $\overline{y} \in Y$ having minimal distance from x_0. This point is characterized by the orthogonality relation

$$< x_0 - \overline{y}, y >= 0 \quad \forall y \in Y.$$

PROOF: We show that $\overline{y} \in C$ is the point of minimum distance iff (*) holds. Suppose indeed (*) holds, and let $y \in C$. Then

$$\|x_0 - y\|^2 = \|x_0 - \overline{y} + \overline{y} - y\|^2 = \|x_0 - \overline{y}\|^2 + \|y - \overline{y}\|^2 - 2 < x_0 - \overline{y}, y - \overline{y} > \geq \|x_0 - \overline{y}\|^2$$

and $\overline{y}$ is the point of minimum distance.

Conversely, let $\overline{y} \in C$ be the point of minimum distance, $y \in C$. Then we have, for $t \in [0,1]$, $ty + (1-t)\overline{y} \in C$ so that

$$(**) \qquad \|x_0 - \overline{y}\|^2 \leq \|x_0 - (ty + (1-t)\overline{y})\|^2 = \|(x_0 - \overline{y}) - t(y - \overline{y})\|^2 = \|x_0 - \overline{y}\|^2 + t^2\|y - \overline{y}\|^2 - 2t < x_0 - \overline{y}, y - \overline{y} >,$$

whence $< x_0 - \overline{y}, y - \overline{y} > \leq \frac{t}{2}\|y - \overline{y}\|^2$ and $(*)$ follows by letting $t \to 0$.

If $C = Y$, with Y a closed vector subspace of X, inequality $(**)$ must hold *for every* $t \in \mathbf{R}$: this is possible iff the scalar product in the left hand side is identically zero (and when y ranges in Y, $y - \overline{y}$ exhausts all element of Y). Q.E.D.

The last corollary is very important: we deduce that every Hilbert space splits into the direct sum of any closed subspace and its orthogonal, with continuous projections.

PROPOSITION: Let $Y \subset X$ be a closed subspace of the Hilbert space X, $p : X \to X$ the map that takes any $x \in X$ to its closest point in the subspace Y. Then p is linear and continuous and its restriction to Y is the identity map. Moreover, $x - p(x)$ is orthogonal to Y, and so we can write $X = Y \oplus Y^\perp$, with continuous projections. Finally, $\|x\|^2 = \|p(x)\|^2 + \|x - p(x)\|^2$ for any $x \in X$.

PROOF: By the previous corollary, $p(x)$ is the unique point in Y such that $< x - p(x), y >= 0$ for every $y \in Y$, i.e. the unique point of Y such that $x - p(x) \in M^\perp$: for this reason, it is called the *orthogonal projection* of x on Y.

Now p coincides with the identity map on Y. We show it is linear: let $x_1, x_2 \in X$, $t \in \mathbf{R}$. Then we have $0 =< x_1 - p(x_1), y >=< x_2 - p(x_2), y >$ for every $y \in Y$, and so

$$< x_1 - tx_2 - (p(x_1) + tp(x_2), y >= 0 \quad \forall y \in Y,$$

whence $p(x_1 + tx_2) = p(x_1) + tp(x_2)$.

If $x \in X$, since $p(x) \in Y$ we get $< x - p(x), p(x) >= 0$, whence

$$\|x\|^2 =< p(x) + (x - p(x)), p(x) + (x - p(x) \geq \|p(x)\|^2 + \|x - p(x)\|^2.$$

So p is continuous, because the identity implies

$$\|p(x)\| \leq \|x\|,$$

i.e. the norm of p is less or equal than 1 (actually, it is exactly 1 because it coincides with the identity map on Y). Q.E.D.

We remark that there is an easy explicit formula for the orthogonal projection on a subspace of finite dimension:

REMARK: If Y is a finite dimensional subspace of X, and $\{e_1, \ldots, e_n\}$ is an orthonormal basis of Y, then we have

$$p(x) = \sum_{i=1}^{n} < x, e_i > e_i.$$

Moreover $\|p(x)\|^2 = \sum_{i=1}^{n}(< x, e_i >)^2$. Indeed, we only need to verify that $x - p(x)$ is orthogonal to every vector in Y: it is of course enough to check this on the basis vectors. Now

$$< x - p(x), e_j >=< x, e_j > - \sum_{i=1}^{n} < x, e_i >< e_i, e_j >= 0,$$

as we wanted. The expression for the norm of $p(x)$ follows immediately from the orthonormality of the basis vectors e_i.

Notice that this result does not depend on the completeness of X: in the projection theorem, completeness was needed to *prove* the existence of a point of minimum distance. Here, we explicitely exhibit this point!

We next characterize the dual of a Hilbert space: for every continuous linear functional $T \in X'$ there exists a unique $y \in X$ such that $T(x) =< y, x >$ for every $x \in X$. In particular, the dual of X is isometrically isomorphic to X:

THEOREM (Riesz representation theorem): Let X be a Hilbert space. Define the application

$$\begin{aligned} \Phi : X \quad &\rightarrow \quad X' \\ y \quad &\mapsto T_y \end{aligned}$$

where, by definition, $T_y(x) :=< y, x >$ for every $x \in X$. Then Φ is an isometric isomorphism between X and X'.

PROOF: From the Cauchy-Schwarz inequality we get

$$T_y(x) =< y, x >\leq \|y\| \, \|x\|,$$

and the linear functional T_y is continuous with norm $\leq \|y\|$. On the other hand, $T_y(\frac{y}{\|y\|}) = \|y\|$, whence $\|T_y\|_{X'} = \|y\|$.

So the linear map $\Phi : X \to X'$ is a well defined isometry.

To conclude, we have just to show that Φ is surjective: for every $T \in X'$ there is $y \in X$ such that $T = T_y$.

Let $Y = \ker(T)$. In case $Y = X$, we obviously have $y = 0$: we can thus suppose Y is a closed, proper subspace of X. Let then $x_0 \in X \setminus Y$, $\overline{y}$ the orthogonal projection of x_0 on Y. For every fixed $x \in X$ we have

$$x - \frac{T(x)}{T(x_0 - \overline{y})}(x_0 - \overline{y}) \in Y.$$

This vector must then be orthogonal to $x_0 - \overline{y}$:

$$< x_0 - \overline{y}, x - \frac{T(x)}{T(x_0 - \overline{y})}(x_0 - \overline{y}) >= 0,$$

whence with easy computations

$$T(x) =< x, T(x_0 - \overline{y})\frac{x_0 - \overline{y}}{\|x_0 - \overline{y}\|^2} >,$$

and our claim is proved with $y = T(x_0 - \overline{y})\frac{x_0 - \overline{y}}{\|x_0 - \overline{y}\|^2}$. Q.E.D.

Before we proceed, we need to define the sum of an arbitrary (not necessarily countable) family of nonnegative numbers:

DEFINITION: Let $\{t_\alpha\}_{\alpha \in I}$ be a family of nonnegative real numbers. We define

$$\sum_{\alpha \in I} t_\alpha = \sup\{\sum_{\alpha \in J} t_\alpha : \ J \subset I, \ J \ finite \ set\}.$$

The family $\{t_\alpha\}$ is said to be *summable* if the sum is finite.

An equivalent definition is the following: the sum is the integral of $\{t_\alpha\}$ with respect to the counting measure on I.

Remark that if the set I is countable and $\{\alpha_n\}_{n \in \mathbf{N}}$ is an enumeration, then

$$\sum_{\alpha \in I} t_\alpha = \sum_{n=1}^{\infty} t_{\alpha_n}$$

(and in particular the sum does not depend on the enumeration chosen).

REMARK: If $\{t_\alpha\}_{\alpha \in I}$ is summable, then the set $I' = \{\alpha \in I : \ t_\alpha > 0\}$ is at most countable.

Indeed, for every fixed $n = 1, 2, 3, \ldots$, the set $I_n = \{\alpha \in I : t_\alpha > 1/n\}$ is finite.

REMARK: If $\{c_\alpha\}_{\alpha\in I}$ is a family of real numbers such that $\sum_{\alpha\in I}|c_\alpha| < +\infty$, the sum $\sum_{\alpha\in I} c_\alpha$ is a well defined real number.

An easy way to define this sum is to take the *integral* of $\{c_\alpha\}$ w.r.t. the counting measure on I. Or, equivalently, we can enumerate the non-zero terms and compute the sum of the series.

DEFINITION: Let I be a set of indices. Denote by $\ell^2(I)$ the set of families of real numbers $\{c_\alpha\}_{\alpha\in I}$ such that the sum

$$\sum_{\alpha\in I} c_\alpha^2$$

is finite. This is a Hilbert space with the inner product

$$< \{a_\alpha\}, \{b_\alpha\} >:= \sum_\alpha a_\alpha b_\alpha,$$

where the sum in the r.h.s. is absolutely convergent thanks to the Hölder inequality in ℓ^2.

We now give the fundamental definition of the abstract Fourier series of an element x of a Hilbert space X, with respect to some fixed orthonormal family of vectors.

PROPOSITION (Bessel inequality): Let X be a Hilbert space, $\{e_\alpha\}_{\alpha\in I}$ an orthonormal family of elements of X (i.e. $\|e_\alpha\| = 1$ for every $\alpha \in I$ and $< e_\alpha, e_\beta >= 0$ whenever $\alpha, \beta \in I$, $\alpha \neq \beta$). If $x \in X$, we define its Fourier coefficients *w.r.t. $\{e_\alpha\}$ as the real numbers*

$$c_\alpha =< x, e_\alpha >, \quad \alpha \in I.$$

Then the following Bessel inequality holds

$$\sum_{\alpha\in I} |c_\alpha|^2 \leq \|x\|^2.$$

In particular, at most countably many Fourier coefficients are non zero.

PROOF: Let $J \subset I$ be any finite set of indices. Then the vector $\sum_{\alpha\in J} < x, e_\alpha > e_\alpha$ is the orthogonal projection of x on the space spanned by $\{e_\alpha\}_{\alpha\in J}$ and

$$\|x\|^2 = \|\sum_{\alpha\in J} < x, e_\alpha > e_\alpha\|^2 + \|x - \sum_{\alpha\in J} < x, e_\alpha > e_\alpha\|^2$$

whence (using orthonormality):

$$\|x\|^2 \geq \|\sum_{\alpha\in J} < x, e_\alpha > e_\alpha\|^2 = \sum_{\alpha\in J} \| < x, e_\alpha > e_\alpha\|^2 = \sum_{\alpha\in J} c_\alpha^2.$$

Taking the supremum over all finite subsets $J \subset I$ we get our thesis. Q.E.D.

15 Lecture of november 3, 2015 (2 hours)

Bessel inequality ensures that the Fourier coefficients $\{c_\alpha\}$ of $x \in X$, w.r.t. an orthonormal family $\{e_\alpha\}_{\alpha\in I}$, belong to the space $\ell^2(I)$. Conversely, every element of $\ell^2(I)$ coincides with the Fourier coefficients of some element of the space X:

THEOREM: Let X be a Hilbert space, $\{e_\alpha\}_{\alpha\in I}$ a fixed orthonormal family. For every $\{c_\alpha\}_{\alpha\in I} \in \ell^2(I)$ there exists an element $x \in X$ such that

$$< x, e_\alpha >= c_\alpha \quad \forall \alpha \in I.$$

In other words, the linear map

$$\begin{aligned} \Psi : X &\to \ell^2(I) \\ x &\mapsto \{< x, e_\alpha >\}_{\alpha\in I} \end{aligned}$$

is surjective.

PROOF: The coefficients c_α are non zero at most for a countable family of indices $I' \subset I$. Choose an enumeration of I':

$$I' = \{\alpha_k : \ k = 1, 2, 3 \dots\}.$$

Put then

$$x_n = \sum_{k=1}^{n} c_{\alpha_k} e_{\alpha_k}.$$

By the orthonormality of e_α we have

$$\|x_n - x_{n+h}\|^2 = \sum_{k=n}^{n+h} c_{\alpha_k}^2,$$

whence $\{x_n\}$ is a Cauchy sequencey (because the series $\sum\limits_{k=1}^{\infty} c_{\alpha_k}^2$ converges). Then $x_n \to x \in X$. By the continuity of the scalar product,

$$< x, e_\alpha >= \lim_{n\to+\infty} < x_n, e_\alpha >= c_\alpha$$

(consider the cases $\alpha \in I'$ and $\alpha \in I \setminus I'$ separately). Q.E.D.

REMARK: In the proof of the previous theorem, the point x was found as the sum of the series $\sum_{k=1}^{\infty} c_{\alpha_k} e_{\alpha_k}$. We would very much like to write, for any $x \in X$,

$$x = \sum_{\alpha \in I} c_\alpha e_\alpha.$$

This is true if the orthonormal set is *maximal*, a statement which follows immediately from the following theorem.

More generally, given any orthonormal system $\{e_\alpha\}_{\alpha \in I}$ in X, $x \in X$ and $c_\alpha =< x, e_\alpha >$, we will see that the *Fourier series*

$$\sum_{\alpha \in I} c_\alpha e_\alpha$$

is well defined and converges to an element of X: precisely, it converges to the orthogonal projection $p(x)$ of x on the *closure* of the space spanned by the vectors $\{e_\alpha\}$.

The next result ensures that the application Ψ defined in the last theorem, is an isometric isomorphism as soon as the orthonormal system $\{e_\alpha\}$ is *maximal*. In that case, given $x \in X$ we can always write

$$x = \sum_{\alpha \in I} c_\alpha e_\alpha,$$

where $c_\alpha =< x, e_\alpha >$ are the Fourier coefficients of x.

THEOREM (Abstract Fourier series): Let X be a Hilbert space, $\{e_\alpha\}_{\alpha \in I}$ be an orthonormal family in X. Then the following facts are equivalent:

(i) The family $\{e_\alpha\}_{\alpha \in I}$ is maximal: if we add any vector of X to the family, it is no longer orthonormal;

(ii) The space spanned by $\{e_\alpha\}_{\alpha \in I}$ is dense in X;

(iii) Parseval identity holds: for every $x \in X$, if we denote by $c_\alpha =< x, e_\alpha >$ its Fourier coefficients, then

$$\|x\|^2 = \sum_{\alpha \in I} |c_\alpha|^2.$$

In particular, the map Ψ defined in our previous theorem is an isometric isomorphism between X and $\ell^2(I)$. By the injectivity of Φ, the Fourier series

$$\sum_{\alpha \in I} c_\alpha e_\alpha$$

defined in the proof of the theorem converges to x, and so its sum does not depend on the enumeration chosen for the non zero Fourier coefficients.

PROOF: We show that $(i) \Rightarrow (ii)$: suppose by contradiction that $Y = \text{span}\{e_\alpha\}$ is not dense, and let $x_0 \in X \setminus \overline{Y}$. Then, if $p(x_0)$ is the orthogonal projection of x_0 on $\overline{Y}$, $x_0 - p(x_0)$ is a non zero vector which is ortogonal to all e_α, against the maximality hypothesis.

We then show $(ii) \Rightarrow (iii)$: given $\varepsilon > 0$, by (ii) for every $x \in X$ we can find a finite linear combination $\lambda_1 e_{\alpha_1} + \lambda_2 e_{\alpha_2} + \ldots + \lambda_N e_{\alpha_N}$ such that

$$\|x - \lambda_1 e_{\alpha_1} + \lambda_2 e_{\alpha_2} + \ldots + \lambda_N e_{\alpha_N}\|^2 < \varepsilon.$$

This implies

$$\|x - c_{\alpha_1} e_{\alpha_1} - c_{\alpha_2} e_{\alpha_2} - \ldots - c_{\alpha_N} e_{\alpha_N}\|^2 < \varepsilon$$

(by the minimality property of the orthogonal projection $p(x)$ on $Y = \text{span}\{e_{\alpha_1}, \ldots, e_{\alpha_N}\}$), whence $\varepsilon > \|x - p(x)\|^2 = \|x\|^2 - \|p(x)\|^2 = \|x\|^2 - \sum_{i=1}^N c_{\alpha_i}^2$ and thus

$$\|x\|^2 \leq \sum_{\alpha \in I} c_\alpha^2 + \varepsilon.$$

We already know that $\sum_{\alpha \in I} c_\alpha^2 \leq \|x\|^2$ (Bessel inequality), so Parseval identity is proved because ε is arbitrary.

Finally, we have to prove that $(iii) \Rightarrow (i)$: let $x_0 \in X$ be orthogonal to all the vectors e_α. By the Parseval identity we have $\|x\| = 0$, so $x = 0$ and the orthonormal family $\{e_\alpha\}$ is maximal. Q.E.D.

DEFINITION: A maximal orthonormal set in a Hilbert space is called a *Hilbert basis*. It is easy to check that a Hilbert basis always exists (Zorn lemma): in particular, every Hilbert space X is isomprphic and isometric to $\ell^2(I)$ for a suitably chosen set of indices I.

ESXERCISE: Let X be a Hilbert space, $\{e_\alpha\}_{\alpha \in I}$ a (not necessarily maximal) orthonormal family. Show that for every $x \in X$ the sum of the series

$$\sum_{\alpha \in I} < x, e_\alpha > e_\alpha$$

is well defined. *(HINT: Consider the subspace $Y = \overline{\text{span}\{e_\alpha\}_{\alpha \in I}}$. Show that the series converges to the orthogonal projection of x on $Y \ldots$)*

You probably wonder how the abstract theory we just discussed is related with the *Fourier series* in the traditional, trigonometric sense! Here is the answer:

REMARK: Consider the space

$$L^2(2\pi) = \{u : \mathbf{R} \to \mathbf{R} : \; u \; measurable \; 2\pi - periodic, \; \int_{-\pi}^{\pi} u^2(x)\, dx < +\infty\},$$

with the usual equivalence relation identifying functions which are a.e. equal. This is a Hilbert space with the scalar product

$$< u, v >= \int_{-\pi}^{\pi} u(x)v(x)\, dx.$$

Consider the following family of functions in $L^2(2\pi)$:

$$\mathcal{F} = \{\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \sin nx, \frac{1}{\sqrt{\pi}} \cos nx : \; n = 1, 2, \ldots\}.$$

One easily checks that this family is orthonormal, and that the abstract Fourier series of $u \in L^2(2\pi)$ w.r.t. this orthonormal set is precisely the usual Fourier series.

Moreover, we will prove that the family $\mathcal{F}$ is maximal. As a consequence, the classical Fourier series of a function in $L^2(2\pi$ converges in L^2 to u.

Notice that because the above family of functions is a Hilbert basis of $L^2(2\pi)$, then this space is *separable*. Recall that a topological space is separable if it has a countable dense subset.

PROPOSITION: A Hilbert space X is separable iff it has a countable Hilbert basis.

PROOF: If X has a countable Hilbert basis, the space generated by this basis is dense in X. Consider the set of *linear combinations with rational coefficients* of the basis elements: this is a dense countable set.

Conversely, let $\{x_n\} \subset X$ be a dense countable set. Apply the Gram-Schmidt orthogonalization process to this set: we obtain a sequence of orthonormal vectors $\{e_k\}$ which spans a subspace of X cointaining all vectors x_n, i.e. a dense subspace: we have a countable Hilbert basis. Q.E.D.

The fact that the orthonormal system in $L^2(2\pi)$ given by

$$\mathcal{F} = \left\{ \frac{1}{\sqrt{2\pi}}, \; \frac{1}{\sqrt{\pi}} \cos nx, \; \frac{1}{\sqrt{\pi}} \sin nx, \; n = 1, 2, \ldots \right\}$$

is *maximal*, and hence a Hilbert basis, comes form the *Stone-Weierstrass theorem*, which ensures that every continuous and 2π-periodic function can be approximated in the uniform norm with trigonometric polynomials (which are, by definition, linear combinations of elements of $\mathcal{F}$).

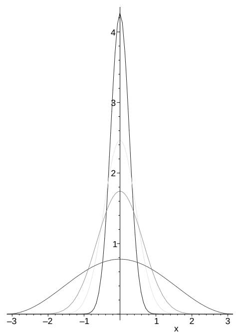
Since we will see very soon that every element of $L^2(2\pi)$ can be approximated in the L^2 norm with continuous functions, if follows that it can also be approximated with linear combinations of elements of $\mathcal{F}$: the space spanned by our orthonormal system is thus dense in $L^2(2\pi)$, and $\mathcal{F}$ is maximal.

We will see now the proof of Stone-Weierstrass theorem for trigonometric polynomials. We will need the fact that the set of trigonometric polynomials is an algebra: the product of two of those is still a trigonometric polynomial. This is easily checked if we express the sin and cos functions in terms of complex exponentials: as a consequence, notice that any polynomials in $\sin x$, $\cos x$ is indeed a trigonometric polynomial.

THEOREM (Stone-Weierstrass): Let $u : \mathbf{R} \to \mathbf{R}$ *be a* 2π*-periodic, continuous function. Then for every* $\varepsilon > 0$ *there exists a trigonometric polynomial* $v(x)$ *such that* $\|u - v\|_\infty < \varepsilon$*.*

PROOF: For every natural n, consider the following trigonometric polynomial: $\phi_n(t) = c_n \left(\frac{1+\cos t}{2}\right)^n$, where the constants c_n are chosen in such a way that $\int_{-\pi}^{\pi} \phi_n(t)\, dt = 1$.

If we draw the graph of these functions, we see non-negative periodic functions which "concentrate" around the points $2k\pi$:



We will see next time how these functions will allow us to construct the desired approximations of u with trigonometric polynomials.

We define the trigonometric polynomials approximating u as follows:

$$u_n(x) = \int_{-\pi}^{\pi} u(x+t)\phi_n(t)\, dt.$$

These functions are obtianed essentially by computing a "weighted average" of u: we will see that $u_n \to u$ uniformly.

Before we do that, we need to show that u_n are indeed trigonometric polynomials, which is not at all clear from the definition… To see that, it is enogh to change variables in the integral, by putting $s = x + t$: recalling that all functions involved are periodic, we get:

$$u_n(x) = \int_{-\pi}^{\pi} u(s)\phi_n(s - x)\, ds.$$

Since ϕ_n are trigonometric polynomials, by using the addition formulas for sin and cos, and the linearity of the integral, we can patently see that the functions u_n are indeed trigonometric polynomials!

We now need a simple estimates of the constants c_n appearing in the definition of ϕ_n: we have

$$\frac{1}{c_n} = \int_{-\pi}^{\pi} \left(\frac{1+\cos t}{2}\right)^n dt \geq \int_{-1/\sqrt{n}}^{1/\sqrt{n}} \left(\frac{1+\cos t}{2}\right)^n \geq \frac{2}{\sqrt{n}} \left(\frac{1+\cos(\frac{1}{\sqrt{n}})}{2}\right)^n.$$

The quantity between brackets in the last expression converges to $e^{-1/4}$, so that $c_n \leq k\sqrt{n}$ for n large enough, with k a suitable positive constant.

Tomorrow, we will use this estimate to show that $u_n \to u$ uniformly.

16 Lecture of november 4, 2015 (2 hours)

To show that $u_n \to u$ uniformly, we use insted the original definition of u_n. Since the functions ϕ_n are non negative with integral 1, we get

$$(*)|u_n(x) - u(x)| = |\int_{-\pi}^{\pi} (u(x+t) - u(x))\phi_n(t)\, dt| \leq \int_{-\pi}^{\pi} |u(x+t) - u(x)|\phi_n(t)\, dt.$$

Let M be an upper bound for $|u|$, and remark that u is uniformly continuous: for every $\varepsilon > 0$ we find $\delta > 0$ such that $|x - y| < \delta$ implies $|u(x) - u(y)| < \varepsilon$.

Now, split the r.h.s. integral in (*) on the sets $[-\delta, \delta]$ and $[-\pi, -\delta] \cup [\delta, \pi]$. By our choice of δ we get

$$\int_{-\delta}^{\delta} |u(x+t) - u(x)|\phi_n(t)\, dt \leq \varepsilon \int_{-\delta}^{\delta} \phi_n(t)\, dt < \varepsilon.$$

On the other hand

$$\int_{\delta}^{\pi} |u(x+t) - u(x)|\phi_n(t)\, dt \leq 2Mc_n \int_{\delta}^{\pi} \left(\frac{1+\cos t}{2}\right)^n dt \leq$$
$$2Mk\pi\sqrt{n} \left(\frac{1+\cos(\delta)}{2}\right)^n,$$

and the last quantity vanishes as $n \to +\infty$, uniformly in x (because it does not depend on x!). The integral on $[-\pi, -\delta]$ is estimated in the same way: for n large enough we thus get $|u_n(x) - u(x)| < 2\varepsilon$ for every x. Q.E.D.

To finish our discussion on Hilbert spaces, we study weak convergence in this setting. By Riesz's representation theorem, weak convergence in a Hilbert space X reads as follows: if $\{x_n\} \subset X$, then

$$(x_n \rightharpoonup x) \Longleftrightarrow_{Def} (< y, x_n > \to < y, x > \quad \forall y \in X).$$

The following are some interesting facts about weak convergence in a Hilbert spcae:

PROPOSITION: Let X be a Hilbert space. Then

(i) If $x_n \rightharpoonup x$, then $\{x_n\}$ is bounded and $\|x\| \leq \liminf_{n \to +\infty} \|x_n\|$.

(ii) If $\{x_n\}$ is such that for every $y \in X$ the limit $T(y) := \lim\limits_{n \to +\infty} < x_n, y >$ exists and is finite,then there is a unique $x \in X$ such that $x_n \rightharpoonup x$.

(iii) If $x_n \rightharpoonup x$ and $y_n \to y$ (strong convergence), then

$$< x_n, y_n > \to < x, y > .$$

(iv) We have $x_n \to x$ iff $x_n \rightharpoonup x$ and $\|x_n\| \to \|x\|$.

PROOF: We already know (i) in a general Banach space. For (ii), apply the Banach Steinhaus theorem to the functionals $T_n(y) := < x_n, y >$, and Riesz representation theorem to the limit functional $T(y)$.

To prove (iii), write

$$< x_n, y_n > - < x, y > = (< x_n, y_n > - < x_n, y >) + (< x_n, y > - < x, y >).$$

The second bracket converges to zero by definition of weak convergence. To estimate the first bracket, notice that $\{x_n\}$ is bounded in norm by (i), and apply the Cauchy-Schwarz inequality to $< x_n, y_n - y >$: the first bracket also goes to 0 and (iii) is proved.

An implication of (iv) is obvious. For the other, write

$$\|x_n - x\|^2 = \|x_n\|^2 - 2 < x_n, x > + \|x\|^2$$

and apply weak convergence and convergence of the norms. Q.E.D.

We now see the proof of the Banach Alaoglu theorem in a separable Hilbert space:

THEOREM: Let X be a separable Hilbert space, $\{x_n\}_{n\in\mathbf{N}}$ a bounded sequence in X. Then there exists $x \in X$ and a subsequence x_{n_k} of x_n such that $x_{n_k} \rightharpoonup x$.

PROOF: Let $\{e_j\}_{j\in\mathbf{N}}$ be a Hilbert basis in X. Then for every n we can write

$$x_n = \sum_{j=1}^{\infty} < x_n, e_j > e_j.$$

By the boundedness hypothesis, there exists $C > 0$ such that

$$\sum_{j=1}^{\infty} < x_n, e_j >^2 \leq C \quad \forall n \in \mathbf{N}.$$

In particular, for every fixed j we have $| < x_n, e_j > | \leq \sqrt{C}$ for every n.

Consider now the Fourier coefficients $< x_n, e_1 >$: they form a bounded sequence in $\mathbf{R}$, and we can extract a subsequence $x_n^{(1)}$ of x_n such that $< x_n^{(1)}, e_1 > \to \overline{c}_1 \in \mathbf{R}$.

The real sequence $< x_n^{(1)}, e_2 >$ is also bounded: we extract a subsequence $x_n^{(2)}$ of $x_n^{(1)}$ in such a way that $< x_n^{(2)}, e_2 > \to \overline{c}_2 \in \mathbf{R}$.

Proceeding in this way, we construct by recursion a sequence of subsequences $x_n^{(k)}$ such that $x_n^{(k)}$ is a subsequence of $x_n^{(k-1)}$, and such that the Fourier coefficients satisfy

$$\lim_{n\to+\infty} < x_n^{(k)}, e_j > = \overline{c}_j, \ j = 1, 2, \ldots, k.$$

Take the diagonal sequence defined by $\tilde{x}_n = x_n^{(n)}$. It is a subsequence of $\{x_n\}$ with the property that

$$\lim_{n\to+\infty} < \tilde{x}_n, e_j > = \overline{c}_j \quad \forall j \in \mathbf{N}.$$

By the linearity of the limit and of the scalar product, if we put $Y = \mathrm{span}\{e_j : \ j \in \mathbf{N}\}$ (a dense subspace), we have that for every $y \in Y$ the following limit exists and is finite

$$\lim_{n\to+\infty} < \tilde{x}_n, y > = T(y),$$

and $T : Y \to \mathbf{R}$ is obviously linear. Moreover, T is bounded (as a pointwise limit of norm-bounded functionals), and so can be extended to a bounded linear functional T defined on the whole space X. Let $\overline{x} \in X$ be such that $T(y) = < \overline{x}, y >$ for every $y \in X$ (this exists by Riesz representation theorem...and its Fourier coefficients are clearly $\overline{c}_j$). Our construction ensures that $<\tilde{x}_n, y> \to <\overline{x}, y>$ for every $y \in Y$: let us show that the same holds for every $y \in X$.

Indeed, let $y \in X$, $\varepsilon > 0$. Since Y is dense in X, there exists $\tilde{y} \in Y$ such that $\|\tilde{y} - y\| < \varepsilon$. Hence

$$
\begin{aligned}
&<\tilde{x}_n, y> - <\overline{x}, y> = \\
&<\tilde{x}_n, y - \tilde{y}> + (<\tilde{x}_n, \tilde{y}> - <\overline{x}, \tilde{y}>) - <\overline{x}, y - \tilde{y}> .
\end{aligned}
$$

The modulus of the quantity between brackets is less than ε for large enough n. The other two terms are estimated by $C\varepsilon$: take for instance the first, by Cauchy-Schwarz we have

$$
|<\tilde{x}_n, y - \tilde{y}>| \leq \|\tilde{x}_n\| \, \|y - \tilde{y}\| < C\varepsilon.
$$

We then conclude that

$$
|<\tilde{x}_n, y> - <\overline{x}, y>| < (2C+1)\varepsilon
$$

for large enough n. Q.E.D.

REMARK: The last theorem is easily extended to a non-separable Hilbert space X. Indeed, if $\{x_n\}$ is a bounded sequence in X, define $Z = \overline{\text{span}\{x_n : \ n \in \mathbf{N}\}}$. This is obviously a separable Hilbert space (the linear combinations with coefficients in $\mathbf{Q}$ of the vectors x_n are a countable dense subset): by our previous result we find $\overline{x} \in Z$ and a subsequence x_{n_k} such that $<x_{n_k}, y> \to <\overline{x}, y>$ for every $y \in Z$ as $k \to +\infty$. But the same holds for every $y \in Z^\perp$ (because all scalar products are zero!): it thus holds for every $y \in X$, because $X = Z \oplus Z^\perp$: $x_{n_k} \rightharpoonup \overline{x}$ in X.

REMARK: The argument we used in a separable Hilbert space applies, with few modifications, in the case of a *reflexive Banach space* whose dual is *separable* (it can be shown that this hypothesis is equivalent to ask that the space is *reflexive and separable*). In this case, we must replace the Hilbert basis $\{e_j\}$ with a countable family of elements of the dual space which generate a dense subspace: indeed, in the proof the orthonormality of the basis vectors was not used in any essential way!

17 Lecture of november 9, 2015 (2 hours)

Among the most important function spaces we met in this course are the Lebesgue spaces $L^p(\Omega)$. We will now study some properties of these spaces, which are very important for the applications: in particular, we will see that every function in $L^p(\Omega)$ (for finite p) can be approximated with regular functions.

The following theorem highlights a surprising relation between measurable functions and continuous functions:

THEOREM (Lusin): Let $u : \Omega \to \mathbf{R}$ be a measurable function, with Ω a bounded and Lebesgue-measurable set. Then, for every $\varepsilon > 0$, there is a compact set $K \subset \Omega$ such that $|\Omega \setminus K| < \varepsilon$ and such that the restriction of u to K is continuous.

PROOF: For $j = 1, 2, 3, \ldots$, write $\mathbf{R} = \bigcup_{i=1}^{+\infty} I_{ij}$, with I_{ij} disjoint intervals with length less than $1/j$. Fix also points $y_{ij} \in I_{ij}$.

Let then $A_{ij} = u^{-1}(I_{ij})$: those are pairwise disjoint measurable sets, whose union is Ω. By regularity of Lebesgue measure, we can find compact sets $K_{ij} \subset A_{ij}$ such that $|A_{ij} \setminus K_{ij}| < \frac{\varepsilon}{2^{i+j}}$.

Obviously, $|\Omega \setminus \bigcup_{i=1}^{\infty} K_{ij}| < \frac{\varepsilon}{2^j}$, and by continuity of the measure on decreasing sequences of sets we can choose $N_j \in \mathbf{N}$ such that

$$|\Omega \setminus \bigcup_{i=1}^{N_j} K_{ij}| < \frac{\varepsilon}{2^j}.$$

Define $K_j = \bigcup_{i=1}^{N_j} K_{ij}$: this is a compact set. We then define $u_j : K_j \to \mathbf{R}$ by $u_j(x) = y_{ij}$ for $x \in K_{ij}$: we obtain a continuous function (it is constant on K_{ij}, and we have only a finite number of these sets, which are at a positive distance from each other) with the property that $|u_j(x) - u(x)| < 1/j$ for every $x \in K_j$.

If we then define $K = \bigcap_{j=1}^{\infty} K_j$, we get a compact set satisfying $|\Omega \setminus K| < \varepsilon$ over which $u_j \to u$ uniformly. It follows that the restriction of u to K is continuous, being the uniform limit of continuous functions. Q.E.D.

REMARK: Lusin's Theorem does not contraddict the fact that there are measurable functions which are everywhere discontinuous: we are not saying that points of K are continuity points for u, but only for its restriction!

The following is a well-known extension theorem:

THEOREM (Tietze): Let $K \subset \mathbf{R}^n$ be a compact set. If $u : K \to \mathbf{R}$ is continuous, there exists a continuous function $\tilde{u} : \mathbf{R}^n \to \mathbf{R}$ extending u (i.e., such that $u(x) = \tilde{u}(x)$ for all $x \in K$) and such that $\|\tilde{u}\|_\infty = \|u\|_\infty$. Moreover, if $K \subset \Omega$, with Ω open in $\mathbf{R}^n$, we can also require that $u \in C^0_C(\Omega)$.

PROOF: Put $M = \|u\|_\infty$ and define the compact sets $K_1 = u^{-1}([-M, -M/3])$, $K_2 = u^{-1}([M/3, M])$: suppose for a moment they are both nonempty, and let $\delta > 0$ be their mutual distance. Then the function

$$\tilde{u}_1(x) = \min\{M/3, -M/3 + \frac{2M}{3\delta}\text{dist}(x, K_1)\}$$

is continuous, everywhere defined and takes values between $-M/3$ e $M/3$. Moreover, on K_1 it takes the value $-M/3$, on K_2 the value $M/3$. It follows that $|\tilde{u}_1(x) - u(x)| \leq \frac{2}{3}M$ for every $x \in K$. If K_1 is empty, we obtain the same result by putting $\tilde{u}_1(x) = M/3$ (constant function). A similar argument works if K_2 is empty.

We repeat the same construction for the function $u_2 = u - \tilde{u}_1$: we find a continuous function $\tilde{u}_2$ which is defined everywhere, with $\|\tilde{u}_2\|_\infty \leq \frac{2}{9}M$ and such that $\|u - \tilde{u}_1 - \tilde{u}_2\|_\infty < \frac{4}{9}M$. Proceeding in the same way, we find a sequence $\tilde{u}_k$ of continuous functions such that $\|\tilde{u}_k\|_\infty \leq \frac{2^{k-1}}{3^k}M$ and such that

$$(*)\ \|u - \tilde{u}_1 - \tilde{u}_2 - \ldots - \tilde{u}_k\|_\infty < \frac{2^k}{3^k} \textit{ in } K.$$

The series of continuous functions

$$\sum_{k=1}^{\infty} \tilde{u}_k(x)$$

converges uniformly in $\mathbf{R}^n$ to some function $\tilde{u}$ (because the series of the norms converges). By using our estimates of the terms in this sum, we immediately see that the norm of $\tilde{u}$ is less or equal to M. Moreover, by (*) $\tilde{u}$ coincides with u on the points of K.

Let now $\Omega \supset K$ be open. Take two bounded open sets Ω', Ω'' such that $K \subset \Omega' \subset\subset \Omega'' \subset\subset \Omega$, and add the compact set $\overline{\Omega}'' \setminus \Omega'$ to K, with the prescription that $u = 0$ on that set. Apply the theorem in the form proved above, and modify the extension $\tilde{u}$ so that it is zero on $\Omega \setminus \Omega''$. Q.E.D.

The following result ensures the density of continuous functions in L^p:

THEOREM (Density of continuous functions in L^p): Let $1 \leq p < +\infty$, Ω be an open set in $\mathbf{R}^n$. Then continuous functions with compact support are dense in $L^p(\Omega)$.

PROOF.: We have to show that given $u \in L^p(\Omega)$ and $\varepsilon > 0$, we can find $v \in C^0_C(\Omega)$ such that $\|u - v\|_{L^p} < \varepsilon$.

Suppose first Ω is bounded and $\|u\|_\infty = M < +\infty$: we will remove these restrictions later.

By Lusin's theorem, we find a compact set $K \subset \Omega$ such that $|\Omega \setminus K| < (\frac{\varepsilon}{2M})^p$ and the restriction of u to K is continuous. By the Tietze extension theorem, we find $v \in C^0_C(\Omega)$ such that $\|v\|_\infty = M$ and $v \equiv u$ on K. Then we get

$$\|u - v\|_{L^p(\Omega)} = \|u - v\|_{L^p(\Omega\setminus K)} \leq \|u\|_{L^p(\Omega\setminus K)} + \|v\|_{L^p(\Omega\setminus K)} < \varepsilon,$$

as we wanted.

If Ω is bounded, but u is unbounded, remark that the truncated functions $u_M(x) = \max\{-M, \min\{M, u(x)\}\}$ converge to u in the L^p norm as $M \to +\infty$ (dominated convergence theorem). Finally, if Ω is unbounded consider the functions

$$u_R(x) = \begin{cases} u(x) & if\ |x| < R, \\ 0 & if\ |x| \geq R. \end{cases}$$

By the dominated convergence theorem, we see that $u_R \to u$ in L^p as $R \to +\infty$. The functions u_R are supported in $\Omega \cap B_R(0)$, which is a bounded open set, so the previous result applies. Q.E.D.

18 Lecture of november 10, 2015 (2 hours)

REMARK: The density result is of course false for $p = +\infty$. Indeed, continuous functions are a closed proper subspace of $L^\infty(\Omega)$.

Here are some consequences of the density of continuous functions in L^p.

REMARK: $L^p(\Omega)$ is separable for $1 \leq p < +\infty$: it is easy to construct a countable set of functions which is dense in the subspace of continuous functions. For instance, take the *step functions* whose steps are intervals with rational endpoints and have rational heights. By using the uniform continuity, every continuous function with compact support is arbitrarily close to a step function of this kind!

On the other hand, you can show as an exercise that $L^\infty(\Omega)$ is *not* separable.

EXERCISE: Another important consequence of the density of continuous functions is the following *continuity of translations* in $L^p(\mathbf{R}^n)$ $(1 \leq p < +\infty)$: let $u \in L^p(\mathbf{R}^n)$, $y \in \mathbf{R}^n$. Define $u_y(x) := u(x-y)$ (translated function). Show that

$$\lim_{y\to 0} \|u - u_y\|_{L^p} = 0.$$

(HINT: If $u \in C^0_C(\mathbf{R}^n)$, the result is an easy consequence of uniform continuity. A generic function can be approximated by continuous, compactly supported functions.)

We show a "better" version of the density theorem, where we show that C_C^∞ is a dense subspace of $L^p(\Omega)$. This result can be proved by applying a procedure called *regularization by convolution*: essentially, it is the same trick we already used to prove the completeness of the trigonometric system in $L^2(2\pi)$.

We begin with a regularity result:

LEMMA (Regolarity of the convolution product): Let $u \in L^1_{loc}(\mathbf{R}^n)$, $\phi \in C^1_C(\mathbf{R}^n)$. *Then the function*

$$v(x) = \int_{\mathbf{R}^n} u(z)\phi(x-z)\,dz$$

is of class C^1 *and*

$$\frac{\partial v}{\partial x_i}(x) = \int_{\mathbf{R}^n} u(z)\frac{\partial \phi}{\partial x_i}(x-z)\,dz.$$

By iterating this result, if $\phi \in C_C^\infty$, *we get* $v \in C^\infty$.

PROOF: Since the integrands depends from x in a C^1 way, this is just a theorem about differentiation under the sign of integral. It is an easy enough consequence of the dominated convergence theorem.

We begin by showing that v is continuous: let indeed $x \in \mathbf{R}^n$, $y \in \mathbf{R}^n$ with $|y| \leq 1$. Let then K be a compact set containing the support of $\phi(x+y-\cdot)$ for every y as above, M be the uniform norm of ϕ.

Then

$$|v(x+y) - v(x)| \leq \int_K |u(z)||\phi(x+y-z) - \phi(x-z)|\,dz.$$

By continuity of ϕ, the integrand converges pointwise to 0 as $y \to 0$. Convergence is dominated by $2M|u|\mathbf{1}_K$: we thus infer that v is continuous.

Next, we show that v is differentiable and that partial derivatives are as in the thesis: continuity of partial derivatives will then follows by repeating the above argument with ϕ replaced by $\frac{\partial \phi}{\partial x_i}$. For fixed $i = 1, \ldots n$ and for all $|h| < 1$ we have

$$\frac{v(x+he_i) - v(x)}{h} = \int_K u(z)\frac{\phi(x+he_i-z) - \phi(x-z)}{h}\,dz.$$

As $h \to 0$, the integrand converges pointwise to $u(z)\frac{\partial \phi}{\partial x_i}(x-z)$ and convergence is dominated by $L|u(z)|\mathbf{1}_K$, with L the Lipschitz constant of ϕ, as we wanted. Q.E.D.

An L^p function is approximated by a sequence of regular functions, obtained by computing the convolution product of the original function with some C^∞_C maps called *mollifiers*.

THEOREM (Regularization by convolution): Let $u \in L^p(\mathbf{R}^n)$ *with* $1 \le p < +\infty$*. Then there exists a sequence* $\{u_k\} \subset C^\infty(\mathbf{R}^n)$ *such that* $u_k \to u$ *in* $L^p(\mathbf{R}^n)$*. This sequence can be obtained in such a way that* $\{u_k\} \subset C^\infty_C(\mathbf{R}^n)$*.*

PROOF: Let $\phi : \mathbf{R}^n \to \mathbf{R}$ be a C^∞ function such that $\phi(x) \ge 0$, $\phi(x) = \phi(-x)$ for every x and such that

$$\mathrm{spt}\, \phi \subset B_1(0), \qquad \int\limits_{\mathbf{R}^n} \phi(x)\, dx = 1.$$

Such a function is called a *mollifier* or *bump function*: for instance, we can take $\phi(x) = c \exp(-\frac{1}{1-|x|^2})$ for $|x| < 1$, $\phi(x) = 0$ for $|x| \ge 1$, where c is a positive constant chosen in such a way that the integral is 1.

We then define $\phi_k(x) = k^n \phi(kx)$: those functions share the main qualitative properties of ϕ, but concentrate more and more around the origin, because $\mathrm{spt}\, \phi_k \subset B_{1/k}(0)$.

Consider now the sequence of functions

$$u_k(x) = \int_{\mathbf{R}^n} u(x-y)\phi_k(y)\, dy.$$

Whith a change of variables, the last expression can also be written $u_k(x) = \int_{\mathbf{R}^n} u(z)\phi_k(x-z)\ dz$ and by the Lemma, we immediatly see that $u_k \in C^\infty(\mathbf{R}^n)$: this process is called the *regularization by convolution* of u.

We show that $u_k \to u$ in L^p: one has

$$\|u_k - u\|^p_{L^p(\mathbf{R}^n)} = \int_{\mathbf{R}^n} \left| \int_{\mathbf{R}^n} (u(x-y) - u(x))\phi_k(y)\, dy \right|^p dx.$$

In the inner integral, write $\phi_k(y) = \phi_k(y)^{1/p}\phi_k(y)^{1-1/p}$ and use Hölder inequality: since ϕ_k has integral 1, that integral is less or equal than

$$\left(\int_{\mathbf{R}^n} |u(x-y) - u(x)|^p \phi_k(y)\, dy \right)^{1/p}.$$

By replacing this upper bound in the above expression we thus obtain:

$$\|u_k - u\|^p_{L^p(\mathbf{R}^n)} \le \int_{\mathbf{R}^n} \int_{\mathbf{R}^n} |u(x-y) - u(x)|^p \phi_k(y)\, dy\, dx =$$
$$\int_{B_{1/k}(0)} \phi_k(y) \int_{\mathbf{R}^n} |u(x-y) - u(x)|^p\, dx\, dy,$$

where we applied Fubini theorem and the properties of the support of ϕ_k.

Fix $\varepsilon > 0$: since in the double integral we have $|y| \leq 1/k$, continuity of translations in L^p ensures that for large enough k the inner integral is less than ε: the whole expression is estimated with ε, and $u_k \to u$, as we wanted.

If we wish the approximating functions to be compactly supported, we observe that fore every $\varepsilon > 0$ there exists $N \in \mathbf{N}$ such that $\|u_N - u\|_{L^p} < \varepsilon$, where $u_N = u \cdot \mathbf{1}_{B_N(0)}$ (compare the proof of the theorem about density of continuous functions). The function u_N is identically 0 outside the ball of radius N, so the functions $(u_N)_k$ obtained regularizing by convolution are compactly supported (precisely, their support is contained within the ball centered in 0 and with radius $N + 1/k$). For k large enough we have $\|(u_N)_k - u_N\|_{L^p} < \varepsilon$, whence $\|u - (u_N)_k\| < 2\varepsilon$. Q.E.D.

REMARK: If Ω is open in $\mathbf{R}^n$ and $u \in L^p(\Omega)$ $(1 \leq p < +\infty)$, then there is a sequence $u_k \in C_C^\infty(\Omega)$ converging to u in L^p. Indeed, for every $\varepsilon > 0$ we find $\tilde{u} \in C_C^0(\Omega)$ such that $\|u - \tilde{u}\|_p < \varepsilon$. We regularize $\tilde{u}$ by convolution, obtaining a sequence of functions $\tilde{u}_k$ converging to $\tilde{u}$ in L^p: for k large enough we have $\|tildeu_k - \tilde{u}\|_p < \varepsilon$, so that $\|tildeu_k - u\|_p < 2\varepsilon$. Moreover, for k large enough $\tilde{u}_k$ is compactly supported in Ω: indeed, it is easy to check (by looking at the definition of the regularized functions as convolution products) that the support of $\tilde{u}_k$ is contained in a neighborhood of radius $1/k$ of the support of $\tilde{u}$.

In the following lectures we will see some complements in measure theory.

We will first introduce some conditions, which ensure that a given measure has regularity properties similar to those of Lebesgue measure: in particular, we are interested in the possibility of approximating the measure of a set by means of open and/or compact sets.

In the following, we will assume that X is a *locally compact and separable* metric space.

DEFINITION: If X is as above, *Borel σ-algebra* $\mathcal{B}$ is defined as the smallest σ-algebra containing the open sets of X. An outer measure (resp. measure) μ is said to be *Borel* if Borel sets are μ-measurable.

An outer measure (resp. measure) μ is *Borel regular* if every set (resp. measurable set) A is contained in a Borel set B such that $\mu(A) = \mu(B)$.

Finally, μ is a *Radon measure* if it is Borel-regular and $\mu(K) < +\infty$ for every compact set K.

19 Lecture of november 11, 2015 (2 hours)

A Radon measure is regular in the same sense of Lebesgue measure:

THEOREM (Approximation of the measure with open, closed, compact sets): Let μ be a Borel outer measure on X. Suppose further there is a sequence of open sets $\{V_j\}$ such that $X = \bigcup V_j$ and $\mu(V_j) < +\infty$ (a sort of strengthened σ-finiteness). Then for every $A \subset X$ we have

$$\begin{aligned} (*)\ \mu(A) &= \inf\{\mu(U):\ U\ \textit{open},\ U \supset A\}, \\ (**)\ \mu(A) &= \sup\{\mu(C):\ C\ \textit{closed},\ C \subset A\}. \end{aligned}$$

If μ is only a Borel measure, the same relations hold for A a Borel set. If μ is a Radon measure, then the above open sets V_j always exist. Morevoer,

$$(***)\ \mu(A) = \sup\{\mu(K):\ K\ \textit{compact},\ K \subset A\}.$$

PROOF: We first show (*) for a Borel set A.

Suppose $\mu(X) < +\infty$: we will remove later this additional hypothesis. Define $\mathcal{A} = \{A\ \textit{Borel}:\ (*)\ \textit{holds}\}$. We show that $\mathcal{A}$ is closed under countable unions and intersections. Indeed, if $A = \bigcup\limits_{n=1}^{\infty} A_n$ with $A_n \in \mathcal{A}$, ther for every $\varepsilon > 0$ we can find open sets U_n such that $A_n \subset U_n$ and $\mu(U_n \setminus A_n) < \varepsilon/2^n$. Then $U = \bigcup_{n=1} U_n$ is an open set containing A and

$$\mu(U \setminus A) \le \mu\left(\bigcup_{n=1}^{\infty}(U_n \setminus A_n)\right) < \varepsilon,$$

whence $A \in \mathcal{A}$. If on the other hand $B = \bigcap\limits_{n=1}^{\infty} A_n$, then $B \subset \bigcap_{n=1}^{\infty} U_n$ and we immediately check that $\mu(\bigcap_{n=1}^{\infty} U_n \setminus B) < \varepsilon$. If we define $V_N = \bigcap_{n=1}^{N} U_n$, we have a decreasing sequence of open sets with finite measure, all containing B, such that $\mu(V_N \setminus B) \to \mu(\bigcap_{n=1}^{\infty} U_n \setminus B)$. So for large enough N we have $\mu(V_N \setminus B) < \varepsilon$ and $B \in \mathcal{A}$.

Obviously, the family $\mathcal{A}$ contains all open sets in X. As it is closed under countable intersections, it also contains the closed sets: a closed set C in a metric space can be expressed as a countable intersection of open sets by

$$C = \bigcap_{n=1}^{\infty}\{x \in X:\ \mathrm{dist}(x, C) < 1/n\}.$$

Define now $\mathcal{A}' = \{A \in \mathcal{A} : \; A^C \in \mathcal{A}\}$: this is clearly a σ-algebra, and it contains the open sets. So $\mathcal{A}'$ is the Borel σ-algebra.

In case $\mu(X) = +\infty$, we use the open sets V_j in the hypothesis: given a Borel set A, we apply our previous result to the finite measures $\mu|_{V_j}$ (defined by $\mu|_{V_j}(A) = \mu(A \cap V_j))$ and for every $\varepsilon > 0$ we find open sets U_j such that $U_j \cap V_j \supset A \cap V_j$ and $\mu(U_j \cap V_j) < \mu(A \cap V_j) + \varepsilon/2^j$. We immediately see that the open set $U = \bigcup_{j=1}^{\infty}(U_j \cap V_j)$ contains A and approximates its measure within ε.

So (*) holds for Borel sets if we just have a Borel measure. If the measure is also Borel regular, then (*) holds for every set.

(**) follows immediately by taking the complements.

Finally, if $\mu(K) < +\infty$ for every compact K and we recall that X is separable and locally compact, it is easy to see that the sequence V_j in the statement exists: indeed, in a separable metric space the topology has a countable basis. Thank to the local compactness, this can be replaced by a countable basis of relatively compact open sets.

(***) easily follows, because every closed set is the union of an increasing sequence of compact sets (use in a suitable way the closure of the relative comapct open sets in the argument above...) Q.E.D.

When we construct (outer) measures on a metric space, it is useful to have a criterion ensuring that Borel sets are measurable:

THEOREM (Caratheodory criterion): Let μ be an outer measure on X such that $\mu(A \cup B) = \mu(A) + \mu(B)$ whenever $\operatorname{dist}(A, B) > 0$. *Then μ is a Borel measure.*

PROOF: It is enough to show that every closed set C is measurable, i.e. for every $T \subset X$ we have

$$\mu(T) \geq \mu(T \setminus C) + \mu(T \cap C).$$

Consider the closed sets $C_j = \{x \in X : \; \operatorname{dist}(x, C) \leq 1/j\}$: since $T \setminus C_j$ has a strictly positive distance from C, we get $\mu(T) \geq \mu((T \setminus C_j) \cup (T \cap C)) = \mu(T \setminus C_j) + \mu(T \cap C)$.

To conclude, we just have to show that $\mu(T \setminus C_j) \to \mu(T \setminus C)$ as $j \to +\infty$. On the other hand, if we define

$$R_k = \{x \in T : \; \frac{1}{k+1} < \operatorname{dist}(x, C) \leq \frac{1}{k}\}$$

then $T \setminus C = (T \setminus C_j) \cup (\bigcup_{k=j}^{\infty} R_k)$ and we can conclude thanks to countable subadditivity, provided we show that

$$\lim_{j \to +\infty} \sum_{k=j}^{\infty} \mu(R_k) = 0.$$

This is true because the series $\sum_{k=1}^{\infty}\mu(R_k)$ converges (the *j-th remainder* of a convergent series goes to 0 as $j \to +\infty$). Consider indeed any finite sum of *even* terms of the series: using the additivity of measure on sets at a positive distance from each other, and monotonicity, we get $\sum_{k=1}^{N}\mu(R_{2k}) = \mu(\bigcup_{k=1}^{N} R_{2k}) \le \mu(T)$. A similar bound holds of course for any finite sum of *odd* terms, so the partial sums of the series are bounded from above by $2\mu(T)$ and the series converges (notice indeed that if we have $\mu(T) = +\infty$, we have nothing to prove!). Q.E.D.

EXAMPLE: The previous theorem shows for instance that the *Hausdorff measures* are Borel measures.

The α-dimensional Hausdorff measure of $A \subset \mathbf{R}^n$ is defined as follows:

$$\begin{aligned}
\mathcal{H}^{\alpha}(A) &= \lim_{\delta\to 0^+} \mathcal{H}^{\alpha}_{\delta}, \text{ where} \\
\mathcal{H}^{\alpha}_{\delta}(A) &= c(\alpha)\inf\{\sum_{k=1}^{\infty}(\mathrm{diam}(A_k))^{\alpha} :\ A_k \text{ closed},\ \mathrm{diam}(A_k) \le \delta,\ A \subset \bigcup_{k=1}^{\infty} A_k\}.
\end{aligned}$$

Here, $c(\alpha)$ is a renormalization constant which, for integer α, gives the Lebesgue measure of the α-dimensional ball of diameter 1. This function is extended to non-integer values of α by using Euler's Γ function[15].

It is easy to check that $\mathcal{H}^1(C)$ gives the correct length of a rectifiable curve C, while $\mathcal{H}^2(S)$ gives the area of a regular surface S. Moreover, one can show (but it is not easy!) that $\mathcal{H}^3$ coincides with the 3-dimensional Lebesgue measure.

The Hausdorff measure makes sense for every real value of α, and appears for instance in the definition of the *Hausdorff dimension* of a set:

$$\dim_{\mathcal{H}}(A) = \inf\{\alpha > 0 :\ \mathcal{H}^{\alpha}(A) = 0\} = \sup\{\alpha > 0 :\ \mathcal{H}^{\alpha}(A) = +\infty\}.$$

Typically, fractal sets have non-integer Hausdorff dimension!

REMARK: We remark that the proof of the Lusin theorem and the density of continuous functions in $L^p(\mathbf{R}^n)$ depend essentially on the possibility of approximating the measure of a given set with open and compact sets. We just saw that this is true also for Radon measures on a locally compact and separable metric space: checking that continuous and compactly supported functions are dense in $L^p(\mu)$ is now a lengthy but easy exercise!

[15]Precisely: $c(\alpha) = \frac{\Gamma(1/2)}{\Gamma(\alpha/2+1)2^{\alpha}}$, where Euler's Γ function is defined by $\Gamma(t) = \int_0^{+\infty} s^{t-1}e^{-s}\,ds$.

20 Lecture of november 16, 2015 (2 hours)

Let μ be a measure on a set X, whose σ-algebra of measurable set is $\mathcal{S}$. We can build a number of new measures as follows: given a measurable $u : X \to [0, +\infty]$, we define a measure ν on the σ-algebra $\mathcal{S}$ by

$$(****)\ \nu(A) = \int_A u(x)\, d\mu(x) \quad \forall A \in \mathcal{S}.$$

It is easy to verify that ν is a positive measure. Moreover, it is clear that $\nu(A) = 0$ whenever $\mu(A) = 0$: we express this fact by saying that ν is *absolutely continuous* with respect to μ.

DEFINITION: Given two measures μ, ν on the same σ-algebra $\mathcal{S}$, we say that ν is absolutely continuous with respect to μ (and we write $\nu << \mu$) if $A \in \mathcal{S}$, $\mu(A) = 0$ implies $\nu(A) = 0$.

If μ is a finite measure, then all measures which are absolutely continuous with respect to μ can be written as in (****):

THEOREM (Radon-Nikodym): Let μ be a finite measure on X (i.e. $\mu(X) < +\infty$), ν another finite measure defined on the same σ-algebra $\mathcal{S}$ and such that $\nu << \mu$. Then there exists a function $w \in L^1(\mu)$, $w \geq 0$, such that

$$\nu(A) = \int_A w(x)\, d\mu(x) \quad \forall A \in \mathcal{S}.$$

PROOF of the Radon-Nikodym Theorem: Consider the measure $\rho = \mu + \nu$. Osserve that, by the absolute continuity of ν w.r.t. μ, two functions which are a.e. equal with respect to μ or ρ are a.e. equal with respect to ν.

If $u \in L^1(\rho)$ define $T(u) := \int_X u\, d\nu$. This is a linear functional: moreover, by the Cauchy-Schwarz inequality we get

$$T(u) \leq \int_X |u|\, d\nu \leq \int_X |u|\, d\rho \leq \|u\|_{L^2(\rho)} \rho(X)^{1/2}.$$

This means that $T \in (L^2(\rho))'$: by Riesz representation theorem (in a Hilbert space), there exists a unique function $v \in L^2(\rho)$ such that

$$(I)\ \int_X u\, d\nu = \int_X vu\, d\rho \quad \forall u \in L^2(\rho).$$

We would like to write

$$\int_X (1-v)u\,d\nu = \int_X uv\,d\mu \quad \forall u \in L^2(\rho)$$

and to choose $u = \mathbf{1}_E \frac{1}{1-v}$: if we knew that this function belong to $L^2(\rho)$, we would have our thesis with $w = v/(1-v)$. But in general this is not true... and we also risk dividing by 0: we need a more solid argument!

By applying (I) to the function $u = \mathbf{1}_E$, with E measurable, we obtain $\nu(E) = \int_E v\,d\rho$. Since $0 \le \nu(E) \le \rho(E)$, we also get

$$(II)\ 0 \le \frac{1}{\rho(E)} \int_E v\,d\rho \le 1 \quad \forall E \in \mathcal{S},\ \rho(E) > 0.$$

It follows that $0 \le v(x) \le 1$ for ρ-almost every $x \in X$ (if $E_n = \{x \in X : v(x) \ge 1 + 1/n\}$ had a positive measure, the central term in (II) would be strictly greater than 1... With a similar argument, v cannot be strictly negative on a set with positive measure).

(I) then becomes

$$(III)\ \int_X (1-v)u\,d\nu = \int_X uv\,d\mu \quad \forall u \in L^2(\rho).$$

If $A = \{x \in X : v(x) = 1\}$, (III) with $u = \mathbf{1}_A$ gives $\mu(A) = 0$, whence $\nu(A) = 0$. Outside this set of measure zero, the function $1/(1-v)$ is well defined...but not necessarily in $L^2(\rho)$.

But fixed a set $E \in \mathcal{S}$, for every $n \in \mathbf{N}$ the functions $v_n(x) = (1 + v(x) + v^2(x) + \ldots + v^n(x))\mathbf{1}_E(x)$ belong to $L^2(\rho)$. Pluggiong these in (III) we get

$$\int_E (1 - v^{n+1}(x))\,d\nu(x) = \int_E (v(x) + v^2(x) + \ldots + v^n(x))\,d\mu(x).$$

The left hand side converges to $\nu(E)$ by the monotone convergence theorem (the integrands grow to 1 for a.e. $x \in X$)... The integrands in the r.h.s. grow to $w(x) = \frac{v(x)}{1-v(x)}$, and by Beppo Levi the integrals converge to $\int_E w(x)\,d\mu(x)$. So we have $\nu(E) = \int_E w(x)\,d\mu(x)$. Summability of w comes from the fact that ν is a finite measure. Q.E.D.

EXERCISE: Show that the Radon-Nikodym theorem is still true if μ and ν are *σ-finite measures*. In that case, the function w in the thesis is not necessarily summable. Show then that the theorem is false if the measures are not σ-finite: take for instance ν the Lebsegue measure on $\mathbf{R}$, μ the counting measure: we have $\nu << \mu$, but the function w in the statement of the Radon-Nikodym theorem cannot exist.

We will now introduce *signed measures.*

DEFINITION (Signed measure): Let $\mathcal{S}$ be a σ-algebra of subsets of X. A (finite) *signed measure* is a function $\mu : \mathcal{S} \to \mathbf{R}$ such that $\mu(\emptyset) = 0$ and $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$ whenever $A_n \in \mathcal{S}$ for every n and $A_m \cap A_n = \emptyset$ for $m \neq n$ (countable additivity).

Notice that the requirement that the measure be countably additive is actually very strong: if we change the order of the sets A_n, their union does not change. This implies that the series in the r.h.s. must be *absolutely convergent.*

Obviously, a signed measure does not enjoy of the monotonicity property: in general $A \subset B$ does not imply $\mu(A) \leq \mu(B)$. Nevertheless, it is easy to verify that the usual properties of continuity of the measure on increasing and decreasing sequences of sets still hold.

DEFINITION (Positive and negative sets): Given a signed measure μ, a measurable set $P \in \mathcal{S}$ is called *positive* if $\mu(E) \geq 0$ for every $E \in \mathcal{S}$, $E \subset P$. Likewise, a measurable set is called *negative* if every measurable subset has measure ≤ 0. A *null set* is a measurable set whose measurable subsets have all measure 0: it is both a positive and a negative set.

THEOREM (Hahn decomposition of a signed measure): Let μ be a signed measure on X (whose σ-algebra of measurable sets is $\mathcal{S}$). Then there exist a positive set $P \in \mathcal{S}$ and a negative set $N \in \mathcal{S}$ such that $P \cap N = \emptyset$ and $P \cup N = X$. Such a decomposition of X is called a Hahn decomposition*: it is unique up to null sets.*

Before we prove the theorem, let us derive one of its most important consequences! Every signed measure is the difference of two finite, positive measures supported in disjoint sets:

DEFINITION (Positive, negative, total variation of a signed measure, Jordan decomposition): Let μ be a signed measure on X, $X = P \cup N$ a Hahn decomposition for μ. For every $E \in \mathcal{S}$ define

$$
\begin{array}{lll}
\mu^+(E) = & \mu(E \cap P) & \textit{(Positive variation of } \mu\textit{)}, \\
\mu^-(E) = & -\mu(E \cap N) & \textit{(Negative variation of } \mu\textit{)}, \\
|\mu|(E) = & \mu^+(E) + \mu^-(E) & \textit{(Total variation of } \mu\textit{)}.
\end{array}
$$

These are obviously positive measures, and $\mu = \mu^+ - \mu^-$ (Jordan decomposition of the measure μ).

The following is a simple characterization of the total variation measure: it is precisely the *smallest* positive measure which is bigger or equal than the modulus of μ.

PROPOSITION: If μ is a signed measure, then for every $A \in \mathcal{S}$ one has

$$|\mu|(A) = \sup\{\sum_{n=1}^{\infty} |\mu(E_n)| : \ E_n \in \mathcal{S}, \ A = \bigcup_{n=1}^{\infty} E_n, \ E_n \cap E_m = \emptyset \ per \ m \neq n\}.$$

PROOF (skipped in class...): Let A, E_n be as above, $X = P \cup N$ a Hahn decomposition for μ, μ^+ e μ^- its variations. Then

$$\begin{aligned}\sum_{n=1}^{\infty} |\mu(E_n)| = \sum_{n=1}^{\infty} |\mu^+(E_n) - \mu^-(E_n)| \leq \\ \sum_{n=1}^{\infty} (\mu^+(E_n) + \mu^-(E_n)) = \sum_{n=1}^{\infty} |\mu|(E_n) = |\mu|(A).\end{aligned}$$

The supremum is actually a maximum: it suffices to decompose A in the two sets$A \cap P$ e $A \cap N$. Q.E.D.

21 Lecture of november 17, 2015 (2 hours)

We are ready to prove the Hahn decomposition.

PROOF of the Theorem on Hahn decomposition: Uniqueness of the Hahn decomposition up to null sets is obvious... much less is existence! We proceed in several steps.

CLAIM I: for every fixed measurable set M we have $\sup\{\mu(E) : \ E \in \mathcal{S}, E \subset M\} < +\infty$

To prove this, suppose by contradiction we have $\sup\{\mu(E) : \ E \in \mathcal{S}, E \subset M\} = +\infty$: we show that there are two disjoint meaurable sets A and B such that $A \cup B = M$, $|\mu(A)| \geq 1$ and $\sup\{\mu(E) : \ E \in \mathcal{S}, E \subset B\} = +\infty$.

Indeed, thanks to our hypothesis that $\sup\{\mu(E) : \ E \in \mathcal{S}, E \subset M\} = +\infty$ we can choose a measurable set B such that $\mu(B) > 1 + |\mu(M)|$, and we set $A = M \setminus B$. Then $\mu(M) = \mu(A) + \mu(B) > \mu(A) + 1 + |\mu(M)|$, whence $\mu(A) < -1$: both A and B have a measure whose modulus is bigger than 1. Now,

$$\begin{aligned}&\sup\{\mu(E) : \ E \in \mathcal{S}, E \subset B\}, \\ &\sup\{\mu(E) : \ E \in \mathcal{S}, E \subset A\}\end{aligned}$$

are certainly not both finite, otherwise the same would be true for the same sup made over all measurable subsets of M: our claim is proved by interchanging the roles of A and B if necessary.

The same procedure is then applied to B, which can be decomposed in two sets with similar properties: iterating this step, we are able to construct two sequence of measurable sets A_n, B_n such that $A_n \cap B_n = \emptyset$, $A_n \cup B_n = B_{n-1}$,$|\mu(A_n)| \geq 1$ and $\sup\{\mu(E) : \; E \in \mathcal{S}, \; E \subset B_n\} = +\infty$. In particular, A_n are pairwise disjoint and the absolute value of their measure is ≥ 1. By countable additivity we have

$$\sum_{n=1}^{\infty} \mu(A_n) = \mu\left(\bigcup_{n=1}^{\infty} A_n\right) \in \mathbf{R},$$

a contradiction because the terms of the series do not converge to 0!

This step is by far the most delicate in the proof of the theorem: we will now be able to obtain our statement pretty quickly.

CLAIM II: for every $A \in \mathcal{S}$ and every $\varepsilon > 0$ we can find $B \in \mathcal{S}$, $B \subset A$ such that $\mu(B) \geq \mu(A)$ and $\mu(E) > -\varepsilon$ for every $E \subset B$, $E \in \mathcal{S}$.

Basically, we claim we can find an "almost positive" subset of A, whose measure is $\geq \mu(A)$...

Let indeed $c = \sup\{\mu(C) : C \subset A, \; C \in \mathcal{S}\}$: obviously $\mu(A) \leq c < +\infty$ (by Claim I), and so we can find a measurable subset $B \subset A$ such that

$$\mu(B) \geq \max\{\mu(A), c - \varepsilon/2\}.$$

This set has the required properties: if we had $E \subset B$ with $\mu(E) \leq -\varepsilon$, then $\mu(B \setminus E) = \mu(B) - \mu(E) \geq c + \varepsilon/2$, against the definition of supremum.

In our third step our "almost positive" set becomes positive:

CLAIM III: if $A \in \mathcal{S}$, there exists a positive set $B \subset A$ such that $\mu(B) \geq \mu(A)$.

Apply indeed Claim II with $\varepsilon = 1/n$: we find a decreasing sequence of measurable sets $A \supset A_1 \supset A_2 \supset A_3 \supset \ldots$ such that $\mu(A_n) \geq \mu(A)$ and $\mu(E) > -1/n$ for every measurable E, $E \subset A_n$.

Define $B = \bigcap_{n=1}^{\infty} A_n$. Then $B \subset A$ and, by continuity of the measure on decreasing sequences, $\mu(B) \geq \mu(A)$. Moreover, B is a positive set: if $E \subset B$, then E is also a subset of A_n for every n and so $\mu(E) > -1/n$.

With Claim III, we are now able to construct our Hahn decomposition: let $s = \sup\{\mu(A) : \; A \in \mathcal{S}\}$. Choose a sequence A_n of measurable sets such that $\mu(A_n) \to s$: by Claim III we can replace each of the sets A_n with a positive subset B_n such that $\mu(B_n) \geq \mu(A_n)$, so that $\mu(B_n) \to s$. We then define

$$P_n = \bigcup_{k=1}^{n} B_k :$$

this is an increasing sequence of positive sets such that $\mu(P_n) \to s$. We then put

$$P = \bigcup_{n=1}^{\infty} P_n, \quad N = X \setminus P.$$

By construction, $\mu(P) = s$. P is also positive: if $E \subset P$, $E \in \mathcal{S}$, then the sets $E \cap P_n$ are an increasing sequence of positive sets whose union is E, so that $\mu(E) \geq 0$. Finally, N is negative: if we had $E \subset N$ measurable such that $\mu(E) > 0$, then $\mu(P \cup E) = \mu(P) + \mu(E) > s$, a contradiction with the definition of s. Q.E.D.

The following is very easy... but also very useful!

EXERCISE (Radon-Nikodym theorem for signed measures): Let μ be a finite positive measure on X, ν a *signed* measure such that $\nu << \mu$. Then there exists a function $v \in L^1(\mu)$ such that

$$\nu(E) = \int_E v(x)\, d\mu(x) \quad \forall E \in \mathcal{S}.$$

(HINT: Let $X = P \cup N$ be a Hahn decomposition. Apply the Radon-Nikodym theorem to the positive measures ν^+ e ν^-, which are concentrated on disjoint sets...)

By means of the Radon-Nikodym theorem for signed measures, we are finally able to prove that the dual space of $L^p(\mu)$ is $L^q(\mu)$, for $1 \leq p < +\infty$ and for a finite positive measure μ[16]:

THEOREM: Let μ be a finite positive measures on X, $1 \leq p < +\infty$. Then for every $T \in (L^p(\mu))'$ there exists a unique function $v \in L^q(\mu)$ (with q the conjugate exponent of p) such that

$$T(u) = \int_X u(x)v(x)\, d\mu(x) \quad \forall u \in L^p(\mu).$$

Moreover, $\|T\| = \|v\|_{L^q}$.

[16]Having proved this, the theorem is easily extended to the case where the measures are σ-finite

PROOF: We already proved that the map $\Phi : L^q \to (L^p)'$ sending every $v \in L^q(\mu)$ into the functional $T_v : u \mapsto \int_X uv \, d\mu$ is a linear isometry. We only have to show that Φ is surjective.

Let then $T \in (L^p(\mu))'$. Define $\nu(E) = T(\mathbf{1}_E)$ (notice that $\mathbf{1}_E \in L^p$ because μ is finite): we claim that ν is a signed measure on X, $\nu << \mu$.

Indeed, we obviously have $\nu(E) = 0$ whenever $\mu(E) = 0$. Moreover, if A and B are measurable and disjoint, then $\mathbf{1}_{A\cup B} = \mathbf{1}_A + \mathbf{1}_B$ whence $\nu(A \cup B) = \nu(A) + \nu(B)$ by the linearity of the functional.

Let's verify that μ is countably additive: let $A = \bigcup_{n=1}^{\infty} A_n$, wher A_n are measurable and pairwise disjoint. One immediately checks that

$$\mathbf{1}_A(x) = \sum_{n=1}^{\infty} \mathbf{1}_{A_n}(x),$$

and the sequence of partial sums is dominated by $\mathbf{1}_A$: the above series converges in $L^p(\mu)$.

Then, by continuity of T we have $\nu(A) = \sum_{n=1}^{\infty} \nu(A_n)$ and ν is indeed a measure.

The Radon-Nikodym theorem gives us a function $v \in L^1(\mu)$ such that

$$(A)\ T(\mathbf{1}_E) = \nu(E) = \int_E v(x) \, d\mu(x) \quad \forall E \in S,$$

whence

$$(B)\ T(s) = \int_X s(x)v(x) \, d\mu(x) \quad \forall s \text{ simple}.$$

We need to prove that $v \in L^q$: if this is true, we can replace the simple function with any $u \in L^p$ because simple functions are dense in this space[17]. Indeed, it is enough to take a sequence s_n of simple function converging to u in L^p: by the continuity of T we have $T(s_n) \to T(u)$, on the other hand by Hölder we have $\int\limits_X s_n(x)v(x) \, d\mu(x) \to \int\limits_X u(x)v(x) \, d\mu(x)$.

We now check that actually $v \in L^q(\mu)$, thus concluding the proof. Let us begin with the case $p = 1$. We know from (A) that

$$|\int_E v(x) \, d\mu(x)| \leq \|T\|\mu(E) \quad \forall E \in \mathcal{S},$$

whence $\mu(\{x : \ v(x) > \|T\| + 1/n\}) = 0$ for every n (otherwise the inequality would fail), and similarly $\mu(\{x : \ v(x) < -\|T\| - 1/n\}) = 0$ whence $\|v\|_\infty \leq \|T\|$.

[17] Every bounded function can be approached uniformly with simple functions. Moreover, every function $u \in L^p$ can be approached in L^p with a sequence of bounded function (take for instance $u_n(x) = \max\{-n, \min\{u(x), n\}\}$.

In the case $1 < p < +\infty$, (B) holds for every $s \in L^\infty(\mu)$ (because, as we said above, every bounded function can be approximated uniformly with simple functions). For $n \in \mathbf{N}$ let $E_n = \{x \in X : |v(x)| \leq n\}$ and define $s_n(x) = \mathbf{1}_{E_n}(x)|v(x)|^{q-1}\text{sgn}(v(x))$. These functions are in $L^\infty(\mu)$, and $|v(x)|^q = |s_n(x)|^p$ on E_n. We then get from (B)

$$\int_{E_n} |v(x)|^q \, d\mu(x) = \int_X s_n(x)v(x) \, d\mu(x) = T(s_n) \leq \|T\| \left(\int_{E_n} |v(x)|^q \, d\mu(x) \right)^{1/p},$$

i.e.

$$\left(\int_{E_n} |v(x)|^q \, d\mu(x) \right)^{1/q} \leq \|T\|.$$

Passing to the limit as $n \to +\infty$ and using the monotone convergence theorem we get $\|v\|_{L^q} \leq \|T\|$. Q.E.D.

22 Lecture of november 18, 2015 (2 hours)

We now prove a small but important result we will need in the following:

LEMMA (fundamental lemma of the Calculus of Variations): Let $w \in L^1([a, b])$ be such that $\int_a^b w\phi \, dx = 0$ for every $\phi \in \mathcal{C}^1_C([a, b])$. Then $w = 0$ a.e.

PROOF: Approximate the function sgn $w(x)$ with a sequence ϕ_n of function in $\mathcal{C}^1_C$, taking values in the inverval $[-1, 1]$ and converging a.e. By the dominated convergence theorem we get then

$$0 = \int_a^b w\phi_n \, dx \to \int_a^b |w(x)| \, dx,$$

whence $w = 0$ a.e. Q.E.D.

To study problems involving differential equations (both O.D.E.s and P.D.E.s), we need spaces of functions which are *differentiable* (in some appropriate sense), and which have good compactness properties.

The *Sobolev spaces* $W^{1,p}([a, b])$ are a family of spaces modelled on L^p which fulfill perfectly both requirements. Before we give the definition, we need the following important notion:

DEFINITION (Weak derivative): Let $u \in L^1([a, b])$. A function $v \in L^1([a, b])$ is a *weak derivative* of u if

$$\int_a^b u(x)\phi'(x) \, dx = -\int_a^b v(x)\phi(x) \, dx \quad \forall \phi \in \mathcal{C}^1_0([a, b]).$$

Notice that if $u \in \mathcal{C}^1$, then its derivative is also a weak derivative: this is just the integration by parts formula!

Converserly, we will see that if u and v are continuous functions, then u is differentiable with derivative v (this is the classical du Bois-Reymond lemma in the Calculus of Variations).

Moreover, it is easy to see that the weak derivative is unique whenever it exists: if v and $\tilde{v}$ are weak derivatives of u, it follows from the definition that $w = v - \tilde{v}$ satisfies the hypothesis of the fundamental lemma of the calculus of variations. It follows that $v = \tilde{v}$ a.e.

By the uniqueness of weak derivative, we are allowed to denote it by u'.

DEFINITION (Sobolev spaces $W^{1,p}$): If $1 \leq p \leq +\infty$ we define the Sobolev space

$$W^{1,p}([a,b]) = \{u \in L^p([a,b]) : \textit{ there exists } u' \in L^p([a,b]) \textit{ weak derivative of } u\}.$$

On the space $W^{1,p}$ one usually puts one of the following two equivalent norms:

$$\|u\|_{W^{1,p}} = \|u\|_{L^p} + \|u'\|_{L^p} \textit{ or } \|u\|_{W^{1,p}} = (\|u\|_{L^p}^p + \|u'\|_{L^p}^p)^{1/p}.$$

We will use indifferently the first or the second. The second is more appropriate in case $p = 2$, because it is induced by a scalar product, thus making $W^{1,2}$ a Hilbert space.

REMARK: It is easy to verify that $W^{1,p}$ is a Banach space: if $\{u_k\} \subset W^{1,p}$ is a Cauchy sequence, then both $\{u_k\}$ and $\{u_k'\}$ are Cauchy sequences in L^p. As L^p is complete, we get u, v such that $u_k \to u$, $u_k' \to v$ in L^p.

We have to show that $u \in W^{1,p}$ and $v = u'$. Indeed, by definition of weak derivative we have

$$\int_a^b u_k \phi' \, dx = -\int_a^b u_k' \phi \, dx \quad \forall \phi \in \mathcal{C}_0^1.$$

Passing to the limit as $k \to +\infty$ we get

$$\int_a^b u \phi' \, dx = -\int_a^b v \phi \, dx,$$

as we wanted. Notice that the same argument works also if we only have $u_k \rightharpoonup u$, $u_k' \rightharpoonup v$ in L^p: we will use this remark later on.

The notion of Sobolev spaces is easily extended to functions defined on $\mathbf{R}^n$: if $\Omega \subset \mathbf{R}^n$ is open, $u \in L^1(\Omega)$, we say that $v \in L^1(\Omega)$ is the weak derivative of u with respect to x_i if and only if

$$\int_\Omega u(x)\frac{\partial\phi(x)}{\partial x_i}\,dx = -\int_\Omega v(x)\phi(x)\,dx \qquad \forall\phi\in C_0^1(\Omega).$$

We denote the weak derivative (if any) by $\frac{\partial u}{\partial x_i}$ (there is uniqueness of the weak derivative as in the 1-dimensional case). The Sobolev spaces are then defined in the obvious way:

$$W^{1,p}(\Omega) = \{u\in L^p(\Omega):\ \exists \text{ weak derivatives } \frac{\partial u}{\partial x_i}\in L^p(\Omega),\ i=1,\dots,n\}.$$

These are Banach spaces with the norm

$$\|u\|_{W^{1,p}(\Omega)} = \|u\|_{L^p} + \sum_{i=1}^n \|\frac{\partial u}{\partial x_i}\|_{L^p}.$$ [18]

But in dimension 1 Sobolev functions are much better than in higher dimension: they are continuous (after possibly changing them on a set of measure 0) and their weak deriviative coincides almost everywhere with their classic derivative.

We need the following definition:

DEFINITION (Absolutely continuous functions): The space of absolutely continuous functions is the set of all primitives of L^1 functions:

$$AC([a,b]) = \{u:[a,b]\to\mathbf{R}:\ \exists v\in L^1([a,b])\ s.t.\ u(x)=u(a)+\int_a^x v(t)\,dt\,\forall x\in[a,b]\}.$$

A theorem in real analysis which is not simple (nor overly difficult, to be fair...) tells us that AC functions are differentiable a.e., and $u'(x)=v(x)$ for a.e. x: so the fundamental theorem of calculus holds, in the sense that u is a primitive of its derivative.

Moreover, the following ε-δ characterization of AC functions holds:
A function $u:[a,b]\to\mathbf{R}$ is absolutely continuous if and only if for every $\varepsilon>0$ there exists $\delta>0$ such that, for every finite collection $[a_1,b_1]$,

[18] Or with the equivalent norm

$$\|u\|_{W^{1,p}(\Omega)} = \left(\|u\|_{L^p}^p + \sum_{i=1}^n \|\frac{\partial u}{\partial x_i}\|_{L^p}^p\right)^{1/p},$$

which is a Hilbert norm for $p=2$.

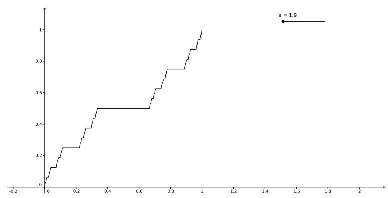
$[a_2, b_2]$,...,$[a_N, b_N]$ *of pairwise disjoint subintervals of* $[a, b]$ *with* $\sum_{i=1}^{N}(b_i - a_i) < \delta$, *one has*

$$\sum_{i=1}^{N} |u(b_i) - u(a_i)| < \varepsilon.$$

This characterization shows that AC function satisfy a property which is a stronger version of uniform continuity: an absolutely continuous function is in particular uniformly continuous. Moreover, it is an easy exercise to check from the characterization that the *product* of two AC functions is still in AC.

Due to time constraints, we will omit the proof of the characterization, and also of the differentiability a.e. of AC functions.

REMARK: From the fact that $u : [a, b] \to \mathbf{R}$ is *continuous* and *differentiable a.e.* we cannot conclude that $u \in AC([a, b])$. A famous counterexample is the so-called *Cantor's staircase*, a function which is continuous and increasing in the interval $[0, 1]$, whose image is the whole interval $[0, 1]$... and whose derivative is 0 almost everywhere. Obviously, such a function is *not* a primitive of its derivative! Here is how Cantor's staircase looks like (click on the image to see an animation, with a zoom on a portion of the graph to see the finer structure of the function):



We will now prove that the space of absolutely continuous functions *coincides*, with the Sobolev space $W^{1,1}([a, b])$. Precisely, each absolutely continuous functions belongs to the Sobolev space and, conversely, given $u \in W^{1,1}$ there exists an absolutely continuous function which coincide with u *almost everywhere.*

We need two lemmas:

LEMMA 1 (du Bois-Reymond): If the weak derivative of $u \in W^{1,1}([a, b])$ *is* 0, *then* u *is a.e. equal to a constant.*

PROOF: Let $\psi \in \mathcal{C}^0([a, b])$: define $w(x) = \psi(x) - \frac{1}{b-a}\int_a^b \psi(t)\, dt$ and

$$\phi(x) = \int_a^x w(t)\, dt.$$

Then $\phi \in \mathcal{C}_0^1([a,b])$ and by our hypothesis on u we get

$$0 = \int_a^b u(x)\phi'(x)\, dx =$$
$$\int_a^b [u(x)\psi(x) - u(x)\frac{1}{b-a}\int_a^b \psi(s)\, ds]\, dx =$$
$$\int_a^b [u(x) - \frac{1}{b-a}\int_a^b u(s)\, ds]\psi(x)\, dx.$$

By the fundamental lemma of the Calculus of Variations, this implies

$$u(x) = \frac{1}{b-a}\int_a^b u(s)\, ds \quad per\ q.o\ x \in [a,b].$$

Q.E.D.

23 Lecture of november 24, 2015 (2 hours)

LEMMA 2: If $u \in AC([a,b])$, then $u \in W^{1,1}([a,b])$. Moreover, the pointwise derivative of u (which is defined a.e.) is also the weak derivative of u.

PROOF: From the definition of AC we know that u and u' are both in L^1. Let then $\phi \in \mathcal{C}_0^1([a,b])$: obviously, $\phi \in AC$.

Then the product $u\phi$ is also absolutely continuous and we have $(u\phi)' = u'\phi + u\phi'$ a.e. By integrating we get

$$0 = \int_a^b (u\phi)'\, dx = \int_a^b (u'\phi + u\phi')\, dx,$$

and the weak derivative of u is u'. Q.E.D.

We finally prove that $W^{1,1}$ is essentially the same as AC:

THEOREM: Let $u \in W^{1,1}([a,b])$. Then there exists $\tilde{u} \in AC([a,b])$ such that $u(x) = \tilde{u}(x)$ for a.e. x. So, after possibly changing u in a set of measure 0, the weak derivative of u coincides with its classical derivative[19].

PROOF: Define $w(x) = \int_a^x u'(t)\, dt$. This is an absolutely continuous function which, by LEMMA 2, belongs to $W^{1,1}$ and whose weak derivative is u'. Then the weak derivative of the function $u - w$ is 0, whence, by LEMMA 1,

[19] Recall the statement of LEMMA 2.

$u(x) - w(x) = c$ a.e., with c a constant. We can then define $\tilde{u}(x) = c + w(x)$. Q.E.D.

REMARK: If $1 < p \leq +\infty$, the functions in $W^{1,p}$ are also in $W^{1,1}$. Then if $u \in W^{1,p}$ there is no loss of generality in supposing that u is absolutely continuous: it is enough to choose the appropriate element in the equivalence class of u in L^p. In this sense, the space L^p is the space of those AC functions, whose derivatives belong to L^p. These functions, besides being AC, are also Hölder continuous with exponent $1 - 1/p$: if $x, y \in [a, b]$ we have by Hölder inequality

$$|u(x) - u(y)| = |\int_x^y u'(t)\, dt| \leq \|u'\|_{L^p} |x - y|^{1-1/p}.$$

Functions in $W^{1,\infty}$ are lipschitz continuous: if we always choose the AC representative in each equivalence class, we can say that $W^{1,\infty}$ *coincides* with the space of lipschitz continuous functions. Indeed, every lipschitz continuous function is absolutely continuous, and if u has Lipschitz constant L, then its incremental quotients are bounded by L... and so $|u'(x)| \leq L$in all differentiability points: the derivative of a Lipschitz continuous function is in L^∞.

The remark we just made is key for the following important compactness result:

THEOREM (weak compactness in $W^{1,p}$): Let $\{u_n\} \subset W^{1,p}([a, b])$, $1 < p < +\infty$ (and suppose we have chosen the AC representative of each u_n). If there exists a constant $C > 0$ such that $\|u_n'\|_{L^p} \leq C$ for every n, and one of the two following conditions holds:

(i) $|u_n(a)| \leq C$

(ii) $\|u_n\|_{L^p} \leq C$

then there exists $u \in W^{1,p}$ and a subsequence $\{u_{n_k}\}$ such that $u_{n_k} \to u$ uniformly, $u_{n_k}' \rightharpoonup u'$ weakly in L^p as $k \to +\infty$.

PROOF: By previous remark and the equiboundedness of the derivatives in L^p, all our function satisfy the following Hölder continuity estimate:

$$(*) |u_n(x) - u_n(y)| \leq C|x - y|^{1-1/p} \quad \forall x, y \in [a, b].$$

In particular, the functions u_n are equicontinuous.

Suppose now (i) holds: by using $(*)$ we have for every x and n

$$|u_n(x)| \leq |u_n(a)| + |u_n(x) - u_n(a)| \leq C + C(b-a)^{1-1/p}$$

and the functions u_n are also equibounded.

By the Ascoli-Arzelà theorem, and weak compactness in the reflexive space L^p, we find $u \in C^0$, $v \in L^p$ and a subsequence $\{u_{n_k}\}$ such that $u_{n_k} \to u$ uniformly, $u'_{n_k} \rightharpoonup v$ weakly in L^p. As we remarked earlier (in proving the completeness of the spaces $W^{1,p}$), this implies that $u \in W^{1,p}$ and $v = u'$.

We still have to prove that the same holds when we replace (i) with (ii). But we actually have $(ii) \Rightarrow (i)$: indeed we have, for every $x \in [a,b]$, $u(a) = u(x) - \int_a^x u'(x)\,dx$, so that $|u(a)| \leq |u(x)| + \int_a^b |u'(x)|\,dx$. Integrating both sides we get:

$$|u(a)| \leq \frac{1}{b-a}\int_a^b |u(x)|\,dx + \int_a^b |u'(x)|\,dx,$$

and we conclude by using Hölder's inequality. Q.E.D.

REMARK: A stronger result holds for $p = +\infty$: if $\{u_n\} \subset W^{1,\infty}$ we can use the theorem for every finite p to find a convergent subsequence. Moreover, this sequence is equilipschitz (because derivatives are equibounded in L^∞), so the limit is also lipschitz continuous. Derivatives converge weakly in L^p for finite p, but also weakly* in L^∞[20].

On the other hand, the compactness theorem is false for $p = 1$: it is easy to construct a sequence of functions which is bounded in the $W^{1,1}$ norm, which converges to a *discontinuous* function: for instance, take the following functions on $[-1,1]$:

$$u_n(x) = \begin{cases} -1 & se \ -1 \leq x \leq -1/n \\ nx & se \ -1/n < x < 1/n \\ 1 & se \ 1/n \leq x \leq 1 \end{cases}$$

COROLLARY (weak convergence in $W^{1,p}$): Let $1 < p < +\infty$, $\{u_n\} \subset W^{1,p}([a,b])$ be a sequence such that $u_n \rightharpoonup u$, $u'_n \rightharpoonup u'$ weakly in L^p, with $u \in W^{1,p}$. Then u_n converges uniformly to u.

[20] Weak* convergence can be defined in a space which is the *dual* X' of a Banach space X: instead of testing the weak convergence of a sequence in X' on *every* linear functional $S \in X''$, we only test on the elements of $J(X)$. In other words, $T_k \rightharpoonup^* T$ weakly* in X' if and only if $T_k(x) \to T(x)$ for every $x \in X$: in particular, $u_n \rightharpoonup^* u$ weakly* in L^∞ if and only if $\int_a^b u_n v\,dx \to \int_a^b uv\,dx$ for every $v \in L^1$. There is also a compactness result for the weak* convergence: if X is *separable*, then bounded sequences in X' are weakly* compact.

PROOF: We know that sequences converging weakly in L^p are norm-bounded.

Let now $\{u_{n_k}\}$ be a fixed subsequence of $\{u_n\}$: this is equibuonded in the $W^{1,p}$ norm, and by the compactness theorem we can extract a further subsequence $\{u_{n_{k_h}}\}$ converging uniformly to a continuous function (with weakly convergent derivatives). Now, the limit is necessarily u by uniqueness of the weak limit.

Since the subsequence we started with is arbitrary, it follows that *the whole sequence* converges uniformly to u[21]. Q.E.D.

EXAMPLE: The situation is not so good in dimension n. A function $u \in W^{1,p}(\Omega)$, with Ω a open subset of $\mathbf{R}^n$, is in general not even continuous.

For instance, consider the following function on the unit open ball of $\mathbf{R}^2$:

$$u(x,y) = \frac{1}{(x^2+y^2)^{1/4}}.$$

This function is clearly discontinuous at $(0,0)$, but we will see in a moment that $u \in W^{1,p}(B_1((0,0))$ for $1 \leq p < 4/3$. By using slightly more sophisticated examples, one can show that there are discontinuos Sobolev functions in $W^{1,p}$ for every $1 \leq p \leq n$, where n is the dimension of the ambient euclidean space.

Passing to polar coordinates (ρ, θ) and integrating, we immediately see that $u \in L^p(B_1(0))$ for $p < 4$. Moreover, we have

$$|\nabla u(x,y)| = \frac{1}{2}\rho^{-3/2}.$$

By integrating over the unit ball, one sees that this function is in $L^p(B_1(0))$ for $p < 4/3$.

To show tha $u \in W^{1,p}$, we need to verify that the pointwise derivatives of u are also its weak derivatives. But this is easily proved by approximating u with the C^1 functions

$$u_n(x,y) = \frac{1}{(x^2+y^2+\frac{1}{n})^{1/4}}.$$

Indeed, this sequence of regular functions converges to u in the $W^{1,p}$ norm, and so $u \in W^{1,p}$.

24 Lecture of november 25, 2015 (2 hours)

Luckily for us, even if Sobolev functions are not necessarily continuous, there are important tools that allow for their easy handling of Sobolev.

[21]This is a general fact, valid in any metric space

For instance, one can prove that smooth functions are dense in $W^{1,p}$:

THEOREM (Meyers-Serrin): for any domain $\Omega \subset \mathbf{R}^n$ and for every $1 \leq p \leq +\infty$, the space $C^\infty(\Omega) \cap W^{1,p}(\Omega)$ is dense in $W^{1,p}(\Omega)$ with respect to the Sobolev norm.

We will omit the proof of this result, which is rather technical. In the special case, $\Omega = \mathbf{R}^n$, the proof is easy enough and is obtained simply by regularizing by convolution.

Notice that the theorem says that a function in $W^{1,p}(\Omega)$ can be approximated in norm with functions in $C^\infty(\Omega)$, which in general are not continuous up to the boundary.

Another important fact is that Sobolev functions have typically a higher summability than L^p. To understand this, we begin by introducing some important subspaces:

DEFINITION: Let $\Omega \subset \mathbf{R}^n$ be a domain, $1 \leq p \leq +\infty$. The space $W_0^{1,p}(\Omega)$ is the closure in the $W^{1,p}$ norm of $C^1_C(\Omega)$.

Morally, $W_0^{1,p}$ is the subspace of those Sobolev functions which are "zero at the boundary", but this statement would make no immediate sense, since for regular domains Ω, the boundary $\partial\Omega$ is a set of measure 0.

One can prove that $W_0^{1,p}(\mathbf{R}^n) = W^{1,p}(\mathbf{R}^n)$, but this is not the case for bounded domains Ω. Indeed, the following important result holds:

THEOREM (Sobolev embedding Theorem, first version): Suppose $1 \leq p < n$ (where n is the dimension of the ambient space). There is a constant $C > 0$, depending on p but not on u nor Ω, such that

$$(**) \qquad \|u\|_{L^{p^*}(\Omega)} \leq C\|\nabla u\|_{L^p(\Omega)} \qquad \forall u \in W_0^{1,p}(\Omega),$$

where $p^* = \frac{np}{n-p}$ is called Sobolev exponent.

This theorem is a sort of *regularity theorem* ensuring that functions in $W_0^{1,p}$ have a *higher summability* than L^p (indeed, $p^* > p$).

The Sobolev exponent may seem mysterious, but it is easy to realize it is the *unique* for which inequality $(**)$ can be true for every $u \in C^1_c(\mathbf{R}^n)$. Precisely, suppose we have

$$\|u\|_{L^q(\mathbf{R}^n)} \leq C\|\nabla u\|_{L^p(\mathbf{R}^n)} \qquad \forall u \in C^1_c(\mathbf{R}^n).$$

Then necessarily $q = p^*$.

To show this, fix $u \in C^1_c(\mathbf{R}^n)$, $u \neq 0$. The inequality must hold, with the same constant, also for every function of the form $v(x) = u(rx)$ with $r > 0$. A simple change of variables lead to

$$\|u\|_{L^q(\mathbf{R}^n)} \leq C r^{1+\frac{n}{q}-\frac{n}{p}} \|\nabla u\|_{L^p(\mathbf{R}^n)}.$$

The exponent of r in the last inequality must be 0: otherwise we have a contradiction, because by letting $r \to 0$ or $r \to +\infty$, we obtain that the right hand side converges to 0, while the left hand side is strictly positive. The exponent is 0 exactly when $q = p^*$.

To prove the Sobolev embedding theorem, we use the following lemma:

LEMMA (Gagliardo): Let $f_1, \ldots, f_n : \mathbf{R}^{n-1} \to \mathbf{R}$ be non negative measurable functions. Then the following inequality holds:

$$\int_{\mathbf{R}^n} \prod_{i=1}^{n} f_i(\hat{x}_i)\, dx \leq \prod_{i=1}^{n} \left(\int_{\mathbf{R}^{n-1}} f_i^{n-1}(\hat{x}_i)\, d\hat{x}_i \right)^{\frac{1}{n-1}}.$$

PROOF (Not seen in class): The proof goes by induction over n: for $n = 2$ it is Fubini's Theorem. The inductive step is a bit technical but not difficult: one has to use Hölder's inequality and its straightforward generalization saying that

$$\int f_1 \cdot f_s \cdot \ldots \cdot f_n \, d\mu \leq \|f_1\|_{L^n} \cdot \|f_2\|_{L^n} \cdot \ldots \cdot \|f_n\|_{L^n}.$$

Suppose the inequality is truee for n, let us prove it for $n+1$:

$$\int_{\mathbf{R}^{n+1}} \prod_{i=1}^{n+1} f_i(\hat{x}_i)\, dx = \int_{\mathbf{R}^n} f_{n+1}(\hat{x}_{n+1})\, d\hat{x}_{n+1} \int_{\mathbf{R}} \prod_{i=1}^{n} f_i(\hat{x}_i)\, d\hat{x}_{n+1} \leq$$

$$\int_{\mathbf{R}^n} f_{n+1}(\hat{x}_{n+1}) \prod_{i=1}^{n} \left(\int_{\mathbf{R}} f_i^n(\hat{x}_i)\, dx_{n+1} \right)^{1/n} d\hat{x}_{n+1} \leq$$

$$\left(\int_{\mathbf{R}^n} f_{n+1}^n(\hat{x}_{n+1})\, d\hat{x}_{n+1} \right)^{1/n} \left(\int_{\mathbf{R}^n} \left(\prod_{i=1}^{n} \int_{\mathbf{R}} f_i^n(\hat{x}_i)\, dx_{n+1} \right)^{\frac{1}{n-1}} d\hat{x}_{n+1} \right)^{\frac{n-1}{n}} \leq$$

$$\prod_{i=1}^{n+1} \left(\int_{\mathbf{R}^{n+1}} f_i(\hat{x}_i)\, d\hat{x}_i \right)^{1/n}$$

where the last inequality comes from the inductive hypothesis. Q.E.D.

We can now prove the embedding theorem:

PROOF OF SOBOLEV'S EMBEDDING FOR $p < n$: By definition of the space $W_0^{1,p}(\Omega)$, it is clearly enough to prove the result for functions $u \in C_C^1(\mathbf{R}^n)$. We begin by proving the inequality for $p = 1$ (in which case we have $1^* = \frac{n}{n-1}$).

Now, for every $x \in \mathbf{R}^n$ and $i = 1, \ldots, n$ we have

$$u(x) = \int_{-\infty}^{x_i} \frac{\partial u}{\partial x_i} \, dx_i.$$

Thus

$$|u(x)| \leq \int_{\mathbf{R}} \left| \frac{\partial u}{\partial x_i} \right| \, dx_i \leq \int_{\mathbf{R}} |\nabla u| \, dx_i,$$

and

$$\int_{\mathbf{R}^n} |u(x)|^{\frac{n}{n-1}} \, dx \leq \int_{\mathbf{R}^n} \prod_{i=1}^{n} \left(\int_{\mathbf{R}} |\nabla u(x)| \, dx_i \right)^{\frac{1}{n-1}} \, dx.$$

In the productory, each of the expressions in brakets is independent on x_i thanks to the integration: we can apply Gagliardo's Lemma and obtain

$$\int_{\mathbf{R}^n} |u(x)|^{\frac{n}{n-1}} \, dx \leq \prod_{i=1}^{n} \left(\int_{R^{n-1}} d\hat{x}_i \int_{\mathbf{R}} |\nabla u(x)| \, dx_i \right)^{\frac{1}{n-1}} = \left(\int_{\mathbf{R}^n} |\nabla u| \, dx \right)^{\frac{n}{n-1}},$$

which is the result for $p = 1$.

Let then $1 < p < n$. Apply the inequality with $p = 1$ to the function $v(x) = |u(x)|^{1+r}$, where $r > 0$ is to be chosen later. We have $|\nabla v(x)| = (r+1)|u(x)|^r |\nabla u(x)|$. Then, by using Hölder's inequality we obtain:

$$\left(\int_{\mathbf{R}^n} |u(x)|^{\frac{(r+1)n}{n-1}} \, dx \right)^{\frac{n-1}{n}} \leq (r+1) \int_{\mathbf{R}^n} |u(x)|^r |\nabla u(x)| \, dx \leq$$

$$(r+1) \|\nabla u\|_{L^p} \left(\int_{\mathbf{R}^n} |u(x)|^{\frac{pr}{p-1}} \, dx \right)^{\frac{p-1}{p}}.$$

We can now choose r in such a way that the exponents of $|u(x)|$ on both sides of the inequality are equal: with easy computations, we find Sobolev's inequality. Q.E.D.

If $p = n$ and the domain Ω is bounded, we can apply the theorem to all smaller exponents and we find that a function $u \in W_0^{1,p}(\Omega)$ belongs to $L^q(\Omega)$ for every $q < +\infty$. There are examples showing that such a function is not necessarily in L^∞.

For $p > n$ things are even better: functions in $W_0^{1,p}(\Omega)$ are Hölder continuous.

THEOREM (Morrey): If $p > n$, there is a constant $C > 0$, depending only on p and n, such that

$$[u]_\alpha \leq \|\nabla u\|_{L^p(\Omega)} \qquad \forall u \in W_0^{1,p}(\Omega),$$

where $\alpha = 1 - n/p$ and

$$[u]_\alpha = \sup\left\{ \frac{|u(x) - u(y)|}{|x - y|^\alpha} \;:\; x, y \in \Omega, x \neq y \right\}.$$

This is understood in the sense that in the equivalence class of u in $L^p(\Omega)$, there is a Hölder continuous function satisfying the estimate.

DIM.: As before, it is enough to prove the estimate for functions $u \in C_C^1(\mathbf{R}^n)$: indeed, every function in $W_0^{1,p}(\Omega)$ can be approximated in the $W^{1,p}$ norm with functions of this kind, and by passing to subsequences we can have convergence a.e. We then pass to the limit in the estimate.

Let then $u \in C_0^1(\Omega)$, and let $x, y \in \mathbf{R}^n$ be distinct points. Let $\delta = |x - y|$ and define $S = B_\delta(x) \cap B_\delta(y)$. For every $z \in S$ one has $|u(x) - u(y)| \leq |u(x) - u(z)| + |u(z) - u(y)|$. Integrate both sides over S with respect to z: we get

$$(*) \qquad |S||u(x) - u(y)| \leq \int_S |u(z) - u(x)|\, dz + \int_S |u(z) - u(y)|\, dz.$$

We clearly have $|S| = \kappa\delta^n$, with κ a constant independent on δ. Let us estimate the first integal on the right hand side: for the other, the computation is similar. We have

$$|u(z) - u(x)| = \left| \int_0^1 \frac{d}{dt} u(x + t(z - x))\, dt \right| =$$
$$\left| \int_0^1 \nabla u(x + t(z - x)) \cdot (z - x)\, dt \right| \leq \delta \int_0^1 |\nabla u(x + t(z - x)|\, dt,$$

and, by putting $w = x + t(z - x)$ and using Hölder:

$$\int_S |u(z) - u(x)|\, dz \leq \delta \int_0^1 dt \int_S |\nabla u(x + t(z - x)|\, dz \leq$$
$$\delta \int_0^1 t^{-n}\, dt \int_{B_{t\delta}(x)} |\nabla u(w)|\, dw \leq \delta \int_0^1 t^{-n} (\omega_n t^n \delta^n)^{1-1/p}\, dt \|\nabla u\|_{L^p(\mathbf{R}^n)}.$$

Here, ω_n is the measure of the unit ball in $\mathbf{R}^n$. The integral in t is finite (the exponent is $-n/p > -1$).

Inequality $(*)$ then becomes

$$\kappa\delta^n|u(x) - u(y)| \leq C\delta^{1+n-n/p}\|\nabla u\|_{L^p(\mathbf{R}^n)},$$

as we wanted. Q.E.D.

A natural question is now whether or not Sobolev's embedding hold in the space $W^{1,p}(\Omega)$ (and not only in the subspace $W_0^{1,p}(\Omega)$). The answer is affirmative if the domain Ω is regular enough (otherwise there are counterexamples).

Suppose for simplicity Ω is a domain of class C^∞. Then the following holds:

THEOREM (extension of Sobolev functions): Let Ω be a bounded domain of class C^∞, Ω' another domain with $\Omega \subset\subset \Omega'$. Then there exists $C > 0$, depending only on p, n, Ω and Ω', such that every $u \in W^{1,p}(\Omega)$ has an extension $\tilde{u} \in W_0^{1,p}(\Omega')$ such that

$$\|\tilde{u}\|_{W^{1,p}(\Omega')} \leq C\|u\|_{W^{1,p}(\Omega)}.$$

We omit the proof, which is not really complicated. One needs a localization argument which uses a partition of 1 and an extension by reflexion on the local charts.

If we apply the theorems by Sobolev and Morrey to the extended function $\tilde{u}$, we easily obtain:

THEOREM (Sobolev-Morrey embedding in $W^{1,p}(\Omega)$): Let Ω be a regular, bounded domain as in the extension theorem. If $1 \leq p < n$, there exists a constant $C > 0$ (depending only on p, n and Ω) such that

$$\|u\|_{L^{p^*}(\Omega)} \leq C\|u\|_{W^{1,p}(\Omega)} \qquad \forall u \in W^{1,p}(\Omega).$$

If $p > n$, every function in $u \in W^{1,p}(\Omega)$ is Hölder continuous with exponent $\alpha = 1 - n/p$ and there exists $C > 0$ (depending only on p, n and Ω) such that

$$[u]_\alpha \leq C\|u\|_{W^{1,p}(\Omega)} \qquad \forall u \in W^{1,p}(\Omega).$$

An important compactness result you'll see better with Giandomenico is:

THEOREM (Rellich): Let Ω be a bounded, regular domain, $q < p^*$. Every bounded sequence in $W^{1,p}(\Omega)$ has a subsequence which converges strongly in $L^q(\Omega)$.