Notes on Measure Theory, MA 813, Autumn 2008

Gopala Krishna Srinivasan

Zu meiner Frau gewidmet

Historical sign posts in the development of 19th and 20th century analysis

We attempt here a thumb-nail sketch of the evolution of the theory of functions of a real variable. While the theory of functions of a complex variable advanced rapidly in the nineteenth century, progress in the theory of functions of a real variable has been relatively slow. For example the great scholar Ampère in 1806 tried to establish that an "arbitrary" function of a real variable had a derivative everywhere save along a set of "exceptional points". Needless to say his attempt failed and so did those of others in the decades to follow until K. Weierstrass (1861) put an end to these by exhibiting a continuous nowhere differentiable function. A few years earlier Riemann stated that the sum of the following series

$$\sum_{n=1}^{\infty} \frac{\sin n^2 x}{n^2}$$

is not a differentiable function though unveiling the complete truth about this function took more than a century [5] in the works of G. H. Hardy (1916) and later by Gerver (1970). The function is differentiable precisely at points $p\pi/q$ where p and q are odd integers.

Power series representations dominated eighteenth century analysis and well into the nineteenth century. A function was understood to mean piecewise analytical expressions for which differentiability "almost everywhere" is well neigh obvious and so it is not clear what Ampère was really trying to prove. In the year 1822 the revolutionary work of Joseph Fourier appeared in which Fourier attempts at expressing "arbitrary" functions in terms of series of sines and cosines. Although his treatment was for the most part formal aimed at applications to mathematical physics, the work stirred considerable interest among mathematicians in the decades to follow. Most notable among them being the works of Dirichlet and Riemann. The first serious attempt at developing a rigorous theory was due to Dirichlet [3]. In this connection work of Schlömilch [10] ought to be mentioned which appeared four years before Riemann's seminal paper.

Riemann begins with a review of the works of his predecessors and proceeds to develop the theory further for which purpose he establishes a rigorous theory of the integral which is the theory taught in the undergraduate curricula. The theory of Fourier series now became a part of analysis and was soon to be at the service of other parts of mathematics such as number theory[1]. By this time the notion of a function had already begun widening and examples like the one given by Dirichlet

$$f(x) = \begin{cases} x & x \in \mathbb{Q} \\ 1 - x & x \in \mathbb{R} - \mathbb{Q} \end{cases}$$

became widely known though many decades passed before they were generally accepted and viewed without skepticism by the mathematical community. Indeed as late as 1893, C. Hermite

[1] Five years down the road Riemann used it to give one of the proofs of his famous functional equation

in a letter to Stieltjes, writes in connection with certain series expansions with non-smooth sums [6] (Volume 2, p. 318):

“ *Mais cet développements, si élégantes, sont frappés de malédiction ; leurs dérivées d'ordre* $2m+1$ *et* $2m+2$ *sont des séries qui n'ont acune sens. L'Analyse retire d'une main ce qu'elle donne de l'autre. Je me détourne avec effroi et horreur de cette plaie lamentable des fonctions continues qui n'ont point de dérivées...* ”

It became gradually clear in the later part of the nineteenth century that the resolution of several questions in the theory of Fourier series present formidable difficulties and as such fall outside the scope of the Riemannian theory. In the course of investigation of these, set theory and point-set topology developed. The notion of measure also underwent a metamorphosis over a period starting with contributions from Stolz (1884), Cantor, Harnack and finally E. Borel who was the first to employ infinite coverings of the set in order to define measure. Notable in this transition are the works of E. Borel and René Baire in the year 1898 which are precursors to the ultimate establishment of Lebesgue's theory. Interestingly in 1898 Lebesgue published a paper on a new proof of Weierstrass's approximation theorem which can easily be considered as the best among all known proofs. The ideas centering around these nascent theories of point set topology and measure theory were subjected to deeper analysis by the “Russian school” led Luzin and his students[2]. The work of Henri Lebesgue (1902) is a denouement of the efforts of many mathematicians over a century and elegantly resolves many important issues in classical analysis. Lebesgue authored a series of six papers in 1901 and his thesis appeared in 1902. To quote J. C. Burkill [1] of this great work:

“One of the finest (thesis) which a mathematician has ever written”

It was only slowly that Lebesgues ideas were assimilated in the main stream of Analysis. On this point the interesting article of J. L. Doob [4] may be consulted. Lebesgue's ideas were further generalized in the works of Radon who carved an abstract theory of measure. The appearance of A. N. Kolmogorov's foundations of probability theory and Tate's thesis in number theory brought measure theory and Fourier analysis as focal points in the main-stream of mathematics. After the establishment of Laurant Schwartz's theory of distributions, their more exotic cousins known as *currents* introduced by de Rham became indispensible objects of study in differential geometry, partial differential equations and several complex variables. The study of currents led to the development of a distinguished branch of measure theory called Geometric measure theory, [11] dealing with regularity properties of sets and maps, has gained prominance in the latter of the twentieth century.

We have only given a very sketchy account of the the developments in real variable theory. The book by Hawkins [5] gives a delightful account on the historical aspects of the development of the classical theory of measure and integral on the real line. The paper by R. Cooke [2] contains many references and focuses on those parts of measure theory that received impetus from Fourier analysis on the real line. The book of Riesz and Nagy [9] is still the best source

[2]The greatest is indisputably A. N. Kolmogorov.

for a treatment, along classical lines, of the theory of functions of a real variable including Lebesgue's theory of integration. In particular the authors provide an elementary proof of the differentiability almost everywhere of a monotone function and the theorem that almost all points of a set have Lebsegue density one.

References

[1] J. C. Burkill, *Henri Lebesgue 1875-1941,* Biographical memoirs of the fellows of the Royal Society, **4** (1944) 483-490.

[2] R. Cooke, *Uniqueness of trigonometric series and descriptive st theory,* Archive for the history of exact sciences, **45**, 281-334 (1993).

[3] L. Dirichlet, Gesammelte Abhandlunden, Chelsea Pub. Co., New York,

[4] J. L. Doob, *Probability vs. measure,* Paul Halmos-Celebrating 50 years of mathematics, Ed., J. H. Ewing and F. W. Gehring, Springer-Verlag, 1991.

[5] T. Hawkins, *Lebesgue's theory of integration - Its origin and development,* Second edition, AMS Chelsea Publishing, Providence, Rhode Island, 1975.

[6] C. Hermite and T.-J. Stieltjes,

[7] a. Pinkus, *Weierstrass and approximation theory,*

[8] B. Riemann,

[9] F. Riesz and B. Nagy, *Functional analysis,*

[10] O. Schlömich, *Analytische Studien,* Jena 1847

[11] L. Simon, *Lectures on geometric measure theory,*

Note on Caratheodory's definition of measurablility

The definition of measurability that is universally adopted today is due to Constantin Caratheodory. It is difficult to find an adequate or convincing motivation of this concept. Here we shall try to follow Caratheodory closely [1] to see how he was led to this definition. Let us recall

Definition Given an outer measure μ^* on the set of all subsets of a given set X, we say that $A \subset X$ is measurable (with respect to μ^*) when the following holds for all subsets W:

$$\mu^*(W) = \mu^*(W \cap A) + \mu^*(W \cap A^c). \tag{1}$$

In his formulation Caratheodory works with metric outer measures which is defined below:

Definition: An outer measure on a metric space (X, d) is said to be a metric outer measure if

$$\mu^*(A \cup B) = \mu^*(A) + \mu^*(B) \tag{2}$$

whenever $d(A, B) > 0$.

Note: the Lebesgue outer measure is an example of a metric outer measure and so are Hausdorff measures.

Caratheodory examines the extent to which (2) can be weakened and shows that (2) holds whenever there exists an open set H such that $B \subset H$ and $A \subset H^c$. The condition $d(A, B) > 0$ is seen to be a special case since

$$d(A, B) > 0 \Rightarrow d(A, \overline{B}) > 0$$

and the required open set is $X - \overline{B}$. Setting $W = A \cup B$ we recast (2) as

$$\mu^*(W) = \mu^*(W \cap H) + \mu^*(W \cap H^c). \tag{3}$$

Caratheodory proves (3) for all sets W and all open sets H. Note that (3) is the same as the Caratheodory's condition (1) for open sets H. Proceeding with the conviction that "good" sets are those that can be "approximated" by open sets then (3) should hold for all "good" sets and thus (3) is the natural replacement for (2) that does not refer to any metric on the space, leading to the definition of measurability. We now prove the theorem of Caratheodory that open sets are measurable.

Theorem: Given a metric outer measure on metric space X, all open subsets of X are measurable.

Proof: We need to show that for an arbitrary W,

$$\mu^*(W) \geq \mu^*(W \cap H) + \mu^*(W \cap H^c)$$

and the only non-trivial case is when $\mu^*(W) < \infty$ and $H \neq X$. Let H_m denote the set of all points x such that

$$d(x, H^c) > \frac{1}{m}$$

and K_m be the compliment of H_m. It is clear that

$$H_n \subset H_{m+1}, \quad K_m \supset K_{m+1}, \quad \bigcup_{m=1}^{\infty} H_m = H$$

and further

$$d(H_m, H^c) = \frac{1}{m}, \quad d(H_m, K_{m+1}) = \frac{1}{m} - \frac{1}{m+1} = \frac{1}{m(m+1)}.$$

Let B be any subset of H with finite outer measure and $B_m = B \cap H_m$. So $\{B_m\}$ is a monotone increasing sequence and has union B whereby the sequence of numbers $\{\mu^*(B_m)\}$ is monotone increasing and bounded above by $\mu^*(B)$. .

Claim: $\lim_{m\to\infty} \mu^*(B_m) = \mu^*(B)$.

Denote the limit by λ and observe that $\lambda \leq \mu^*(B)$. To get the opposite inequality one would think of a splitting of B into onion rings but the usual proofs given in the context of measures (that is invoking additivity and not sub-additivity) do not work. The proof given by Caratheodory involves a little twist wherein he allows the onion rings to "eventually" overlap but then these overlaps must eventually be small.

Since

$$B_{n+1} - B_n = B \cap (H_{n+1} - H_n) = B \cap H_{n+1} \cap H_n^c \subset K_n,$$

whereas $B_{n-1} \subset H_{n-1}$ the sets B_{n-1} and $B_{n+1} - B_n$ are separated by positive distance. Hence

$$\mu^*((B_{n+1} - B_n) \cup B_{n-1}) = \mu^*((B_{n+1} - B_n)) + \mu^*(B_{n-1})$$

But $(B_{n+1} - B_n) \cup B_{n-1} \subset (B_{n+1} - B_n) \cup B_n = B_{n+1}$ whereby we get the key inequality

$$\mu^*(B_{n+1} - B_n) + \mu^*(B_{n-1}) \leq \mu^*(B_{n+1}). \tag{4}$$

The reader is encouraged to draw the picture depicting the onion ring and a disjoint inner shell. Now

$$\mu^*(B) = \mu^*(\bigcup_{j=n}^{\infty}) \leq \mu^*(B_n) + \mu^*(B_{n+1} - B_n) + \mu^*(B_{n+2} - B_{n+1}) + \dots$$

where we have used non-overlapping rings. Unfortunately we CANNOT write

$$\mu^*(B_{n+1} - B_n) \leq \mu^*(B_{n+1}) - \mu^*(B_n)$$

We using instead the inequality (4) to get the overlapping-rings estimate:

$$\mu^*(B) \leq \mu^*(B_n) + (\mu^*(B_{n+1}) - \mu^*(B_{n-1})) + (\mu^*(B_{n+2}) - \mu^*(B_n)) + \dots.$$

The series telescopes and we are left with

$$\mu^*(B) \leq \lim_{p\to\infty} \Big(\mu^*(B_{n+p}) + \mu^*(B_{n+p+1}) - \mu^*(B_{n-1})\Big) = 2\lambda - \mu^*(B_{n-1}) \tag{5}$$

The over-estimate caused by the overlapping onion rings produced the extra $\lambda - \mu^*(B_{n-1})$ but this is vanishingly small as n becomes large[3] Letting $n \to \infty$ in (5) gives the desired result.

Now let W be arbitrary with $\mu^*(W) < \infty$ and take $B = H \cap W$. Then $B_n = H_n \cap B = H_n \cap W$. One sees at once that $d(W - B, B_n) \geq 1/n$ and so

$$\mu^*(W) \geq \mu^*((W - B) \cup B_n) = \mu^*(W - B) + \mu^*(B_n)$$

Letting $n \to \infty$ we get

$$\mu^*(W) \geq \mu^*(W - B) + \mu^*(B) = \mu^*(W \cap H^c) + \mu^*(W \cap H)$$

as advertised.

References

[1] C. Caratheodory, *Vorlesungen über reelle Funktionen,* Chelsea publishing comapny, New York, 1948.

[3]The onion rings become become indefinitely thin as we go out.

Lebesgue measure on $\mathbb{R}$ and $\mathbb{R}^n$

The Lebesgue and the Lebesgue Stieltjes measures are among the most important examples of measures encountered in analysis and their study deserves a special chapter. We follow the tradition of constructing the Lebesgue measure on the real line through the Caratheodory process after defining the outer measure. The procedure suggests an obvious generalization to $\mathbb{R}^n$ but some technicalities arise that are easy to cope with in case of Lebesgue measures. It would be quite sticky to follow this approach to construct the multi-dimensional Lebesgue Stieltjes measures. We shall return to these issues in a later chapter.

Later in the chapter we prove an important covering lemma due to Vitali and illustrate its power in the analysis of certain finer properties of sets and functions namely the theorems of Lebesgue on metric density of sets and the differentiability of monotone functions. We shall return to these circle of ideas later when we discuss the maximal function of Hardy and Littlewood.

Lebesgue measure on the real line:

Non-measurable sets: Note that the set E we have constructed has the property that

$$(E - E) \cap \mathbb{Q} = \emptyset \text{ and } \bigcup_{q \in \mathbb{Q}} (E + q) = \mathbb{R}.$$

These follow immediately from elementary properties of cosets in an abelian group. One can use this to prove that E is non-measurable. Another proof of the existence of non-measurable sets using the continuum hypothesis for is available in C. Goffman [1]. This proof is quite interesting and employs a property of perfect sets known as the Cantor-Bendixon theorem[4].

Inner Measures: This is occasionally useful as the exercise 5 below shows. However, we shall discuss this only in the context of Lebesgue measure on the interval $I = [0, 1]$.

Definition: Given a subset A of $[0, 1]$, the inner measure $\mu_*(A)$ is defined as

$$\mu_*(A) = 1 - \mu^*(I - A)$$

The following properties of inner measure is left as an exercise.

Exercises

1. Show that when $A \subset B$, $\mu_*(A) \leq \mu_*(A)$.
2. Prove that $\mu_*(A)$ is the supremum of $\mu^*(K)$ as K varies over all compact subsets of A.

[4]That every closed subset of $\mathbb{R}$ is a union of a perfect set and a set that is at-most countable.

3. Deduce that the inner measure is countably super-additive that is

$$\mu_*\Big(\bigcup_{i=1}^{\infty} A_i\Big) \geq \sum_{i=1}^{\infty} \mu_*(A_i)$$

when the sets A_i are pairwise disjoint.

4. Show that a set $A \subset I$ is measurable if and only if $\mu_*(A) = \mu^*(A)$.

5. Prove that every set $A \subset I$ with positive outer measure contains a non-measurable subset. Hint: Take the standard non-measurable set E we constructed in class and look the intersection of its rational translates with A.

Lebesgue measure on $\mathbb{R}^n$:

Here we discuss the construction and basic properties of Lebesgue measures in $\mathbb{R}^n$. Given two points $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{R}^n$ we define $\mathbf{a} < \mathbf{b}$ if $a_j < b_j$ for $j = 1, 2, \ldots, n$. If $\mathbf{a} < \mathbf{b}$ then $R(\mathbf{a}, \mathbf{b})$ denotes the open rectangular box

$$R(\mathbf{a}, \mathbf{b}) = \prod_{j=1}^{n} (a_j, b_j) \tag{400}$$

with ends $\mathbf{a}$ and $\mathbf{b}$ and $\overline{R}(\mathbf{a}, \mathbf{b})$ its closure. Note that every open covering of a set by open boxes has a countable subcover and that every set A admits such a cover. We shall in what follows consider only countable coverings. Define $v(R(\mathbf{a}, \mathbf{b}))$ to be $\prod(b_j - a_j)$ and abbreviate $R(\mathbf{a}, \mathbf{b})$ to R when there is no need to specify the ends $\mathbf{a}$ and $\mathbf{b}$. Now we define the outer measure $\mu^*(A)$ of a subset A of $\mathbb{R}^n$ as

$$\mu^*(A) = \inf\Big\{ \sum_{p=1}^{\infty} v(R_p) : \ A \subset \bigcup_{p=1}^{\infty} R_p \Big\} \tag{401}$$

That is, over all possible countable coverings of A through open boxes. For the empty set we assign the value zero.

Theorem: The set function μ^* is a metric outer measure on $\mathbb{R}^n$.

The proof is exactly similar to the one dimensional case and we omit it. One then defines measurability in the sense of Caratheodory and being a metric outer measure all Borel sets are measurable. The class of all measurable sets is the Lebesgue $\sigma-$algebra and μ^* restricted to the class of measurable sets is denoted by μ. Since all Borel sets are measurable, all continuous functions are measurable and hence so are all the Baire functions. The integral of a measurable function f with respect to this measure is then the Lebesgue integral.

It would seem at first as though there is nothing further that needs to be done. However it must be remembered that the only effective procedure for computing any integral is to reduce it ultimately to a Riemann integral and employ the usual techniques of calculus[5]. The only essential step in relating the Lebesgue and Riemann integrals is the following non trivial result. It is exactly here that the analogous approach to Lebesgue-Stieltjes measures gets particularly unpleasant.

Theorem: The outer measure of $\overline{R}(\mathbf{a},\mathbf{b}) = \prod(b_j - a_j)$.

This result would be fairly trivial if we have $\mu^*(R) = v(R)$ for open rectangles but all we get from the definition is that $\mu^*(R) \leq v(R)$. It is clear that if R and S are two open boxes and $R \subset S$ then

$$v(R) \leq v(S)$$

Another important property of the function v is the following described in the context of $\mathbb{R}^2$. Suppose that

$$\Big\{a = t_0 < t_1 < \ldots < t_n = b\Big\} \times \Big\{c = s_0 < s_1 < \ldots < s_m = d\Big\}$$

is a partition of the open rectangle $(a,b)\times(c,d)$ and R_{ij} denotes the (i,j)-th open sub-rectangle then it is immediate[6] that

$$v(R) = \sum_{i,j} v(R_{ij}) \tag{402}$$

although the R_{ij} together do not cover R. Now we turn to the

Proof of the theorem: As in the one-dimensional case we see directly from the definition

$$\mu^*(\overline{R}(\mathbf{a},\mathbf{b})) \leq \prod(b_j - a_j)$$

To get the reverse inequality we work in the plane though the ideas are simple and extend routinely to any number of dimensions. Let $\overline{R}$ be the rectangle $[a,b] \times [c,d]$ and let $\epsilon > 0$ be arbitrary. There is a covering by open rectangles $\Big\{R_j\Big\}$ such that

$$\mu^*(\overline{R}) + \epsilon > \sum_{j=1}^{\infty} v(R_j) \geq \sum_{j=1}^{N} v(R_j) \tag{403}$$

where we assume by compactness that the first N of the rectangles cover $\overline{R}$. Now pick a Lebesgue number δ for the cover and partition $\overline{R}$ into equal closed squares say

$$\Big\{a = t_0 < t_1 < \ldots < t_n = b\Big\} \times \Big\{c = s_0 < s_1 < \ldots < s_m = d\Big\}$$

with the diagonal of each square less than δ. We refer to the squares as $S_{11}, S_{12}, \ldots, S_{nm}$. We assume that R_1 contains along with the bottom-left square S_{11} all the squares S_{ij} with

[5] Such as the fundamental theorem of calculus at least for nice enough functions, the change of variables formula and the Fubini theorem.

[6] This is merely a trivial algebraic identity that is recast using the symbol v.

$(i, j) = (1, 1), \ldots, (k_1, l_1)$ and none others. Now the union of the squares contained in R_1 is a closed rectangle C_1 contained in R_1. Now,

$$v(R_1) \geq v(\text{int}C_1) = \sum_1 v(\text{int}S_{ij})$$

by formula (402) where the subscript 1 under the summation sign indicates that we sum over those indices (i, j) such that $S_{ij} \subset R_1$. Since all squares S_{ij} that are contained in R_1 have been accounted for, R_1 goes out of reckoning and we proceed collect all the squares S_{ij} contained in R_2 and call their union C_2 so that

$$v(R_2) \geq v(\text{int}C_2) = \sum_2 v(\text{int}S_{ij})$$

the subscript 2 under the summation sign indicating that the sum is over those indices (i, j) such that $S_{ij} \subset R_2$. Proceeding thus we must in finitely many steps exhaust all the squares S_{ij} and finally upon summing, we get

$$\mu^*(\overline{R}) + \epsilon > \sum_{j=1}^{N} v(R_j) \geq \sum_{j=1}^{N} v(\text{int}(C_j)) = \sum_1 \sum_2 \ldots \sum_N v(\text{int}S_{ij})$$

Invoking the formula (402) again we see that

$$\mu^*(\overline{R}) + \epsilon > (b - a)(c - d).$$

Since $\epsilon > 0$ is arbitrary we get the desired inequality and thereby the proof is complete.

Corollary: $\mu^*(R) = v(R)$ for any open box R.

Proof: For $L = [a, b] \times \{c\}$ we see that $\mu^*(L) = 0$ and we use now the sub-additivity and monotonicity of outer measure. Another way would be take a sequence of closed boxes whose union in R and use the sub-additivity of μ^*.

Theorem: Suppose that $\overline{R}$ is a closed bounded box and $f : \overline{R} \longrightarrow \mathbb{R}$ is a bounded Riemann integrable function then the Riemann integral of f agrees with the Lebesgue integral of f.

The proof is exactly the same as for the one dimensional case.

Corollary: If f is continuous and has compact support then the Riemann integral of f over any box containing the support of f agrees with the Lebesgue integral of f over $\mathbb{R}^n$.

Corollary: Suppose that $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ and $\displaystyle\int_{\mathbb{R}^n} |f(x)|dx$ exists as an improper Riemann integral then $f \in L^1(\mathbb{R}^n)$ and the improper Riemann integral of f agrees with the Lebesgue integral of f

Uniqueness of translation invariant measures: We shall prove here the important result that the Lebesgue measure is the only translation invariant measure on the Borel sigma algebra on $\mathbb{R}^n$ that assumes value 1 on the unit cube $[0,1]^n$. This implies the higly non-trivial property of rotation invariance of the measure as well. We follow the presentation in L. Schwartz [2].

Theorem: The lebesgue measure on $\mathbb{R}^n$ is the unique translation invariant measure defined on the sigma algebra of Borel sets of $\mathbb{R}^n$ that assigns value 1 to the unit cube $[0,1]^n$.

Proof: It is evident that the Lebesgue measure μ possesses these properties. Assume that ν is such a measure defined on all Borel sets. We proceed to prove first that

$$\nu([0,1)^n) = 1$$

where, given vectors $\mathbf{a}$ and $\mathbf{b}$ with $\mathbf{a} < \mathbf{b}$, $[\mathbf{a}, \mathbf{b})$ denotes the half-open box $\prod [a_j, b_j)$. To see this, define F_n and E_n by

$$F_p = [0,1)^{n-1} \times [0, \frac{1}{p}), \quad E_p = \{0\}^{n-1} \times \{0, \frac{1}{p}, \frac{2}{p}, \dots, \frac{p-1}{p}\}$$

Then $[0,1)^n$ is the disjoint union of the translates of F_p through points in E_p. As a result,

$$p\nu(F_n) = \nu([0,1)^n) \leq \nu([0,1]^n) = 1$$

whereby $\nu(F_n) \leq 1/p$ implying that $\nu([0,1)^{n-1} \times \{0\}) = 0$. We deduce from this that the hyperplanes $x_j = c$ have measure zero and in particular

$$\nu([0,1]^n) = \nu((0,1)^n) = 1, \quad \nu([\mathbf{a}, \mathbf{b}]^n) = \nu((\mathbf{a}, \mathbf{b})^n) = 1$$

In the next step we decompose $[0,1)^n$ into p^n sub-cubes. Let Q_p denote the basic cube $[0,1/p)^n$ and consider the translates of Q_p through the points of the lattice

$$R_p = \{0, \frac{1}{p}, \dots, \frac{p-1}{p}\}^n$$

These translates are pairwise disjoint and cover $[0,1)^n$ so that

$$1 = \nu([0,1)^n) = \sum_{x \in R_p} \nu(x + Q_p) = p^n \nu(Q_p),$$

wherby $\nu(Q_p) = p^{-n}$. Consider the semi-open interval with end-points $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{Q}^n$ with $\mathbf{a} < \mathbf{b}$. Taking a common denominator N for the coordinates of $\mathbf{a}$ and $\mathbf{b}$ we write

$$\mathbf{a} = \Big(\frac{p_1}{N}, \frac{p_2}{N}, \dots, \frac{p_n}{N}\Big), \quad \mathbf{b} = \Big(\frac{q_1}{N}, \frac{q_2}{N}, \dots, \frac{q_n}{N}\Big).$$

The interval $[\mathbf{a}, \mathbf{b})$ is eveidently the disjoint union of the intervals $Q_N + \{x\}$ as $\{x\}$ varies over the lattice

$$S_N = \prod_{i=1}^{n} \Big\{ \frac{r}{N} : r \in \mathbb{Z},\ p_i \leq r < q_i \Big\}$$

Consequently,

$$\nu([\mathbf{a},\mathbf{b})) = \sum_{x\in S_N} \nu(Q_p + x) = \prod_{i=1}^{n}(q_i - p_i)/N^n = \prod_{i=1}^{n}(b_i - a_i) = \mu([\mathbf{a},\mathbf{b}))$$

Now, for an arbitary interval $[\mathbf{a},\mathbf{b})$, select sequences $\mathbf{a}_k$ and $\mathbf{b}_k$ of rational vectors such that

$$\mathbf{a}_k \longrightarrow \mathbf{a}, \quad \mathbf{b}_k \longrightarrow \mathbf{b}$$

and the intervals $[\mathbf{a}_k,\mathbf{b}_k)$ increase to $[\mathbf{a},\mathbf{b})$. Then

$$\nu([\mathbf{a},\mathbf{b})) = \lim_k \nu([\mathbf{a}_k,\mathbf{b}_k)) = \lim_k \mu([\mathbf{a}_k,\mathbf{b}_k)) = \mu([\mathbf{a},\mathbf{b})).$$

The measure ν agrees with the Lebesgue measure μ on boxes and hence on all Borel sets.

Corollary: The Lebesgue measure on $\mathbb{R}^n$ is rotation invariant.

Proof: If ϕ is a rotation, then being a homeomorphism, carries Borel sets to Borel sets. Define a set function ν on Borel sets as

$$\nu(A) = \alpha\mu(\phi(A))$$

where μ denotes the Lebesgue measure. Clearly ν is a measure and we show that it is transaltion invariant. Well,

$$\nu(x + A) = \alpha\mu(\phi(x) + \phi(A)) = \alpha\mu(\phi(A) = \nu(A).$$

Now we show that $\mu(\phi([0,1]^n))$ is not zero. Suppose the contrary, then we see by putting together 2^n translates of $[0,1]^n$ that $\mu(\phi([-1,1]^n)) = 0$. But $[-1,1]^n$ contains the open unit ball in $\mathbb{R}^n$ and this in turn contains the open cube $(0,1)^n$. Thus the Lebesgue measure of the open unit cube would be zero which is a contradiction.

Now take $\alpha = \mu(\phi([0,1]^n))^{-1}$ and we see that ν is a translation invariant measure defined on all Borel sets and assigns the value 1 to the unit cube $[0,1]^n$. By uniqueness, we conclude $\nu = \mu$ that is to say

$$\mu(A) = \alpha\mu(\phi(A)).$$

Choosing A to be the unit ball of $\mathbb{R}^n$ we deduce that $\alpha = 1$ completing the argument.

The Vitali covering theorem and its applications:

The Vitali covering theorem is a powerful tool that enables us to discuss the finer properties of sets and functions. Some of its applications are rather spectacular as we shall see. We first define a Vitali cover. This theorem is one of a series of delicate covering theorems each adapted to cope with specific (but typical) issues that repeatedly occur in analysis.

Definition: A Vitali cover of $E \subset \mathbb{R}^k$ is a covering of E by non-degenerate closed cubes such that for each $x \in E$ and $\epsilon > 0$ there is a cube Q of the covering with $\mu(Q) < \epsilon$ and $x \in Q$.

Theorem (Vitali covering theorem): If E is a subset of $\mathbb{R}^k$ of finite outer measure and $\mathcal{S}$ is a Vitali covering of E then there is a countable collection of pairwise disjoint cubes $Q_1, Q_2, \ldots$ in $\mathcal{S}$ such that

$$\mu^*(E - \bigcup_{j=1}^{\infty} Q_n) = 0.$$

In particular given any $\eta > 0$ there is a finite subcollection $Q_1, Q_2, \ldots, Q_n$ of pairwise disjoint cubes such that

$$\mu^*(E - \bigcup_{j=1}^{n} Q_n) < \eta.$$

Remark: We are not assuming that the set E is measurable.

Proof (Banach): Pick an open set G containing E with $\mu(G) < \infty$ and retain only those cubes that are contained in G. This is also a Vitali cover and we may assume that $\mathcal{S}$ is this subcover. Pick any $Q_1 \in \mathcal{S}$ and if $\mu(E - Q_1) = 0$ we are done. Else we continue inductively and assume that pairwise disjoint cubes $Q_1, Q_2, \ldots, Q_n$ have been selected. If the finte collection $Q_1, \ldots, Q_n$ covers E the process stops and the proof is complete. In the contrary case there are non-degenerate cubes of arbitrarily diameters contained in the open set $G - \bigcup_{j=1}^{n} Q_j$. The number

$$k_n = \sup\Big\{\mathrm{diam}(Q) : Q \in \mathcal{S} \text{ and } Q \subset \mathbb{R}^k - \bigcup_{j=1}^{n} Q_j\Big\}$$

is positive and is also finite since $\mu(G) < \infty$. Select a cube Q_{n+1} such that

$$\mathrm{diam}(Q_{n+1}) > \frac{1}{2} k_n, \quad Q_{n+1} \cap \Big(\bigcup_{j=1}^{n} Q_j\Big) = \emptyset.$$

Claim: The countable collection of cubes $Q_1, Q_2, \ldots, Q_n, \ldots$ satisfies the requirements.

Well, the cubes Q_j are pairwise disjoint by construction and the conclusion is trivial in case the process terminated in a finite numebr of steps. We assume then that the collection

$Q_1, Q_2, \dots$ consists of infinitely many cubes and observe that since their union is contained in the set G of finite measure,

$$\sum_{j=1}^{\infty} \mu(Q_j) < \infty$$

Let $\eta > 0$ be arbitrary and $m_0 \in \mathbb{N}$ be such that $\sum_{m_0+1}^{\infty} \mu(Q_j) < \eta/5$. We shall estimate the outer measure of the set

$$R = E - \bigcup_{j=1}^{m_0} Q_j.$$

Let $x \in \mathbb{R}$ and Q be a non-degenerate cube of the covering containing x and disjoint from $Q_1, \dots, Q_{m_0}$. This cube Q cannot be disjoint from all the Q_j for in that case,

$$\operatorname{diam}(Q) \leq k_j < 2 \operatorname{diam}(Q_{j+1}) \longrightarrow 0, \quad \text{as } j \to \infty.$$

giving a contradiction since Q is non-degenerate. Let n be the least natural number such that $Q \cap Q_n \neq \emptyset$. then $Q \cap Q_j = \emptyset$ for $j = 1, 2, \dots, n-1$. By the choice of $K - n$, we get

$$\operatorname{diam}(Q) \leq k_{n-1} < 2 \operatorname{diam}(Q_n) \tag{1}$$

For each m let C_m be the cube with the same center as Q_m but with five times the diameter. Then the inequality (1) shows that C_n engulfs Q and so contains x. We conclude thereby

$$R \subset \bigcup_{j=m_0+1}^{\infty} C_j$$

and consequently

$$\mu^*(R) < \sum_{j=m_0+1}^{\infty} \mu(C_j) = 5 \sum_{j=m_0+1}^{\infty} \mu(Q_j) < \eta.$$

The first application will be to

Theorem (Lebesgue): Let ν be a regular Borel measure on $\mathbb{R}^n$. Then ν is differentiable almost everywhere with respect to the Lebesgue measure.

Proof: Let us prove that the set

$$A = \Big\{x : \overline{D}\nu(x) > \underline{D}\nu(x)\Big\}$$

has zero oute rmeasure by showing that for each pair of rational numbers r and s with $r < s$, the sets

$$A_{r,s} = \Big\{x : \overline{D}\nu(x) > s > r > \underline{D}\nu(x)\Big\}$$

have outer measure zero. It suffices to show that each bounded subset B of $A_{r,s}$ has zero outer measure. We suppose the contrary and to arrive at a contradiction, let $\epsilon > 0$ and G be an open set containing B such that

$$\mu^*(B) + \epsilon > \mu(G).$$

For each $x \in B$ choose a sequence of closed non-degenerate cubes $\{Q_k(x)\}$ centered at x with diameters shrinking to zero such that $Q_k(x) \subset G$ and

$$\nu(Q_k(x)) < r\mu(Q_k(x)).$$

The family $\{Q_k(x)\}_{x,k}$ is a Vitali cover for B and so we can extract a countable subcover of pairwise disjoint cubes $\{R_j\}$ that cover a subset B_0 of B such that $\mu^*(B - B_0) = 0$. Now

$$\sum_{j=1}^{\infty} \nu(R_j) < r \sum_{j=1}^{\infty} \mu(R_j) < r\mu\Big(\bigcup_{j=1}^{\infty} R_j\Big) < r\mu(G) < r(\mu^*(B) + \epsilon)$$

Now using the lemma on the pair $\bigcup_{j=1}^{\infty} R_j$ and B_0 with the condition $\overline{D}\nu(x) > s$ on B_0 we get

$$\nu\Big(\bigcup_{j=1}^{\infty} R_j\Big) \geq s\mu^*(B_0) = s\mu^*(B),$$

using countable subadditivity of μ^* coupled with the fact that $\mu^*(B - B_0) = 0$. Finally we see that

$$s\mu^*(B) < r(\mu^*(B) + \epsilon).$$

Since $\epsilon > 0$ is arbitrary and $\mu^*(B) > 0$ we get $s \leq r$ which is a contradiction.

Exercises:

1. Show that Vitali's theorem remains true if balls are used in place of cubes.

2. Show that if E is a union of closed cubes/balls of positve diameter, then E is measurable.

Handling simultaneously improper Riemann integrals and Lebesgue integrals:

1. Show that $\int \frac{\sin x}{x} dx$ is not Lebesgue integrable but that $\int_0^\infty e^{-xt}\Big(\frac{\sin x}{x}\Big)dx$ is for every $t > 0$. Denoting by $J(t)$ the latter, show that $lim_{t\to 0+}J9t)$ exists. Is it equal to $\int_0^\infty \Big(\frac{\sin x}{x}\Big)dx$?

2. Subject the integrals $\int_0^\infty \sin x^2 dx$ and $\int_0^\infty \cos x^2 dx$ to an analysis similar to the previous exercise. Specifically for the integrals $P(t) = \int e^{-xt}\sin x^2 dx$ and $Q(t) = \int e^{-xt}\cos x^2 dx$ show that $\lim_{t\to 0+} P(t) = \int_0^\infty \sin x^2 dx$ and $\lim_{t\to 0+} Q(t) = \int_0^\infty \cos x^2 dx$. Obtain also a pair of first order differential equations satisfied by these functions.

3. Referring to the previous exercise, examine the behavior of $P(t)$ and $Q(t)$ as $t \longrightarrow \infty$ and explain the difference if any.

   Note: $P(t)$ and $Q(t)$ are the Laplace transforms of $\sin x^2$ and $\cos x^2$. The former is expected to decay faster at infinity since $(\sin x^2)\chi_{[0,\infty)}$ is Lipschitz at the origin but $(\cos x^2)\chi_{[0,\infty)}$ is not.

4. Determine the values of c for which $\|x\|^c$ is in $L^1(\mathbb{R}^n)$ and compute the integral in case it is finite.

5. Compute $\int_{\mathbb{R}^n} \exp(-\|x\|^2)dx_1dx_2\ldots dx_n$

6. Compute $\int_{\mathbb{R}^3} \exp(-\|x\| - i\xi\cdot x)dx_1dx_2dx_3$

References

[1] C. Goffman, *Real functions,* Reinehart and comapny, New York, 1960.

[2] L. Schwartz, *Analyse III Calcul intégral*, Herman, Paris 1993.

[3] D. Strook, *Concise introduction to the theory of integration*, World Scientific, Singapore, 1990.

[4] A. E. Taylor, *General theory of functions and Integration,* Dover, New York 1965.

Additivity of the integral: We have defined the integral in two ways denoting them by $I_1(f)$ and $I_2(f)$ respectively and have shown that the two agree when the measure space is σ-finite and also whenever $I_2(f)$ is finite. We now prove the finite additivity of the integral. It appears that the proof of finite additivity for $I_2(f)$ is quite troublesome. We wish to make our development as transparent as possible and so do not want to base our proof of such a fundamental a result on convergence theorems of Lebesgue theory. It may be noted that the additivity of the integral was used right in the beginning of our proof of the dominated convergence theorem.

Theorem: Suppose that f and g are two non-negative measurable functions then

$$\int_X (f+g)d\mu = \int_X f d\mu + \int_X g d\mu \tag{1}$$

Proof: When $\mu(X) < \infty$ and f and g are bounded on X, the proof parallels the one given in the theory of Riemann integrals. We now turn to the general case but assuming that the measure space is σ-finite. Note that if one of the summands on the right hand side of (1) is infinite then the result is trivial and so assume that both summands on the right hand side are finite. Let B be an arbitrary measurable subset of X with finite measure on which f and g bounded. Then

$$\int_B (f+g)d\mu = \int_B f d\mu + \int_B g d\mu \tag{2}$$

and the left hand side of the last equation is less than $\int_X (f+g)d\mu$ so that

$$\int_B f d\mu + \int_B g d\mu \leq \int_X (f+g)d\mu. \tag{3}$$

Taking supremum over all such B we get the inequality

$$\int_X f d\mu + \int_X g d\mu \leq \int_X (f+g)d\mu \tag{4}$$

To get the reverse inequality, let $\epsilon > 0$ be arbitrary. There exists a measurable subset B on which $f+g$ is bounded and

$$\int_X (f+g)d\mu - \epsilon < \int_B (f+g)d\mu \leq \int_B f d\mu + \int_B g d\mu$$

since the result is assumed for bounded functions on subsets on a finite measure. The right hand side however is less than

$$\int_X f d\mu + \int_X g d\mu$$

and $\epsilon > 0$ being arbitrary, we get

$$\int_X (f+g)d\mu \leq \int_X f d\mu + \int_X g d\mu \tag{5}$$

and the proof is complete.

Case when f and g are not necessarily positive: The equation

$$\int_X (f+g)d\mu = \int_X f d\mu + \int_X g d\mu \tag{6}$$

holds when

(i) Both summands on the right hand side are finite

(ii) Both are infinite but of the same sign, that is to say when both equal $+\infty$ or both equal $-\infty$.

(iii) One of them is finite and the other is infinite.

Proof: Let $h = f + g$ and we have with the usual notations

$$(f^+ - f^-) + (g^+ - g^-) = h^+ - h^-$$

which means

$$f^+ + g^+ + h^- = f^- + g^- + h^+ \tag{7}$$

Case (i) follows immediately from this upon integration and rearranging. we turn to cases (ii) and (iii). Assume that the integrals of f^+ or g^+ equal $+\infty$ while the integrals of f^- and g^- are finite. From (7) we get

$$h^+ + (f^- + g^-) \geq f^+ + g^+$$

On integration we get that $\int_X h^+ d\mu = +\infty$. We now have to show that the integral of h^- is finite. For this we consider the set $E = \{h^- > 0\}$. On E, the function h^+ vanishes and (7) reads

$$f^+ + g^+ + h^- = f^- + g^-$$

which further implies $h^- \leq f^- + g^-$ whereby we conclude that h^- has finite integral.

Dealing with the additivity of I_2: $I_2(f)$ is the integral of f as defined in Rudin's book. For non-negative functions f and g,

$$I_2(f+g) = I_2(f) + I_2(g)$$

When at-least one the summands on the right hand side is infinite, the result is trivial. Assume that both are finite. Then the functions must both vanish outside a $\sigma-$finite subset Y and so must the sum $f + g$. The result to be proved is the additivity of the integral over the subset Y. But then we have that $I_2(f) = I_1(f)$, $I_2(g) = I_1(g)$ and $I_2(f+g) = I_1(f+g)$.

The L^p spaces

Convex functions: Convexity plays a central role in many parts of analysis. Many important and frequently used inequalities such as the inequality of arithmetic and geometric means are special cases of convexity of the exponential function. Yet another case in point is the characterization of the Gamma function due to Bohr and Mollerup [4]. We gather here a few basic facts about convex functions of one variable that we shall occasionally use. The reference for this is W. Rudin's *principles* [4]. We begin with the definition

Definition: A real valued function defined on an interval I is said to be convex if

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y), \quad x, y \in I,\ 0 \leq \lambda \leq 1.$$

The parabola $y = x^2$ serves as an excellent paradigm for a convex function. The reader is encouraged to draw the figure of this parabola and take three points P, Q and R in that order along the curve and conjecture the following:

Basic slope lemma: Suppose that P, Q and R are three points on the graph of a convex function with Q between P and R, and P to the left of Q. Then the following hold:

(i) Slope PQ $\leq$ Slope QR.

(ii) Slope PQ $\leq$ Slope PR.

(iii) Slope PR $\leq$ Slope QR.

Proof: Let p, q and r be in the interval I with $p < q < r$. The proof proceeds along predictable lines. Then (it may be useful to recall the section formula from elementary coordinate geometry)

$$q = \Big(\frac{r-q}{r-p}\Big)p + \Big(\frac{q-p}{r-p}\Big)r,$$

so that

$$f(q) \leq \Big(\frac{r-q}{r-p}\Big)f(p) + \Big(\frac{q-p}{r-p}\Big)f(r), \tag{1}$$

Now we subtract off $f(p)$ on both sides after noting that

$$f(p) = \Big(\frac{r-q}{r-p}\Big)f(p) + \Big(\frac{q-p}{r-p}\Big)f(p)$$

and we get

$$f(q) - f(p) \leq \Big(\frac{q-p}{r-p}\Big)(f(r) - f(p))$$

from which (ii) follows. Now to prove (i) we subtract off $f(q)$ from either side of (1) after writing

$$f(q) = \Big(\frac{r-q}{r-p}\Big)f(q) + \Big(\frac{q-p}{r-p}\Big)f(q)$$

and we get

$$0 \le \Big(\frac{r-q}{r-p}\Big)(f(p)-f(q)) + \Big(\frac{q-p}{r-p}\Big)(f(r)-f(q)),$$

which after rearrangement gives (i). Finally we subtract off $f(r)$ from either side of (1) after writing

$$f(r) = \Big(\frac{r-q}{r-p}\Big)f(r) + \Big(\frac{q-p}{r-p}\Big)f(r)$$

to get

$$f(q) - f(r) \le \Big(\frac{r-q}{r-p}\Big)(f(p)-f(r)),$$

which after rearrangement proves (iii).

We now proceed to derive all the basic properties of convex functions out of the basic slope lemma.

Theorem: Suppose $f : I \longrightarrow \mathbb{R}$ is convex then,

(i) f is continuous on I

(ii) The left and right hand derivatives of f exist at each point of I.

(iii) At each point $p \in I$, $f'(p-) \le f'(p+)$.

(iv) If $p < q$ then $f'(p+) \le f'(q-)$.

(iv) If f has one sided derivatives at each point of I and satisfies (iv) then f is convex.

(v) If f'' exists and is non-negative then f is convex. Conversely if f is convex and twice differentiable then f'' is non-negative.

Proof: Let $p \in I$ and we pick q, r ans s in the interval I such that $s < p < q < r$. The basic slope lemma gives

$$\frac{f(s)-f(p)}{s-p} \le \frac{f(q)-f(p)}{q-p} \le \frac{f(q)-f(r)}{q-r}$$

Fix s, r and let $q \to p+$ to get continuity from the right. Repeat the argument with $s < q < p < r$ and we get continuity from the left. To prove (ii),

Theorem (Support theorem):

Theorem (Jensen's Inequality): Let X be a measure space with probability measure μ and $f \in L^1(\mu)$. Let ϕ be convex on an interval containing the image of f. Then

$$\phi\Big(\int_X f d\mu\Big) \le \int_X \phi \circ f d\mu$$

Proof: Note that the result is a generalization of the inequality

$$\Big|\int_X f d\mu\Big| \leq \int_X |f| d\mu.$$

For linear ϕ the result is trivial and for general ϕ one simply invokes the support theorem.

Completeness of L^p:

Hilbert proved in 1907 that L^2 is complete and the general case was settled a year later.

Convexity and duality in L^p spaces:

We shall first show that if $p^{-1}+q^{-1}=1$ then every element g of L^q defines a continuous linear form on L^p via the prescription

$$f \mapsto \int_X fg d\mu, \quad f \in L^p(\mu) \tag{500}$$

The converse is true if $q \neq 1$ and the standard proofs use the Radon-Nikodym theorem. However we prove a partial converse here namely, if a continuous linear form on L^p is given by (500) for some measurable function g, then g must be in L^q. This result is then used to prove the converse of Hölder's inequality, a result that is often useful (see for instance Hörmander [3], p [//]).

There is an alternate argument that avoids the Radon-Nikodym theorem but employs certain ideas of convexity from functional analysis, specifically the uniform convexity of the unit ball in L^p.

Theorem: Let X be a measure space with measure μ. Let p and q be conjugate exponents that is, $p, q \in [1, \infty]$ and $p^{-1} + q^{-1} = 1$. Each $g \in L^q$ defines a continuous linear form $T_g : L^p \longrightarrow \mathbb{R}$ given by (500). Moreover $\|T_g\| = \|g\|_q$ when $1 \leq q < \infty$. The last result also holds when $q = \infty$ provided the measure space is $\sigma-$finite.

Proof: Linearity of T_g is clear and by Hölder's inequality we get $|T_g(f)| \leq \|f\|_p \|g\|_q$ from which continuity of T_g follows and also

$$\|T_g\| = \sup\{|T_g(f)| : \|f\|_p \leq 1\} \leq \|g\|_q \tag{501}$$

We need to show that equality holds. We may clearly assume that $\|g\| \neq 0$. We define sgn $g(x)$ to be the sign of $g(x)$ when $g(x) \neq 0$ and zero if $g(x) = 0$.

Case $q < \infty$: The function $|g|^{q-1}$ is in L^p and if we take

$$f = (\text{sgn } g)|g|^{q-1}\|g\|^{1-q} \in L^p(\mu) \tag{502}$$

then $T_g(f)$ is easily seen to be $\|g\|_q$. On the other hand it is easily checked that $\|f\|_p = 1$ so that the supremum is (501) is attained at this f and equals $\|g\|_q$ and the proof is complete in this case.

Case $q = \infty$: Pick α arbitrary and $0 < \alpha < \|g\|_\infty$. We shall construct an $f \in L^1(\mu)$ such that $|T_g(f)| \geq \alpha$. To do this, let A be an an arbitrary measurable subset A finite positive measure such that $|g| \geq \alpha$ on A and

$$f = (\text{sgn } g)\chi_A/\mu(A) \in L^1(\mu).$$

It is immediately checked that $|T_g(f)| \geq \alpha$ and so $\|T_g\| \geq \alpha$. Since $\alpha < \|g\|_\infty$ is arbitrary we get the desired result.

The idea of constructing the f in (502) has already been encountered in the proof of the Minkowski's inequality. We shall use it again to prove the following partial converse of the previous result. Before stating the result, let us assume that T is a continuous linear form on $L^p(\mu)$ and try to obtain a candidate for g such that

$$Tf = \int_X fg d\mu$$

It is natural to apply this to χ_A with $\mu(A) < \infty$ to get

$$T(\chi_A) = \int_A g d\mu$$

Our problem is then to recover the point function g from the set function

$$A \mapsto T(\chi_A) = \int_A g d\mu$$

It is here that the Radon-Nikodym theorem is needed. The following result states that once the measurable function g is given then it must be in $L^q(\mu)$.

Theorem: Suppose that X is a $\sigma-$ finite measure space and p, q are conjugate exponents and g is a measurable function such that

$$M = \sup\Big\{ \int_X fg d\mu \ : \ \|f\|_p \leq 1 \Big\}, \tag{503}$$

Then,

(i) $M < \infty$ implies $g \in L^q$ and $\|g\|_q = M$.

(ii) If M is infinite then $\|g\|_q = +\infty$.

Note: The hypothesis states that $\int_X fg d\mu$ exists (though possibly infinite) for every $f \in L^p$.

Proof: We begin with Hölder's inequality with non-negative f

$$\Big|\int_X fgd\mu\Big| \le \|f\|_p\|g\|_q \le \|g\|_q,$$

which holds even if $\|g\|_q = \infty$ provided $\|f\| \neq 0$. We break the proof into three cases.

1. The case $q < \infty$ and $\|g\|_q < \infty$. This is exactly what was done in the last part of the previous theorem.

2. Case $q < \infty$ and $\|g\|_q = \infty$. We show that $M = +\infty$ as well by constructing a function $f \in L^p$ with $\|f\|_p = 1$ such that $\int fgd\mu$ is arbitrarily large. Write

$$X = \bigcup_{j=1}^{\infty} E_j, \quad \mu(E_j) < \infty, \quad E_j \subset E_{j+1},$$

and define $g_n(x) = g(x)$ when $x \in E_n$ and $|g(x)| \le n$ and zero elsewhere. Then

$$|g_n|^{q-1}\text{sgn } g \in L^r, \quad r \ge 1$$

and in particular is in L^p. Normalize this by dividing by $\|g_n\|^{q-1}$ and take

$$f_n = \|g\|_q^{1-q}|g_n|^{q-1}\text{sgn } g$$

Then we have on the one hand

$$\int_X |f_n|^p d\mu = \frac{1}{\|g_n\|^q}\int_{E_n} |g_n|^q d\mu \le 1.$$

and on the other hand

$$\int_X f_n g d\mu = \int_{E_n} |g_n|^q\|g_n\|^{1-q} d\mu = \|g_n\|_q \longrightarrow +\infty$$

by monotone convergence theorem.

3. The case $q = \infty$ is again handled exactly as in the previous theorem.

Theorem (Converse of Hölder's inequality): Let X be a sigma finite measure space. Suppose that g is a measurable function such that for every $f \in L^p(\mu)$ the product fg is in $L^1(\mu)$ then $g \in L^q(\mu)$ where $p^{-1} + q^{-1} = 1$.

Proof: We have the linear map

$$f \mapsto \int_X fgd\mu \tag{504}$$

but we do not yet know that this is continuous. The result is trivial if $g = 0$ almost everywhere. Replacing f by f sgn g we may assume that $g \ge 0$. As before let $X = \bigcup E_j$ where $E_j \subset E_{j+1}$ and $\mu(E_j) < \infty$ for all j. This time define $g_n(x) = g(x)\chi_{E_n}$ if $0 \le g(x) \le n$ and $g(x) = n\chi_{E_n}$

whenever $g(x) \geq n$. Then $\{g_n\}$ is a monotone increasing sequence of non-negative functions. Since each g_n is in $L^q(\mu)$ the linear maps T_n given by

$$T_n : f \mapsto \int_X f g_n d\mu, \quad f \in L^p(\mu)$$

are continuous and

$$|T_n f| \leq \Big| \int_X f g_n d\mu \Big| \leq \int_X |f| g d\mu.$$

Thus for each $f \in L^p(\mu)$,

$$\sup_n |T_n f| < \infty$$

By Banach-Steinhaus' theorem it follows that

$$\sup_n \|T_n\| < \infty$$

But $\|T_n\| = \|g_n\|_q$ and by monotone convergence theorem,

$$\sup_n \|g_n\|_q^q = \int_X |g|^q.$$

In other words, $g \in L^q(\mu)$.

Uniform convexity and reflexivity of L^p spaces:

Note that if $p = 2$ then the conjugate exponent is also 2 and Hölder's inequality becomes Cauchy-Schwarz' inequality

$$\Big| \int_X f g d\mu \Big| \leq \sqrt{\int_X |f|^2 d\mu} \sqrt{\int_X |g|^2 d\mu}, \quad f, g \in L^2(\mu)$$

The integral of the product fg defines an inner-product on $L^2(\mu)$ namely

$$< f, g >= \int_X f g d\mu.$$

If we consider complex valued functions then we must replace g with $\overline{g}$. Here we shall work only with real-valued functions.

Definition: Functions f and g in $L^2(\mu)$ are said to be orthogonal if $< f, g >= 0$. A family $\mathcal{F}$ of functions in L^2 is said to be an orthonormal basis if

1. Any two members of $\mathcal{F}$ are orthogonal
2. Each element of $\mathcal{F}$ has unit norm.
3. The linear span of $\mathcal{F}$, consisting of finite linear combinations of members of $\mathcal{F}$ with real coefficients is a dense subset of $L^2(\mu)$.

Theorem: If $L^2(\mu)$ is separable then there exists a countable orthonormal basis for $L^2(\mu)$.

For a proof see Rudin [//]. Important examples of such spaces are $L^2(\Omega)$ where Ω is an open subset of $\mathbb{R}^n$ or a compact box. A few examples of orthonormal bases will appear in the exercises.

We gather here a few properties of the inner-product whose proofs are trivial.

Theorem: (i) If f and g are orthogonal then

$$\|f\|^2 + \|g\|^2 = \|f+g\|^2$$

(ii) For any pair $f, g \in L^2(\mu)$, the parallelogram law hold:

$$\left\|\frac{f+g}{2}\right\|^2 + \left\|\frac{f-g}{2}\right\|^2 = \frac{1}{2}(\|f\|^2 + \|g\|^2)$$

One would like to inquire if there are analogues of the parallelogram law for other L^p spaces. although there are no corresponding identities when $p \neq 2$, inequalities are available. We state these now

Clarkson's first inequality: Suppose $2 \leq p < \infty$ and $f, g \in L^p(\mu)$,

$$\left\|\frac{f+g}{2}\right\|_{L^p}^p + \left\|\frac{f-g}{2}\right\|_{L^p}^p \leq \frac{1}{2}(\|f\|_p^p + \|g\|_p^p)$$

Proof: It suffices to prove that for all $a, b \in \mathbb{R}$,

$$\left|\frac{a+b}{2}\right|^p + \left|\frac{a-b}{2}\right|^p \leq \frac{1}{2}(|a|^p + |b|^p).$$

Using calculus, the reader can check that

$$\alpha^p + \beta^p \leq (\alpha^2 + \beta^2)^{p/2}.$$

Taking $\alpha = |a+b|/2$ and $\beta = |a-b|/2$ and appealing to the convexity of the function

$$t \mapsto |t|^{p/2}, \quad t \in \mathbb{R},$$

we get the desired result.

We now state without proff Clarkson's second inequality. The proof is non-trivial and is available in Hewitt and Stromberg [2].

Theorem (Clarkson's second inequality): Suppos $1 < p < 2$ and q is the conjugate exponent namely $p^{-1} + q^{-1} = 1$ then for $f, g \in L^p(\mu)$ we have the inequality

$$\left\|\frac{f+g}{2}\right\|_p^q + \left\|\frac{f-g}{2}\right\|_p^q \leq \left(\frac{1}{2}\|f\|_p^p + \frac{1}{2}\|g\|_p^p\right)^{1/(p-1)}$$

We define now the notion of uniform convexity of Banach spaces.

Definition (Uniform convexity): A Banach space is said to be uniformly convex if given any $\epsilon > 0$ there is a $\delta > 0$ such that

$$\|x\| \leq 1, \ \ \|y\| \leq 1 \text{ and } \|x - y\| \geq \epsilon \quad \text{implies} \quad \left\| \frac{x+y}{2} \right\| < 1 - \delta$$

It follows at once from the Clarkson's inequalities that the spaces $L^p(\mu)$ are uniformly convex when $1 < p < \infty$. It is an exercise to check that the spaces $L^1[0,1]$, $L^\infty[0,1]$ and $C[0,1]$ are not uniformly convex.

The importance of uniform convexity stems from the following fact that we state without proof (see Brezis [//]).

Theorem: A uniformly convex Banach space is separable.

Theorems of Lusin and Riesz:

In this section we prove an important theorem due to Lusin on the approximations of L^p functions by continuous functions. We prove the result on the real line with Lebesgue measure but the proof proceeds along general lines and is applicable in the case of regular Borel measures on locally compact Hausdorff spaces. Given a topological space X, we denote by $C_c(X)$ the space of all real valued continuous functions on X having compact support.

Theorem (Lusin): (i) The space $C[a,b]$ is dense in $L^p[a,b]$ for every $p \in [1, \infty)$.

(ii) The space $C_c(\mathbb{R})$ is dense in $L^p(\mathbb{R})$ for every $p \in [1, \infty)$.

Proof: (i) Let $f \in L^p[a,b]$ and $\epsilon > 0$ be arbitrary. Define $E_N = \{x : |f(x)| > N\}$ and seeing that $\mu(E_N) < N^{-p}\|f\|_p$, there is some N such that $\mu(E_N) < \delta$ for any preassigned $\delta > 0$. The δ we take is such that

$$\int_A |f|^p d\mu < (\epsilon/3)^p, \qquad \text{whenever} \quad \mu(A) < \delta.$$

By dividing the interval $[-N, N]$ into $2^{n+1}N$ equal subintervals $\{(a_j, b_j]\}_j$, defining the simple function s_n as taking the constant value a_j on the piece $E_N^c \cap f^{-1}(a_j, b_j]$ and zero outside A we see that

$$\sup_{E_N^c} |s_n(x) - f(x)| \longrightarrow 0$$

and hence

$$\|f - s_n\|_p^p = \int_{E_N^c} |f - s_n|^p d\mu + \Big(\frac{\epsilon}{3}\Big)^p, \qquad \text{for all } n.$$

By uniform convergence we can find an m such that

$$\sup_{E_N^c} |s_m(x) - f(x)|^p < \frac{1}{\mu(E_N^c)} \Big(\frac{\epsilon}{3}\Big)^p$$

and hence

$$\|f - s_m\| \leq 2^{1/p}\Big(\frac{\epsilon}{3}\Big) < \frac{2\epsilon}{3}.$$

all that remains now is to show that there is a continuous function g on $[a, b]$ such that

$$\|g - s_m\|_p < \frac{\epsilon}{3}.$$

It clearly suffices to prove this for characteristic functions of a measurable set C with finite measure. That is given $\eta > 0$ we find a continuous function h on $[a, b]$ such that

$$\|\chi_C - h\| \leq \eta.$$

To do this, pick a compact set K and an open set G such that $K \subset C \subset G$ and $\mu(G - C) < \eta^p$ and define $h = 0$ outside G and $h = 1$ on K. By Tietze's extension theorem[7] we can extend h continuously to $[a, b]$ with the extension preserving the bounds 0 and 1. Then

$$\|h - \chi_C\|_p \leq \mu(G - K)^p < \eta.$$

To prove (ii) we reduce it to the case (i). Let $f \in L^p(\mathbb{R})$ and $\epsilon > 0$ be arbitrary. Let us break the domain into a countable union of an increasing family of measurable sets $\{X_j\}$ each having finite measure[8]. By the dominated convergence theorem,

$$\lim_{n\to\infty} \int_{X_n} |f|^p d\mu = \int |f|^p d\mu,$$

we see that

$$\int_{X_N^c} |f|^p d\mu < \Big(\frac{\epsilon}{4}\Big)^p$$

for some N. Now since $\mu(X_N)$ is finite we can for any preassigned $\delta > 0$ find a compact set $K \subset X_N$ such that $\mu(X_N - K) < \delta$ and the δ is chosen such that

$$\int_A |f|^p d\mu < (\epsilon/4)^p, \qquad \text{whenever} \quad \mu(A) < \delta.$$

Now assume $K \subset [a, b]$ and we have so far

$$\int_{\mathbb{R}-[a,b]} |f|^p d\mu < 2(\epsilon/4)^p$$

by (i) choose a continuous function g on $[a, b]$ such that

$$\|f\chi_{[a,b]} - g\| < \epsilon/4.$$

Define g to be zero outside $(a - \eta, b + \eta)$ and use Tietze's theorem to extend g continuously to the whole real line preserving the bounds. The η in the last line is chosen as $\eta = M^p/2$ where M is the supremum of $|g|$ on $[a, b]$. Putting all the pieces together,

$$\|f - g\|_p^p < 2(\epsilon/4)^p + \|f\chi_{[a,b]} - g\|^p + 2\eta M^p < 4(\epsilon/4)^p,$$

or $\|f - g\|_p < \epsilon$ completing the argument.

[7] If one is interested only on the real line this could be done without appealing to the general Tietze's theorem from topology. See exercise [//].

[8] In general since $f \in L^p$, f must vanishes outside a $\sigma-$ finite subset

Corollary: Suppose given $f \in L^1[a, b]$ then given any $\delta > 0$ there is a continuous function g on $[a, b]$ such that $\mu\{x : f(x) \neq g(x)\} < \delta$. Further if $|f(x)| < M$ almost everywhere then g can be chosen such that $\sup |g| < M$.

Proof: First choose a sequence of continuous functions $\{g_n\}$ converging to f in L^1. Riesz' theorem gives a subsequence converging to f pointwsie almost-everywhere. Replacing the original sequence by this subsequence we assume $\{g_n\}$ converges pointwise almost-everywhere. Egoroff's theorem gives a measurable subset A_δ such that $\mu(A_\delta^c) < \delta/2$ and $g_n \to f$ uniformly on A_δ. We choose a compact set $K_\delta \subset A_\delta$ such that $\mu(A_\delta - K_\delta) < \delta/2$. Since $g_n \to f$ uniformly on A_δ and hence also on K_δ we see that f restricted to K_δ is continuous. By Tietze's extension theorem, the function $f\Big|_{K_\delta}$ has a continuous extension to $[a, b]$. Denoting this extension by g we have finally a $g \in C[a, b]$ such that

$$\mu(\{x : f(x) \neq g(x)\}) \leq \mu(K_\delta) < \delta.$$

The last part also follows from Tietze's theorem.

Definition (Regular Borel measures): A measure μ on a locally compact Hausdorff space is said to be a regular Borel measure if the following conditions hold:

(i) All Borel sets are measurable

(ii) Given any measurable set E of finite measure and any $\epsilon > 0$ there is an open set G containing E such that $\mu(G) < \mu(E) + \epsilon$.

(iii) Given any measurable set E of finite measure and any $\epsilon > 0$ there is a compact set $K \subset E$ such that $\mu(E) - \epsilon < \mu(K)$.

The Lebesgue measure on $\mathbb{R}^n$ is an example of a regular Borel measure and so are the Dirac measures. Any positive linear combination of Dirac measures is a regular Borel measure. We shall see later that they exist in great profusion. The proof of Lusin's theorem we have just discussed holds for all regular Borel measures.

Theorem (Lusin): If μ is a regular Borel measure on a locally compact Hausdorff space X then all continuous functions are measurable and the space $C_c(X)$ is dense in $L^p(\mu)$ for every $p \in [1, \infty)$.

Theorem: If $f(x) \in L^p(\mathbb{R}^n)$ then $f(-x)$ and $f(x + h)$ both lie in $L^p(\mathbb{R}^n)$ and

$$\|f(x)\| = \|f(-x)\| = \|f(x + h)\|$$

Proof: Clearly it suffices to discuss the case $p = 1$. It is easy to check that $\mu^*(A) = \mu^*(-A)$ for any set A from which we deduce the stated result for characteristic functions thereby for simple functions and upon passing to limits, for all L^1 functions.

Here is an application of Lusin's theorem that is so frequently used that it merits the status of a theorem.

Theorem: Suppose that $1 \leq p < \infty$ and $f \in L^p(\mathbb{R}^n)$ then

$$\lim_{h \to 0} \|f(x+h) - f(x)\|_p = 0$$

Proof: Let g be a continuous function with compact support K and L be a compact set such that

$$d(K, L^c) \geq 1$$

then both $g(x)$ and $g(x+h)$ vanish when x lies in L^c and $|h| \leq 1/2$. By uniform continuity of g on L, given any $\epsilon > 0$ there is $\delta > 0$ such that

$$|g(x+h) - g(x)| < \epsilon(\mu(L))^{-1/p}, \quad |h| \leq \delta.$$

Now,

$$\|g(x+h) - g(x)\|_p^p = \int_L |g(x+h) - g(x)|^p d\mu < \epsilon^p, \quad |h| < \delta$$

For $f \in L^p$ and $\epsilon > 0$ arbitrary, choose $g \in C_c(\mathbb{R}^n)$ such that $\|f - g\| < \epsilon/3$. Then by translation invariance of Lebesgue measure, $\|f(x+h) - g(x+h)\| < \epsilon/3$. For g there is a $\delta > 0$ such that $\|g(x+h) - g(x)\| < \epsilon/3$ whenever $|h| < \delta$ and consequently,

$$\|f(x+h) - f(x)\| \leq \|f(x+h) - g(x+h)\| + \|g(x+h) - g(x)\| + \|g(x) - f(x)\| < \epsilon.$$

Separability of L^p spaces: We now prove an important result on the separability of L^p spaces. For this we need in addition the second countability of the topological space. The real line with discrete topology is not second countable but the counting measure is a regular Borel measure. The corresponding L^p spaces are not separable.

Theorem: Let X be a locally compact second countable Hausdorff space and μ be a regular Borel measure on X. Then the spaces $L^p(X)$ is separable for $1 \leq p < \infty$.

Proof: We have already seen in the course of the proof of Lusin's theorem that the simple functions vanishing outside sets of finite measure are dense in L^p for $1 \leq p < \infty$. These can in turn be approximated by linear combinations of characteristic functions with rational coefficients. We merely have to obtain a countable subset of functions in L^p that approximate characteristic functions of sets with finite measure. To do this pick a countable base $\mathcal{S}$ of open sets with compact closures. We can assume that the family $\mathcal{S}$ is closed under finite unions. Let C be any measurable set with finite measure and $\epsilon > 0$ be arbitrary. There exists a compact set K and an open set G such that $K \subset C \subset G$ and $\mu(G - K) < \epsilon^p$. Since our family $\mathcal{S}$ is closed under finite unions there is an open set $L_i \in \mathcal{S}$ containing K. Since closure of L_i is compact there is an $L_{ij} \in \mathcal{S}$ such that

$$K \subset \overline{L}_i \subset L_{ij} \subset \overline{L}_{ij} \subset G. \qquad (*)$$

If u is any continuous function such that $0 \leq u \leq 1$ and $u = 1$ on L_j and zero outside L_{ij} then

$$\|\chi_C - u\|_p \leq (\mu(G - C))^{1/p} < \epsilon.$$

Such a function is called an Urysohn function corresponding to the pair $\{L_i, L_{ij}\}$ where the sets L_i and L_{ij} belong to $\mathcal{S}$ and satisfy $*$. We select one such Urysohn function for each such pair and the countable dense set we seek is the family of finite linear combinations of these Urysohn functions with rational coefficients.

We now state a very important result in measure theory known as the Riesz representation theorem. For proof we refer to Folland [//]. Note that the space $C_c(X)$ is an ordered vector space which enables us to talk of positive linear forms. specifically a linear form $\phi : C_c(X) \longrightarrow \mathbb{R}$ is said to be positive if

$$\phi(f) \geq 0, \quad \text{whenever} \ \ f \geq 0.$$

Theorem (Riesz representation theorem): Let X be a locally compact Hausdorff space and ϕ be a positive linear form on $C_c(X)$ then there exists a regular Borel measure μ on X such that

$$\phi(f) = \int_X f d\mu, \quad f \in C_c(X).$$

Theorem (Vitali-Caratheodory): The proof is relegated to the exercises.

Notes and comments: Our treatment of uniform convexity follows Brezis [//]. In this book the readr will find a proof of the duality of L^p and L^q spaces where p and q are conjugate exponents with $1 < p < \infty$, without the Radon-Nikodym theorem. The proof employs Clarkson's inequalities. Brezis's book also contains compactness criteria for L^p spaces. The exercise [//] is also adapted from Brezis. For a proof of the Riesz representation theorem see the book by Folland [//]. A version of the Vitali-Caratheodory theorem was proved by Vitali in 1905 and completed by Caratheodory in 1918. For further details and references to original papers the reader must consult Saks [//].

Result of exercise [//] is due to Paul Montel[9]. In this connection see the book of D. S. Mitrinović[10]

[9] Sur les fonctions convexes et les fonctions sousharmoniques, Jour. de math. pures et appl., (**9**)**7**, 29-60, 1928.

[10] Analyitic Inequalities, Springer Verlag, Berlin, 1970.

Exercises:

1. Prove that if f is continuous on a compact subset $K \subset \mathbb{R}$ and vanishes outside an open set G containing K then f extends continuously on the whole real line such that the extension has the same bounds as the given function.

2. Show that if ψ is a strictly increasing convex function and ϕ is convex then $\psi \circ \phi$ is convex.

3. Suppose $\{\phi_\alpha\}_\alpha$ is a family of convex functions on an interval I and $\phi = \sup_\alpha$ is finite on I then ϕ is also convex. Deduce that $-\log^+$ is convex on the real line.

4. Show that if f is convex on an interval I and $x_1, \ldots, x_k$ are k points in I and $\lambda_1, \ldots, \lambda_k$ are k non-negative reals such that $\lambda_1 + \lambda_2 + \ldots + \lambda_k = 1$ then

$$f(\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_k x_k) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2) + \ldots + \lambda_k f(x_k)$$

5. Show that $\exp x$ is a convex function on the real line

6. Prove the inequality of means, namely if $a_1, a_2, \ldots, a_n$ are n positive numbers then

$$(a_1 \cdot a_2 \cdot \ldots \cdot a_n)^{1/n} \leq \frac{1}{n}\Big(a_1 + a_2 + \ldots + a_n\Big)$$

7. Show that if $u_1, u_2, \ldots, u_n$ are n non-negative numbers and $p_1, p_2, \ldots, p_n$ positive reals such that $\frac{1}{p_1} + \frac{1}{p_2} + \ldots + \frac{1}{p_n} = 1$ then

$$u_1 \cdot u_2 \cdot \ldots \cdot u_n \leq \frac{{u_1}^{p_1}}{p_1} + \ldots + \frac{{u_n}^{p_n}}{p_n}$$

8. Use Jensen's inequality to show that

$$\sin x > \frac{2x}{\pi}, \quad 0 < x < \pi/2$$

This is known as Jordan's inequality.

9. Given a positive function f in an interval I, prove that if $e^{cx} f(x)$ is convex on I for every $c \in \mathbb{R}$ then $\log f(x)$ is convex. Prove also that $f(x)$ must be monotone increasing. (P. Montel) Note: Convexity of $\log f(x)$ is already a very powerful condition since log is a highly concave function. In particular if $\log f(x)$ is convex then $f(x)$ must be convex. The above problem is from the Berkeley Problems in Math.

10. Show that $\exp(\exp x)$, $\sec x$ and $\cosh x$ are examples of functions whose logarithms are convex. The following exercise provides yet another.

11. Show directly from the integral or using the Gauss' product formula that $\log\Gamma(x)$ is convex.

    Note: The Bohr Mollerup theorem states that if $f(x)$ satisfies the condition $f(x+1) = xf(x)$ and $\log f(x)$ is convex then $f(x) = c\Gamma(x)$ for some constant c.

12. Use the preceding exercise and Jensen's inequality to show that if $x_1, x_2, \ldots, x_n$ are n positive numbers such that $x_1 + x_2 + \ldots + x_n = nx$ then

$$\Gamma(x_1)\Gamma(x_2)\ldots\Gamma(x_n) \geq \Gamma(x)^n$$

    This exercise is from Mitrinović [//], p. 285.

13. Show that the counting measure μ on $\mathbb{R}$ is a regular Borel measure with respect to the discrete topology. Show that $L^p(\mu)$ spaces are not separable.

14. Consider the set theoretic inclusion $L^q[0,1] \subset L^p[0,1]$ where $q > p \geq 1$. Is the set $L^q[0,1]$ closed in the metric space $L^p[0,1]$? Is it open? Is it dense? Is it of the first category?

Exercises

1. Suppose that $A \subset \mathbb{R}$ is measurable, show that given any $\epsilon > 0$, there exists an open set G containing A such that $\mu^*(G - A) < \epsilon$. Further show that if $\mu^*(A) < \infty$ then there is a finite collection of pairwise disjoint open intervals such that
$$\mu^*(G \Delta A) < \epsilon$$

2. Show that if A is measurable with finite outer measure then given any $\epsilon > 0$, there is a step function g such that $\|g - \chi_A\|_{L^1} < \epsilon$. Deduce that there is a continuous function h such that $\|h - \chi_A\|_{L^1} < \epsilon$. Show that if s is a simple function on $[a, b]$ then there exists a continuous function ϕ such that $\|\phi - s\|_{L^1} < \epsilon$ and finally if $f \in L^1[a, b]$, there is a continuous function ϕ such that $\|\phi - f\|_{L^1} < \epsilon$. Note: This is Luzin's theorem on closed bounded subintervals of $\mathbb{R}$. However this proof relies heavily on the fact that an open subset of $\mathbb{R}$ is a countable disjoint union of pairwise disjoint intervals. We cannot expect such a simple construction in multi-dimensions or locally compact Hausdorff spaces. We must rely on Tietze's extension theorem.

3. Assume that f is in $L^1(\mathbb{R})$. For each $x \in \mathbb{R}$ prove that
$$\lim_{t \longrightarrow 0+} \frac{1}{\sqrt{4\pi t}} \int_{\mathbb{R}} f(y) \exp(-(x-y)^2/4t) dy = f(x)$$
Hint: First assume that f is continuous with compact support.

4. Given that I and J are intervals in $\mathbb{R}$ and f is continuous on $I \times J$ such that
(i) $\dfrac{\partial f}{\partial y}$ exists on the rectangle $I \times J$.
(ii) There exists a measurable function $g \in L^1(I)$ such that $\left|\dfrac{\partial f}{\partial y}\right| \leq g$ on $I \times J$. Show that
$$\frac{d}{dy} \int_I f(x, y) dx = \int_I \frac{\partial f}{\partial y} dx$$

5. Use the preceding to prove that if f is bounded measurable on $[-\pi, \pi]$ then
$$\lim_{n \to \infty} \int_{-\pi}^{\pi} f(x) \sin nx dx = 0$$
This is called the Riemann Lebesgue lemma.

6. For a continuous function $f(x)$ re-prove the preceding directly in the context of theory of Riemann integrals without any reference to measure theory. Hint[11]: Consider the substitution $x = y + \frac{\pi}{n}$ after extending f periodically.

[11] This was Riemann's original proof.

7. Prove that if $f \in L^1(\mathbb{R})$ and the function $\widehat{f}$ is defined as

$$\widehat{f}(\xi) = \int_{\mathbb{R}} f(x) \exp(-ix\xi) dx$$

has the property that $\widehat{f}(\xi) \longrightarrow 0$ as $|\xi| \longrightarrow \infty$.

8. Compute using exercise 4 to justify differentiating under the integral sign,

$$J(\xi) = \int_{\mathbb{R}} \exp(-x^2/a) \cos x\xi \;\; dx$$

Hint: Find an ODE for $J(\xi)$.

9. Suppose that n_k is an increasing sequence of natural numbers and E is the set of points in $[-\pi, \pi]$ such that $\lim\limits_{k\to\infty} \sin n_k x$ exists, then show that E has Lebesgue measure zero.

10. Suppose given a sequence of real-valued functions $\{a_n(k,x)\}_n$ each term $a_n(x,k)$ is a function of $k \in \mathbb{N}$ and a parameter x that is to be regarded here as a real or complex constant satisfying the following conditions:

    (i) $\lim\limits_{n\to\infty} a_n(k,x)$ exists and equals $a_k(x)$ say.

    (ii) There is a sequence of constants $\{b_k\}_k$ such that $|a_n(k,x)| \leq b_k$, for all $k, n \in \mathbb{N}$

    (iii) $\sum_{k=1}^{\infty} b_k$ converges.

    Then, the following holds:

$$\lim_{n\to\infty} \sum_{k=1}^{n} a_n(k,x) = \sum_{k=1}^{\infty} a_k(x).$$

More generally show that

$$\lim_{n\to\infty} \sum_{k=1}^{F(n)} a_n(k,x) = \sum_{k=1}^{\infty} a_k(x).$$

where $F : \mathbb{N} \longrightarrow \mathbb{N}$ is such that $F(n) \longrightarrow \infty$ when $n \longrightarrow \infty$.

Hint: Use Lebesgue's dominated convergence theorem on the sequence of functions $a_n(k)$ by giving $\mathbb{N}$ the counting measure.

Note: This result appeared first in the classic treatise of Tannery and Molk [//], and Bromwich included it in his famous book [1] and called it Tannery's theorem. The proofs in these classical works are of course elementary and do not employ Lebesgue theory. The next two exercises are illustrations on the use of Tannery's theorem.

11. Prove that for any x real or complex,

$$\lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots.$$

12. Deriving the series expansion for $\sinh x$.

References

[1] T. J. I. Bromwich, *Infinite series,*

[2] E.Hewitt and K. Stromberg, Real and abstract analysis, springer Verlag,

[3] L. Hörmander, *Analysis of linear partial differential operators - I*, Springer Verlag, 1990.

[4] W. Rudin, *Principles of mathematical analysis,* Third edition, 1976.

Rudiments on Hilbert space

Test - I, Open book, open notes:

(i) Attempt any four questions.

(ii) Do the work initially in rough and submit only a neatly written final version.

1. You can use the result of any question (even if you have not attempted it) to answer another question.

1. Examine whether the sequence $\{\sin n^2 x\}$ converges to zero in measure on the interval $[0, \pi]$.

2. Let S^1 denote the circle $\{z \in \mathbb{C} : |z| = 1\}$ and $\theta \in \mathbb{R} - \mathbb{Q}$. Denote by T the function $T(z) = e^{2\pi i\theta} z$, the rotation through an angle that is an irrational multiple of π. Show that the orbit $S = \{\zeta, T\zeta, T^2\zeta, \ldots\}$ is dense in S^1 for every $\zeta \in S^1$. Deduce that the fractional parts $\{n\theta\}$ are dense in $[0, 1]$.

   Hint: Suppose that the orbit is not dense then there is a maximal open arc $l \subset S^1$ that does not contain any point of S. The sets l and $T^{-k}l$ have the same measure for every k.

3. Determine $\limsup_{n\to\infty} \sin n$

4. Prove that

$$\int_0^\infty t^{a-1} e^{-t} dt = \lim_{n\to\infty} \frac{n! n^{a-1}}{a(a+1)\ldots(a+n-1)}$$

   Hint: Apply the monotone convergence theorem to $\left(1 - \dfrac{t}{n}\right)^n \chi_{[0,n]}$

5. Show that every subset of $[0, 1]$ of positive outer measure contains a non-measurable subset.

   Hint: Take the standard non-measurable set $E \subset [0, 1]$ and E_j be their rational translates. Use countable sub-additivity of outer measure to show that $\mu^*(A \cap E_j) > 0$ for some j whereas from the countable super-additivity of inner measures, $\mu_*(E_k) = 0$ for every k.

Test - I, Solutions:

1. Examine whether the sequence $\{\sin n^2 x\}$ converges to zero in measure on the interval $[0, \pi]$.

   The straight forward method is to estimate. Let $\epsilon > 0$ be arbitrary. We need to estimate the measure of the set
$$E_n = \Big\{x : |\sin n^2 x| \geq \epsilon\Big\}$$
   Now, the function $|\sin n^2 x|$ is periodic with period π/n^2 and we split the interval $[0, \pi]$ into n^2 equal subintervals. On $[0, \pi/n^2]$ we have the piece
$$\frac{\sin^{-1}\epsilon}{n^2} < x < \frac{\pi - \sin^{-1}\epsilon}{n^2}$$
   and the measure of this subinterval is $(\pi - 2\sin^{-1}\epsilon)/n^2$ and so the measure of E_n is exactly $\pi - 2\sin^{-1}\epsilon$ which does not go to zero as $n \longrightarrow \infty$.

   Aliter: Suppose it converges in measure then there is a subsequence $\{\sin_k^2 x\}$ converging point-wise almost every-where. One of the students had the following cute idea. To use dominated convergence theorem to arrive at a contradiction but unfortunately could not execute it successfully. Now $(\sin n_k^2 x)^2$ converges to zero point-wise almost everywhere and so by dominated convergence theorem,
$$\lim_{k\to\infty}\int_0^\pi \sin^2 n_k^2 x dx = \frac{\pi}{2} \neq 0$$

2. Suppose the orbit of a point is not dense in S^1 then there is an open arc not containing any point of the orbit. Let l be a maximal such arc and look at the sequence
$$l, T^{-1}l, T^{-2}l, \ldots$$
   Non of these contain any point of the orbit and they all have the same measure and this measure is positive. Hence they cannot be pairwise disjoint. Assume then $T^{-i}l \cap T^{-j}l \neq \emptyset$ where $j < i$. Then $T^{-k}l \cap l \neq \emptyset$ $(k = i - j)$. This intersection is an open arc and so by maximality and the fact that l and $T^{-k}l$ have the same length, it must be that $T^{-k}l = l$. Since the endpoints of l and $T^{-k}l$ must correspond, it follows that $\exp(2\pi i k\theta) = 1$ or that θ must be rational which is a contradiction.

   To prove that if $\theta \in \mathbb{R}$ is irrational, the fractional parts $\{n\theta\}$ are dense in $[0, 1]$ we observe that $t \mapsto \exp(2\pi i t)$ is a homeomorphism from $(0, 1)$ onto $S^1 - \{1\}$. This last result is a famous result called Kronecker's theorem.

3. To prove that $\limsup\limits_{n\to\infty} \sin n = 1$. It is clear that the limit superior is less than or equal to 1. Let $\epsilon > 0$ be arbitrary. We have to show that there are infinitely many values of n such that
$$1 - \epsilon < \sin n \leq 1.$$

Equivalently that there are infinitely many pairs (n, k_n) such that

$$\lim_{n\to\infty} \sin(n - 2\pi k_n) = 1,$$

which in turn is equivalent to

$$\lim_{n\to\infty} (n - 2\pi k_n) = \pi/2.$$

Dividing by 2π gives

$$\lim_{n\to\infty} (\frac{n}{2\pi} - k_n) = 1/4.$$

Taking k_n to be the integer part of $\frac{n}{2\pi}$ gives

$$\left\{\frac{n}{2\pi}\right\} = \frac{1}{4}.$$

By Kronecker's theorem (exercise 2), we can surely find infinitely many such values of n since $1/2\pi$ is irrational.

4. To prove that

$$\int_0^\infty t^{a-1}e^{-t}dt = \lim_{n\to\infty} \frac{n!n^{a-1}}{a(a+1)\dots(a+n-1)}$$

This was Gauss' product formula for the Gamma function. Gauss employed it as the definition of the Gamma function and derived its princi[pal properties in his great memoir on the hypergoemetric function of 1812.

We apply the monotone convergence theorem to the sequence $\left(1 - \frac{t}{n}\right)^n \chi_{[0,n]}(t)$.. we must show that

$$\left(1 - \frac{t}{n}\right)^n \leq \left(1 - \frac{t}{n+1}\right)^{n+1}, \quad 0 < t < n.$$

One way to do this is to differentiate the function $a \mapsto \left(1 - \frac{t}{a}\right)^a$ and examine the sign of the derivative. Another method is to apply the arithmetic mean-geometric mean inequality to the $n+1$ numbers

$$1, 1 - \frac{t}{n}, 1 - \frac{t}{n}, \dots, 1 - \frac{t}{n}.$$

The monotone convergence theorem gives

$$\int_0^\infty e^{-t}t^{a-1}dt = \lim_{n\to\infty} \int_0^n \left(1 - \frac{t}{n}\right)^n t^{a-1}dt.$$

Repeated integration by parts gives the desired result.

Aliter: One of the students has correctly applied the dominated convergence theorem and the argument proceeds as follows. Observe that

$$1 - t < e^{-t}, \quad t > 0.$$

Replace t by t/n and raise both sides to the power n to get

$$\left(1 - \frac{t}{n}\right)^n < e^{-t}.$$

Multiply by t^{a-1} and appeal to the dominated convergence theorem.

5. To prove that every $A \subset [0,1]$ of positive outer measure contains a non-measurable subset. Let E be the standard non-measurable set constructed in class and we may assume without loss of generality that it is contained in $[0,1/2]$ (well, if a point in it is greater than $1/2$ then subtract off a suitable rational to bring it inside $[0,1/2]$. Let E_j, $j = 1, 2, 3, \ldots$ be its rational translates. Then

$$\bigcup_j (A \cap E_j) = A$$

so that

$$0 < \mu^*(A) \leq \sum_j \mu^*(A \cap E_j)$$

showing that at-least one $\mu^*(A \cap E_j)$ must be strictly positive. We show that for every j $\mu_*(E_j) = 0$ from which we would get $\mu_*(A \cap E_j) = 0$. Again, we only look at the rational translates through rationals less than $1/2$ so that over these

$$\bigcup_j E_j \subset [0,1]$$

Using the super-additivity of inner measures,

$$1 = \mu_*([0,1]) \geq \sum_j \mu_*(E_j)$$

Now since all the E_j have the same inner measure (by translation invariance) we see that $\mu_*(E_j) = 0$ for every j.

Lebesgue Stieltjes measures on the real line

This was merely defined earlier in class for completeness and we did not work with it. The definition I gave was straight out of the book by A. Friedman [1]. There are a few points in that need to be addressed. There are two equivalent[12] ways in which this is done in the literature. The standard and perhaps more appropriate procedure is to use coverings by open-closed intervals (see [//], [//] or [//]) while some authors use coverings by open intervals to define the outer measure. The latter approach calls for some caution and the purpose of this note is to elaborate these points. Owing to its importance we have isolated the discussion on Lebesgue Stieltjes measures as a separate chapter.

The idea of using open coverings to define outer measure gets hopelessly complicated in dealing with Stieltjes measures on $\mathbb{R}^n$ though the standard Lebesgue measure can still be dealt with using open coverings. This is carried out for example in A. E. Taylor's book [//]. The treatment may be simplified using the Lebesgue number for coverings as we show below. The multi-dimensional Lebesgue Stieltjes measures is taken up in a later chapter after some set theoretic preliminaries.

Definition: Let F be a monotone right continuous function defined on the real line. For a covering of a subset A by a sequence of open intervals $\Big\{(a_n, b_n)\Big\}_n$ we consider $\sum_{n=1}^{\infty}(F(b_n) - F(a_n))$ and define

$$\mu_F^*(A) = \inf \sum_{n=1}^{\infty}(F(b_n) - F(a_n)),$$

the infimum being taken over all possible coverings by open intervals.

The set function μ_F^* is called the Lebesgue Stieltjes outer measure. It is easy to check that this is an outer measure and indeed a metric outer measure. All this parallels closely the proofs of the corresponding results for Lebesgue measure on the real line. We now come to some important and serious points that need more careful attention. First we make the following

Remark: It is clear that for an open interval (a, b),

$$\mu_F^*(a, b) \leq F(b) - F(a).$$

However, we shall show presently that strict inequality may hold. The reader ought to first contemplate on how he would establish directly from the infimum definition that the usual Lebesgue outer measure of (a, b) is $b - a$.

Theorem: The outer measure of $[a, b]$ equals $F(b) - F(a-)$, where $F(a-)$ denotes as usual the limit of $F(x)$ as $x \to a$ from the left.

[12] We shall always assume that the monotone functions used are right continuous.

Proof: Consider the covering by the single open interval $(a - \frac{1}{n}, b + \frac{1}{n})$

$$\mu_F^*([a,b]) \leq F(b + \frac{1}{n}) - F(a - \frac{1}{n}) \longrightarrow F(b) - F(a-)$$

To get the reverse inequality, let $\epsilon > 0$ and $\Big\{(a_n, b_n)\Big\}$ be a sequence of open intervals covering $[a,b]$ such that

$$\mu_F^*([a,b]) + \epsilon > \sum_{n=1}^{\infty}(F(b_n) - F(a_n))$$

We extract a finite subcover $\Big\{(a_1,b_1),(a_2,b_2),\ldots,(a_N,b_N)\Big\}$ with the property that none of the intervals are contained in the union of others and ordered so that

$$a_1 < a < b_1, a_2 < b_1 < b_2, \ldots, a_N < b_{N-1} < b < b_N$$

Noting that when $a < c$, $F(a) \leq F(c-)$ we get

$$\begin{aligned}
\mu_F^*([a,b]) + \epsilon &> F(b_1) - F(a_1) + F(b_2) - F(a_2) + \ldots + F(b_N) - F(a_N) \\
&\geq F(b_1) - F(a-) + F(b_2) - F(a_2) + \ldots + F(b_N) - F(a_N) \\
&\geq F(a_2) - F(a-) + F(b_2) - F(a_2) + \ldots + F(b_N) - F(a_N) \\
&= F(b_2) - F(a-) + F(b_3) - F(a_3) + \ldots + F(b_N) - F(a_N) \\
&\geq \ldots \quad \ldots \quad \ldots \\
&\geq F(b_N) - F(a-) \geq F(b) - F(a-).
\end{aligned}$$

Corollary: $\mu_F^*((a,b]) = F(b) - F(a)$

Proof: First using monotonicity of outer measure, we get

$$\mu_F^*((a,b]) \geq \mu_F^*([a + \frac{1}{n}, b]) = F(b) - F(a + \frac{1}{n}) \longrightarrow F(b) - F(a)$$

On the other hand invoking the definition of the outer measure we get for a cover consisting of the single open interval $(a, b + \frac{1}{n})$,

$$\mu_F^*((a,b]) \leq F(b + \frac{1}{n}) - F(a) \longrightarrow F(b) - F(a).$$

Corollary: $\mu_F^*(\{a\}) = F(a) - F(a-)$.

Proof: Let $\epsilon > 0$ be arbitrary. Then there is a covering by open intervals $\Big\{(a_n, b_n)\Big\}$ such that

$$\mu_F^*(\{a\}) + \epsilon > \sum_{n=1}^{\infty}(F(b_n) - F(a_n)) > F(b_1) - F(a_1),$$

where we assume that $a \in (a_1, b_1)$. But then $F(b_1) \geq F(a)$ and $F(a_1) \leq F(a-)$ from which follows

$$\mu_F^*(\{a\}) \geq F(a) - F(a-)$$

To get the reverse inequality,

$$\mu_F^*(\{a\}) \leq F(a + \frac{1}{n}) - F(a - \frac{1}{n}) \longrightarrow F(a) - F(a-).$$

Atoms: A point of a measure space having positive measure is called an atom. For example the Lebesgue measure on the real line has no atoms. We have just proved that the atoms of the Lebesgue Stieltjes measure defined by a right continuous function F are precisely the discontinuities of F.

Theorem: $\mu_F^*((a, b)) = F(b-) - F(a)$

Proof: There seems to be no direct way of arriving at this straight from the definition. Infact this seems to be the case for the usual Lebesgue measure as well. We use the fact that (being a metric outer measure) all Borel sets are measurable.

$$\mu_F^*((a, b)) = \mu_F^*((a, b]) - \mu_F^*(\{b\}) = F(b) - F(a) - (F(b) - F(b-)) = F(b-) - F(a).$$

In particular it must be noted that $\mu_F^*((a, b))$ is strictly smaller than $F(b) - F(a)$ in case the point b is a point of discontinuity of F.

Integration with respect to Lebesgue Stieltjes measure: We now drop the superscript $*$ from μ_F^* while dealing with μ_F^*-measurable subsets. The reader must review the theory of Riemann Stieltjes integrals as laid down in [4]. In particular it is known that if F is monotone and ϕ is continuous on $[a, b]$ then

$$\int_a^b \phi dF$$

exists. The value of the integral is the common value of $\inf U(\phi, \Pi, F)$ and $\sup L(\phi, \Pi, F)$ where the infimum and supremum are taken over all partitions $\Pi = \{a = t_0, t_1, \ldots, t_n = b\}$ of the closed interval $[a, b]$ and $U(\phi, \Pi, F)$ and $L(\phi, \Pi, F)$ are the upper and lower Riemann sums given by

$$U(\phi, \Pi, F) = \sum_{j=1}^{n} \Big(\sup_{t_{j-1} \leq t \leq t_j} \phi(t) \Big) (F(t_j) - F(t_{j-1}))$$

$$L(\phi, \Pi, F) = \sum_{j=1}^{n} \Big(\inf_{t_{j-1} \leq t \leq t_j} \phi(t) \Big) (F(t_j) - F(t_{j-1}))$$

Beyond $[a, b]$ extend F as a constant function taking value $F(b)$ on $[b, \infty)$ and value $F(a)$ on $(-\infty, a]$. We shall continue to assume that F is right continuous $[a, b]$ and hence on the whole real line. It may be useful to recall the general definition of the Lebesgue integral for a bounded function on a finite measure space in terms of upper and lower Lebesgue-Darboux sums. The numbers $U(\phi, \Pi, F)$ and $L(\phi, \Pi, F)$ are seen to be special cases of Lebesgue-Darboux sums for the measurable partition[13] $\{\{a\}, (a, t_1], (t_1, t_2], \dots, (t_{n-1}, b]\}$ since

$$F(t_j) - F(t_{j-1}) = \mu_F^*((t_{j-1}, t_j])$$

and hence thanks to our general theorem on the integral of bounded measurable functions on a finite measure space, the Riemann Stieltjes integral $\int_a^b \phi dF$ agrees with the Lebesgue integral of ϕ with respect to the measure μ_F namely

$$\int_a^b \phi dF = \int_{[a,b]} \phi d\mu_F$$

It is known [?] that the computation of the Stieltjes integral $\int_a^b \phi dF$ reduces to a combination of computing Riemann integrals and summing infinite series.

Distribution functions: Aside from integrals reducible to infinite series, the only efficient method known to compute integrals is the fundamental theorem of calculus for Riemann integrals and integrals reducible to Riemann integrals on the real line. Multiple integrals in elementary calculus courses are evaluated as repeated integrals thereby reducing the problem to computing Riemann integrals over the real line. It is thus remarkable that the integral

$$\int_X f d\mu$$

over fairly arbitrary measure spaces can be reduced to Riemann Stieltjes integrals as we now proceed to show. Indeed this reduction is one of the versatile tools in the probabilist's tool-kit.

Definition: Let X be a measure space and $f \in L^1(\mu)$ be a real valued function. The distribution function F corresponding to f is the real valued function

$$F(t) = \mu\Big(\{x : f(x) < t\}\Big)$$

The properties of the distribution function are summarized in the following

Theorem: The distribution function F is monotone increasing bounded right continuous function.

Proof: Denoting by E_t the set $\{x : f(x) < t\}$ and integrating over this set,

$$\int_{E_t} f d\mu \leq t\mu(E_t)$$

[13]The singleton $\{a\}$ has measure zero since F is continuous at a.

More exercises:

1. Show that if $P(x)$ is a polynomial and $a > 0$ then $e^{-ax^2}P(x) \in L^p(\mathbb{R})$ for every $p \in [1, \infty]$. Is this dense in $L^p(\mathbb{R})$? Let $H_n(x)$ denote the polynomial

$$\exp(x^2)\Big(\frac{d}{dx}\Big)^n \exp(-x^2)$$

is an orthogonal family of functions in $L^2(\mathbb{R})$ with respect to the measure $e^{-x^2}dx$. Prove that the linear span of the set $\Big\{H_n(x)e^{-x^2}\Big\}_{n=0}^{\infty}$ is dense in $L^2(\mathbb{R})$.

Note: The polynomials $H_n(x)$ are called the Hermite polynomials.

2. Show that the polynomials $P_n(x) = \dfrac{1}{2^n n!}\Big(\dfrac{d}{dx}\Big)^n(1-x^2)^n$ form an orthogonal family in $L^2[-1, 1]$. Show that the vector space generated by them dense in $L^2[-1, 1]$.

Note: A lovely proof of this is available in Hewitt and Stromberg [//].

3. Show that if X is a finite measure space and $1 \le p < q \le \infty$ then $L^q \subset L^p$ and that the inclusion is strict. Infact show using Hölder's inequality that

$$\|f\|_p \le \|f\|_q\Big(\mu(X)\Big)^{\frac{1}{p}-\frac{1}{q}}$$

and hence that the mean values $M_p(f) = \Big(\dfrac{1}{\mu(X)}\displaystyle\int_X |f|^p d\mu\Big)^{1/p}$ increase with p.

4. **Interpolation inequality:** Let $f \in L^p(\mu) \cap L^q(\mu)$ for some $p < q$ then show that for every r with $p < r < q$, $f \in L^r(\mu)$. The measure space is not assumed to be finite. Prove using Hölder's inequality that if

$$\frac{1}{r} = \frac{\alpha}{p} + \frac{1-\alpha}{q}, \quad 0 < \alpha < 1,$$

then,

$$\|f\|_r \le \|f\|_p^{\alpha}\;\|f\|_q^{1-\alpha}$$

In other words the function

$$\xi \mapsto \log\|f\|_{1/\xi}$$

is a convex function of ξ on the interval $[1/q, 1/p]$.

5. Show that when $1 \le p < q \le \infty$ the spaces L^p and L^q are distinct in either of the following two cases.

(i) [0, 1] with Lebesgue measure and

(ii) $\mathbb{R}$ with Lebesgue measure.

More over show that in case of (ii) neither space is contained in the other.

6. Show that $L^\infty[0, 1]$ is NOT equal to $\displaystyle\bigcap_{1 \le p < \infty} L^p[0, 1]$

7. Show that when $1 \leq p < q$, the subset $L^q[0,1]$ is dense and of the first category in $L^p[0,1]$. Hint: Look at the sets $E_n = \{f \in L^q : \|f\|_q \leq n\}$.

8. Do problem 15 on page 59 of Rudin's Real and complex analysis (Third edition)

9. Do problem 9 on page 32 of Rudin's book cited above.

10. Prove that if $f \in L^p[0,1]$ with $1 \leq p < \infty$ is such that $\int_0^1 f(x)x^n dx = 0$ for all n then $f = 0$ almost everywhere. Is the result true when $p = \infty$?

11. Suppose that $f \in L^2[0,1]$ and $k(x,y)$ is continuous on $[0,1] \times [0,1]$ then the map

$$Kf : x \mapsto \int_0^1 k(x,y)f(y)dy$$

is continuous on $[0,1]$. Thus K maps $L^2[0,1]$ into $C[0,1]$. Show that K is a continuous function with respect to the L^2 norm on the domain and the L^∞ norm on the co-domain. Show that K maps the closed unit ball of $L^2[0,1]$ onto a compact set in $C[0,1]$.

12. Under the set up of the previous exercise, noting that $C[0,1] \subset L^2[0,1]$, is it true that K is continuous as a map from $L^2[0,1]$ into $L^2[0,1]$? Is the image of the closed unit ball compact in $L^2[0,1]$?

13. Suppose that $f \in L^1(\mathbb{R})$, use the method of variation of parameters to obtain the solution with zero initial conditions of the ODE

$$y'' + k^2 y = f(x), \quad k > 0,$$

as an integral. Show that this solution lies in $L^\infty(\mathbb{R}) \cap C(\mathbb{R})$. After subtracting off $A\cos kx + B\sin kx$ for suitable A and B, there exist solutions $y_1(x)$ and $y_2(x)$ of the ODE such that

$$y_1(x) = e^{ikx} + o(1), \quad y_2(x) = e^{-ikx} + o(1).$$

Show that these must be linearly independent.

14. Let $q(x) \in L^1(\mathbb{R})$ and consider the ordinary differential equation

$$y'' + q(x)y = -k^2 y$$

Show that when $|k| > \|q\|_{L^1}$, for every choice of constants c_1 and c_2 there is a solution $y(x)$ in $L^\infty(\mathbb{R})$ such that $y(0) = c_1$ and $y'(0) = c_2$ by obtaining an integral operator and appealing to the fixed point property. Show that through suitable choices of c_1 and c_2 we can find linearly independent solutions $y_1(x)$ and $y_2(x)$ such that

$$y_1(x) = e^{ikx} + o(1), \quad y_2(x) = e^{-ikx} + o(1), \quad x \longrightarrow \infty.$$

Note: When $q \in L^1 \cap L^2$ one would like to know if there are solutions in L^2. This is an important but difficult question with applications to scattering theory and the study of the Korteweg de Vries PDE

$$u_t = u_{xxx} + uu_x.$$

15. Suppose that $f \in L^1[0,1]$ and $\|f\| = 1$. Show that there exists g and h in $L^1[0,1]$ such that $f = \frac{1}{2}(g+h)$, $\|g\| = \|h\| = 1$ and $\mu(\{x : g(x) \neq h(x)\}) > 0$.

16. Suppose that $(X, \mathcal{B}, \mu)$ is a finite measure space and $\{f_n\}$ is a sequence of functions in $L^2(\mu)$ converging to f weakly. That is for each $g \in L^2(\mu)$,

$$\int_X f_n g d\mu \longrightarrow \int_X f g d\mu$$

Assume that $\|f_n\|$ converges to $\|f\|$. Then show that $\{f_n\}$ converges to f in L^2.

17. If f_n converges to f in L^p, $1 \leq p < \infty$ and $g_n \longrightarrow g$ pointwise almost everywhere and $\|g\|_\infty \leq M$ for all n then show that $f_n g_n \longrightarrow fg$ in L^p.

18. If f_n is a sequence in L^p $(1 < p < \infty)$ converging pointwise almost-everywhere to f and $\|f\|_p \leq M < \infty$ for all n then show that f_n converges to f weakly.

19. Show that if f_n is a sequence of functions in L^p converging to f weakly then

    (i) There exists $M < \infty$ such that $\|f_n\|_p \leq M$ for all n.

    (ii) $\|f\|_p \leq \liminf \|f_n\|_p$

    Hint: For the first part use Banach-Steinhaus.

20. Prove that if $1 < p < \infty$ and f_n is a sequence of functions in L^p converging weakly to f and $\|f_n\|_p \longrightarrow \|f\|_p$ then f_n converges to f in L^p. Hint: May assume that $\|f\|_p = 1$ and then the preceding exercise shows that $\left\|\frac{f_n + f}{2}\right\|_p \longrightarrow 1$. Invoke uniform convexity of the unit ball.

Weierstrass' approximation theorem and Fourier series

This is a short chapter in which we prove the classical theorem of Weierstrass on polynomial approximations of continuous functions by polynomials. There are numerous proofs of this result but the proof given here is due to Weierstrass himself using the heat kernel[14]. The method of proof is very important in several applications particularly in the theory of differential equations and Fourier analysis. Another elementary proof is sketched in exercise [//]. The theorem of Weierstrass is then used it to establish a few basic results in the theory of Fourier series. We end the chapter with a theorem of H. Weyl that sharpens Kronecker's theorem on the density in $[0, 1]$ of the fractional parts $\{n\theta\}_n$, where θ is an irrational number.

The Weierstrass' aprroximation theorem: Let f be a continuous function f on a closed bounded interval I. Given any $\epsilon > 0$, there is a polynomial P such that

$$\sup_{x\in I} |f(x) - P(x)| < \epsilon.$$

Proof: We begin by extending f continuously on the whole real line with support in $[-c, c]$. The extension is uniformly continuous and M will be used for the supremum of $|f|$ on tbe real line. We begin by recalling the definition of the heat kernel $G(x, f)$:

$$G(u,t) = \frac{1}{\sqrt{4\pi t}} \exp(-u^2/4t),$$

which has the property that $\int_{\mathbb{R}} G(u,t)du = 1$. The substitution $u - x = \sqrt{4t}v$ gives

$$\lim_{t\to 0+} \frac{1}{\sqrt{\pi}} \int_{\mathbb{R}} f(x + \sqrt{4t}v) \exp(-v^2) dv = f(x) \tag{550}$$

The idea is to replace the exponential in the last integral by its Taylor polynomial. However some care is needed as the power series expansion for the exponential function does not converges uniformly on the whole real line but only on compact subsets. We first choose $h > 0$ such that

$$\int_{\mathbb{R}-[-h,h]} e^{-u^2} du < \frac{\epsilon}{8M},$$

whereby we get for for all $x \in [-c, c]$ and $t > 0$ the estimate

$$\Big|\frac{1}{\sqrt{\pi}} \int_{\mathbb{R}-[-h,h]} (f(x + \sqrt{4t}v) - f(x))e^{-v^2} dv\Big| < \frac{\epsilon}{4}.$$

Now since

$$\frac{1}{\sqrt{\pi}} \int_{\mathbb{R}} e^{-u^2} f(x)du = f(x),$$

[14]Actually Weierstrass [5] uses more general kernels, first citing the heat kernel and then abstracting the requisite prperties.

we get upon splitting the domain of integration into two pieces,

$$\sup_{x\in[-c,c]}\Big|f(x)-\frac{1}{\sqrt{\pi}}\int_{\mathbb{R}} f(x+\sqrt{4t}v)e^{-v^2}dv\Big| \le \sup_{x\in[-c,c]}\frac{1}{\sqrt{\pi}}\int_{-h}^{h}|f(x+\sqrt{4t}v)-f(x)|e^{-v^2}dv+\frac{\epsilon}{4} \tag{551}$$

By uniform continuity of f there is a $\delta > 0$ such that

$$|f(x)-f(y)| < \frac{\epsilon}{4}, \quad |x-y| < \delta.$$

If we take t_0 such that $|h\sqrt{4t_0}| < \delta$ then (551) yields

$$\sup_{x\in[-c,c]}\Big|f(x)-\frac{1}{\sqrt{\pi}}\int_{\mathbb{R}} f(x+v\sqrt{4t_0})e^{-v^2}dv\Big| < \frac{\epsilon}{2} \tag{552}$$

whereby we get the estimate

$$\sup_{x\in[-c,c]}\Big|f(x)-\frac{1}{\sqrt{\pi}}\int_{-h}^{h} f(x+v\sqrt{4t_0})e^{-v^2}dv\Big| < \frac{\epsilon}{2}+\frac{M}{\sqrt{\pi}}\int_{\mathbb{R}-[-h,h]} e^{-v^2}dv < \frac{\epsilon}{2}+\frac{\epsilon}{8}. \tag{553}$$

If we choose a Taylor polynomial $P(v)$ of e^{-v^2} such that

$$\sup_{[-h,h]}\Big|e^{-v^2}-P(v)\Big| \le \frac{\epsilon}{8Mh},$$

then we get for all $x\in[-c,c]$,

$$\Big|\frac{1}{\sqrt{\pi}}\int_{-h}^{h} f(x+v\sqrt{4t_0})P(v)dv-\frac{1}{\sqrt{\pi}}\int_{-h}^{h} f(x+v\sqrt{4t_0})e^{-v^2}dv\Big| < \frac{\epsilon}{4}.$$

Using this in conjunction with (553) gives

$$\sup_{x\in[-c,c]}\Big|f(x)-\frac{1}{\sqrt{\pi}}\int_{-h}^{h} f(x+v\sqrt{4t_0})P(v)dv\Big| < \epsilon \tag{554}$$

The change of variables $x+v\sqrt{4t_0}=y$ shows that the expression

$$\frac{1}{\sqrt{\pi}}\int_{-h}^{h} f(x+v\sqrt{4t_0})P(v)dv$$

is a polynomial $Q(x)$ which satisfies by (554)

$$\sup_{x\in[-c,c]}\Big|f(x)-Q(x)\Big| < \epsilon$$

Remark: We note that a similar argument would yied the multi-dimensional version of the Weierstrass' approximation theorem which states that given a continuous function f on a closed bounded box J in $\mathbb{R}^n$, for every $\epsilon > 0$ there exists a polynomial P in n variables such that

$$\sup_{x\in J}|f(x)-P(x)| < \epsilon.$$

One would use for this purpose the multi-dimensional heat kernel

$$G_n(x,t) = (4\pi t)^{-n/2}\exp(-\|x\|^2/4t)$$

Definition (Trigonometric polynomials:) A trigonometric polynomial is a finite linear combination

$$a_0 + \sum_{j=1}^{n}(a_j \cos jx + b_j \sin jx).$$

If a_n or b_n is non-zero we say that n is the degree of the trigonometric polynomial. The set of all trigonometric polynomials is a vector subspace of the space of all continuous $2\pi-$periodic functions on the real line.

It is readily checked that $\cos^k x$ and $\sin^k x$ are trigonometric polynomials for every non-negative integer k. Also $T(x+\alpha)$ and $T(x)\sin x$ are trigonometric polynomials whenever $T(x)$ is. The reader can check that neither $\sin^{1/3} x$ nor $\sin\sqrt{3}x$ is a trigonometric polynomial.

We shall now prove the following version of the Weierstrass' approximation theorem for continuous $2\pi-$periodic functions of the real line.

Theorem: Suppose that f is a continuous $2\pi-$periodic function on the real line. Then given any $\epsilon > 0$, there is a trigonometric polynomial T such that

$$\sup_{x\in\mathbb{R}} |f(x) - T(x)| < \epsilon.$$

Proof: First we prove the result for an even continuous $2\pi-$periodic function f. We first consider the restriction of f to $[0,\pi]$. Applying the Weierstrass' approximation theorem on $[-1,1]$ to the continuous function $f \circ \cos^{-1}$ we get a polynomial P such that

$$\sup_{|y|\leq 1} |f(\cos^{-1} y) - P(y)| < \epsilon,$$

which is equivalent to

$$\sup_{0\leq x\leq \pi} |f(x) - P(\cos x)| < \epsilon.$$

Since both $f(x)$ and $P(\cos x)$ are even functions the estimate is valid on the whole interval $[-\pi,\pi]$ and so by 2π-periodicity, on the whole real line. For an arbitrary $2\pi-$periodic function f both $(f(x) + f(-x))\sin^2 x$ and $(f(x) - f(-x))\sin x$ are even $2\pi-$periodic functions and so there exist trigonometric polynomials $T_1(x)$ and $T_2(x)$ such that

$$\sup_{x\in\mathbb{R}} |(f(x) + f(-x))\sin^2 x - T_1(x)| < \frac{\epsilon}{4}, \quad \sup_{x\in\mathbb{R}} |(f(x) - f(-x))\sin x - T_2(x)| < \frac{\epsilon}{4}$$

From the second we get,

$$\sup_{x\in\mathbb{R}} |(f(x) - f(-x))\sin^2 x - T_2(x)\sin x| < \frac{\epsilon}{4}$$

Hence

$$\sup_{x\in\mathbb{R}} |f(x)\sin^2 x - \frac{1}{2}(T_1(x) + T_2(x)\sin x)| < \frac{\epsilon}{4} \tag{555}$$

If we apply the last result to $f(x+\frac{\pi}{2})$ in place of $f(x)$ we get a trigonometric polynomial $T_3(x)$ such that

$$\sup_{x\in\mathbb{R}} |f(x+\frac{\pi}{2})\sin^2 x - T_3(x)| < \frac{\epsilon}{4}$$

and so writing $T_4(x)$ in place of $T_3(x-\frac{\pi}{2})$, we get

$$\sup_{x\in\mathbb{R}} |f(x)\cos^2 x - T_4(x)| < \frac{\epsilon}{4} \tag{556}$$

Adding (555) and (556) we get finally

$$\sup_{x\in\mathbb{R}} |f(x) - \frac{1}{2}(T_1(x) + T_2(x)\sin x) - T_4(x)| < \frac{\epsilon}{2}$$

Corollary: If $1 \le p < \infty$ then the space of trigonometric polynomials is dense in $L^p[-\pi,\pi]$.

Proof: The proof is easy except for an annoying "end-correction" we need to make in order to pass on from continuous functions to $2\pi-$periodic continuous functions. Let $f \in L^p[-\pi,\pi]$ and $\epsilon > 0$ be arbitrary. Luzin's theorem gives us a continuous function $g \in [-\pi,\pi]$ such that

$$\|f-g\|_p < \epsilon/3$$

The problem is that if $g(\pi) \neq g(-\pi)$ then g would not extend continuously to the whole real line. To remedy this we do an end-correction on g modifying it on the two ends. Let M be an upperbound for $|g|$ on $[-\pi,\pi]$ and $\delta < \epsilon^2/36(2M)^p$. Then define h such that $h = g$ on $[-\pi+2\delta, \pi-2\delta]$ and zero on $[-\pi,-\pi+\delta]\cup[\pi-\delta,\pi]$. Extend h continuously to the whole of $[-\pi,\pi]$ with the same bound M. Now

$$\int_{-\pi}^{\pi} |g(x)-h(x)|^p dx \le 4\delta(2M)^p < \epsilon^2/9,$$

whereby

$$\|f-h\|_p < 2\epsilon/3$$

Since h vanishes at the endpoints $\pm\pi$ we can extend h continuously as a $2\pi-$ periodic function on the whole real line. Appealing now to the previous theorem, there is a trigonometric polynomial $T(x)$ such that

$$\|h-T\|_p < \epsilon/3,$$

which gives finally

$$\|f-T\| < \epsilon.$$

Theorem of H. Weyl on equidistribution of $\{n\alpha\}$:

The theorem of Kronecker asserts that the sequence of fractional parts $\{n\alpha\}$, where $\alpha \in \mathbb{R}-\mathbb{Q}$ is dense in $[0,1]$. This theorem has been considerably sharpened by H. Weyl namely the points of the sequence appear in any given interval with an asymptotic frequency equal to the length of the interval. We shall provide a proof of Weyl's result. Our proof follows closely the presentation in K. Jacobs [//].

We begin with a few preliminaries on the equivalence of the measure spaces $(S^1, \mathcal{B}, \tilde{\mu})$ and $([0,1), \mathcal{B}, \mu)$ where S^1 denotes the unit circle in the compelx plane and $\mathcal{B}$ denotes the Borel sigma algebra in both cases, μ denotes the Lebesgue measure on $[0,1)$ and $\tilde{\mu}$ unique regular Borel measure on S^1 such that

$$\tilde{\mu}(A) = (2\pi)^{-1}\text{length}(A).$$

The factor $(2\pi)^{-1}$ is a normalizing factor due to which, S^1 receives unit measure. The map

$$\phi : [0,1) \longrightarrow S^1$$

given by $\phi(t) = \exp(2\pi i t)$ is bijective continuous, and maps $(0,1)$ homeomorphically onto $S^1 - \{1\}$. The map is thus "almost" a homeomorphism. If A is a Borel subset of S^1 then $\phi^{-1}(A - \{1\})$ is a Borel subset of $(0,1)$ and so $\phi^{-1}(A)$ is a Borel subset of $[0,1)$. Moreover we see that

$$\mu(\phi^{-1}(A)) = \tilde{\mu}(A)$$

since this is true for all open arcs $A \subset S^1$. Note that if f is any real valued Borel measurable function on S^1 then $f \circ \phi^{-1}$ is also Borel measurable. Further we see that if f is intergable then so is $f \circ \phi^{-1}$ and

$$\int_{[0,1)} f d\mu = \int_{S^1} f \circ \phi^{-1} d\tilde{\mu}.$$

In view of these observations, we shall make no distinction between the measure spaces $(S^1, \mathcal{B}, \tilde{\mu})$ and $([0,1), \mathcal{B}, \mu)$ and denote $f \circ \phi^{-1}$ simply by f. Let us now consider ϕ as a map from $\mathbb{R} \longrightarrow S^1$. If F is a continuous function on S^1 then $F \circ \phi$ is a continuous periodic function on the real line with period one. The converse is also true and is left as an exercise.

For the rest of the section we shall always assume that $\alpha \in \mathbb{R} - \mathbb{Q}$ and that $T : S^1 \longrightarrow S^1$ denotes the rotation of the circle:

$$T(\zeta) = (\exp 2\pi i \alpha)\zeta, \quad \zeta \in S^1.$$

Then, we may view T may be viewed as a map $\tilde{T} : [0,1) \longrightarrow [0,1)$ given by

$$\tilde{T}(x) = (x + \alpha)\text{mod}(1),$$

which can be expressed in terms of ϕ as $\tilde{T} = \phi \circ T \circ \phi^{-1}$. Hereafter we shall make no distiction between T and $\tilde{T}$ and denote them both by T.

Theorem: For any continuous periodic function f on the real line with period one,

$$\lim_{n\to\infty} \frac{1}{n}\Big(f(x) + f(Tx) + \ldots + f(T^{n-1}x)\Big) = \int_0^1 f(x)dx \qquad (580)$$

the convergence being uniform.

Proof: We use the notation $T_n f$ to denote

$$\frac{1}{n}\Big(f(x) + f(Tx) + \ldots + f(T^{n-1}x)\Big)$$

and it is immediate that both sides of (580) are linear in f. The result is trivial for constant functions and for $f(x) = \exp(2\pi ikx)$ with $k \neq 0$, we see that

$$T^j f(x) = \exp(2\pi ij\alpha)\exp(2\pi ikx)$$

and the series $f(x) + f(Tx) + \ldots + f(T^{n-1}x)$ is a finite geometric series whose sum is

$$\frac{e^{2\pi ikx}(1 - \exp(2\pi ikn\alpha))}{1 - \exp(2\pi ik\alpha)},$$

from which follows

$$\sup_x |T_n(x)| \leq \frac{2}{n}\frac{1}{|1 - \exp(2\pi ik\alpha)|} \longrightarrow 0.$$

The integral on the right hand side of (580) is obviously zero and the result holds in this case. By linearity the result holds for all finite linear combinations of $\exp(2\pi ikx)$ which will be referred to as trigonometric poynomials of period one. These are dense in the space of all continuous functions on the real line with period one. Now let f be a continuous function with period one and $\epsilon > 0$ be arbitrary. Pick a trigonometric polynomial P such that

$$\sup_x |f(x) - P(x)| < \epsilon/3.$$

Then, we see immediately that

$$\sup_x |T_n f(x) - T_n P(x)| < \epsilon/3.$$

Applying the result for P, there exists and $n_0 \in \mathbb{N}$ suhc that for all $n > n_0$,

$$\sup_x \left|T_n P(x) - \int_0^1 P(x)dx\right| < \epsilon/3.$$

Now a simple 3ϵ argument gives, for $n > n_0$,

$$\sup_x \left|T_n f - \int_0^1 f(x)dx\right| < \frac{2\epsilon}{3} + \int_0^1 |f(x) - P(x)|dx < \epsilon,$$

completing the proof.

We now wish to extend this further but we can no longer expect the convergence to be uniform. Exercise [//] concerns convergence in L^1 of $T_n f$ in case $f \in L^1[0,1]$. It is true, but much harder to prove that $T_n f$ converges pointwise for every $f \in L^1$. We shall however do this for Riemann integrable functions.

Theorem: For any Riemann integrable function f on $[0,1]$,

$$\lim_{n\to\infty} \frac{1}{n}\Big(f(x) + f(Tx) + \ldots + f(T^{n-1}x)\Big) = \int_0^1 f(x)dx \tag{581}$$

pointwise almost everywhere.

Proof: We begin with the case when f is the characteristic function of an interval. For an interval $[a, b] \subset (0, 1)$, let $\{f_n\}$ be an increasing sequence of continuous functions with support in (a, b) converging to the characteristic function f of (a, b) and $\{g_n\}$ be a decreasing sequence of continuous functions with support in $(0, 1)$ converging to the characteristic function g of $[a, b]$. Then we have the inequalities

$$T_n f_k \leq T_n f \leq T_n g_k. \tag{582}$$

Let $\epsilon > 0$ be arbitrary and choose k such that

$$\int_0^1 |f_k(x) - g_k(x)| dx < \frac{\epsilon}{6} \tag{583}$$

Since the functions f_k and g_k vanish at the ends 0 and 1 they extend continuously as one periodic functions on the whole real line whereby

$$\lim_n T_n f_k = \int_0^1 f_k(x) dx, \quad \lim_n T_n g_k = \int_0^1 g_k(x) dx$$

uniformly on $[0, 1]$. Choose an n_0 such that for $n \geq n_0$,

$$\sup_x \left| T_n f_k - \int_0^1 f_k(x) dx \right| < \epsilon/6, \quad \sup_x \left| T_n g_k - \int_0^1 g_k(x) dx \right| < \epsilon/6$$

Observe that since f and g differ only at two points, $T_n f(x) = T_n g(x)$ except on a countable set $E \subset [0, 1]$. On E^c we have the inequalities,

$$T_n f_k \leq T_n f = T_n g \leq T_n g_k.$$

For $x \in E^c$ and $n \geq n_0$, the difference $|T_n f_k - T_n g_k|$ is estimated by

$$\left| T_n f_k(x) - \int_0^1 f_k \right| + \left| T_n g_k(x) - \int_0^1 g_k \right| + \int_0^1 |f_k - g_k| < \frac{\epsilon}{2}.$$

From this we get for $x \in E^c$ and $n \geq n_0$,

$$\begin{aligned} \left| T_n f(x) - \int_0^1 f \right| &\leq |T_n f_k(x) - T_n f(x)| + \left| T_n f_k(x) - \int_0^1 f_k \right| + \int_0^1 |f_k - f| \\ &\leq |T_n f_k(x) - T_n g_k(x)| + \frac{\epsilon}{6} + \int_0^1 |f_k - g_k| \\ &\leq \frac{\epsilon}{2} + \frac{\epsilon}{6} + \frac{\epsilon}{6} < \epsilon. \end{aligned}$$

Now if the left end a of the interval is 0 then we choose the same approximating sequence f_n from below but choose the decreasing sequence $\{g_n\}$ of continuous functions in such a way that

$$g_k(0) = g_k(1), \quad \text{for all } k$$

with g_k converging pointwise to the characteristic function of $[a,b] \cup \{1\}$ and the above argument would go through. The case $b = 1$ is similar. By passing to finite linear combinations we see that (581) holds for all step functions.

An arbitrary Riemann integrable function can be approximated above and below by step functions whose difference has arbitrarily small integral. So we go through the above steps with f_k and g_k as step functions instead of continuous functions in (582)-(583) to get the result for Riemann integrable functions.

For further discussion on Weyl's theorem see the book by K. Chandrasekharan [1].

Basic theory of Fourier series:

Let $f \in L^1[-\pi, \pi]$ and consider the problem of expressing f as a trigonometric series

$$a_0 + \sum_{j=1}^{\infty} (a_j \cos jx + b_j \sin jx) \tag{600}$$

or equivalently in complex form

$$\sum_{k=-\infty}^{\infty} c_k e^{ikx}. \tag{601}$$

To begin with let us assume that the series (600) converges to $f(x)$ uniformly which would require that $f(x)$ be continuous. Multiplying (600) by $\cos kx$ and integrating over $[-\pi, \pi]$ we find that

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx dx, \quad k = 1, 2 \dots \tag*{(602)$_1$}$$

and similarly,

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx dx, \quad k = 1, 2 \dots \tag*{(602)$_2$}$$

The constant a_0 is given by

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) dx, \tag*{(602)$_3$}$$

and the coefficients in the complex form (601) are given by

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} dx, \quad k = 0, \pm 1, \pm 2, \dots \tag*{(602)$_4$}$$

We now proceed to the general definition:

Definition: The Fourier series of a function f in $L^1[-\pi, \pi]$ is defined to be the infinite series (600)-(601), where the coefficients are given by equations (602).

To indicate the relationship between the series and the function we shall write

$$f(x) \sim a_0 + \sum_{j=1}^{\infty} (a_j \cos jx + b_j \sin jx) \tag{603}$$

The symbol $\sim$ is not to be interpreted in any other way except that the coefficients a_0, a_k and b_k are obtained through equations (602). The central problem in the theory of Fourier series is to investigate the convergence of the series (603). Though we investigate several types of convergence, we shall be only touching upon the most elementary parts of the theory. The interested reader would do well to begin with the book [6].

Corollary: If $f \in L^1[-\pi,\pi]$ and all the Fourier coefficients of f vanish then $f = 0$ almost everywhere.

Proof: By Fejer's theorem the arithmetic means $\sigma_n f$ of the partial sums of the Fourier series converges to f in L^1. But $\sigma_n f = 0$ for every n and so $f = 0$ almost everywhere.

The Poisson summation formula: This is important result has several applications in number theory. We merely state the important theorem relegating the proof to the list of exercises. To begin with we take a continuously differentiable function f on the real line that satisfies an estimate of the form

$$f(x) \leq \frac{A}{1+|x|^\alpha}, \quad x \in \mathbb{R} \tag{650}$$

where A and α are positive constants and $\alpha > 1$. This shows in particular that $f \in L^1(\mathbb{R})$. We consider the function Φ_f given by

$$\Phi_f(x) = \sum_{n=-\infty}^{\infty} f(x + 2\pi n) \tag{651}$$

It is easy to see that Φ_f is a continuous $2\pi-$periodic function on the real line. The complex Fourier coefficients c_k of Φ_f turn out to be

$$c_k = \frac{1}{2\pi}\widehat{f}(k),$$

where $\widehat{f}$ is the Fourier transform of f namely

$$\widehat{f}(\xi) = \int_{\mathbb{R}} f(x)e^{-ix\xi}dx$$

If Φ_f is of class C^1 then the Fourier series of Φ_f converges pointwise to Φ_f and we get the following important result:

Theorem (Poisson summation formula): Assume that f is a continuously differentiable function on the real line satisfying an estimate (650) and that the function $\sum_{n=-\infty}^{\infty} f(x+2\pi n)$ is continuously differentiable. Then

$$\sum_{n=-\infty}^{\infty} f(x + 2\pi n) = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} \widehat{f}(n)e^{inx}$$

Odds and ends:

The subject of Fourier analysis is old and flourishing still. Many seemingly innocent questions have turned out to be rather deep and answered fairly recently. The purpose of this section is to provide the reader some examples of these theorems. We also gather here a few facts that are old and well-known but do not fit in with the theme of this book.

A question that arises naturally is whether a convergent series of sines and cosines is the Fourier series of an L^1 function. This is not the case and the following counter example-example was known long ago.

Theorem: The series $\sum_{n=2}^{\infty} \frac{\sin nx}{\ln n}$ is convergent but not the Fourier series of any L^1 function.

The convergence is not difficult to establish and is relegated as an exercise [//]. That it is not a Fourier series of any L^1 function follows from the following theorem of G. H. Hardy.

Theorem (Hardy): Suppose that $\{a_n\}$ is a sequence of numbers such that $a_n = O(1/n)$ and the arithmetic means

$$\sigma_n = \frac{1}{n}(s_0 + s_1 + \ldots + s_{n-1})$$

of partial sums $s_n = a_0 + a_1 + \ldots a_{n-1}$ converges then the series $\sum a_n$ converges.

Proof: See Whittaker and Watson [8] (pp 156-157) for the proof due to J. E. Littlewood.

We have seen that there exist continuous functions whose Fourier series diverges at a given point. The first such example is due to Paul du Bois Reymond. We have proved this using the Banach-Steinhaus' theorem. The proof can easily be modified so as to establish the existence of a continuous function whose Fourer series diverges at every point of a prescribed countable subset of $[-\pi, \pi]$.

The Lebesgue constants were so named by L. Fejer who also gave the principal term in the asymptotic expansion

$$L_n = \frac{4}{\pi^2} \ln n + O(1)$$

For a detailed commentary on the life and works of Fejer see the two articles of J. -P. Kahane and K. Tandori in the volume [//] dedicated to the 100th anniversary of Riesz and Fejer.

An old conjecture of N. Luzin states that if $f \in L^2[-\pi, \pi]$ then the Fourier series of f converges pointwise almost everywhere. The conjucture remain unsettled for more than 50 years. Meanwhile Kolmogorov proved in in 1928 (??) that there exist functions in $L^1[-\pi, \pi]$ whose Fourier series diverges everywhere. Luzin's conjecture was finally settled by Carleson in 1966. Hunt extended it to all $L^p[-\pi, \pi]$ with $p > 1$. A modern account of the Carleson-Hunt theorem is the book by Reyna [//].

A comprehensive account of Fourier series is the two volume set of A. Zygmund. The books of Katznelson is also delightful. The book of Pinsky [6] cited earlier is qutie accessible and can be read with profit.

References

[1] K. Chandrasekharan, *Introduction to analytic number theory,* Springer Verlag,

[2] Y. Katznelson, *An introduction to harmonic analysis,* second edition, Dover publications, New York, 1976.

[3] L. Lorch, *Asymptotic expressions for some integrals which include certain Lebesgue and Fejr constants,* Duke math. journal, **20**, (1953) 89-103

[4] B. Sz.-Nagy and J. Szabados (ed), *Functions, series, operators, Volume I,* Proceedings of the international conference organized by János Bolyai Mathematical Society, North Holland publishing company, Budapest 1983.

[5] A. Pinkus, *Weierstrass and approximation theory,*

[6] M. Pinsky,

[7] Juan Arias de Reyna, *Pointwise convergence of Fourier series,* Springer lecture notes in mathematics 1785, Springer Verlag,

[8] E. T. Whittaker and G. N. Watson, *A course of modern analysis,* Cambridge university press, Cambridge,

Exercises:

1. Given that f is a continuous function on the real line with two periods τ_1 and τ_2. What can you say about f if the periods τ_1 and τ_2 are linearly independent over $\mathbb{Q}$?

2. Is the function $\sin x + \sin \sqrt{3}x$ periodic?

3. Suppose we are given two trigonometric polynomials

$$T_1(x) = a_0' + \sum_{j=1}^{N} (a_j' \cos jx + b_j' \sin jx), \quad T_2(x) = a_0'' + \sum_{j=1}^{N} (a_j'' \cos jx + b_j'' \sin jx).$$

   such that $T_1(x) = T_2(x)$ for all $x \in [-\pi, \pi]$. Show that $a_j' = a_j''$ for $j = 0, 1, 2, \ldots, N$ and $b_j' = b_j''$ for $j = 1, 2, \ldots, N$.

4. Note that $\sin^2(x/2)$ is a trigonometric polynomial. Is there a $k \in \mathbb{N}$ such that $\sin^k(x/3)$ is a trigonometric polynomial?

5. Is $\sin \alpha x$ a trigonometric polynomial when α is irrational? What about $\sin^{1/3} x$?

6. Show that in a metric space every closed set is a G_δ and use this fact to show that the characteristic function of a closed set is the pointwise limit of a decreasing family of continuous functions. State and prove an analogous result for characteristic functions of open sets.

7. Let T denote the irrational rotation of the circle S^1 and for any $f \in L^1[0,1)$, show that

$$\frac{1}{n}\Big(f + f\circ T + \ldots + f\circ T^{n-1}\Big) \longrightarrow \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\exp(2\pi i\theta))d\theta,$$

in L^1. Is the analogous statement for other L^p classes true?

8. Show that the function $\psi(x)$ given by

$$\psi(x)$$

is an even periodic with period one. Show that it is differentiable infinitely often. Find its value at 1/2. Obtain the Fourier coefficients of $\psi(x/2\pi)$.

9. Consider the function $f(x) = x(\pi - x)$ on $[0,\pi]$ and on $[-\pi, 0]$, define $f(x) = -f(-x)$. Determine the Fourier coefficients of f and examine whether the series converges pointwise everywhere. Use it to compute the sum of the series

$$\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{(2k-1)^3}$$

10. Suppose that f isa continuous $2\pi-$ periodic function then prove that the partial sums of the Fourier series converge at a point of differentiability of the function.

11. Show that if f is continuous $2\pi-$periodic function on the real line then the Fejer means $\sigma_n f$ converge to f uniformly.

12. Suppose that f is a Borel measurable function on $[-\pi,\pi]$ such that $|f(x)| \le 1$. Prove that $|\sigma_n f(x)| \le 1$. Further show that if $\sigma_n f(x)| = 1$ for some n and x then f is a constant. (Hoffman)

13. Show that if $f \in L^\infty[-\pi,\pi]$ then the Fejer means of f converge to f weakly. That is for every $g \in L^1[-\pi,\pi]$,

$$\lim_{n\to\infty}\int_{-\pi}^{\pi} \sigma_n f(x)g(x)dx \longrightarrow \int_{-\pi}^{\pi} f(x)g(x)dx.$$

14. Use Parseval's formula on the function $e^{i\alpha x}$ and show that

$$\sum_{n=-\infty}^{\infty} \frac{1}{(n+\alpha)^2} = \frac{\pi^2}{\sin^2 \pi\alpha}$$

Hence determine the value of the intergal $\displaystyle\int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2} dx$

15. Use Parseval's formula on the characteristic function of $[-\alpha, \alpha]$ to determine $\displaystyle\sum_{n=1}^{\infty} \frac{\sin^2 \alpha n}{n^2}$

16. Compute the Fourier coefficients of $\sqrt{|\cos x|}$ and $|\cos x|^{1/3}$. Integrate $\left(z + \frac{1}{z}\right)^a z^n$ along a semicircle and use Cauchy's theorem from complex analysis.

17. Show by imitating one of the proofs of Riemann Lebesgue lemma that if f is $2\pi-$periodic and Hölder continuous of exponent $\alpha > 0$ then the Fourier coefficients decay like $1/n^{\alpha}$.

18. Discuss the decay properties of the complex Fourier coefficients $\{c_k\}$ of the function $x^{-\alpha}$ $(0 < \alpha < 1)$ defined on $[0, 2\pi]$ (extended as a $2\pi-$periodic function on the real line). Determine exponents p such that the function $x^{-\alpha} \in L^p[0, 2\pi]$ and exponents q such that the series

$$\sum_{k=-\infty}^{\infty} |c_k|^q$$

converges.

19. Let f be a holomorphic function on the unit disc $D = \{z \in \mathbb{C} : |z| < 1\}$ and assume that f is continuous on the closed disc $\overline{D}$. Then prove that

$$\lim_{r \to 1-} f(re^{it}) = f(e^{it}),$$

almost everywhere on $[-\pi, \pi]$. Hint use Parseval formula on the circles $|z| = r$. More generally prove this when f is holomorphic on the open unit disc D such that

$$\sup_{0<r<1} \int_{-\pi}^{\pi} |f(re^{it})|^2 dt < \infty.$$

Note: The space of all holomorphic functions f on the disc D satisfying the condition

$$\sup_{0<r<1} \int_{-\pi}^{\pi} |f(re^{it})|^p dt < \infty.$$

is denoted by $H^p(D)$. These spaces are called Hardy spaces.

Trouble here!! WE are not discussing the Poisson KErnel. Can this be done without it?? This point needs checking.

20. Let H_s be the subspace of $L^2[-\pi.\pi]$ defined as follows. We identify a function f in $L^2[-\pi, \pi]$ by its complex Fourier series namely

$$f = \sum_{-\infty}^{\infty} c_n e^{inx}.$$

Here the equality is to be understood in the L^2 sense that is,

$$\lim_{N \to \infty} \Big\| \sum_{n=-N}^{N} c_n e^{inx} - f \Big\|_2 = 0$$

is the limit in L^2 of the sequence of partial sums of the series appearing on the right hand side. For $s \geq 0$, consider the space

$$H_s = \Big\{ \sum_{-\infty}^{\infty} c_n e^{inx} : \sum_n |c_n|^2 n^{2s} < \infty \Big\}$$

(i) Show that the spaces H_s are Hilbert spaces with respect to the inner-product

$$< \sum_{-\infty}^{\infty} a_n e^{inx}, \quad \sum_{-\infty}^{\infty} b_n e^{inx} > \;=\; a_0 \bar{b}_0 + \sum_{-\infty}^{\infty} a_n \bar{b}_n |n|^{2s}$$

(ii) Show that $H_s \subset H_t \subset L^2$ when $0 < t < s$ and the respective inclusion maps are continuous.

(iii) Show that if $s > 1/2$ then $H_s \subset C[-\pi, \pi]$ and that the inclusion map is a compact operator. Further if $s > m + 1/2$ where $m \in \mathbb{N}$ then $H_s \subset C^m[-\pi, \pi]$.

21. Prove that for arbitrary sequences $\{a_n\}$ and $\{b_n\}$ of complex sequences,

$$\sum_{j=1}^{n} a_j b_j = -s_0 b_1 + s_1(b_1 - b_2) + s_2(b_2 - b_3) + \ldots + s_{n-1}(b_{n-1} - b_n) + s_n b_n,$$

where $s_n = a_1 + a_2 + \ldots + a_n$ and $s_0 = 0$. Deduce that if the sequence of partial sums of $\sum a_n(x)$ is uniformly bounded and b_n decreases down to zero monotonically then the series $\sum b_n a_n(x)$ converges uniformly. Discuss for convergence the series $\sum \frac{\sin nx}{\ln(n+1)}$

22. (i) Determine the Fourier series of the square wave train which is the odd $2\pi-$periodic function $f(x)$ such that $f(x) = 1$ on the interval $(0, \pi)$ and vanishing at 0 and $\pm\pi$.

(ii) Show that the Fourier series of the square wave train converges uniformly on compact subsets of $(-\pi, 0) \cup (0, \pi)$.

(iii) Let $F(x)$ denote the integral of f over the interval $[-\pi, x]$. Then $F(\pi) = F(-\pi) = 0$, F is differentiable on $(-\pi, 0) \cup (0, \pi)$ and $F' = f$ except at 0 and $\pm\pi$.

(iv) Deduce that even if the complex Fourier coefficients c_n of a function satisfy $|c_n| = O(n^{-2})$, the function need not be of class C^1.

23. Suppose that $\{c_n\}$ is the sequence of complex Fourier coeficients of an L^2 function then $\displaystyle\sum_n \frac{c_n}{n}$ converges. Use Hardy's theorem to prove this when $\{c_n\}$ is the sequence of complex Fourier coeficients of an L^1 function

24. Suppose that $f \in L^1[-\pi, \pi]$ and $\displaystyle\int_{-\pi}^{\pi} f(x)dx = 0$. Then determine the complex Fourier coefficients of the function F given by $F(x) = \displaystyle\int_{-\pi}^{x} f(t)dt$. Note that F extends extends as a continuous $2\pi-$periodic function on the real line.

The Poisson summation formula:

11.

Product measures and integral operators

Minkowski's inequality for integrals:

We provide here a useful generalization of the classical Minkowski inequality where the sums are replaced by integrals. Recall the familar inequality

$$|\sum_{n=1}^{\infty} a_n| \leq \sum_{n=1}^{\infty} |a_n|$$

for convergent series and the corresponding one for L^1 functions

$$\Big|\int_X f d\mu\Big| \leq \int_X |f| d\mu$$

generalize immediately to vector-valued functions where f takes values in $\mathbb{R}^N$.

Theorem: Let $(X, \mathcal{B}, \mu)$ and $(Y, \mathcal{A}, \nu)$ be sigma finite measure spaces and $1 \leq p \leq \infty$. If $f(x, y)$ is measurable with respect to the product measure then

$$\Big\|\int_X f(x,y) d\mu\Big\|_{L^p(\nu)} \leq \int_X \|f(x,y)\|_{L^p(\nu)} d\mu.$$

This is to be interpreted in the sense that the integrand on the right hand side is measurable and whenever the integral on the right is finite then the integrand on the left is finite for almost every y. For other values of y we redefine the integrand as zero and the inequality holds. and the inequality holds.

Proof: We assume first that $1 < p < \infty$ and measurability follows from Fubini's theorem. We may assume that $f \geq 0$ so that absolute value signs may be dropped. Recalling the proof of the classical Minkowski's inequality, we write the p-th power of left as

$$\begin{aligned}\int_Y \Big(\int_X f(x,y) d\mu\Big)^p d\nu &= \int_Y \Big(\int_X f(x,y) d\mu\Big)^{p-1} \Big(\int_X f(x,y) d\mu\Big) d\nu \\ &= \int_X \Big\{\int_Y f(x,y) \Big(\int_X f(x,y) d\mu\Big)^{p-1} d\nu\Big\} d\mu\end{aligned}$$

Applying the Hölder's inequality to the inner intergal we arrive at

$$\Big\|\int_X f(x,y) d\mu\Big\|_{L^p(\nu)}^p \leq \int_X \|f(x,y)\|_{L^p(\nu)} d\mu \Big[\int_Y \Big(\int_X f(x,y) d\mu\Big)^p d\nu\Big]^{(p-1)/p}$$

Observe that the second factor is precisely

$$\Big\|\int_X f(x,y) d\mu\Big\|_{L^p(\nu)}^{p-1}$$

and if this were zero the inequality is trivial and if it were fintie and non zero we could divide through it and get the desired inequality. In particular this proves the result for bounded functions vanishing outside rectangles of finite measure. In the general case we simply take an increasing sequence of such functions and appeal to the monotone convergence theorem.

Intergal operators between L^p spaces:

As an application of the Fubini-Tonelli theorem we state and prove a result on integral operators between L^p spaces. Operators of this type arise naturally and we have seen some instances of this earlier in the list of exercises [//]. The Fourier transform ranks among the most important of examples. Given $f \in L^1(\mathbb{R})$, the Fourier transform $\widehat{f}$ is the L^∞ function

$$\widehat{f} = \int_{\mathbb{R}} f(x) e^{-ix\xi} dx, \tag{727}$$

whereby we get a continuous linear map $L^1 \longrightarrow L^\infty$ given by $f \mapsto \widehat{f}$. The transformation (727) is a special case of the following

$$K : f \mapsto \int_X k(x,y) f(x) d\mu, \quad f \in L^p(\mu). \tag{728}$$

The function $k(x,y)$ appearing in (728) is required to measurable on a product space $X \times Y$, and is called ther kernel of the operator K. For example in the case of the Fourier transform, the kernel is in $L^\infty(\mathbb{R} \times \mathbb{R})$. We shall see in a later that the Fourier transform is an operator from L^2 onto itself[15].

The first result gives sufficient conditions when the domain and target are the same space.

Theorem: Suppose that (X, μ) and (Y, ν) are sigma finite measure spaces and $k(x,y)$ is a measurable function on $X \times Y$ with respect to the product measure. Assume that

$$\sup_y \int_X |k(x,y)| d\mu = M_2 < \infty, \quad \sup_x \int_Y |k(x,y)| d\nu = M_1 < \infty \tag{730}$$

Then the map K given by

$$K : f \mapsto \int_X k(x,y) f(x) d\mu$$

defines a bounded linear operator from $L^p(\mu)$ to $L^p(\nu)$ for any $p \in [1, \infty]$. The norm of the operator is $M_2^{1/p} M_1^{(p-1)/p}$.

Proof: Assume that $1 < p < \infty$, that f and g are non-negative. In the inequality

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

we take $a = f$ and $b = g$, multiply by $|k|$ and integrate over X and then over Y, to get

$$\Big| \int_Y d\nu \int_X |k(x,y)| f(x) g(y) d\mu \Big| \leq \frac{M_1 \|g\|_q^q}{q} + \frac{1}{p} \int_Y d\nu \int_X |f|^p |k(x,y)| d\mu,$$

which on using Tonelli's theorem yields

$$\Big| \int_Y d\nu \int_X |k(x,y)| f(x) g(y) d\mu \Big| \leq \frac{M_2 \|f\|_p^p}{p} + \frac{M_1 \|g\|_q^q}{q}$$

[15]However the intergal does not make sense for functions in L^2.

Now if we replace f by tf and g by $t^{-1}g$ we get for all $t > 0$,

$$\Big|\int_Y d\nu \int_X |k(x,y)|f(x)g(y)d\mu\Big| \le \frac{M_2\|f\|_p^p t^p}{p} + \frac{M_1\|g\|_q^q}{t^q q}$$

Since the left hand side is independant of t we get

$$\Big|\int_Y d\nu \int_X |k(x,y)|f(x)g(y)d\mu\Big| \le \inf_{t>0}\Big(\frac{M_2 t^p\|f\|_p^p}{p} + \frac{M_1\|g\|_q^q}{t^q q}\Big)$$

Using calculus, the infimum is easily found to be $M_2^{1/p}M_1^{1/q}\|f\|_p\|g\|_q$.

We conclude that given $f \in L^p(\mu)$, for each $g \in L^q(\nu)$, the product

$$g(y)\Big(\int_X k(x,y)f(x)d\mu\Big) \in L^1(\nu)$$

and moreover

$$\int_Y g(y)\Big(\int_X |k(x,y)|f(x)d\mu\Big)d\nu \le M_2^{1/p}M_1^{1/q}\|f\|_p\|g\|_q$$

The converse of Hölder's inequality enables us to conclude that

$$\Big\|\int_X k(x,y)f(x)d\mu\Big\|_{L^p(\nu)} \le M_2^{1/p}M_1^{1/q}\|f\|_{L^p(\mu)}$$

from which we get the result when $1 < p < \infty$. We now turn to the extreme cases.

The case $p = 1$: Let $f \in L^1(\mu)$. Applying Tonelli's theorem we get

$$\int_Y d\nu \int_X |k(x,y)||f(x)|d\mu = \int_X |f(x)|d\mu \int_Y |k(x,y)|d\nu \le M_2\|f\|.$$

In other words,

$$\Big\|\int_X k(x,y)f(x)d\mu\Big\|_{L^1(\nu)} \le M_2\|f\|_{L^1(\mu)}.$$

The case $p = \infty$: Here the condition $M_2 < \infty$ is superfluous and the proof is completely trivial.

Another interesting case is when the kernel is in $L^2(X \times X)$. The resulting operator is a bounded operator on L^2 which is also compact in case L^2 is separable. This is left as an exercise [//]. We have proved just one result though several other conditions are available in the literature. We recommend the book by Edwards [1] from which we shall state the following.

Theorem: The proof is not hard and the reader can look it up on p [//] in Edwards [1].

Convolutions:

This is an extremely important operation on the space of functions. Given two measurable functions f and g on $\mathbb{R}^n$ assume that the integral

$$\int_{\mathbb{R}^n} f(x-y)g(y)dy$$

exists for almost al $x \in \mathbb{R}^n$. The integral above is called the convolution of f and g and denoted by $f * g(x)$. This is in some sense an average of the translates of f with respect to a weight function g. It turns out that if both f and g are in L^1 then the convolution is also in L^1 and due to the averaging process is better behaved than either of the factors. We now examine the conditions under which the convolution exists. It is trivial to see that if $f * g(x)$ exists then so does $g * f$ and they are equal.

Theorem (Young): Suppose that $f \in L^p(\mathbb{R}^n)$ and $g \in L^1(\mathbb{R}^n)$ with $1 \leq p < \infty$, then $f * g \in L^p(\mathbb{R}^n)$.

Proof: If we regard $f(x-y)$ as a kernel function we see at once that

$$\sup_x \int |f(x-y)|dy = \sup_y \int |f(x-y)|dx = \|f\|_{L^1},$$

so that by theorem [//] we conclude that $f * g \in L^p(\mathbb{R}^n)$ whenever $g \in L^p(\mathbb{R}^n)$.

References

[1] R. E. Edwards, *Functional analysis,* Holt, Rinehart and Winston Inc., New York, 1965.

Kolmogorov's compactness criterion for $L^p(\Omega)$:

This is a useful compactness criterion akin to the Ascoli-Arzela theorem. This is quite different in spirit from the theorem of Vitali on equi-integrability that we discussed earlier. While the theorem of Vitali is useful for dealing with general measures, the present criterion is somewhat specialized for $L^p(\Omega)$ where Ω is an open set in $\mathbb{R}^n$. It relies heavily on regularization through convolutions and Luzin's theorem. We begin by recalling a few preliminary notions from general topology.

Definition: Given a metric space X and a subset A, an $\epsilon-$net for A is a subset S of X such that any point of A is within a distance of ϵ from a point of S. That is to say balls of radius ϵ centered at points of S cover A.

According to this definition A itself is an $\epsilon-$net and is clearly not an interesting one. Note that the members of the net need not be in A. We are primarily interested in finite $\epsilon-$nets that is for every $\epsilon > 0$ we need a finte collection of balls of radius ϵ that cover A. The proof of the following result is left as an exercise.

Theorem (Hausdorff): Let X be a complete metric space and $A \subset X$. Then A is relatively compact in X if and only if for every $\epsilon > 0$ there is a finite $\epsilon-$net for A.

Given a function $f \in L^P(\mathbb{R}^n)$ and $h \in \mathbb{R}^n$, we denote by $\tau_h f$ the $h-$translate of f namely,

$$\tau_h f : x \mapsto f(x+h)$$

We now tun to the statement and proof of Kolmogorov's theorem.

Theorem: Let Ω be a bounded open subset of $\mathbb{R}^n$ and ω be an open subset with $\overline{\omega} \subset \Omega$. Given a family $\mathcal{F}$ of functions in $L^P(\Omega)$ which are bounded in L^p norm and satisfy the following condition:

For every $\epsilon > 0$, there is a $\delta > 0$ with $\delta < \text{dist}(\omega, \Omega^c)$ such that

$$\|\tau_h f - f\|_{L^p(\omega)} < \epsilon, \quad h \in \mathbb{R}^n, \ \ |h| < \delta,$$

for every $f \in \mathcal{F}$. Then the family of restrictions $\mathcal{F}\Big|_\omega$ is compact in $L^P(\omega)$.

Remark: The condition described above is an analogue of the equi-continuity condition in the Ascoli-Arzela theorem. We shall refer to this also as the equi-continuity hypothesis.

Proof: Extend each $f \in \mathcal{F}$ to the whole of $\mathbb{R}^n$ be setting them equal to zero outside Ω and we shall continue to use the same letter f for the extended function. We proceed in three steps. In what follows ρ denotes a fixed C^∞ function that is positive, supported on the unit ball and whose integral over the ball is one and ρ_m dentoes the function $m^n\rho(mx)$ which is supported in the ball of radius $1/m$ centered at the origin and has integral one.

(i) We have

$$\|\rho_m * f - f\|_{L^p(\omega)} < \epsilon, \quad m > 1/\delta$$

for all $f \in \mathcal{F}$. Observe that the measure $\rho_m(x)dx$ is a probability measure on which we can apply Jensen's inequality and the result follows on using the equi-continuity hypothesis.

(ii) Applying Hölder's inequality to the convolutions $\rho_m * f$ we get

$$\sup_x |\rho_m * f(x)| \le \|\rho_m\|_{L^\infty} \|f\|_{L^1} \le C_m \|f\|_{L^1}, \quad f \in \mathcal{F}$$

Note that the functions vanish outside the bounded set Ω and so are bounded in in L^1 norm as well. On the other hand using the mean value theorem on ρ_m we get,

$$|\rho_m * f(x_1) - \rho_m * f(x_2)| \le |x_1 - x_2| \Big(\sup_x |D\rho_m(x)| \Big) \|f\|_{L^1}$$

which shows that the family $\{\rho_m * f : f \in \mathcal{F}\}$ is relatively compact in $C(\overline{\omega})$ and hence relatively compact in $L^p(\omega)$ as well.

(iii) Let $\epsilon > 0$ be arbitrary and fix an $m > 1/\delta$. Since $\{\rho_m * f : f \in \mathcal{F}\}$ has compact closure in $L^p(\omega)$ we can find a finite $\epsilon-$net for it. This serves as a finte $2\epsilon-$net for $\mathcal{F}$ by the estimate obtained in step (i) and the proof is complete.

The following corollary gives the compactness condition on $L^p(\Omega)$ itself rather than having to restrict the family of functions to ω.

Corollary: Let Ω be a bounded open subset of $\mathbb{R}^n$ and $\mathcal{F}$ is a family of functions in $L^P(\Omega)$, bounded in L^p norm such that

(i) For every $\epsilon > 0$ and every $\omega \Subset \Omega$, there is a $\delta > 0$ with $\delta < \text{dist}(\omega, \Omega^c)$ such that

$$\|\tau_h f - f\|_{L^p(\omega)} < \epsilon, \quad h \in \mathbb{R}^n, \ |h| < \delta, \quad f \in \mathcal{F}$$

(ii) For every $\epsilon > 0$ there is an $\omega \Subset \Omega$ such that

$$\|f\|_{L^p(\Omega - \omega)} < \epsilon, \quad f \in \mathcal{F}$$

Then $\mathcal{F}$ is relatively compact in $L^p(\Omega)$.

Proof: Let $\epsilon > 0$ be arbitrary and take a $\omega \subset \Omega$ such that $\|f\|_{L^p(\Omega-\omega)} < \epsilon$ for all $f \in \mathcal{F}$. By the theorem the family of restrictions $\mathcal{F}\Big|_\omega$ is relatively compact in $L^p(\omega)$ and so there is a finite $\epsilon-$net $g_1, g_2, \dots g_N$ for this family. Denoting by χ the characteristic function of ω we see that $\chi g_1, \chi g_2, \dots \chi g_N$ is a $2\epsilon-$net for the family $\mathcal{F}$.

Notes and comments:

The criterion of Kolmogorov appeared in [2]. Kolmogorov's arguments appear in exercises below and relies on estimates for mean values of functions over balls that require the assumption $p > 1$. Tamarkin [4] extended the results to unbounded domains. The formulation of the theorem given here is due to M. Riesz [3] who observed that the result so reformulated subsists for $p = 1$ also. The proof using regularization is from Brezis [1].

References

[1] H. Brezis, *Analyse fonctionnelle théorie et applications,* Masson, Paris, 1983.

[2] A. Kolmogoroff, *Über die Kompaktheit der Funktionenmengen bei der Konvergenz im Mittel,* Nachrichten Ges. Wiss. Göttingen, 1931, 60-63.

[3] M. Riesz, *Sur les ensembles compacts de fontions sommables,* Collected works p 136-142, Sprinegr Verlag,

[4] J. D. Tamarkin, *On the compactness of the spaces L_p,* Bull. American Math. Soc, **38** (1932), 79-84.

Exercises:

Averaging over balls:

1. Prove that if $f \in L^p(\mathbb{R}^n)$ and vanishes outside a bounded set, then the averages

$$f_\epsilon(x) = \frac{1}{\mu(B_\epsilon(x))} \int_{B_\epsilon(x)} f d\mu$$

are continuous functions of x for each fixed $\epsilon > 0$. Here $B_\epsilon(x)$ denotes a ball of radius ϵ centered at x.

Hint: The key point here is to estimate the volume of the difference of two balls of radius ϵ centered at points x' and x''.

2. Is the above result true when $p = 1$?

Exercises:

1. Show that a separately continuous function $f(x,y)$ defined on $\mathbb{R}\times\mathbb{R}$ is measurable. Hint: First work on the square $[0,1]\times[0,1]$ and show that f is a limit of a sequence of piecewise linear functions in one of the variables.

2. Can you describe a topology on $\mathbb{R}\times\mathbb{R}$ such that continuity with respect to this topology is precisely separate continuity?

3. Suppose that f is $2\pi-$periodic on the real line and $f \in L^2[-\pi,\pi]$, show that the function $\phi(x)$ given by
$$\phi(x) = \int_{-\pi}^{\pi} f(x+t)f(t)dt$$
is continuous and $2\pi-$ periodic. Determine the Fourier coefficients of g.

4. Suppose that $f \in L^p(\mathbb{R})$ and $g \in L^{(p-1)/p}(\mathbb{R})$, where $1 < p < \infty$, then show that
$$F(t) = \int_{-\infty}^{\infty} f(x+t)g(x)dx$$
is continuous on $\mathbb{R}$. Further show that $F(t) \longrightarrow 0$ as $t \longrightarrow \infty$.

5. Show that if $f \in L^p(0,\infty)$ and $1 < p < 2$, the function ϕ given by
$$\phi(y) = \int_0^{\infty} f(x)\frac{\sin xy}{\sqrt{x}}dx$$
satisfies the estimate
$$\phi(y) = o(y^{\frac{1}{p}-\frac{1}{2}}), \quad \text{as } y \to 0+$$

6. Let $f \in L^1(0,\infty)$ and $n \geq 1$ be an integer. Define $I_n(f)$ inductively as follows:
$$I_1(f) = \int_0^x f(t)dt, \;\; I_2(f) = I_1(I_1(f)), \ldots.$$
Prove that
$$I_n(f) = \frac{1}{(n-1)!}\int_0^x (x-t)^{n-1}f(t)dt.$$

7. Define for $\alpha > 0$ and $f \in L^1(\mathbb{R})$, the operator $I_\alpha(f)$ as
$$I_\alpha(f) = \frac{1}{\Gamma(\alpha)}\int_0^x (x-t)^{\alpha-1}f(t)dt$$
Prove that $I_\alpha(I_\beta(f)) = I_{\alpha+\beta}(f)$. These are the integrals of fractional order defined by Riemann and Liouville.

8. Show that if $f \in L^p(\mathbb{R})$ then $I_\alpha(f)$ is continuous if $\alpha > 1/p$. This must remind the reader of the Sobolev lemma. (See Titchmarsh, p 398)

Averaging over balls:

11. Prove that if $f \in L^p(\mathbb{R}^n)$ and vanishes outside a bounded set, then the averages

$$f_\epsilon(x) = \frac{1}{\mu(B_\epsilon(x))} \int_{B_\epsilon(x)} f d\mu$$

are continuous functions of x for each fixed $\epsilon > 0$. Here $B_\epsilon(x)$ denotes a ball of radius ϵ centered at x.

Hint: The key point here is to estimate the volume of the difference of two balls of radius ϵ centered at points x' and x''.

12. Is the above result true when $p = 1$?

Odds and Ends

Here may be found an assortment of problems, historical details and other paraphernalia that are interesting but could not find place in the text.

Works of F. Riesz and M. Riesz: The name of Riesz has appeared in several places and refers to either of the two brothers F. Riesz or the younger M. Riesz. Both of them were prolific and their contributions have enriched the field of analysis. A vivid account of the life and works of F. Riesz are the ones due to P. Halmos and B. Nagy [//]. The book on Functional analysis by F. Riesz with Nagy [//] is a veritable delight written in the spirit of classical analysis. The book originally written in French (1952) was translated into English by Leo Boron. F. Riesz and A. Haar were the editors of the journal "Acta Scientiarum Mathematiccarum" about which B. Nagy has interesting reminiscences ([//], p. 72).

Halmos's account besides being scholarly is quite piquant, particularly his comment on Riesz' reaction to Egoroff's theorem. Another interesting nugget out of Halmos' lecture is Riesz' proof (in a letter to G. H. hardy dated 1930) of the following version of the arithmetic-geometric mean inequality or rather the Jensen's inequality.

Theorem: Suppose that f is a positive function on $[0, 1]$ then

$$\int_0^1 \log f(x)dx \leq \log \int_0^1 f(x)dx.$$

Problems: Prove the converse of Jensen's inequalty.

References

[1] B. Sz.-Nagy and J. Szabados (ed), *Functions, series, operators, Volume I,* Proceedings of the international conference organized by János Bolyai Mathematical Society, North Holland publishing company, Budapest 1983.