STANFORD ANALYSIS PH.D. QUAL 2015

ZULFIKAR MOINUDDIN AHMED

1. Problem I.1

Suppose that A is a Borel set in $\mathbf{R}$ with the property that if $x \in A$ and if y is any real number such that the decimal expansions of x and y differ by at most finitely many places then $y \in A$. Prove that the Lebesgue measure of either A or $\mathbf{R} - A$ is zero.

If q has a decimal expansion that differs from zero only in finitely many places then obviously $q \in \mathbf{Q}$. Let $Q_f \subset \mathbf{Q}$ be such real numbers. It is evident that

$$A = \bigcup_{x \in A} (x + Q_f)$$

If $A = \mathbf{R}$ then obviously Lebesgue measure of the complement is zero. We assume that there is a real number b such that $b \notin A$.

Next we prove that either there is a nonempty interval $J \subset A$ with $\mu(J) > 0$ or $\mu(A) = 0$.

If there is an interval $J \subset A$ with $\mu(J) > 0$, then there is interval $J_0 \subset J$ with rational endpoints $J_0 = [j_1/m, j_2/m]$ with $m, j_1, j_2 \in \mathbf{Z}$. Since $\mathbf{Z} \subset Q_f$, we have

$$J_k = J_0 + \frac{j_2 - j_1 + k}{m} \subset A$$

and

$$\mathbf{R} = \bigcup_{k \in \mathbf{Z}} J_k \subset A$$

so $\mu(\mathbf{R} - A) = \mu(\emptyset) = 0$.

Now consider the situation where A contains no nonempty intervals at all.

We want to examine a general situation: if a Borel set $B \subset \mathbf{R}$ does not contain any open intervals at all then what can we say about it? We would complete the problem if we can prove $\mu(B) = 0$ in that case. I am not yet sure about this.

I don't really understand Borel sets so well, so I will take the following path. I will go to the *metric topology* of $[0, 1]$ and prove that every closed set $C \subset [0, 1]$ that does not contain any nonempty open interval $I \subset [0, 1]$ is a *finite set* of points. In order to see this, let $U = [0, 1] - C$ and cover it with $\{U_a\}$ consisting of open intervals $U_a = (u_a, v_a)$ and use Heine-Borel theorem to take a finite subcover. Then we have finite set of points $P = \{u_1, v_1, \ldots, u_N, v_N\} \subset C$.

At this point we want to show that $C = P$. Let's see how this could be done. Suppose $y \in C$ and $y \notin P$. Then

$$y \in \bigcap_{j=1}^{N} [0, u_j] \cup [v_j, 1]$$

Date: December 28, 2021.

Now let us do a manipulation without proof or justification:

$$\bigcap_{j=1}^{N}(A_j \cup B_j) = \bigcup_{1\le j,k\le N}(A_j \cap B_k)$$

I am aware this could be wrong, but I proceed anyway.

$$y \in \bigcup_{1\le j,k\le N}([0,u_j]\cap[v_k,1]) = \bigcup_{1\le j,k\le N}[v_k,u_j]$$

There thus exists a $1 \le k, j \le N$ such that $y \in [v_k, u_j]$. The endpoints are in P, so $v_k < y < u_j$. We claim then that $(v_k, u_j) \subset C$ which contradicts assumption that C contains no nonempty open intervals. Therefore $C = P$.

This is the *metric topology* case. Next we want to prove that if $B \in [0,1]$ is Borel, and B contains no nonempty open interval then it is a countable set.

1.1. **Abstract Sigma Algebra Idea.** Let $\mathcal{B}$ be the Borel sigma-algebra for $[0,1]$. We will consider a sub sigma-algebra $\mathcal{B}_0$ which with contain $[0,1], \emptyset$. Otherwise it will be abstractly defined as having the property "does not contain interior". Clearly open sets without interior are empty. We showed in the last sections that closed sets without interior are finite. Therefore the only nontrivial members of $\mathcal{B}_0$ are countable.

2. Problem 2

Let T be a nonzero compact operator on a Hilbert space H.
(a) Give an example of T with spectrum 0.
(b) Prove that this is impossible for T self-adjoint.

Let's start (b) off assuming the spectral theorem for compact selfadjoint operators. Then we have an orthonormal complete set u_j with $\lambda_j \to 0$ and $Tu_j = \lambda_j u_j$. If $\lambda_1 = 0$ then T is the zero operator. So assuming spectral theorem for compact selfadjoint operators (b) is easy to verify.

Let's look at (a). I don't know an example, so let's just look at non-diagonalisable small matrices for inspiration. I had a proof of the spectral theorem for compact self-adjoint operators directly by projection into subspace $\{u \in H : \|Tu\| \ge \epsilon\}$. For (a) I suspect that if we see how to get small matrices with only zero eigenvalues we might be able to produce an infinite dimensional example.

I will be quite honest. Non-selfadjoint compact operators are not exactly something I have any familiarity with at all. In my Four-Sphere Theory I prove that *resolvents of molecular Hamiltonians* are compact and self-adjoint, something that is spectacular for physics. Years ago, when Tosio Kato proved that molecular Hamiltonians are essentially self-adjoint in Schroedinger Theory, it was against the explicit prophesies of the great Jon Von Neumann. That story is told beautifully by Barry Simon in *Fifty Years of Perturbation Theory.* Now non-selfadjoint compact operators is so strange to me that I don't think I've even considered the concept seriously in my life. I am a symmetric-matrix sort of man. I try to stay away from things that are not diagonalisable generally.

I will use a two-by-two example. Let

$$A = \begin{pmatrix} 1 & 21/2 & -1 \end{pmatrix}$$

The characteristic polynomial is

$$\det(A-\lambda) = (1-\lambda)(-1-\lambda) + 1 = -1 + \lambda^2 + 1 = \lambda^2 = 0$$

Here spectrum is just zero. Then we just take any orthonormal basis $\phi_1, \phi_2, \ldots$ of H and define T as A on span of ϕ_1, ϕ_2 and zero for all other basis elements. It is compact because range of $B = \{x \in H : \|x\| \leq 1\}$ by T is the image of the disk in $\mathbf{R}^2$ by A which is a bounded set in two dimensional subspace and so has compact closure.

3. Problem 3

Show that there is a smooth function ϕ on $\mathbf{R}$ such that $\phi \geq 0$ and $\phi(0) > 0$ and the Fourier transform $F\phi$ is non-negative.

We want ϕ to be even for this problem, i.e. $\phi(x) = \phi(-x)$. Suppose ϕ even with support in $[-\pi, \pi]$. Then

$$F\phi(\xi) = \int_{-\pi}^{\pi} e^{-ix\xi}\phi(x)dx = \int_{-\pi}^{\pi} \cos(x\xi)\phi(x)dx + i\int_{-\pi}^{\pi} \sin(x\xi)\phi(x)dx$$

It's easier when all the quantities in the integral are real-valued to see what's going on. We want to use even condition and do a change of variables and divide the second integral into $[-\pi, 0]$ and $[0, \pi]$ and then the terms sum to zero. Similar operation on the first integral produces:

$$F\phi(\xi) = 2\int_{0}^{\pi} \cos(x\xi)\phi(x)dx$$

Now we need to be careful. If $|\xi| \leq 1$ we will certainly have

$$\cos(x\xi) \geq 0$$

and then $\phi \geq 0$ gives us $F\phi(\xi) \geq 0$.

I have doubts about whether it is possible to have $F\phi(\xi) \geq 0$ for all $\xi \in \mathbf{R}$.

Now let's try to construct a ϕ. Take

$$\phi(x) = \begin{cases} e^{1-1/(1-|x|)} & |x| < 1 \\ 0 & \text{otherwise} \end{cases}$$

This is 1 at 0, zero at +/-1, and nonnegative.

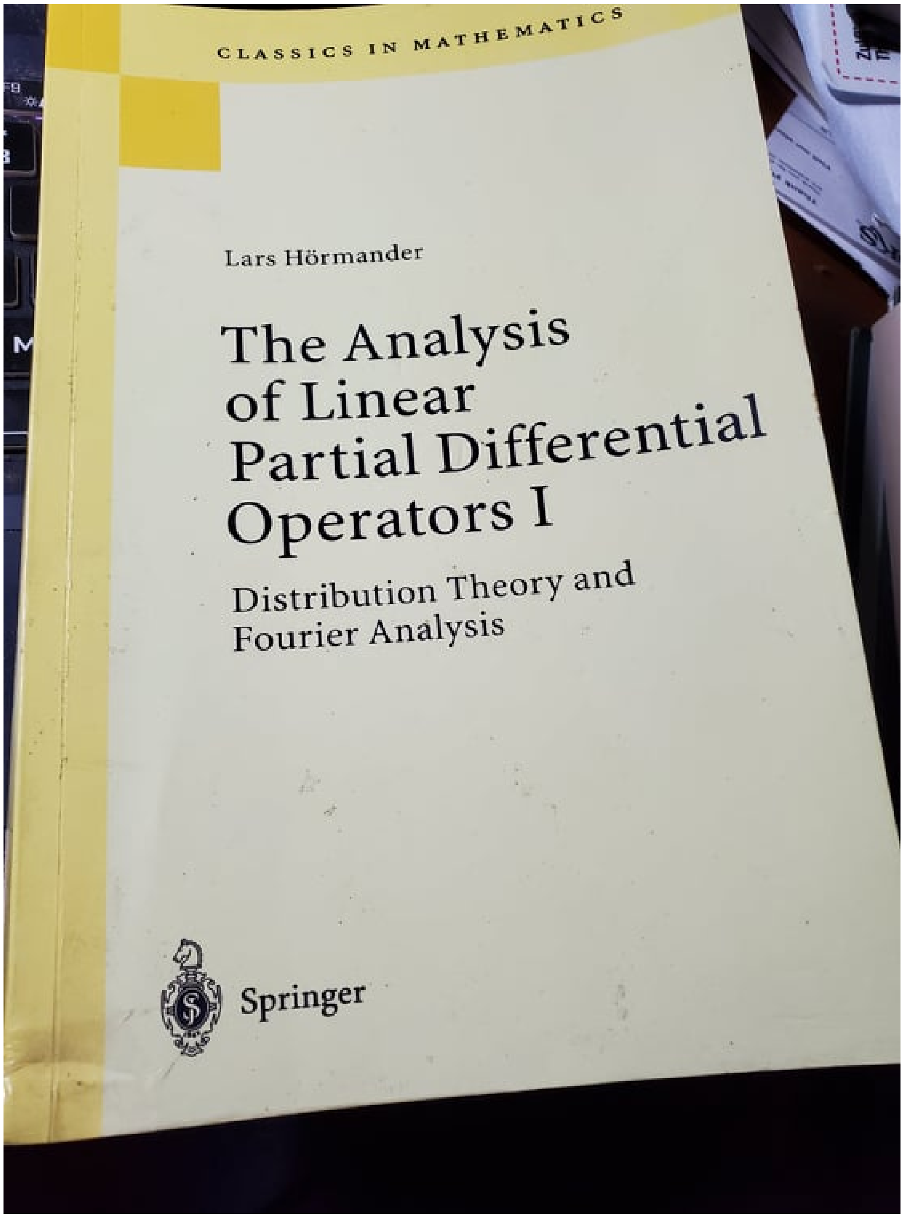
4. Problem 4

Suppose $f \in C^2(\mathbf{R}^2)$.

(a) Prove that mixed partials are equal. (b) Find $f \in C^1(\mathbf{R}^2)$ where the mixed partials exist but are not equal. (c) For $f \in L^1(\mathbf{R}^2)$ show mixed partials are equal in a distributional sense.

Let us take a lazy approach and set up notation. We want to use h_x and h_y for the limiting variables. And we'll use the shorthands ∂_x and ∂_y for $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$.

"That's just syntactic sugar and has no substance, Zulf!" you might say. "And so what if it is?" Let me tell you something. You see this great book?



This is by one of the greatest masters of distribution theory and the deepest practitioner of distribution theory this world has ever known. Look through the book. Do you see him using the verbose language? Very rarely. He uses ∂^α or ∂_j. I don't have his substantial mathematical genius, it is true, but I am quite willing to just do what he does and don't ask too many questions. Who knows if the magic is in the syntactic sugar? As what Spike Lee says in the sneaker commercials about Michael Jordan. "It's gotta be the shoes."

When in doubt, Zulf is quite willing to be superstitious. I'll reverently use shorthand. What are you going to do about it huh? You can't do a thing!

4.1. **What is differentiability?** We will use the definition of the great master Lars Hormander because the world has had very few people who truly understood these issues better. But before that let us remind our dear reader about him. By the way, Lars Hormander is dead; I don't want anyone demanding various sorts of tributes and tithes and sacrifices of my firstborn, so I just give my veneration to dead people to be safe.

That is Hormander in 1969. He had won the Fields medal in 1962 and was the foremost developer of Laurent Schwartz' theory of distributions.

Suppose X and V are Banach spaces. Let $U \subset X$ be open and let $L(U, V)$ be continous functions. We say a function f is differentiable on U if there is an element $f'(x) \in L(U, V)$ such that

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

as h tends to the zero vector in X. By $C^1(X, V)$ we denote the set of continuously differentiable functions from X to V, i.e. functions that are both differentiable and for which the assignment $x \in X \mapsto f'(x) \in L(X,V)$ is continuous. The space $L(X,V)$ is a Banach space with norm

$$\|T\| = \sup_{\|x\|\leq 1} \|Tx\|$$

Now define higher order differentiability by induction. We *define* $f \in C^k(X,V)$ if $f \in C^1(X,V)$ and $f' \in C^{k-1}(X, L(U,V))$.

You don't like this definition huh? You think this definition is a little too complicated for the situation, do you? Well, tough. Zulf is wise and uses the definition that Lars Hormander used.

"Zulf, why don't you use a standard definition?"

"What do you mean standard. Do you understand" – here Zulf grabs both you arms and begins to shake it – "This is *Lars Hormander's definition* for God's sake. Why would you call anything else standard?"

4.2. **Problem 4(c) Is Easy.** When we say $f \in L^1(\mathbf{R}^2)$ is a distribution, we mean the functional L_f on $C_0^\infty(\mathbf{R}^2)$ that is defined by:

$$L_f(\phi) = \int f\phi dx$$

But for any $k \geq 0$ we have

$$\sup_{|\alpha|\leq k} |\partial^\alpha \phi| < \infty$$

We can just let this be a norm on $C_0^\infty(\mathbf{R}^2)$. We have the easy inequality

$$|(\partial^\alpha L_f)(\phi)| \leq \|f\|_{L^1}(\sup_{|\beta|\leq k} |\partial^\beta \phi|)$$

Mixed partials are equal for ϕ so we also have

$$\partial_x\partial_y = L_f(\partial_x\partial_y\phi) = L_f(\partial_y\partial_x\phi) = \partial_y\partial_x L_f(\phi)$$

4.3. **Problem 4(a).** We consider two functions $p(x,y) = (\partial_x f)(x,y)$ and $q(x,y) = (\partial_y f)(x,y)$. The assumptions tell us, by definition of $C^2(\mathbf{R}^2)$ that we have $p, q \in C^2(\mathbf{R}^2)$.

Let us compute the derivative $\partial_y p$.

$$(\partial_y p)(x,y) = \lim_{h\to 0} \frac{p(x,y+h) - p(x,y)}{h}$$

Then

$$p(x,y+h) - p(x,y) = p(x,y+h) - p(x+h,y+h) + p(x+h,y+h) - p(x,y)$$

This gives us

$$(\partial_y p)(x,y) = \lim_{h\to 0} \frac{p(x,y+h) - p(x+h,y+h)}{h} + \lim_{h\to 0} \frac{p(x+h,y+h) - p(x,y)}{h} \tag{1}$$

We could use the Taylor expansion too

$$p(x+h,y+h) = p(x,y) + \partial_x p(x,y)h + \partial_y p(x,y)h + O(|h|^2)$$

So

$$\lim_{h\to 0} \frac{p(x+h,y+h) - p(x,y)}{h} = \partial_x p(x,y) + \partial_y p(x,y)$$

We also have a similar result for $q(x,y) = (\partial_y f)(x,y)$. which is

$$q(x+h,y+h) = q(x,y) + \partial_x q(x,y)h + \partial_y q(x,y) + O(|h|^2)$$

I am stuck now on this problem. I will need to think about this problem more.
Let me try a straightforward approach

$$\partial_x\partial_y f(x,y) = \lim_{h_x\to 0}\lim_{h_y\to 0}\frac{1}{h_x h_y}(f(x+h_x,y+h_y)+f(x,y)-f(x+h_x,y)-f(x,y+h_y))$$

and

$$\lim_{h_y\to 0}\lim_{h_x\to 0}$$

$1\frac{}{h_x h_y(f(x+h_x,y+h_y)+f(x,y)-f(x+h_x,y)-f(x,y+h_y))=\partial_y\partial_x f(x,y)}$
There is some subtlety here that I am still not very clear about at all. Regardless, this switching of limits above needs more care.

4.4. **Problem I.4(b).** I will look this up. I really do not have good ideas for this problem

5. Problem I.5

Let $f(x) = x^4 - 2x^2 + 1$. Prove that there is an alpha $\alpha \in \mathbf{R}$ (and find its value) such that

$$C^{-1}\lambda^{-\alpha} \leq |\int e^{i\lambda f(x)}\sin^4(x)dx| \leq C\lambda^{-\alpha}$$

for some $C > 0$ as $\lambda \to \infty$.

This looks like a stationary phase approximation problem. So we note

$$f(x) = (x^2-1)^2$$

Intuitively the dominant contribution of these sorts of oscillatory integrals are where the phase, i.e. $f(x)$ is stationary, meaning $f'(x_0) = 0$.

Let us put aside the rigorous bounds and just try to understand this problem by determining the critical points of $f(x)$.

$$f'(x_0) = 2x(x^2-1)$$

we have three critical points at $x = 0, -1, 1$. The second derivative is

$$f''(x) = 4x^4 + 2x^2 - 2$$

So $f''(0) < 0$ and there is a maximum there. Taylor expansion tells us

$$f(h) = 1 - h^2 + O(|h|^3)$$

So what happens when we ignore the error term? We can control the integral by

$$|\int e^{-i\lambda(1-x^2)}\sin^4(x)dx|$$

We can pull out the constant part and that has modulus 1 and so the crucial quantity is this:

$$Q(\lambda) = |\int e^{-ix^2\lambda}\sin^4(x)dx|$$

I will return to this problem later.

6. Problem II.1

Suppose $f \in L^1([0,1])$. Suppose for all $\phi \in C_0^\infty(0,1)$ we have

$$\int_0^1 f\partial_n\phi = 0$$

Show that f is a polynomial of degree n. The hint is to approximate f by convolutions.

Let's prove this for smooth functions g first and then we will worry about regularizing L^1 functions later. For that we will use the Lars Hormander regularisation. As my dear reader will realise, I am deeply convinced that Lars Hormander is worthy of emulating on all matters involving derivatives and distributions.

6.1. The Integration By Parts Formula.

What I want to do here is *organise* integration by parts in a reasonable way so that we can see what we are doing. Now my dear reader is obviously extremely familiar with the confusion that ensues when we want to do integration by parts and get all our constants and signs in a jambalaya.

Jambalaya is an American Creole and Cajun rice dish of French (especially Provencal cuisine), African, and Spanish influence, consisting mainly of meat and vegetables mixed with rice. It's quite savory and I recommend that my dear reader taste the wonderful cuisine of the delicate merging together of French, Spanish, and African cultural influences. However, we don't want the jambalaya to show up in our integration by parts formulae. This is what happens when we are not organised. We would like to keep delicassies of Creole and Cajun origin strictly separate from our integration by parts formulae.

Suppose $p \geq 1$. For any $\phi \in C_0^\infty(0,1)$ we claim

$$(p+1)^{-1}\int x^{p+1}\partial_k\phi = (-1)\int x^p\partial_{k-1}\phi$$

This follows from product rule and the fundamental theorem of calculus applied to endpoints where function $x^{p+1}\partial_{k-1}\phi$ vanishes because ϕ has compact support in $(0,1)$.

$$\partial(x^{p+1}\partial_{k-1}\phi) = (p+1)x^p\partial_{k-1}\phi + x^{p+1}\partial_k\phi$$

The outcome of integrating is

$$(-1)\int x^p\partial_{k-1}\phi = (p+1)^{-1}\int x^{p+1}\partial_k\phi \tag{2}$$

Our problem assumes a fixed n and one equation

$$\int f\partial_n\phi = 0$$

We will use the integration by parts formula to move some of the derivatives over to f assuming that it's just a polynomial in x and our great task is to come to a conclusion *sans jambalaya*.

Let us apply (2) with some other parameters

$$(-1)\int x^{p-1}\partial_{k-2}\phi = \frac{1}{p}\int x^p\partial_{k-1}\phi$$

Then can combine this with the original to form the chain of equalities.

$$(-1)^2(p+1)p\int x^{p-1}\partial_{k-2}\phi = \int x^{p+1}\partial_k\phi$$

So we carefully generalise this

$$(-1)^m(p+1)p\cdots(p+2-m)\int x^{p+1-m}\partial_{k-m}\phi = \int x^{p+1}\partial_k\phi$$

which will hold for

$$p+2-m \geq 1$$

and

$$m \leq k$$

With this conclusion, we can then deduce that for all $q \leq n-1$ we have

$$\int x^q\partial_n\phi = 0$$

for all $\phi \in C_0^\infty(0,1)$.

Now suppose that we had a power series expansion of g available and wrote

$$g(x) = \sum_{k=0}^{\infty} a_k x^k$$

From $\int g\partial_n\phi = 0$ for all $\phi \in C_0^\infty(0,1)$ we would like to show $a_k = 0$ for $k \geq n$.

The intuitive idea is of course to find a way to find ϕ that in some sense behave as selectors of x^k from the monomials as though they were a Fourier basis and take L^2 inner products, so that we get something like this – this is not right but important to this problem:

$$g = \sum_k a_k e_k$$

so let $\phi = e_k$ and take inner products

$$0 = \langle g, e_k\rangle = a_k\langle e_k, e_k\rangle = a_k$$

and therefore $a_k = 0$ for $k \geq n$. So even though we cannot exactly do this in this context, the reason the proposition is true is that we can do something that behaves quite similarly and be still rigorous mathematically. I would say this is one of the strange issues of mathematical substance. This inner product vanishing is what is really the reason why the problem proposition is true.

7. Digression On The Phrase "Ethnic Food"

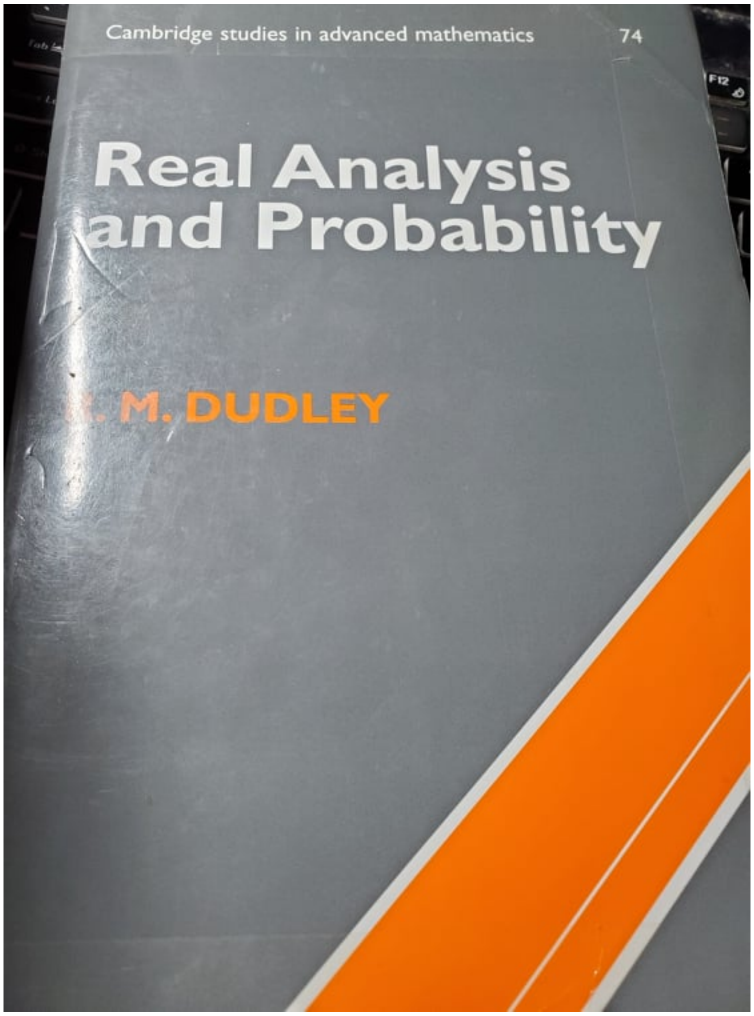
You see, when people do not have refined sense of cultural influences of cuisine. When people think that the difference between Wendy's and McDonalds and cuisine with centuries of development in organic civilisation; when people refer to cuisine in a dismissive manner as 'ethnic food', what they are doing, and usually totally unaware of this, is loudly proclaiming their total lack of cultivation and good taste.

8. Problem II.2

Let H be the subspace of ℓ^2 consisting of (a_n) with $\sum_n n^2|a_n|^2 < \infty$. Show H is first category in ℓ^2.

I will have to look up first category. It should be something like 'nowhere dense' but I always need to double check.

What can I tell you, my brother my killer, what can I possibly say? I guess that I miss you. I guess I forgive you. I'm glad you stood in my way. If you ever come by, for Jane or for me, well your enemy is sleeping and his woman is free. Anyway, I will share with my dear reader the place where I found a good precise definition for first category.



A set A in a topological space S is nowhere dense if for every open subset $U \subset S$ there exists another open subset $V \subset U$ such that $V \subset A = \emptyset$. A set is first category iff it is the union of countably many nowhere dense subsets.

In order to this problem, what I will do is investigate a bit the metric topology of ℓ^2. Suppose $f \in \ell^2$ and $f \notin H$. Let $f = (b_n)$ and we have

$$\sum_n b_n^2 < \infty$$

and

$$\sum_n n^2 b_n^2 = \infty$$

We will have proven the proposition if we show that for every ball B in ℓ^2 we can find some other ball $B_0 \subset B$ with another center such that $B_0 \subset H = \emptyset$.

Yes, and thanks, for the trouble you took from her eyes. I thought it was there for good so I never tried. And Jane came by with a lock of your hair. She said that you gave it to her. That night that you planned to go clear. Dum da da dee da dee.

This problem will involve some care with series convergence it looks like. Let me think about this some more.

9. How Much Does A Generic ℓ^2 Element Shrink To Be in H?

A generic $f \in \ell^2$ with coefficients (b_n) will roughly need to replace all its coefficients with b_n/n to fit into H. And the distance it would move is

$$d^2 = \sum_n b_n^2(1 - 1/n)^2$$

This gives us a relatively simple lower bound

$$1/2\|f\|_2 \leq d$$

Now this is intuition and our task is to take this sort of intuition and prove something solid. I will return to this problem.

10. Problem II.3

(a) Give examples of $A_n \in L(\ell^2, \ell^2)$ bounded operators satisfying the following conditions. (i) $A_n \to 0$ in strong topology but not in norm topology. (ii) $A_n \to 0$ in weak but not strong topology.

I never remember these topologies, so I will get some assistance on these and look at the issues in texts.

(b) The map $A \mapsto A^*$ is continuous in norm and weak topologies but not in the strong topology.

I studied Reed-Simon Vol I for functional analysis in Peter Sarnak's functional analysis course in 1993 and I am sure we covered these then but it has been a long time and I forgot about so many details.

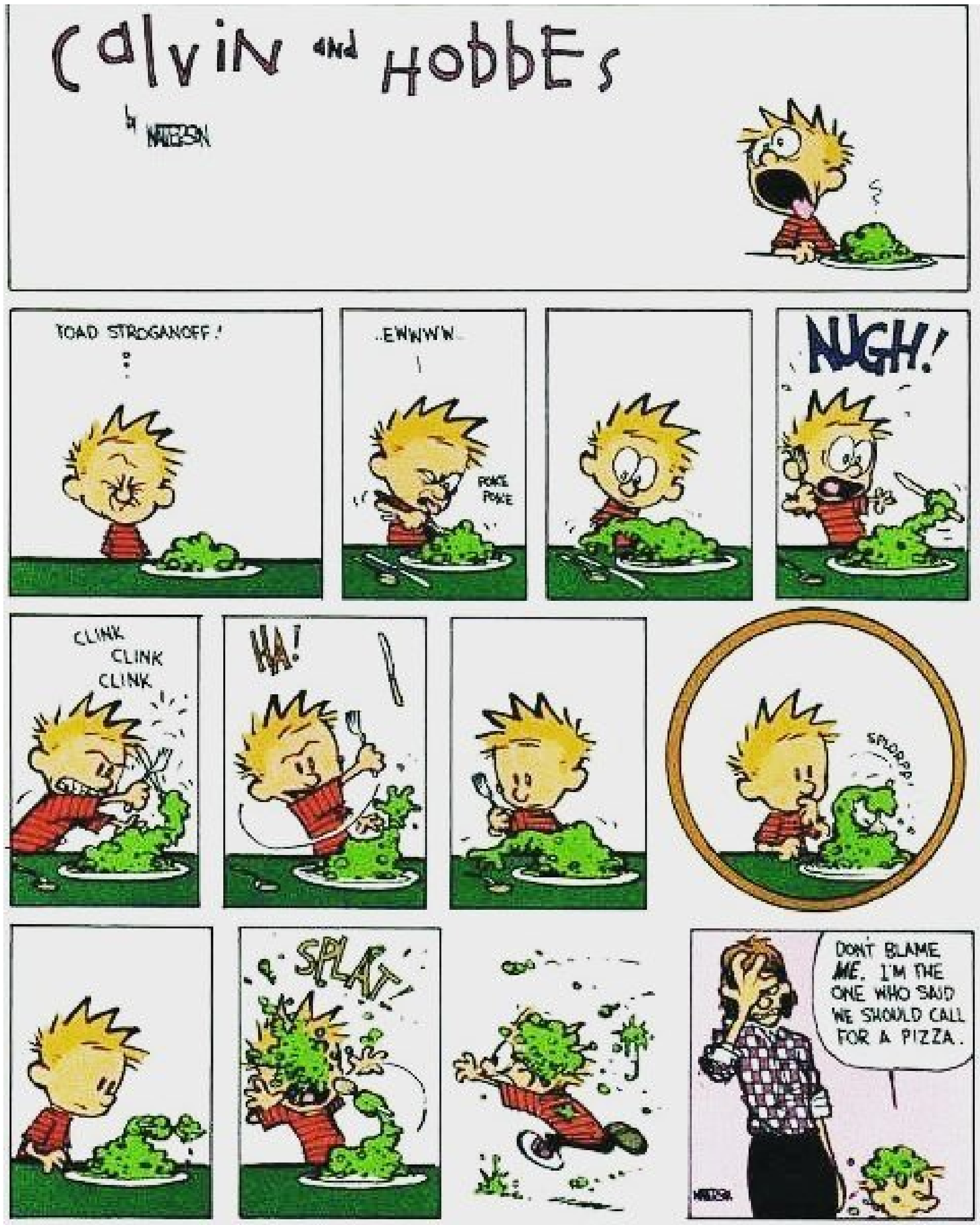
For (a-ii) an example is shifting sequences to the right and zeroing out the left end. So you take

$$A_1(a_n) = (0, a_1, \dots)$$

and $A_n = A_1^n$. Then for a given $x = (a_n)$ we have $A_n x \to 0$ in the weak operator topology but not in the strong operator topology

For (a-i) Terence Tao has a note which I read for illumination [1]. I finally found a good answer [2]. The example is projection rather than shifting on the first n basis vectors, call it P_n and let $A_n = I - P_n$ Then $\|A_n\| = 1$ in norm topology but $A_n \to 0$ in strong topology.

I will admit that these are not clear to me. In 1930 Von Neumann worked on describing neighborhoods in terms of seminorms for the weak operator topology. I am not comfortable with these things at all.



11. Interlude: Discussion Of Deep Differences Between Scientists, Engineers and Mathematicians

All jokes aside from *Big Bang Theory* which is a popular television sitcom, there are reasons for deep differences. I am extremely gifted in Scientific work, not even experimental for more than a decade as I have been working on theoretical physics issues, psychology, universal human moral nature and such. Stanford Mathematicians are professional and would like people doing Mathematical work to have deep mastery of all their tools. This is good. But this is not the way that Scientists can even hope to work. You see, the deep reason for this is that in Mathematics, the problems, ideas, history, are all internally generated in a way. Mathematics does not confront Nature in the same way as Scientists do. As a result, the discipline of Mathematics is much more Theological in Character. The texts and the knowledge that are applied are strongly historical and this is required because progress is not possible without deep understanding of Mathematical issues. In Science this is not only the way, but would be totally suicidal. Scientists have to confront Nature's mystery, and there are no laws of Nature that tell us that being Theological will ever lead to any deep understanding of the external world. Science has an arbiter, and it is Nature. No amount of mastery of things that have been done in the past guarantees any sense because Nature is not contained in all that. Nature is still mysterious to Man, and there are no guarantees that standing on established consensus is actually going to provide any solid ground at all for Science as it re-assures the Mathematicians. Years ago, I was listening to an interview of a great mathematician, Tomasz Mrowka. He turned away from physics and towards Mathematics already quite young because he was more impressed by how much progress was solid in Mathematics. And he has broken great ground. Fate drew me to Scientific problems, and my Princeton years did not predict this. My perspective, at 49 is quite equilibrium and unlikely to change a great deal in the future. But I appreciate gaining deeper understanding for my own motives.

12. Problem II.4

Suppose H_1, H_2 are Hilbert spaces, $A \in L(H_1, H_2)$ and $B \in L(H_2, H_1)$ and compact operators E_1, E_2 with

$$AB = I - E_2$$

and

$$BA = I - E_1$$

Show nullspace of A is finite dimensional and range is closed with finite dimensional orthocomplement.

It is the range of E_1 and E_2 that will be finite dimensional and that is the key to these sorts of finite dimensionality results.

December 28 2021 12:25 PM. In meta I came across a Romanian physicist who liked some of my work, a certain Mr. Radoslav. I am pleased that quantum field theorists are looking at my Four-Sphere Theory. But let me turn to this problem and get a better understanding of the situation. This is an important problem, and I want to have a firmer understanding of it so that I am not repeating the mistakes I have made in this situation, and be a diligent mathematics student.

First of all $AinL(H_1, H_2)$ and I harken back to freshman year linear algebra at Princeton when the world was large and optimism was great, and youth gave special magic to all of existence. The *dimsum theorem.* The dimension of the nullspace and the dimension of the range of a linear transformation has to add up to the dimension of the domain. So let us diligently examine if the obvious analogy will hold for A.

I will first for some understanding of some sort of decomposition

$$H_1 = V \oplus W$$

where W is the nullspace of A and V consists of elements of H_1 that are orthogonal to W. Then we want to examine whether, totally analogously to finite dimensional linear algebra, $A|_W$ is injective and an isomorphism. Then we want to show that this decomposition is respected by E_1 in that E_1 identity on W and does something reasonable on V.

I am a Scientist, so for me the way to make progress in this situation will be chart out some conjectures and attempt then to understand whether they can be proven. The crux will be the behaviour of E_1 on V and W. If we are able to show: $E_1 = 0$ on V and $E_1 = I$ on W that would be good. Given the truth of this *hypothesis* we would then have finite dimensionality of W as follows. Take any orthonormal basis of W without regard to cardinality $e_1, \ldots, e_N, \ldots$. These are all $\|e_j\| = 1$. Since E_1 is compact the orthonormal the closure has to be compact of this bounded set. Therefore it is finite because $\|e_j - e_k\| = 2$ for all j, k. And this would resolve the proposition of the problem.

In order to proceed with the program, let us note that if $\varphi \in W$, then we have

$$(I - E_1)\varphi = BA\varphi = 0$$

In other words, for all elements $\varphi in V$ we have

$$E_1\varphi = \varphi$$

We don't have control of E_1 on V. But fortunately our argument will still yield finite dimensionality of W and that is enough for the proposition in the problem.

Other propositions of the problem is that the *range of A is closed in H_2*.

I won't try to get any efficient proofs today on this problem and just try to have some understanding that is inefficient. Let's assume that A is known to be injective on V, so that A is invertible on $Ran(A) \subset H_2$. Let us then examine what happens to limits on $Ran(A)$. Suppose $\psi_j \in Ran(A)$ and $\psi_j \to \psi$ in H_2. We want to understand whether $\psi \in Ran(A)$ or whether it escapes from $Ran(A)$. With our assumption that A is an injective isomorphism on $V \subset H_1$ we can find unique φ_j such that

$$\psi_j = A\varphi_j$$

And we have $A\varphi_j \to \psi$. Then we apply B to the sequence and obtain

$$BA\phi_j = \varphi_j - E_1\varphi_j$$

So $B\psi_j = \varphi_j - E_1\varphi_j$ and these converge to $B\psi$.

We are a bit closer to our goal, to show if $Ran(A)$ is closed, if we can show

$$B\psi \in V$$

The arguments here are not yet clear because we are not attempting to provide precise mathematical arguments yet. We are, instead, exploring the issues without seeking sharp mathematical arguments and results.

What we would like is to conclude somehow that $B\psi \in V$. We want to, but are not yet sure, use an argument like: $B\psi_j \in V$ which are all orthogonal to W and since V is closed, the known limit, $B\psi \in V$. Now since A is an isomorophism on V with $Ran(A)$ we conclude $\psi \in Ran(A)$.

Let's now assume that we have proven that $Ran(A)$ is closed. Let $H_2 = V_2 \oplus W_2$ where $V_2 = Ran(A)$ and W_2 is its orthocomplement. Then we want to use $E_2 = I$ on W_2 to conclude W_2 is finite dimensional.

13. My Intention Generally For The Stanford Quals Problems

The approach to II.4 above illustrates the discipline – or lack thereof – for my approach to working. My experience with scientific problems had been from industry, from quantitative Finance at Lehman and Gresham in New York, and in the past decade working on developing new closed form stochastic volatility models that are some of the most accurate models of empirical options prices in the markets, to models of signals from time-of-flight mass spectrometers in pharmaceutical biotechnology at Predicant/Biospect in 2003-5 and Medium Frequency Alpha strategies that D. E. Shaw is trading now and I can't get my $120 million obstructed by the racial murderer and horrible evil criminal Bill Gates. My discipline is to *trust my own intuition as a reliable source* and proceed with some simplified model of the situation. The self-trust of a Scientist is not a *posture* but part of being a good scientist. A Scientist who has problems with self-trust is not going to do great science. The approach is to trust my own intuitions, right or wrong, and try to lay down some model of a complex situation without worry about it's actual ability. Then examine the situation over many passes later. So I don't really worry too much if I am wrong; correcting errors is not a matter masochistically punishing myself. It is a foolish Scientist indeed who believes that he will not make errors and a developed Scientist is never embarrassed about errors. In the end no one actually cares about errors and mistakes. The world appreciates successes, and not many people have time for errors. There are some people who believe their job is to find errors in other people's work. They are menial people. The great people are not concerned with the errors; they are ubiquitous and they ought to be because we are not Omniscient and Omnipotent but serious people seeking serious answers and so we accept our errors without too much trouble. It is only with time and repeated scans and deeper intuition that true progress is achieved, and I understand this well and have developed my own ways of handling this over decades.

There is never any perfection in understanding Nature; perfection is relative. My Four-Sphere Theory is a sort of perfection that will be able to displace quantum field theory and relativity and expansionary cosmology, and then demands of perfection in three centuries will see it as primitive and weak. That is just how we, the human race evolve over time in our intellectual development, and it is good.

I am seeking some better understanding of *Nature* always, and that is the constant behind my motivations. This sometimes involves having better understanding of *mathematical situations* but for me proofs of mathematical propositions are quite secondary *by lifetime of cultivation* and not by negligence. Very rarely is rigorous mathematical propositions ever exactly right in scientific theories. In fact the question of when this occurs I believe is a wide open question of extreme interest to all scientists for which very little exists today.

In any case, right now my focus is on these Stanford Analysis Ph.D. Quals questions, and so I want to make some progress on gaining understanding of the issues.

I will take a look next at this issue of convergence of probability measures on $[0,1]$. These situations are not clear to me at all and I do not mind taking a slow route.

Suppose μ_n are a sequence of probability measures on $\mathbf{R}^n$. The notion of convergence that I have some understanding is weak convergence. If there exists a probability measure μ on $\mathbf{R}^n$ with the property that for all bounded (Lebesgue) measurable functions, we have

$$\int f\mu_n(dx) \to \int f\mu(dx)$$

then we say that μ_n converges to μ. This in my mind is *Prokhorov theory*, regardless of whether it is right or not. The main result that I remember about these things is a uniform integrability condition that is the essence of compactness for μ_n. This is covered in Daniel Stroock's book *Probability Theory: An Analytic View* p. 116.

A subset S of $M_1(\mathbf{R}^n)$ sequentially relatively compact iff

$$\lim_{R\to\infty} \sup_{\mu\in S} \mu(\mathbf{R}^n - B(0,R)) = 0$$

This is the heart of convergence of probability measures. Now the situation in problem II.5 is in $[0,1]$ and so the condition above will be trivially satisfied if we just extend all measures to $\mathbf{R}$ by giving zero measures outside $[0,1]$. It looks to me like the II.5 situation is more direct. The other way that weak convergence is established is by proving that the *characteristic functions*, i.e. Fourier transforms, of the measures converge to a function and then proving that that the limit is a Fourier transform of a measure.

14. Problem II.5

Let A_n be the middle-thirds Cantor set after n-th iteration. Show if f is continuous on $[0,1]$ then

$$\frac{1}{m(A_n)}\int_{A_n} f$$

exists where $m(A_n)$ is the Lebesgue measure. Show there exists a Borel measure μ so

$$\lim_{n\to\infty} m(A_n)^{-1}\int_{A_n} f = \int_0^1 f d\mu$$

Here we have one weak convergence of probability measure here, it looks like, and one has to use something like Prokhorov's theory. The major issue here is that we have some probability measures μ_n on $[0,1]$ and we want to prove that they will converge to some Borel measure μ weakly, i.e. that

$$\lim_{n\to\infty}\int f\mu_n(dx) = \int f\mu(dx)$$

for all bounded measurable functions.

My immediate instinct in these situations is to rush to attempting to prove

$$\hat{\mu_n}(\xi)$$

the Fourier transforms of the measures converge. Let's see if this would do any good here.

Let us then just take a look at Fourier transform for measures that are indicators of small compact intervals in $\mathbf{R}$.

We start with

$$\frac{d}{dx}[(-i\xi)^{-1}e^{-ix\xi}] = e^{-ix\xi}$$

and integrate on $[a,b]$ using fundamental theorem of calculus. and obtain:

$$\hat{\rho}(\xi) = \frac{1}{i\xi}(e^{-ia\xi} - e^{-ib\xi}) \tag{3}$$

Then the using some sort of superposition we can take this formula and get Fourier transforms of all sorts of probability measures on $[0, 1]$ where we know which are indicator functions of I_j each of which are disjoint intervals. So that's the first approach. If the corresponding Fourier transforms can be controlled as $n \to \infty$ by analytic means, we'll get a limit of $\hat{\mu_n}$ and we will proclaim that we have in our hands some $\hat{\mu}$ and a probability measure μ on $[0, 1]$ and we're pleased that our approach yielded fruit.

14.1. **Invocation of the Jambalaya Principle.** We remind our dear reader that in Mathematics we deal with the Jambalaya Principle which I shall articulate again. In coding and software engineering you have the Sphaghetti Principle. I love Spaghetti and when I went to visit Rome for some months I made pilgrimmage to all sorts of restaurants to relish the particular divinity of Roman spaghetti which is infinitely superior to the dishes from Italian restaurants in the United States. I do not comprehend the mysteries of Nature that makes things that look similar to the naked eye can be so different, where the Roman Spaghetti are gifts of Olympus to Mankind and the American Italian restaurant fare is not even anywhere as delicate and tasty and refined. But the point is that the Spaghetti Principle says you should keep your Italian fine cuisine far from your code and software design.

In Mathematics, by contrast there is no Spaghetti principle. This has to do with the thankful lack of pointers in Mathematics, especially pointers that point to this thing and that thing and to other pointers that, once you follow them will lead to the Hellenistic Mythological realms and find Labyrinths in Minotaurs, then take a trip to Medieval Japan and by the end you determine what is happening, there will be a huge memory leak, various coders reassuring their supervisors that the whole thing is so solid that their bonuses should justifiably exceed all CEOs on Earth.

In Mathematics something else happens, the Jambalaya Principle, where you have a beautiful idea, something so neat and clean and magnificent, but you do not prepare for the deluge of terms when the countably infinite number of terms will assault you and suddenly you are looking at the sky wondering whether you died by a stampede of reindeers who were moving silently and very fast or whether it was your lack of anticipation of the Jambalaya.

14.2. **Cantor Endpoints Are Series Of 0 and 2.** Now I will introduce some conventions without any proof for the *endpoints* of Cantor Set.

Denote by E the set of endpoints of Cantor sets. The claim is that

$$E = 0.d_{1}2 \cdots : d_j \in \{0, 2\}\}$$

The endpoints of Cantor set are those that in the base-3 fractional representation only contains 0 or 2. I will just assume this without proof. The base-3 representation is precisely the subdivision scheme. I have decades of experience coding, so this sort of translation is routine there.

The purpose is to get some concrete grasp of the endpoints. We can define truncated endpoints up to N digits by $E_N \subset E$ where after N digits the rest are zero. This is a fairly efficient representation of endpoints.

14.3. **The Characteristic Function.** Using our representation scheme, we will have to examine μ_N and see what is new. Being lazy let us start with abstract notation:

$$m\hat{u_{N+1}}(\xi) - \hat{\mu_N}(\xi) = \sum_{a \in E_{N+1} - E_N} F_N(\xi)$$

where we introduce F_N as a symbol for a sum of terms that we want to avoid telling the reader immediately, for fear that the Jambalaya will overwhelm us.

14.4. **Number of Intervals And Their Sizes.** I want to consider again our reference Fourier transform. Let ρ be the probability measure on $[a,b]$ that is

$$\rho(dx) = \frac{1}{b-a} 1_{[a,b]} dx$$

Then the Fourier transform is

$$\hat{\rho}(\xi) = \frac{1}{i\xi}(e^{-ia\xi} - e^{-ib\xi})$$

We want to apply this to all the new intervals that appear in the Cantor set at level $N+1$ and ignore those that appeared at level N. The result will be F_N and we would like an estimate for $|F_N(\xi)|$ perhaps only as a function of N. If we have something that decays fast enough as $N \to \infty$ we will conclude convergence of $\hat{\mu_N}$ and declare a probability measure in the limit.

In order to make progress let's expand the Taylor expansion of $\hat{\rho}(\xi)$ seeking some way to get our desired bound on F_N.

$$(i\xi)|b-a|\hat{\rho}(\xi) = (-i\xi)(b-a) + \frac{1}{2}\xi^2(b^2-a^2) + O(|\xi|^3)$$

That's very nice we have an expansion

$$\hat{\rho}(\xi) = 1 - C(a,b)\xi^2$$

Let's just think for a moment. The F_N will be various weigthed sums of these things where a and b will vary. It is still not intuitively clear why $|F_N|$ will be very small but we're certainly heartened by the Taylor expansion.

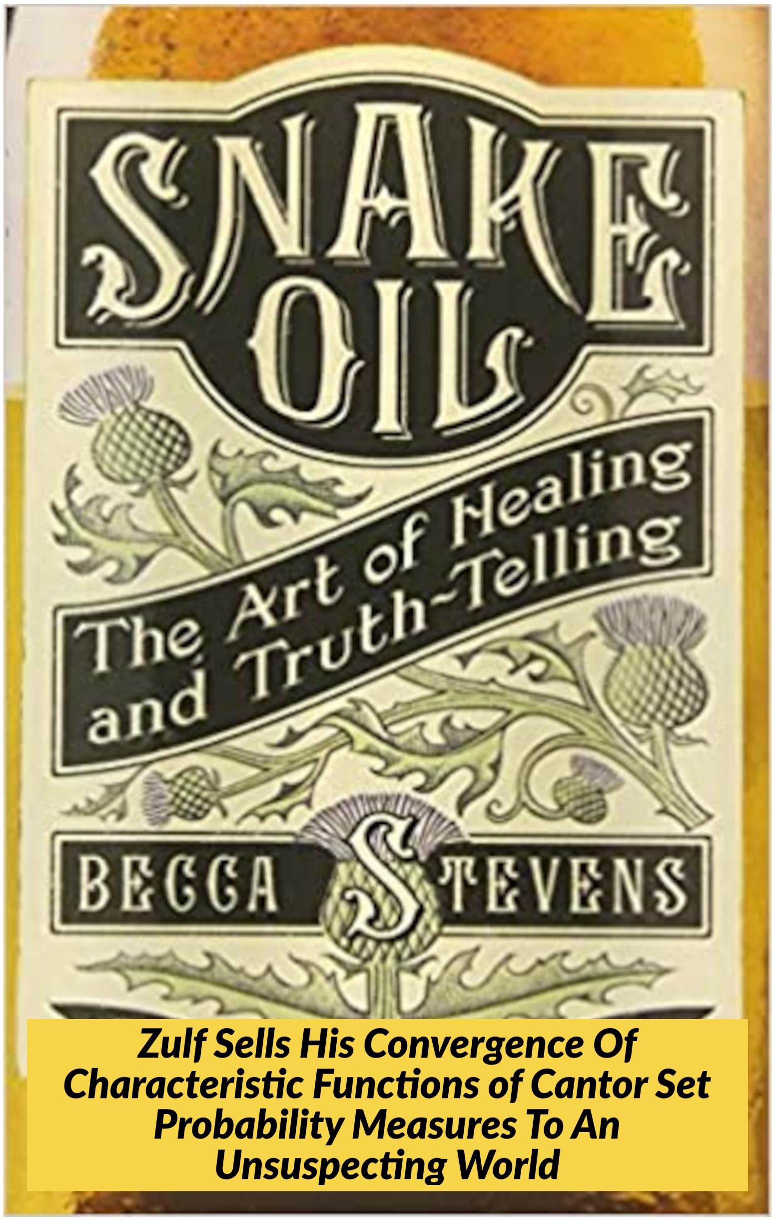
We can formalise one aspect of the main problem of II.5, and that is that in one way or another, we will have

$$\hat{\mu_N}(\xi) = 1 - C(\mu_N)\xi^2 + e_N|\xi|^3$$

where we totally ignore the complexities in $C(\mu_N)$ because the μ_N are all weighted sums of interval probability measures. This observation, while totally not giving us information about the substance of $C(\mu_N)$ still tells us

$$|\mu_{N+1}(\xi) - \mu_N(\xi)| \le |C(\mu_{N+1}) - C(\mu_N)|\xi^2$$

This removes the 1 from both which is crucial for any reasonable answer. Thus the problem of convergence of complicated measures reduces to analysis of whatever the wild things are in $C(\mu_{N+1}) - C(\mu_N)$ which is not a measure-theoretic quantity but simply various sizes and counts of intervals in $E_{N+1} - E_N$. The substantial achievement here is that we have transformed a hard measure-theory problem to one in combinatorial managements of innumerable sizes and lengths and counts of objects and all the measure-theoretic paranoia that accompanies the pathologies that occur routinely in measure theory disappear.



Ladies and Gentlemen, my characteristic functions of Cantor set probability measures indeed has magical properties that will heal all the lepers and give many faithful a path to Kingdom of God where peace reigns and the Angels sing lullabyes to all who are blessed.

The above is *humour*, and I am not in fact going to be producing any path to healing of the lepers and path to the Kingdom of God. I will ask my reader to forgive me and seek those services from someone else.

What I will actually do, instead is to make the conjecture that

$$|C(\mu_{N+1}) - C(\mu_N)| \leq C_0(2/3)^N$$

This will do what? It will give us an estimate

$$|\hat{\mu}_{N+1}(\xi) - \hat{\mu}_N(\xi)| \leq C_0(2/3)^N \xi^2$$

So for bounded intervals $\xi \in [A_0, B_0]$ we will have uniform convergence for $\hat{\mu}_N(\xi)$ to some continuous function.

This is conjectural. Let's just take a look at how fast the convergence is on R to make sure that we like the speed.

```
> ns<-seq(1,50)
> (2/3)^ns
 [1] 6.666667e-01 4.444444e-01 2.962963e-01 1.975309e-01 1.316872e-01
 [6] 8.779150e-02 5.852766e-02 3.901844e-02 2.601229e-02 1.734153e-02
[11] 1.156102e-02 7.707347e-03 5.138231e-03 3.425487e-03 2.283658e-03
[16] 1.522439e-03 1.014959e-03 6.766395e-04 4.510930e-04 3.007287e-04
[21] 2.004858e-04 1.336572e-04 8.910479e-05 5.940319e-05 3.960213e-05
[26] 2.640142e-05 1.760095e-05 1.173396e-05 7.822643e-06 5.215095e-06
[31] 3.476730e-06 2.317820e-06 1.545213e-06 1.030142e-06 6.867615e-07
[36] 4.578410e-07 3.052273e-07 2.034849e-07 1.356566e-07 9.043773e-08
[41] 6.029182e-08 4.019455e-08 2.679636e-08 1.786424e-08 1.190949e-08
[46] 7.939663e-09 5.293109e-09 3.528739e-09 2.352493e-09 1.568329e-09
```

Well it's not instantaneous, but 50 steps get us $\epsilon = 1.6 \times 10^{-9}$ which is pretty good.

I don't know if the conjecture is true yet, but I think it is good to pay attention to the rate of convergence too for these characteristic functions.

14.5. **Size And Counts of Endpoints.** A simple examination of steps in construction tells us if we let $N = 1$ Cantor set be $[0, 1]$ then the number of endpoints introduced is 2^N and the length of the intervals remaining are 3^{-N+1}. These are extremely simple and elementary issues so I won't prove them.

What I will do, instead is count the number of terms. The number of *new terms introduced* between $\hat{\mu}_N$ and $\hat{\mu}_{N+1}$ is 2^N and we don't really want to get too much into listing all of them. Instead we want to get uniform control of $(a + b)/2$ for the endpoints. It suffices just bound them by 1 to get an estimate.

We modify our conjecture to

$$|F_N(\xi)| \leq (2/3)^N \xi^2$$

The intuitive argument is that in the difference $\hat{\mu}_{N+1}(\xi) - \hat{\mu}_N(\xi)$ we have roughly 2^N terms and the lengths of intervals were 3^{-N} for each, and so after all the summing is done we can get the above bound.

15. Zulf Wants Credit For New Discovery From Problem II.5

The problem statement is that certain measures with support on the intervals in middle-thirds Cantor sets converges to a Borel measure, but I do not know if the rate of convergence was known at all. I claim discovery of the rate proportional to $(2/3)^N$. Now I have deviated from Mathematics some two decades ago to go into science in industry without completing my doctorate in Mathematics at Columbia. But I still claim the *discovery* of this rate of convergence $(2/3)^N$ as I am more familiar with scientific discoveries. I can provide a rigorous proof as well since I am now more confident that my approach was successful.

16. Appeal To All Good Men And Women To Destroy The Evil That Is Bill Gates

Never in mythological imagination has there been such a vile disgusting, evil, and harmful being as this *Bill Gates*. I will appeal to every human being on Earth of good character to do everything possible to destroy this disgusting and evil Bill Gates and totally disintegrate his Evil Empire.

Let me remind you that I am American, and I immigrated to New York in 1987 from Bengal and have lived an American life for 34 years.

No man of good Conscience would, after they had discovered such unimaginable and unspeakable evil as Bill Gates would sleep easily at night knowing that they are putting their beloved people, the human race, in danger every moment they have not totally destroyed Bill Gates and exiled him from the land of the living.

I'd like to sing a song and the man in the song is not Bill Gates. He's more the war criminal singing but he is too drab to even have a soul as beautiful as the ageing drummer boy.

I am just an ageing drummer boy
And in the wars I used to play
And I've called the tune
To many a torching session
Now they say I am a war criminal
And I'm fading away
Father please hear my confession!
I have legalized robbery
Called it belief
I have run with the money
I have hid like a thief
Re-written history
With my armies and my crooks
Invented memories
I did burn all the books
And I can still hear his laughter
And I can still hear his song
The man's too big!
The man's too strong!
Well I've tried to be meek
And I have tried to be mild
But I spat like a woman
And I sulked like a child
I have lived behind walls
That have made me alone
Striven for peace
Which I never have known
And I can still hear his laughter
And I can still hear his song...
The man's too big!
The man's too strong!
Well the sun rose on the courtyard
And they all did hear him say
You always was a Judas
But I got you anyway
You may have got your silver
But I swear upon my life
Your sister gave me diamonds
And I gave 'em to your wife
Oh Father please help me
For I have done wrong...
The man's too big!
The man's too strong!

Oh yeah, oh yeah, don't let him rewrite history with his armies and his crooks! Oh yeah! He's legalized robbery and he's called it belief! He's still hear my laughter and he's still hear my song! Oh yeah! Oh yeah!

He spat like a woman, and he sulked like a child! Da da da oh yeah!

He invented memories for the entire world to buy! Oh yeah!

References

[1] https://terrytao.wordpress.com/2009/02/21/245b-notes-11-the-strong-and-weak-topologies/

[2] https://math.stackexchange.com/questions/3005795/comparison-of-bounded-operator-topologies