Comments, Notes, and Questions for Real Analysis II (640:502)

The material that we plan to cover can be conveniently grouped into several units: **Unit 1**. Completeness and its consequences and applications; **Unit 2**. Compactness and its applications (including some aspects of spectral theory of compact linear operators); **Unit 3**. Duality, weak convergence and weak compactness; **Unit 4**. Elements of distribution theory; and **Unit 5**. Additional topics on measures (Lebesgue-Radon-Nikodym Theorem, Hausdorff measure, and area formula for submanifolds in the Euclidean space).

1 Completeness and its consequences and applications

We will first study how completeness,in conjunction with the Baire category theorem, is used to give the three basic theorems involving bounded linear operators: the Open Mapping Theorem, the Closed Graph Theorem, and the Uniform Bounded Principle. More importantly we will illustrate how these theorems are applied. We will introduce Fourier series and Fourier transforms in this context and illustrate how these tools are used. This is also a natural context to discuss the Interpolation Theorem of L^p spaces. We will then discuss how completeness is used in the Dirichlet principle to solve some boundary value problems and in naturally extending the notion of solutions to certain differential equations. Relevant sections from the text are **5.3, 8.1 – 8.5, 6.5**.

Assignment 1. *The *-ed problems are due in class on Thursday, Jan. 28.*
Chapter 5: 27, 29, 30*, 32, 37*, 38*, 40, 42.*

1.1 Fourier series

We now do a quick review of Fourier series and raise some of the issues that need to be addressed. Fourier's original idea was to expand any function $f(x)$ on $[0, 2\pi]$, say, in terms of the functions $\{1, \cos x, \sin x, \cos 2x, \sin 2x, \cdots\}$ in the form of

$$f(x) \sim a_0 + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)) \tag{1}$$

for some choice of the coefficients $\{a_n, b_n\}$, which would be called the Fourier coefficients of $f(x)$. Some immediate questions are:

Question 1. *(a). How to determine the coefficients $\{a_n, b_n\}$ in terms of $f(x)$? (b). In what sense does the series* (1) *converge back to $f(x)$?, and (c). How to reconstruct $f(x)$ from the Fourier coefficients $\{a_n, b_n\}$, if the series* (1) *does not converge pointwise?*

Fourier determined the coefficients $\{a_n, b_n\}$ based on the properties

$$\int_0^{2\pi} \cos(nx)\cos(mx)\,dx = \int_0^{2\pi} \sin(nx)\sin(mx)\,dx = \pi\delta_{nm}, \quad \int_0^{2\pi} \cos(nx)\sin(mx)\,dx = 0,$$

for $n, m = 1, 2, \cdots$, and the ones involving $\cos(nx)$ also allow $n = 0$. In other words, $\{1, \cos x, \sin x, \cos 2x, \sin 2x, \cdots\}$ is a set of orthogonal functions in $L^2[0, 2\pi]$. Then from the Hilbert space point of view,

$$a_n = \frac{(f(x), \cos(nx))}{(\cos(nx), \cos(nx))} = \begin{cases} \frac{1}{\pi}\int_0^{2\pi} f(x)\cos(nx)\,dx, & \text{for } n \neq 0, \\ \frac{1}{2\pi}\int_0^{2\pi} f(x)\,dx, & \text{for } n = 0, \end{cases}$$

and

$$b_n = \frac{(f(x), \sin(nx))}{(\sin(nx), \sin(nx))} = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin(nx)\,dx \quad \text{for } n \neq 0.$$

To answer questions (b) and (c) above, the most natural approach is to form the partial sums

$$S_m f(x) = a_0 + \sum_{n=1}^{m} (a_n \cos(nx) + b_n \sin(nx))$$

and study the convergence of $S_m f(x)$ to $f(x)$ as $m \to \infty$. From our general discussion on Hilbert space, the convergence in (1) holds in $L^2[0, 2\pi]$ sense provided one of the criteria in **Theorem 5.27** holds. The **completeness** criterion would hold if $(f, \cos(nx)) = (f, \sin(nx)) = 0$ for $n = 0, 1, 2, \cdots$ implies $f(x) = 0$ in $L^2[0, 2\pi]$. This is a **uniqueness** criterion on the Fourier coefficients: if $f, g \in L^2[0, \pi]$ have the same Fourier coefficients, then they are the same. Bessel's inequality implies that

$$\frac{a_0^2}{(1,1)} + \sum_{n=0}^{\infty}\left(\frac{a_n^2}{(\cos(nx), \cos(nx))} + \frac{b_n^2}{(\sin(nx), \sin(nx))}\right) = \frac{a_0^2}{2\pi} + \sum_{n=1}^{\infty}\frac{1}{\pi}(a_n^2 + b_n^2) \leq \int_0^{2\pi} |f(x)|^2 dx.$$

So any $f \in L^2[0, 2\pi]$ gives rise to a sequence $\{1, a_1, b_1, a_2, b_2, \cdots\}$ in l^2. **Proposition 5.30** says that this map is a unitary map *onto* l^2, provided we can verify the uniqueness criterion above.

For f in other function spaces such at $C[0, 2\pi]$ or $L^p[0, 2\pi]$ for some $p \geq 1$, the Fourier coefficients can be determined in the same way, and one can ask the same questions on reconstruction, convergence and uniqueness.

Question 2. *(a) Given $f \in C[0, 2\pi]$ or $L^p[0, 2\pi]$ for some $p \geq 1$, what can one say about its Fourier coefficients? Do they form a sequence in some l^q? (b). If there is a positive answer to (a), is the map in (a) onto? invertible? and (c). In what sense can the Fourier coefficients be used to reconstruct f in such settings?*

One basic question in the classical setting is

Question 3. *Given $f \in C[0, 2\pi]$ or $L^p[0, 2\pi]$ for some $p \geq 1$, does the partial sum $S_m f(x)$ converge to $f(x)$ for all or appropriate x?*

Let's first work out $S_m f(x)$ in terms of f:

$$
\begin{aligned}
S_m f(x) &= \int_0^{2\pi} \left\{ \frac{1}{2\pi} + \frac{1}{\pi} \sum_{n=1}^{m} [\cos(nt)\cos(nx) + \sin(nt)\sin(nx)] \right\} f(t)\, dt \\
&= \int_0^{2\pi} \left\{ \frac{1}{2\pi} + \frac{1}{\pi} \sum_{n=1}^{m} \cos n(x-t) \right\} f(t)\, dt \\
&:= \int_0^{2\pi} D_m(x-t) f(t)\, dt,
\end{aligned}
$$

where

$$
D_m(x-t) = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{n=1}^{m} \cos n(x-t) = \frac{\sin(m+\frac{1}{2})(x-t)}{2\pi \sin\frac{x-t}{2}}
$$

is called the Dirichlet kernel. Noting that $\int_0^{2\pi} D_m(t)dt = 1$ for any m, we have

$$
S_m f(x) - f(x) = \int_0^{2\pi} D_m(x-t)\, [f(t) - f(x)]\, dt = \int_{-\pi}^{\pi} D_m(\tau)\, [f(x+\tau) - f(x)]\, d\tau,
$$

here we have tacitly assumed that f is 2π-periodic, or in general, extended the given f into a 2π-periodic function. If f is continuous at x, then we can attempt to break up the above integral into two parts: $|\tau| \leq \tau_0$ for some small τ_0 so that $|f(x+\tau) - f(x)|$ remains small there, and $|\tau| \geq \tau_0$. To make the integral over $|\tau| \geq \tau_0$ small, since $D_m(\tau)$ does not converge to 0 there as $m \to \infty$, one has to make use of the rapid oscillation of $D_m(\tau)$ there when $m \to \infty$; to make the integral over $|\tau| \leq \tau_0$ small, one hopes that $\int_{|\tau| \leq \tau_0} |D_m(\tau)| d\tau$ remains bounded as $m \to \infty$, which is, unfortunately, not the case, despite the fact that $\int_0^{2\pi} D_m(\tau) d\tau = 1$!

One solution to deal with the difficulty is to replace $S_m f(x)$ by

$$
\sigma_m f(x) = (1+m)^{-1} \sum_{k=0}^{m} S_m f(x) = \int_{-\pi}^{\pi} K_m(t) f(x-t) dt, \tag{2}
$$

where

$$K_m(x) = (1+m)^{-1}\sum_{k=0}^{m} D_k(x) = \frac{1}{2(m+1)\pi}\left(\frac{\sin(\frac{m+1}{2}x)}{\sin\frac{x}{2}}\right)^2. \tag{3}$$

Then we still have $\int_{-\pi}^{\pi} K_m(\tau)d\tau = 1$, and now since $K_m(\tau) \geq 0$, we also have $\int_{|\tau|\leq\tau_0} |K_m(\tau)|d\tau \leq 1$ for any $\tau_0 > 0$. Now if f is continuous at x, then for any $\epsilon > 0$, we can find $\tau_0 > 0$ such that $|f(x+\tau) - f(x)| \leq \epsilon$ when $|\tau| \leq \tau_0$, thus

$$\begin{aligned}|\sigma_m f(x) - f(x)| \leq & \left\{\int_{|\tau|\leq\tau_0} + \int_{|\tau|\geq\tau_0}\right\} |K_m(\tau)||f(x+\tau) - f(x)|d\tau \\ \leq & \epsilon\int_{|\tau|\leq\tau_0} |K_m(\tau)|d\tau + \int_{|\tau|\geq\tau_0} |K_m(\tau)||f(x+\tau) - f(x)|d\tau.\end{aligned}$$

Now $K_m(\tau) \to 0$ as $m \to \infty$ uniformly for $|\tau| \geq \tau_0$, so for some $N > 0$, $|K_m(\tau)| \leq \epsilon$ for all $m \geq N$ and $|\tau| \geq \tau_0$, whence,

$$|\sigma_m f(x) - f(x)| \leq \epsilon + 2||f||_u\epsilon,$$

proving that $\sigma_m f(x) \to f(x)$ as $m \to \infty$. One immediate consequence is the uniqueness criterion.

Corollary 1. *Suppose that f is a 2π-periodic continuous function on $[0, 2\pi]$ and all its Fourier coefficients are 0, then $f \equiv 0$.*

The above proof actually also shows that if f is a 2π-period continuous function on $[0, 2\pi]$ then $\sigma_m f$ converges to f uniformly as $m \to \infty$. Since $\sigma_m f(x)$ is a trigonometric polynomial for each m, we have also proved the following

Corollary 2. *Any 2π-periodic continuous function on $[0, 2\pi]$ can be uniformly approximated by trigonometric polynomials.*

It's possible to use Corollary 2 to prove the Weierstrass polynomial approximation theorem, namely, *any continuous function on a finite closed interval can be uniformly approximated by polynomials.*

The discussion above contains two ideas that are worth further exploring: one is the concept of *convolution*, the other is the concept of *good kernels*, or *approximation of identities*—both are taken up in **8.2**.

To go back to the question of characterizing the Fourier coefficients of functions in $C[0, 2\pi]$ or $L^p[0, 2\pi]$, we first have

Proposition 1. *(a). Suppose that f is a C^k 2π-periodic function, then $|a_n|, |b_n| = o(1/n^k)$; in fact, the sequence $\{n^k a_n, n^k b_n\}$ is in l^2.*
(b). (The Riemann-Lebesgue Lemma) Suppose that $f \in L^p[0, 2\pi]$ for some $p \geq 1$, then its Fourier coefficients $a_n, b_n \to 0$ as $n \to \infty$.

The Riemann-Lebesgue Lemma establishes a (bounded) linear map $\mathcal{F}$ from $f \in L^1[0, 2\pi]$ to the sequence of its Fourier coefficients, which lies in c_0, a closed subspace of l^∞. An immediate question refining Question 2 is

Question 4. *Is there a more precise decaying rate for the sequence of the Fourier coefficients of a function in $C[0, 2\pi]$ or $L^p[0, 2\pi]$? Is the map $\mathcal{F} : L^1[0, 2\pi] \mapsto c_0$ onto?*

To minimize the need of having to deal with the Fourier cosine and sine coefficients separately, it is often more convenient to use $\{e^{inx}\}_{n\in\mathbb{Z}}$ to replace $\{1, \cos x, \sin x, \cdots\}$, and allow complex valued functions. Then the L^2 inner product is extended naturally, and $\{e^{inx}\}_{n\in\mathbb{Z}}$ forms an orthogonal basis on $L^2[0, 2\pi]$. Then any $L^1[0, 2\pi]$ function f has a Fourier expansion

$$f \sim \sum_{n\in\mathbb{Z}} \hat{f}(n)e^{inx}, \tag{4}$$

with

$$\hat{f}(n) = \frac{(f, e^{inx})}{(e^{inx}, e^{inx})} = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-inx}dx.$$

Then $c_0 = a_0$, $\hat{f}(n) = (a_n - b_n i)/2$ for $n = 1, 2, \cdots$, and $\hat{f}(n) = (a_{|n|} + b_{|n|}i)/2$ for $n = -1, -2, \cdots$. Furthermore

$$\sum_{n=-m}^{m} \hat{f}(n)e^{inx} = a_0 + \sum_{n=1}^{m} (a_n \cos(nx) + b_n \sin(nx)) = S_m f(x).$$

The text uses $\{e^{2\pi inx}\}_{n\in\mathbb{Z}}$ on $L^2[0, 1]$ instead, and one just modifies the above formulas accordingly. In this framework,

$$\sigma_m f(x) = \sum_{n=-m}^{m} \left(1 - \frac{|n|}{m+1}\right) \hat{f}(n)e^{2\pi nix}.$$

1.2 Uniform boundedness, open mapping and closed graph theorems

The statements and proofs are straight forward—observe carefully where in the statement and proof the completeness enters. The more difficult task is to learn how to apply them. Sometimes it may not be easy to apply them directly, and one needs to go back to the proofs to extract what one needs. We will give an initial illustration of the applications of these ideas in the context of proving that (i) given any point in $[0, 2\pi]$, there is some continuous function whose Fourier series diverges there; and (ii) the map $\mathcal{F}$ sending an $f \in L^1[0, 2\pi]$ to the sequence of its Fourier coefficients in c_0 is not onto.

Exercise 1. *Prove that for any countable sequence $\{x_n\}$ in $[0, 2\pi]$, there is a continuous function f whose Fourier series diverges at each of x_n.*

Exercise 2. *A basic PDE is the Poisson equation $\Delta u(x) = f(x)$, where $\Delta = \sum_{i=1}^n \partial_i^2$ is the Laplace operator. To avoid dealing with boundary conditions, we will just consider the problem on the whole $\mathbb{R}^n$ and require both $u(x)$ and $f(x)$ to decay to 0 as $x \to \infty$. More precisely, given $f \in C_0(\mathbb{R}^n) := \{v \in C(\mathbb{R}^n) : v(x) \to 0 \text{ as } x \to \infty\}$, we look for a solution $u \in C_0^2(\mathbb{R}^n) := \{w \in C^2(\mathbb{R}^n) : \partial_x^\alpha w(x) \to 0 \text{ as } x \to \infty \text{ and for all } |\alpha| \leq 2\}$. Unfortunately, for $n \geq 2$, there exists $f \in C_0(\mathbb{R}^n)$ for which there is no solution to $\Delta u(x) = f(x)$ in $C_0^2(\mathbb{R}^n)$. Prove this. (*Hints: If for every $f \in C_0(\mathbb{R}^n)$ there is a solution $u \in C_0^2(\mathbb{R}^n)$ to $\Delta u(x) = f(x)$, there would exist a constant $C > 0$ such that $||u||_{C_0^2(\mathbb{R}^n)} \leq C||\Delta u||_{C_0(\mathbb{R}^n)}$. Try to construct sequence of functions in $C_0^2(\mathbb{R}^n)$ that would violate this estimate by first looking for radial solutions to $\Delta u(x) = 0$, which unfortunately has a singularity, then modifying the radial solutions.*) The solvability of the Poisson equation has a satisfactory answer if we work with some different choice of function spaces.*

1.3 Convolutions and good kernels

After introducing the notion of convolutions, we will initially focus on Propositions 8.7, 8.8, 8.10, and Theorem 8.14. Note that Theorem 8.14 holds also for convolutions on the round circle $\mathbb{T}$, provided the function $\phi(x)$ used to construct $\phi_t(x)$, considered as a function on $[-\pi, \pi)$, has compact support in (π, π).

Definition 1 (Definition of good kernels). *A family of kernels $\{K_i(x)\}_{i=1}^\infty$ on the circle (or in $\mathbb{R}^n$) is a family of good kernels, if it satisfies*

(a) For all $i \geq 1$,

$$\int K_i(x)dx = 1,$$

(b) There exists $M > 0$ such that for all $i \geq 1$,

$$\int |K_i(x)|dx = 1,$$

(c) For every $\delta > 0$,

$$\int_{\delta \leq |x|} |K_i(x)|dx \to 0, \quad \text{as } i \to \infty.$$

One could also replace the countable parameter i by a continuous parameter.

As in the text, if $\phi \in L^1(\mathbb{R}^n)$ (or $\mathbb{T}^n$) and $\int \phi = 1$, then $\phi_t(x) = t^{-n}\phi(x/t)$ provides a family of good kernels as $t \to 0$. But the Fejer kernel in (3) is not defined this way.

Assignment 2: The *-ed problems are due in class on Thursday, Feb. 4.
Chapter 8: 6, 7, 8*, 14*, 16*, 28, 34, 35*, 36.

1.4 Some situations where Fourier series and transforms arise naturally

Fourier series and transforms arise naturally in finding solutions to standard linear PDEs. We will illustrate how they enter through a construction of solutions to the heat equation either as an initial value problem

$$\begin{cases} \dfrac{\partial u}{\partial t} - \Delta u = 0, \quad \text{for } x \in \mathbb{R}^n \text{ and } t > 0, \\ u(x,0) = f(x); \end{cases} \tag{IVP}$$

or as a boundary value problem (a 1-D case on $[0, l]$, for simplicity)

$$\begin{cases} \dfrac{\partial u}{\partial t} - \Delta u = 0, \quad \text{for } 0 < x < l \text{ and } t > 0, \\ u(0,t) = u(l,t) = 0, \quad \text{for } t > 0, \\ u(x,0) = f(x). \end{cases} \tag{BVP}$$

Our first step is to look for some sample solutions of the PDE using the method of separation of variables, ignoring for the moment the initial or boundary conditions, i.e., we look for solutions of the form $u(x,t) = T(t)X(x)$ and *reduce the construction of solutions of a PDE to some set of ODEs*. Substituting into the heat equation, we are led to $T_t(t)X(x) = T(t)X_{xx}(x)$. We deduce that $X_{xx}(x)/X(x)$ must be a constant, independent of (x,t). Call this constant $-\lambda$. Then we have a set of ODEs

$$X_{xx} = -\lambda X(x), \tag{5}$$

$$T_t(t) = -\lambda T(t). \tag{6}$$

Note that for each ξ, $X(x) = e^{ix\xi}$ is a solution of (5) with $\lambda = \xi^2$, and $T(t) = e^{-\xi^2 t}$ solves (6) for the same λ. So $u_\xi = e^{ix\xi - \xi^2 t}$ is a solution to (5) and (6). By the superposition principle, any finite linear combination of such solutions

$$\sum_{\xi \in \text{a finite set}} c(\xi) e^{ix\xi - \xi^2 t}$$

is a solution of the heat equation.

Question 5. *Does this construction generate all solutions to the Cauchy problem for the homogeneous heat equation?*

To satisfy an arbitrary initial data, it is obviously not enough to take only finite linear combinations of solutions of the form $e^{ix\xi-\xi^2 t}$. For an infinite linear combination, or an integral of the form $\int c(\xi)e^{ix\xi-\xi^2 t}d\xi$, one issue is the convergence of such integrals for $(x,t)\in\mathbb{R}^1\times\mathbb{R}^+$; another issue is how to choose $c(\xi)$ to satisfy a given initial data $u(x,0)$. Although for complex valued ξ, $e^{ix\xi-\xi^2 t}$ is also a solution of the homogeneous heat equation, but such solutions grow exponentially in x as x goes to one end of infinity. So we take $\xi\in\mathbb{R}$. At least formally, the factor $e^{-\xi^2 t}$ helps with the convergence of the integral for $t>0$, but not for $t<0$; also we expect

$$u(x,0)=\int_{\xi\in\mathbb{R}} c(\xi)e^{ix\xi}d\xi. \tag{7}$$

We now ask

Question 6. *Given an initial data $u(x,0)=f(x)$, how do we choose $c(\xi)$ so that the representation* (7) *is valid in some sense?*

We will recognize that (7) is possible if we choose $c(\xi)$ to be the Fourier transform of $u(x,0)$, up to a constant. So we now have an algorithm: construct $c(\xi)$ using $f(x)$, then construct the integral $\int c(\xi)e^{ix\xi-\xi^2 t}d\xi$, and verify that it is a genuine solution to the heat equation and it takes on the initial data $f(x)$ as $t\searrow 0$.

Caution. *We are led to ask whether*

$$\int c(\xi)e^{ix\xi-\xi^2 t}d\xi \to f(x), \quad \textit{as } t\searrow 0, \tag{8}$$

while the convergence of

$$\int c(\xi)e^{ix\xi}d\xi \to f(x), \quad \textit{as } t\searrow 0, \tag{9}$$

is not directly relevant to our task of constructing the solutions to (IVP). You should compare this with the situation of asking for the convergence of the Fourier series

$$S_m f(x)=\sum_{n=-m}^{m}\hat{f}(n)e^{inx}$$

vs the convergence of the Fejer series

$$\sigma_m f(x)=\sum_{n=-m}^{m}\left(1-\frac{|n|}{m+1}\right)\hat{f}(n)e^{inx}.$$

We found the convergence issue for the latter easier to handle than the one for the former.

Our experience in dealing with Fourier series also taught us that, if $H_t(x)$ is a function such that $\hat{H}_t(\xi) = e^{-\xi^2 t}$, then

$$\int c(\xi) e^{ix\xi - \xi^2 t} d\xi = H_t * f(x),$$

and we could answer the convergence more easily using the convolution properties, provided that we have some good control on H_t.

If one is interested in solving the (BVP), then in our sample solutions, we want $X(0) = X(l) = 0$. So the choice of $X(x)$ is limited to satisfy a boundary value problem of ODEs:

$$\begin{cases} X_{xx} = -\lambda X(x), \\ X(0) = X(l) = 0. \end{cases}$$

The choice of λ is now limited to be a discrete set of numbers $\lambda = \left(\frac{n\pi}{l}\right)^2$, for $n = 1, 2, \cdots$, and $X_n(x) = \sin\left(\frac{n\pi x}{l}\right)$. So

$$\sum_{n \in \text{a finite set}} c_n \sin\left(\frac{n\pi x}{l}\right) e^{-\left(\frac{n\pi}{l}\right)^2 t}$$

is a genuine solution of the homogeneous heat equation with the boundary condition $u(0,t) = u(l,t) = 0$ for all $t > 0$. To satisfy an arbitrarily given initial data $u(x,0)$, we again need to make an infinite sum. Formally, we need to choose c_n such that

$$\sum_{n=1}^{\infty} c_n \sin\left(\frac{n\pi x}{l}\right) = u(x,0) \quad \text{on } [0,l]. \tag{10}$$

Again, the factor $e^{-\left(\frac{n\pi}{l}\right)^2 t}$ will help with the convergence for $t > 0$, but not for $t < 0$. Here we see the natural appearance of the series $\{\sin(n\pi x/l)\}_{n=1}^{\infty}$, not the full Fourier basis—they do show up if we ask, instead of the homogeneous Dirichlet boundary condition, but the periodic boundary conditions $u(0,t) = u(l,t)$ and $u_x(0,t) = u_x(l,t)$ for all $t > 0$. It's easy to see that $\{\sin(n\pi x/l)\}_{n=1}^{\infty}$ is orthogonal in $L^2[0,l]$.

Question 7. *Is $\{\sin(n\pi x/l)\}_{n=1}^{\infty}$ complete in $L^2[0,l]$? How do we handle the convergence issues of $\sum_{n=1}^{m} c_n \sin(n\pi x/l)$ as $m \to \infty$ or of*

$$\sum_{n=1}^{\infty} c_n \sin\left(\frac{n\pi x}{l}\right) e^{-\left(\frac{n\pi}{l}\right)^2 t} \quad \textit{as } t \searrow 0, \tag{11}$$

either in some L^p space or pointwise?

Exercise 3. *Choose $l = 2\pi$ and compare $\{\sin(n\pi x/l)\}_{n=1}^{\infty}$ with the classical Fourier basis $\{1, \cos x, \sin x, \cdots\}$ on an interval of length 2π. Could any function in the latter be expanded as an $L^2[0, 2\pi]$ convergent series in the former? How would you determine the coefficients? How would you interpret the convergence result, if any, when x is outside of $[0, 2\pi]$?*

Finally we raise

Question 8. *How do you tackle the (IVP) or (BVP) when the equation is not homogeneous, i.e. when the PDE becomes $u_t(x,t) - \Delta u(x,t) = g(x,t)$ for some given $g(x,t)$? and in the case of (BVP), when the boundary conditions are not homogeneous?*

Instead of tackling the convergence issues of the sine series from scratch, we can reformulate the problem into one where we can apply our knowledge on the classical Fourier series. Note two features in the set $\{\sin(n\pi x/l)\}_{n=1}^{\infty}$: (i) each is an odd function of x, (ii) these functions have a common period $2l$, not l. So it is natural to look at them as functions on $[-l, l]$, or more precisely $2l$-periodic functions; accordingly, we should extend any given function f under consideration on $[0, l]$ first to $[-l, l]$ by odd reflection, and then extend it to f_{ext} as a $2l$-periodic function on $\mathbb{R}$. The convergence issues involving f can be answered by those on f_{ext}. But a simple scaling can translate all our results on the classical Fourier series in terms of $\{e^{inx}\}_{n=-\infty}^{\infty}$ on $[-\pi, \pi]$ to results on $[-l, l]$ in terms of $\{e^{in\pi x/l}\}_{n=-\infty}^{\infty}$, or in the more familiar form $\{1, \cos(\pi x/l), \sin(\pi x/l), \cos(2\pi x/l), \cdots\}$. When f_{ext} is odd, its Fourier series in this basis consists only of the sine series, which is what we needed to investigate when solving the (BVP).

1.5 Passage from Fourier series to Fourier transforms

We will follow the notation of the text to define the Fourier transform of a function $f \in L^1(\mathbb{R}^n)$ by

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-2\pi x \cdot \xi} dx, \tag{12}$$

and the Fourier coefficients of a function $f \in L^1(\mathbb{T}^n)$ by the same formula as (12), replacing $\xi \in \mathbb{R}^n$ by $\kappa \in \mathbb{Z}^n$, and $\mathbb{R}^n$ by $\mathbb{T}^n$. We are more familiar with the algebraic and convergence properties involving the Fourier series, esp. the one-dimensional case. We will sketch here some discussion that relates the Fourier series to Fourier transforms—for simplicity, we will assume $n = 1$.

First we describe Fourier series expansions for any L^1 (or C^k) function f with compact support. Choose $l > 0$ large such that the support of f is included in $(-l, l)$ and let $g(t) = f(lt)$. Then g is compactly supported in $(-1, 1)$, and we understand some conditions under which the Fourier series $\sum_{m=-\infty}^{\infty} \hat{g}(m)e^{2\pi imt}$ converges to $g(t)$. Recall that

$$\hat{g}(m) = \int_{-1}^{1} g(t)e^{-2\pi imt}dt = \frac{1}{l}\int_{-l}^{l} f(x)e^{-2\pi imx/l}dx = \frac{1}{l}\int_{\mathbb{R}} f(x)e^{-2\pi ixm/l}dx = \frac{1}{l}\hat{f}(m/l).$$

If f is C^2, for example, then we have

$$f(x) = \sum_{m=-\infty}^{\infty} \frac{1}{l}\hat{f}(m/l)e^{2\pi ixm/l},$$

where the series converges uniformly, while if f is only known to be in C^0, then we also have

$$f(x) = \lim_{N\to\infty} \sum_{m=-N}^{N} \left(1 - \frac{|m|}{N+1}\right) \frac{1}{l}\hat{f}(m/l)e^{2\pi ixm/l}.$$

Next we derive from the above an expansion (or integral) for f that is not dependent on l. If f is in C^2, then as in the proof for the Riemann-Lebesgue Lemma, we can prove that $|\hat{f}(\xi)|$ has a fast enough decay rate as $\xi \to \infty$ so that $\int_{\mathbb{R}} |\hat{f}(\xi)|d\xi < \infty$ and

$$\lim_{l\to\infty} \sum_{m=-\infty}^{\infty} \frac{1}{l}\hat{f}(m/l)e^{2\pi ixm/l} = \int_{\mathbb{R}} \hat{f}(\xi)e^{2\pi ix\xi}d\xi,$$

proving $f(x) = \int_{\mathbb{R}} \hat{f}(\xi)e^{2\pi ix\xi}d\xi$—you should justify this step. If f is only known to be in C^0, then we can verify that

$$\lim_{N\to\infty} \sum_{m=-N}^{N} \left(1 - \frac{|m|}{N+1}\right) \frac{1}{l}\hat{f}(m/l)e^{2\pi ixm/l} = \lim_{R\to\infty} \int_{-R}^{R} \left(1 - \frac{|\xi|}{R}\right) \hat{f}(\xi)e^{2\pi ix\xi}d\xi,$$

proving that

$$f(x) = \lim_{R\to\infty} \int_{-R}^{R} \left(1 - \frac{|\xi|}{R}\right) \hat{f}(\xi)e^{2\pi ix\xi}d\xi.$$

The justification for the step above is harder than the previous case. In fact it does not seem obvious that the limit on the right exists as $R \to \infty$. We will give a different justification in the lecture. Another, slightly different, approach is as follows. Assume for simplicity that we can take $l = 1$. Then Fejer's theorem says that for any $\epsilon > 0$ there is N_0 such that when $N \geq N_0$,

$$\left| f(x) - \sum_{m=-N}^{N} \left(1 - \frac{|m|}{N+1}\right) \hat{f}(m)e^{2\pi ixm} \right| < \epsilon$$

for all x. For any parameter η, we can apply the same result to $e^{-2\pi ix\eta}f(x)$, and noting that

$$\widehat{e^{-2\pi ix\eta}f}(x)(m) = \hat{f}(m+\eta), \quad \text{by Theorem 8.22 (a)},$$

to obtain

$$\left| e^{-2\pi ix\eta}f(x) - \sum_{m=-N}^{N} \left(1 - \frac{|m|}{N+1}\right) \hat{f}(m+\eta)e^{2\pi ixm} \right| < \epsilon.$$

Here it follows from the proof of Fejer's theorem that we can take the same N_0 uniformly for η. The above estimate implies that

$$\left| f(x) - \sum_{m=-N}^{N} \left(1 - \frac{|m|}{N+1}\right) \hat{f}(m+\eta)e^{2\pi ix(m+\eta)} \right| < \epsilon,$$

from which it follows by integrating over η from $0 \leq \eta \leq 1$ that

$$\begin{aligned}
&\left| f(x) - \sum_{m=-N}^{N} \left(1 - \frac{|m|}{N+1}\right) \int_0^1 \hat{f}(m+\eta)e^{2\pi ix(m+\eta)} d\eta \right| \\
\leq &\int_0^1 \left| f(x) - \sum_{m=-N}^{N} \left(1 - \frac{|m|}{N+1}\right) \hat{f}(m+\eta)e^{2\pi ix(m+\eta)} \right| d\eta \\
\leq &\epsilon.
\end{aligned}$$

Defining $\Phi_N(\xi) = 1 - |m|/(N+1)$ when $m \leq \xi < m+1$, for $m = -N, \ldots, N$, and 0 otherwise, we see that

$$\sum_{m=-N}^{N} \left(1 - \frac{|m|}{N+1}\right) \int_0^1 \hat{f}(m+\eta)e^{2\pi ix(m+\eta)} d\eta = \int_{-N}^{N+1} \Phi_N(\xi)\hat{f}(\xi)e^{2\pi ix\xi} d\xi,$$

thus we have proved

$$\lim_{N\to\infty} \int_{-N}^{N+1} \Phi_N(\xi)\hat{f}(\xi)e^{2\pi ix\xi} d\xi = f(x)$$

uniformly in x. If we have the additional information that $\int_{\mathbb{R}} |\hat{f}(\xi)| d\xi < \infty$, then using $\Phi_N(\xi) \to 1$ as $N \to \infty$, as well as $1 - |\xi|/R \to 1$ as $R \to \infty$ for any fixed ξ, and the dominated convergence theorem we obtain

$$\lim_{N\to\infty} \int_{-N}^{N+1} \Phi_N(\xi)\hat{f}(\xi)e^{2\pi ix\xi} d\xi = \lim_{R\to\infty} \int_{-R}^{R} \left(1 - \frac{|\xi|}{R}\right) \hat{f}(\xi)e^{2\pi ix\xi} d\xi = \int_{\mathbb{R}} \hat{f}(\xi)e^{2\pi ix\xi} d\xi,$$

thus proving $f(x) = \int_{\mathbb{R}} \hat{f}(\xi)e^{2\pi ix\xi} d\xi$ in this case as well.

Assignment 3: Chapter 8: 18*, 22, 25(a), 26, 28*

1.6 Fourier transforms applied to IVP for the heat equation

Back to our earlier discussion on solving the IVP for the heat equation, if we conform with the notation in the text, then for any $\xi \in \mathbb{R}^n$,

$$u(x,t) := e^{2\pi i \xi\cdot x - 4\pi^2|\xi|^2 t} \quad \text{satisfies } u_t - \Delta u = 0.$$

Thus for reasonable $c(\xi)$,

$$u(x,t) := \int_{\mathbb{R}^n} c(\xi) e^{2\pi i \xi\cdot x - 4\pi^2|\xi|^2 t} d\xi \quad \text{satisfies } u_t - \Delta u = 0,$$

at least formally; our question was, given any initial data $u(x,0)$, can we choose $c(\xi)$ such that the $u(x,t)$ defined above solves the homogeneous heat equation, and satisfies the initial condition $\lim_{t\searrow 0} u(x,t) = u(x,0)$ in appropriate sense? It is now clear that we should take $c(\xi) = \widehat{u(x,0)}(\xi)$. This requires technically that $u(x,0) \in L^1(\mathbb{R}^n)$. Then we know $c(\xi) \in C_0(\mathbb{R}^n)$, so the integral defining the $u(x,t)$ above is absolutely convergent at $t>0$, uniformly so at $t > \delta > 0$ for any $\delta > 0$.

Recall that $\widehat{e^{-\pi|x|^2}}(\xi) = e^{-\pi|\xi|^2}$, thus, using (b) of Proposition 8.22, we have

$$H_t(x) = (4\pi t)^{-n/2} e^{-\frac{|x|^2}{4t}} = \int_{\mathbb{R}^n} e^{2\pi i \xi\cdot x - 4\pi^2|\xi|^2 t} d\xi,$$

and

$$u(x,t) = H_t * u(x,0) = \int_{\mathbb{R}^n} c(\xi) e^{2\pi i \xi\cdot x - 4\pi^2|\xi|^2 t} d\xi.$$

We can now see that $\{H_t\}_{t>0}$ forms a family of good kernels, and using our results on convolutions, we conclude that for any $u(x,0) \in L^p(\mathbb{R}^n)$, $1 \le p < \infty$, $u(x,t) = H_t * u(x,0)$ solves $u_t(x,t) - \Delta u(x,t) = 0$ for $t>0$ and $||u(x,t) - u(x,0)||_{L^p(\mathbb{R}^n)} \to 0$ as $t \searrow 0$. For $p = \infty$, $u(x,t) = H_t * u(x,0)$ still solves $u_t(x,t) - \Delta u(x,t) = 0$ for $t>0$, but the sense in which $u(x,t)$ takes on the initial data $u(x,0)$ has to be reformulated; however, if $u(x,0)$ is also assumed to be in C^0, then we know that $u(x,t) \to u(x,0)$ uniformly on compact subsets of $\mathbb{R}^n$ as $t \searrow 0$. Note that the derivation for the solution required $u(x,0) \in L^1(\mathbb{R}^n)$; but after we established $u(x,t) = H_t * u(x,0)$, with good control on H_t, we can bypass the derivation process and obtain results under more flexible conditions on $u(x,0)$. This, of course, is not always possible, and we need to study Fourier transforms of functions in other function spaces.

1.7 Fourier inversion formula

Our key tool is the formula

$$f * H_t = \int_{\mathbb{R}^n} e^{2\pi i \xi\cdot x} h(t\xi) \hat{f}(\xi) d\xi,$$

provided $f, h \in L^1(\mathbb{R}^n)$ and

$$H(x) = h^\vee(x) := \int_{\mathbb{R}^n} e^{2\pi i \xi \cdot x} h(\xi) d\xi.$$

We obtain useful results when, in addition, $H \in L^1(\mathbb{R}^n)$. There are many choices of h that satisfy these conditions. For instance, one could take $h(\xi) = e^{-\pi|\xi|^2}$, then $\{H_t\}$ forms a family of good kernels. There are two types of inversion formulas: one type is under only the condition that $f \in L^1(\mathbb{R}^n)$, and one concludes that

$$||f(x) - \int_{\mathbb{R}^n} e^{2\pi i \xi \cdot x} h(t\xi) \hat{f}(\xi) d\xi||_{L^1(\mathbb{R}^n)} \to 0$$

as $t \searrow 0$; another type is under the condition that both f and $\hat{f}$ are in $L^1(\mathbb{R}^n)$, then

$$\lim_{t \to 0} \int_{\mathbb{R}^n} e^{2\pi i \xi \cdot x} h(t\xi) \hat{f}(\xi) d\xi = \int_{\mathbb{R}^n} e^{2\pi i \xi \cdot x} \hat{f}(\xi) d\xi$$

pointwise by the dominated convergence theorem, so obtain a pointwise inversion formula

$$f(x) = \int_{\mathbb{R}^n} e^{2\pi i \xi \cdot x} \hat{f}(\xi) d\xi$$

for *a.e.* $x \in \mathbb{R}^n$ (why *a.e.*?).

One can verify the condition that $\hat{f} \in L^1(\mathbb{R}^n)$ by assuming that sufficiently many derivatives of f are in $L^1(\mathbb{R}^n)$. However $\hat{f}$ may no longer be in the same class, so the inversion acts on a different class of functions. Even when one takes $f \in C_c^\infty(\mathbb{R}^n)$, then $\hat{f} \in C^\infty(\mathbb{R}^n)$ but no longer has compact support. The Schwartz class, $\mathcal{S}$, has the nice property that $\hat{f} \in \mathcal{S}$ when $f \in \mathcal{S}$. Since $\mathcal{S}$ is also dense in $L^p(\mathbb{R}^n)$ for any $1 \leq p < \infty$, it's often useful to obtain results in the class $\mathcal{S}$ first. The text proves the Plancherel Theorem this way. More prominent use of the class $\mathcal{S}$ will come when discussing the Fourier transforms of tempered distributions. So we will defer the discussion of the topology of $\mathcal{S}$ until it's needed.

1.8 Fourier transforms on $L^2(\mathbb{R}^n)$

The key property that allows the definition of Fourier transforms on $L^2(\mathbb{R}^n)$ is the Plancherel Theorem. Once the Plancherel Theorem is established for a dense subset of $L^2(\mathbb{R}^n)$, the extension of Fourier transforms to $L^2(\mathbb{R}^n)$, and other properties of Fourier transforms, such as (a)-(e) of Theorem 8.22, Lemma 8.25, Theorem 8.26, Corollary 8.27 are all valid, in appropriate sense. For instance, (c) of Theorem 8.22 does not make sense directly, as when $f, g \in L^2$, $f * g \in L^\infty$, and we don't know how to define the Fourier transform of a function in L^∞; however, $\hat{f}\hat{g} \in L^1$ by Hölder's inequality, and we can make sense of its Fourier (or inverse Fourier) transform, and interpret (c) in the sense of Theorem 8.34.

Here is an example illustrating the usefulness of Fourier transforms on $L^2(\mathbb{R}^n)$.

Example 1. *We study the IVP for the Schrodinger equation*

$$\begin{cases} u_t(x,t) = i\Delta u(x,t), \\ u(x,0) = f(x) \end{cases} \tag{13}$$

Again, by formal computation, for any $\xi \in \mathbb{R}^n$,

$$u(x,t) := e^{2\pi i \xi\cdot x - 4\pi^2|\xi|^2 it} \quad \textit{solves } u_t(x,t) = i\Delta u(x,t),$$

thus for reasonable $c(\xi)$,

$$u(x,t) := \int_{\mathbb{R}^n} c(\xi) e^{2\pi i \xi\cdot x - 4\pi^2|\xi|^2 it} d\xi$$

also solves $u_t(x,t) = i\Delta u(x,t)$. However, $e^{2\pi i \xi\cdot x - 4\pi^2|\xi|^2 it}$ no longer has fast decay in ξ; to make sense of the integral, one has to require sufficient decay of $c(\xi)$. To satisfy the initial condition, we also need to choose $c(\xi) = \hat{f}(\xi)$. At least for $f \in \mathcal{S}$, we can justify the formal computations above and verify that $u(x,t)$ defined above is a classical solution to (13)*. Note that for such solutions, we have*

$$||u(x,t)||_{L^2(\mathbb{R}^n)} = ||\widehat{u(\cdot,t)}(\xi)||_{L^2(\mathbb{R}^n)} = ||\hat{f}(\xi)e^{-4\pi^2|\xi|^2 it}||_{L^2(\mathbb{R}^n)} = ||\hat{f}(\xi)||_{L^2(\mathbb{R}^n)} = ||f||_{L^2(\mathbb{R}^n)},$$

and

$$||u(x,t) - f(x)||_{L^2(\mathbb{R}^n)} = ||\widehat{u(\cdot,t)}(\xi) - \hat{f}(\xi)||_{L^2(\mathbb{R}^n)} = ||\hat{f}(\xi)\left[e^{-4\pi^2|\xi|^2 it} - 1\right]||_{L^2(\mathbb{R}^n)} \to 0$$

as $t \to 0$ by the dominated convergence theorem under only the condition that $f \in L^2(\mathbb{R}^n)$.

These estimates allow us to extend the notion of solutions, in some sense, to initial data $f \in L^2(\mathbb{R}^n)$ as follows. For any $f \in \mathcal{S}$, let $S_t[f] := u(x,t)$ denote the solution to (13) defined above. Then $S_t[f]$ is a continuous curve in $L^2(\mathbb{R}^n)$, and $||S_t[f_1] - S_t[f_2]||_{L^2(\mathbb{R}^n)} = ||f_1 - f_2||_{L^2(\mathbb{R}^n)}$ for all t. Thus, for any given $f \in L^2(\mathbb{R}^n)$, when we take a sequence $f_j \in \mathcal{S}$ such that $||f_j - f||_{L^2(\mathbb{R}^n)} \to 0$ as $j \to \infty$, $\{S_t[f_j]\}$ is Cauchy in $L^2(\mathbb{R}^n)$, uniformly over t, and there is a limit $S_t[f] \in L^2(\mathbb{R}^n)$ such that $S_t[f_j] \to S_t[f]$ in $L^2(\mathbb{R}^n)$, uniformly over t, with $S_0[f] = f$. We may take this (continuous) family $S_t[f]$ in $L^2(\mathbb{R}^n)$ as a generalized Schrodinger flow to (13). No derivatives for this family $S_t[f]$ are defined under only the condition $f \in L^2(\mathbb{R}^n)$; yet we can say that $S_t[f]$ satisfies (13) in the following sense. Take any $\eta(x,t) \in C_c^\infty(\mathbb{R}^{n+1})$, and multiply it on both sides of (13) and integrate by parts over $\mathbb{R}^n \times \mathbb{R}^+$, we have

$$\int_{\mathbb{R}^{n+1}} u(x,t)\left[\eta_t(x,t) + i\Delta\eta(x,t)\right] dxdt = -\int_{\mathbb{R}^n} f(x)\eta(x,0)dx. \tag{14}$$

(14) holds if we replace f by f_j, and $u(x,t)$ by $S_t[f_j]$. Passing to the limit (in $L^2_{\text{loc}}(\mathbb{R}^{n+1})$), we see that (14) holds for $u = S_t[f]$.

When the initial data has some further restriction, for instance, when $\hat{f}$ satisfies $|\xi|\hat{f}(\xi) \in L^2(\mathbb{R}^n)$, we can define weak L^2 derivatives for the $S_t[f]$ in the following way.

Definition 2. *$u(x) \in L^2(\mathbb{R}^n)$ is said to have weak L^2 partial derivatives, if there are $L^2(\mathbb{R}^n)$ functions g_j for $j = 1, \cdots, n$, such that for any $\eta \in C_c^\infty(\mathbb{R}^n)$, we have,*

$$\int_{\mathbb{R}^n} u(x)\partial_j\eta(x)dx = -\int_{\mathbb{R}^n} g_j(x)\eta(x)dx, \tag{15}$$

for each $j = 1, \cdots, n$.

If $u(x)$ has weak L^2 partial derivatives, then, for each $j = 1, \cdots, n$, the choice for g_j is unique in $L^2(\mathbb{R}^n)$; and we will simply denote it by $\partial_j u$. We define $H^1(\mathbb{R}^n) = \{f \in L^2(\mathbb{R}^n) : f \text{ has weak } L^2 \text{ partial derivatives}\}$. We can define higher order weak L^2 partial derivatives similarly and define

$$H^k(\mathbb{R}^n) = \{f \in L^2(\mathbb{R}^n) : f \text{ has weak } L^2 \text{ partial derivatives of order } \leq k\},$$

and define an inner product on $H^k(\mathbb{R}^n)$ by

$$(f,g)_{H^k(\mathbb{R}^n)} = \sum_{|\alpha|\leq k} \int_{\mathbb{R}^n} \partial^\alpha f(x)\partial^\alpha g(x)dx,$$

with the induced norm

$$||f||_{H^k(\mathbb{R}^n)} = \left\{\sum_{|\alpha|\leq k} \int_{\mathbb{R}^n} |\partial^\alpha f(x)|^2 dx\right\}^{1/2}.$$

Remark 1. *For a given $u(x) \in L^2(\mathbb{R}^n)$, if there is a sequence $u_j \in \mathcal{S}$ such that $u_j \to u$ in $L^2(\mathbb{R}^n)$, as $j \to \infty$, and $\{\partial_l u_j\}_j$ is Cauchy in $L^2(\mathbb{R}^n)$, then $u(x) \in H^1(\mathbb{R}^n)$, and for each $l = 1, \cdots, n$, $\partial_l u$ is the L^2 limit of $\{\partial_l u_j\}_j$. Likewise, if there is a sequence $u_j \in \mathcal{S}$ such that $u_j \to u$ in $L^2(\mathbb{R}^n)$, as $j \to \infty$, and $\{\partial^\alpha u_j\}_j$ is Cauchy in $L^2(\mathbb{R}^n)$, for all $|\alpha| \leq k$, then $u(x) \in H^k(\mathbb{R}^n)$, and for each $|\alpha| \leq k$, $\partial^\alpha u$ is the L^2 limit of $\{\partial^\alpha u_j\}_j$. In fact, the definition through approximation by functions in $\mathcal{S}$ gives an equivalent definition of $H^k(\mathbb{R}^n)$.*

Remark 2. *Using Plancherel's Theorem and approximation, we find that*

$$\begin{aligned} H^k(\mathbb{R}^n) &= \{f \in L^2(\mathbb{R}^n) : (1+|\xi|^2)^{k/2}\hat{f}(\xi) \in L^2(\mathbb{R}^n)\} \\ &= \textit{completion of } C_c^\infty(\mathbb{R}^n) \textit{ under the norm } ||f||_{H^k(\mathbb{R}^n)} := \left(\sum_{|\alpha|\le k} ||\partial^\alpha f||^2_{L^2(\mathbb{R}^n)}\right)^{1/2} \\ &= \textit{completion of } \mathcal{S}(\mathbb{R}^n) \textit{ under the norm } ||f||_{H^k(\mathbb{R}^n)}, \end{aligned} \tag{16}$$

and an equivalent norm on $H^k(\mathbb{R}^n)$ is

$$||(1+|\xi|^2)^{k/2}\hat{f}(\xi)||_{L^2(\mathbb{R}^n)}.$$

Similar function spaces arise naturally over bounded domains in $\mathbb{R}^n$. More systematic study of these spaces will be made in 9.3.

Back to solutions to (13) when $f \in \mathcal{S}$, we also have

$$||\partial_t^k \partial_x^\alpha u(x,t)||_{L^2(\mathbb{R}^n)} = ||\partial_t^k \widehat{\partial_x^\alpha u(x},t)||_{L^2(\mathbb{R}^n)} = ||(-4\pi^2|\xi|^2 i)^k (2\pi i\xi)^\alpha \hat{f}(\xi)||_{L^2(\mathbb{R}^n)}, \tag{17}$$

for each k and α. So when $f \in H^2(\mathbb{R}^n)$, the $S_t[f]$ we defined actually has weak L^2 derivatives $\partial_t S_t[f]$ and $\partial_x^\alpha S_t[f]$ for any $|\alpha| \le 2$, and the equation in (13) can be taken to be valid as elements in L^2. One key ingredient that makes this process possible is the completeness of L^2.

1.9 Fourier transform on $L^p(\mathbb{R}^n)$ for $1 < p < 2$

On the unit circle $\mathbb{T}$, we have the relation

$$C(\mathbb{T}) \subset L^\infty(\mathbb{T}) \subset L^{p'}(\mathbb{T}) \subset L^2(\mathbb{T}) \subset L^p(\mathbb{T}) \subset L^1(\mathbb{T}),$$

for $1 < p < 2$. So, once the Fourier series for functions in $L^1(\mathbb{T})$ are defined, they are defined for any $L^p(\mathbb{T})$. The L^p spaces on $\mathbb{R}^n$ have no such simple inclusion relations. After we define Fourier transform on $L^1(\mathbb{R}^n)$ and on $L^2(\mathbb{R}^n)$, we can attempt to define Fourier transform on $L^p(\mathbb{R}^n)$ for $1 < p < 2$ by approximations using functions in $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, and noting the relation $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n) \subset L^p(\mathbb{R}^n)$ for $1 < p < 2$. This can be carried out provided we have some kind of a priori inequalities on Fourier transforms of functions in $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. The Riesz-Thorin Interpolation Theorem in 6.5 provides such a tool. Applying this theorem to the setting of Fourier series and Fourier transforms, we obtain the Hausdorff-Young inequality (8.21) and (8.30). It turns out that except in the case $p = 2$, the map $f \in L^p \mapsto \hat{f} \in L^{p'}$ is not onto. Fourier transforms on $L^{p'}$ can not be defined as we did for L^p when $1 \le p < 2$, so there is no straight forward version of the inversion theorem as (8.26). Theorem 8.35, however, can be used as a version of the inversion theorem when $f \in L^p$ with $1 \le p \le 2$.

1.10 Completeness used to solve BVP

Some boundary value problems (BVP) have naturally variational characterizations. In such cases, it is often more useful to find a solution through variational means, rather than attempting to find an explicit solution formula. The Dirichlet problem for the Poisson equation is such an example. Here we are given a function $f(x)$ on a domain Ω, say, in $C(\bar{\Omega})$, and look for a solution $u \in C^2(\bar{\Omega})$ solving

$$\begin{cases} -\Delta u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{18}$$

Define the functional

$$J[u] := \int_\Omega \left(\frac{1}{2} |\nabla u(x)|^2 - f(x)u(x) \right) dx$$

for $u \in X := \{u \in C^2(\bar{\Omega}) : u = 0 \text{ on } \partial\Omega\}$. We can see easily that if $u \in C^2(\bar{\Omega})$ is a solution to (18), then it is a critical point of $J[u]$; and a critical point $u \in C^2(\bar{\Omega})$ of $J[u]$ is a solution to (18). So we attempt to find a solution to (18) by finding a critical point of $J[u]$. A natural approach is to attempt to find a minimizer for $J[u]$. For this purpose, we need to establish two steps:

Step 1. Prove that $\inf_{u \in X} J[u] =: m$ is finite.

Step 2. Prove that there is a $u \in X$ that attains m.

For simplicity, we treat the one-dimensional case $\Omega = [a, b]$ here. Then

$$\int_\Omega f(x)u(x)\, dx \le ||f||_{L^2(\Omega)} ||u||_{L^2(\Omega)} \le \frac{\epsilon}{2} ||u||^2_{L^2(\Omega)} + \frac{1}{2\epsilon} ||f||^2_{L^2(\Omega)}.$$

Using the Wirtinger inequality

$$||u||^2_{L^2(\Omega)} \le (\frac{b-a}{\pi})^2 ||\nabla u||^2_{L^2(\Omega)},$$

we find

$$J[u] \ge \left(1 - \epsilon (\frac{b-a}{\pi})^2 \right) /2 ||\nabla u||^2_{L^2(\Omega)} - \frac{1}{2\epsilon} ||f||^2_{L^2(\Omega)} \tag{19}$$

has a finite lower bound, if we choose $\epsilon > 0$ such that

$$1 - \epsilon (\frac{b-a}{\pi})^2 > 0.$$

This proves that $\inf_{u \in X} J[u] =: m$ is well defined and finite.

Remark 3. *For a bounded domain Ω in higher dimensional Euclidean space $\mathbb{R}^n$, there is an inequality similar to Wirtinger's inequality, called Poincarè's inequality, for some $C > 0$, we have*

$$||u||^2_{L^2(\Omega)} \leq C||\nabla u||^2_{L^2(\Omega)}$$

for all $u \in H^1_0(\Omega)$. One difference between these two inequalities is that we don't have an explicit (optimal) C in general for the Poincarè's inequality.

For Step 2, we take a *minimizing sequence* $\{u_j\}_j$, *i.e.*, $u_j \in X$ such that $J[u_j] \to m$ as $j \to \infty$. The computations in (19) show that $||u_j||_{L^2(\Omega)}$ is bounded. In fact, we know more. Using

$$J[u_j] + J[u_k] - 2J[\frac{u_j + u_k}{2}] = \frac{1}{4}\int_\Omega |\nabla(u_j - u_k)|^2 dx,$$

and $J[u_j] \to m$, $J[u_k] \to m$, $J[(u_j + u_k)/2] \geq m$, as $j, k \to \infty$, we find that $\{\nabla u_j\}_j$ is Cauchy in $L^2(\Omega)$. $\{u_j\}_j$ is also Cauchy in $L^2(\Omega)$ due to Wirtinger's inequality. Thus there is a limiting function $u \in L^2(\Omega)$ such that $u_j \to u$ in $L^2(\Omega)$; furthermore, just as in the entire space situation, since $\{\nabla u_j\}_j$ is Cauchy in $L^2(\Omega)$, we find that the limit u has L^2 weak derivative. In fact, u lies in the completion of X under the norm $||\nabla u||_{L^2(\Omega)}$. Define $H^1_0(\Omega)$ to be the completion of X under the norm $||\nabla u||_{L^2(\Omega)}$. We arrive at the

Conclusion: instead of finding a minimizer for $J[u]$ in X, we found a minimizer for $J[u]$ in its completion $H^1_0(\Omega)$. This u solves (18) in the sense that

$$0 = \frac{dJ[u + t\eta]}{dt}\Big|_{t=0} = \int_\Omega (\nabla u(x) \cdot \nabla \eta(x) - f(x)\eta(x))\, dx \tag{20}$$

for all $\eta \in X$.

We see now that it is relatively easy to extend $J[u]$ to $H^1_0(\Omega)$ and repeat Steps 1 and 2 above in the framework of $H^1_0(\Omega)$ to find a minimizer in the extended function space $H^1_0(\Omega)$, and the price we pay is that we don't know at this point that a minimizer u is in X, the classical function space that we started with. We need to add one more step to complete the program.

Step 3. Investigate whether a solution $u \in H^1_0(\Omega)$ to (20) is a classical solution to (18).

The variational approach can often be applied to nonlinear PDEs. For instance, when $f = f(x, u)$, we can define $F(x, u) = \int_0^u f(x, v)dv$. Then a solution $u \in X$ to

$$\begin{cases} -\Delta u(x) = f(x, u), & \text{in } \Omega, \\ u(x) = 0, & \text{on } \partial\Omega, \end{cases}$$

is a critical point of

$$J[u] := \int_\Omega \left(\frac{1}{2}|\nabla u(x)|^2 - F(x, u(x))\right) dx.$$

Under some restrictions on f, for instance, $|f(x,u)| \leq \lambda|u| + g(x)$ for some small $\lambda > 0$ and $g \in L^2(\Omega)$, we can still carry out Step 1 above. $f(x,u) = \cos u + g(x)$ would be a function satisfying this restriction. More specifically, we have

$$|F(x,u)| \leq \frac{\lambda}{2}|u|^2 + |g(x)||u|,$$

from which we obtain

$$\begin{aligned} J[u] &\geq \int_\Omega \left(\frac{1}{2}|\nabla u(x)|^2 - \frac{\lambda}{2}|u(x)|^2 - \frac{\epsilon}{2}|u(x)|^2 - \frac{1}{2\epsilon}|g(x)|^2\right) dx \\ &\geq \left(\frac{1}{2} - \frac{\lambda+\epsilon}{2}C\right) \int_\Omega |\nabla u(x)|^2 - \int_\Omega |g(x)|^2 dx, \end{aligned}$$

where $C > 0$ is a constant as in Wirtinger's inequality $\int_\Omega |u(x)|^2 \leq C \int_\Omega |\nabla u(x)|^2$ for $u \in X$. This shows that $\inf_{u \in X} J[u] := m$ is finite. But for a minimizing sequence, we no longer can prove easily that $\{\nabla u_j\}_j$ is Cauchy in $L^2(\Omega)$. What we do have easily is that $\{\nabla u_j\}_j$ is bounded in $L^2(\Omega)$. We now face the question of how to extract a subsequence which converges to a minimizer in some sense. We will handle this issue using the idea of compactness.

2 Compactness and its applications

The most familiar compactness criterion to us is the one in $\mathbb{R}^n$: *A subset of* $\mathbb{R}^n$ *is compact iff it is bounded and closed* (Proposition 0.26). Metric spaces suffice for many applications, on which the compactness criteria are given by Theorem 0.25. Applying these criteria to subsets of the function space $C(X)$, where X is a compact Hausdorff space, we obtain Arzelà-Ascoli Theorem (4.43). We will further apply these criteria in the context of $L^p(\Omega)$ for domains $\Omega \subset \mathbb{R}^n$ to obtain compactness criteria for subsets in $L^p(\Omega)$. Folland does not contain this part of the material. A good source is Brezis' text. A new edition is to appear soon, and I have put an earlier French version on reserve.

We often encounter situations where the (metric) compactness criteria can not be verified. The notion of weak (or weak*) compactness turns out to be extremely useful. These notions are defined in 5.4. We will expand the discussion somewhat following Brezis.

2.1 Notion of compactness and sequential compactness

These notions are defined in 4.4, and they are equivalent on metric spaces according to Theorem 0.25. They are, however, not equivalent on general topological spaces. Each is useful when it is available. We will prove a sequential compactness theorem involving weak* convergence. Although it can be included as a consequence of the most general weak* compactness theorem of Alaoglu (5.18), it is worthwhile to understand the idea in this special setting.

2.2 "L^p-version" of Arzelà-Ascoli Theorem

Theorem 1 (M. Riesz-Fréchet-Kolmogrov). *Let $\mathcal{F}$ be a bounded set in $L^p(\mathbb{R}^n)$ with $1 \le p < \infty$, and define $\tau_h f(x) = f(x+h)$ for $h \in \mathbb{R}^n$. Assume that*

$$\lim_{|h|\to 0} \|\tau_h f - f\|_p = 0 \quad \textit{uniformly in } f \in \mathcal{F} \tag{21}$$

namely, $\forall \epsilon > 0\ \exists \delta > 0$ such that $\|\tau_h f - f\|_p < \epsilon\ \forall f \in \mathcal{F}\ \forall h \in \mathbb{R}^n$ with $|h| < \delta$.

Then the closure of $\mathcal{F}|_\Omega$ in $L^p(\Omega)$ is compact for any measurable set $\Omega \subset \mathbb{R}^n$ with finite measure.

Remark 4. *Under the assumptions of the M. Riesz-Fréchet-Kolmogrov theorem, we cannot conclude in general that $\mathcal{F}$ itself has compact closure in $L^p(\mathbb{R}^n)$. Under an additional assumption, however, we have*

Corollary 3. *Let $\mathcal{F}$ be a bounded set in $L^p(\mathbb{R}^n)$ with $1 \le p < \infty$. Assume (21) and also*

$$\forall \epsilon > 0\ \exists \Omega \subset \mathbb{R}^n \textit{ bounded, measurable such that } \|f\|_{L^p(\mathbb{R}^n \setminus \Omega)} < \epsilon\ \forall f \in \mathcal{F}. \tag{22}$$

Then $\mathcal{F}$ has compact closure in $L^p(\mathbb{R}^n)$.

Corollary 4. *Let G be a fixed function in $L^1(\mathbb{R}^n)$ and $\mathcal{B}$ be a bounded set in $L^p(\mathbb{R}^n)$ with $1 \le p < \infty$. Define $\mathcal{F} := G * \mathcal{B}$. Then $\mathcal{F}|_\Omega$ has compact closure in $L^p(\Omega)$ for any measurable Ω with finite measure.*

Corollary 5. *Let $\mathcal{F}$ be a bounded set in $H^1(\mathbb{R}^n)$. Then the closure of $\mathcal{F}|_\Omega$ in $L^2(\Omega)$ is compact for any measurable set $\Omega \subset \mathbb{R}^n$ with finite measure.*

We can define $H^1(\mathbb{T})$ as the completion of $C^1(\mathbb{T})$ under the H^1 norm. We also have

Corollary 6. *Let $\mathcal{F}$ be a bounded set in $H^1(\mathbb{T})$. Then the closure of $\mathcal{F}$ in $L^2(\mathbb{T})$ is compact.*

Using the Fourier series representations of $L^2(\mathbb{T})$ functions, we can give a direct proof to Corollary 6.

Proof. It suffices to prove that any sequence $\{f_j\}$ in $\mathcal{F}$ has a subsequence convergent in $L^2(\mathbb{T})$. Let $C > 0$ be such that $||f||_{H^1(\mathbb{T})} \leq C$ for all $f \in \mathcal{F}$. For $\in H^1(\mathbb{T})$, we still have $\hat{f}'(n) = 2\pi n i \hat{f}(n)$. So we have

$$\sum_n |\hat{f}(n)|^2 + 4\pi^2 \sum_n n^2 |\hat{f}(n)|^2 \leq C^2.$$

It follows that for each n, $\{\hat{f}_j(n)\}_j$ is bounded in $\mathbb{C}$, therefore has a convergent subsequence. We can use a diagonal procedure to pick a subsequence $\{f_{j_k}\}_k$ of $\{f_j\}$ such that for each n, $\{\hat{f}_{j_k}(n)\}_k$ is a convergent sequence in $\mathbb{C}$, and has a limit a_n. $\{a_n\}_n \in l^2$ by Fatou's Lemma. We will prove that $\{\hat{f}_{j_k}(n) - a_n\}_k \to 0$ in l^2. Let me just prove that $\{\hat{f}_{j_k}\}_k$ is Cauchy in l^2. For any $\epsilon > 0$, we can choose $N > 0$ such that $C^2/(4\pi^2 N^2) < \epsilon^2/4$. Since $4\pi^2 \sum_{|n|>N} n^2 |\hat{f}(n)|^2 \leq C^2$ for all $f \in \mathcal{F}$, we find $\sum_{|n|>N} |\hat{f}(n)|^2 < \epsilon^2/4$ for all $f \in \mathcal{F}$. Now

$$\begin{aligned}
\|f_{j_k} - f_{j_l}\|^2_{L^2(\mathbb{T})} &= \|\hat{f}_{j_k} - \hat{f}_{j_l}\|^2_{l^2} \\
&= \sum_{|n|\leq N} |\hat{f}_{j_k}(n) - \hat{f}_{j_l}(n)|^2 + \sum_{|n|>N} |\hat{f}_{j_k}(n) - \hat{f}_{j_l}(n)|^2 \\
&\leq \sum_{|n|\leq N} |\hat{f}_{j_k}(n) - \hat{f}_{j_l}(n)|^2 + \epsilon^2.
\end{aligned}$$

For the finite number of $|n| \leq N$, since $\{\hat{f}_{j_k}(n)\}_k$ is convergent, we can find $L > 0$ such that when $k, l > L$, we have $\sum_{|n|\leq N} |\hat{f}_{j_k}(n) - \hat{f}_{j_l}(n)|^2 < \epsilon^2$, thus $\|f_{j_k} - f_{j_l}\|^2_{L^2(\mathbb{T})} < 2\epsilon^2$ when $k, l > L$, proving that $\{\hat{f}_{j_k}\}_k$ is Cauchy in l^2 and $\{f_{j_k}\}_k$ is Cauchy in $L^2(\mathbb{T})$. $\square$

Assignment 4. Turn in the *ed problems in class on March 9.
Chapter 6, 14, 17, 23, 24, 41*.

A*. Prove the converse to Corollary 3, namely, let $\mathcal{F}$ be a compact subset in $L^p(\mathbb{R}^n)$ with $1 \leq p < \infty$, prove that $\mathcal{F}$ is a bounded subset in $L^p(\mathbb{R}^n)$; in addition, (21) and (22) hold.

B*. Construct a bounded family $\mathcal{F}$ of functions in $L^p(\mathbb{R}^n)$ with $1 \leq p < \infty$ that satisfies (21), but fails to have compact closure in $L^p(\mathbb{R}^n)$. Also construct a bounded family $\mathcal{F}$ of functions in $L^p(\mathbb{R}^n)$ with $1 \leq p < \infty$ that does not satisfies (21), and fails to have compact closure in $L^p(\Omega)$ on any compact domain $\Omega \subset \mathbb{R}^n$.

C. Let Ω be a bounded domain in $\mathbb{R}^n$. Prove that a bounded set in $H_0^1(\Omega)$ (the closure of $C_c^1(\Omega)$ in the H^1 norm) has compact closure in $L^2(\Omega)$.

D*. Let $K(s,t)$ be a bounded measurable function on $[0,1]\times[0,1]$, continuous as a function of s, uniformly in $t\in[0,1]$. Define $\mathcal{K}[u](s)=\int_0^1 K(s,t)u(t)dt$ for $u\in L^2[0,1]$. Prove that $\mathcal{K}$ maps a bounded set in $L^2[0,1]$ into a pre-compact set in $L^2[0,1]$.

2.3 sequential weak (weak*) compactness

In an infinite dimensional normed vector space, a bounded closed set may not be compact, and may not contain any subsequence which converges in norm. The notion of weak (weak*) sequential compactness and weak (weak*) compactness are often useful substitute for compactness in norm.

Definition 3. *A sequence $\{x_j\}_j$ in a normed vector space X is said to converge weakly to $x\in X$, if $\lim_{j\to\infty} l(x_j)=l(x)$ for every $linX^*$. We use $x_j \rightharpoonup x$ to denote the weak convergence of $\{x_j\}_j$ to x.*

Example 2. *The sequence $\{e^{inx}\}_n$ is not Cauchy in $X:=L^2[0,2\pi]$, yet it converges to 0 weakly in X. For, given any $l\in X^*$, which is equal to X in this case, there exists $g\in X$ such that $l(f)=\int_0^{2\pi} f(x)g(x)dx$ for all $f\in X$; and $l(e^{inx})=\int_0^{2\pi} e^{inx}g(x)dx$, which tends to 0 as $n\to\infty$ by Bessel's inequality.*

Example 3. *Let $\{s_j\}_j$ be a sequence in l^p such that* (a) *for each n, $\lim_{j\to\infty} s_j(n)=s(n)$ exists, and* (b) *there is a bound $C>0$ such that $||s_j||_{l^p}\le C$. Assume $\infty>p>1$. Then $\{s(n)\}_n\in l^p$ and $s_j\rightharpoonup s$ as $j\to\infty$.*

The fact that $\{s(n)\}_n\in l^p$ follows from Fatou's lemma and assumption (b). Set $X=l^p$ and let p' denote the conjugate exponent of p. Then $X^*=l^{p'}$, and for any $l\in X^*$, there is a $\{y(n)\}_n\in l^{p'}$ such that $l(s)=\sum_n s(n)y(n)$. Given any $\epsilon>0$. Since $p'<\infty$, there exists N such that $(\sum_{n>N}|y(n)|^{p'})^{1/p'}<\epsilon/3$. For the finite number $1\le n\le N$, we can find J such that $|s_j(n)-s(n)|<\epsilon$ for all $j>J$. Then, for all $j>J$,

$$
\begin{aligned}
|l(s_j)-l(s)| &= |\sum_n (s_j(n)-s(n))y(n)| \\
&\le |\sum_{n=1}^{N}(s_j(n)-s(n))y(n)| + \left(\sum_{n>N}|s_j(n)-s(n)|^p\right)^{1/p}\left(\sum_{n>N}|y(n)|^{p'}\right)^{1/p'} \\
&\le \max_{1\le n\le N}|y(n)|\epsilon + 2C\epsilon/3,
\end{aligned}
$$

proving that $s_j \rightharpoonup s$ as $j \to \infty$. Note that (a) means componentwise convergence of $s_j(n)$, yet we don't know $||s_j - s||_{l^p} \to 0$ and in general can't find J such that $|s_j(n) - s(n)| < \epsilon$ for all $j > J$ and all n, however, the pairing with l in $l^{p'}$ allows us to control the "tail" part $\sum_{n>N}(s_j(n) - s(n))y(n)$ using the smallness of $(\sum_{n>N}|y(n)|^{p'})^{1/p'}$ when N is large.

The converse to the statement in this example also holds, namely, if $s_j \rightharpoonup s$ as $j \to \infty$, then (a) and (b) hold. (a) is verified by picking l to be $e_n = (0, \ldots, 1, 0, \ldots)$, where the 1 is in the nth slot, while (b) is a consequence of the uniform boundedness principle.

In fact, using the idea here and a diagonal process of choosing subsequences, it's easy to see that for any $p > 1$ a sequence bounded in l^p has a subsequence which converges weakly. *This is a useful substitute for Proposition 0.25.*

Question 9. *What happens to our argument above in the case $p = 1$?*

The uniform boundedness principle gives the following general result.

Proposition 2. *A weakly convergent sequence in a normed vector space X is bounded in norm.*

With the help of the Hahn-Banach theorem, we also have

Proposition 3. *Suppose $x_j \rightharpoonup x$ in a normed vector space X, then $||x|| \le \liminf_{j\to\infty} ||x_j||$.*

A very useful weakly sequential compactness criterion is the following

Theorem 2. *In a reflexive Banach space X the closed unit ball is sequentially compact.*

As a consequence, a bounded sequence in a reflexive Banach space X has a weakly convergent subsequence. In particular, this applies to any Hilbert space and to $L^p(\Omega)$ for $1 < p < \infty$. We can use this weakly sequential compactness to complete our proof on the existence of a minimizer at the end of last unit. In fact, the converse to Theorem 2 also holds.

Theorem 3 (Eberlin-Smulyan)**.** *If the closed unit ball in a Banach space X is sequentially compact, then X is reflexive.*

We will not prove either of Theorems 2 or 3. We next introduce the notion of weak* convergence.

Definition 4. *Suppose X is the dual of a normed vector space Y: $X = Y^*$. A sequence $\{l_j\}_j$ in X is said to converge to l weak*, if $l_j(y) \to l(y)$ as $j \to \infty$ for any $y \in Y$.*

Remark 5. *Unless X is reflexive, weak* convergence in X is different from weak convergence. Let $X = l^1$, then $X^* = l^\infty$ and $X = c_0^*$, where c_0 is the space of sequences that converge to 0. Then the sequence $\{e_j\}_j$ in l^1 dose not converge weakly in l^1 (can you supply the detail?). In fact, it does not have any subsequence that converges weakly in l^1. Yet for any $\{c(n)\}_n$ in c_0, $\langle c, e_j \rangle = c(j) \to 0$ as $j \to \infty$. So $\{e_j\}_j \rightharpoonup 0$ weak* in l^1.*

Example 4. *Let $Y = C[-1,1]$. Then $Y^* = X = M[-1,1]$ according to Riesz' theorem (7.18), where $M[0,1]$ denotes the space of finite signed measures on the Borel σ-algebra on $[-1,1]$. Fix a $\phi \in C_c[-1,1]$ with $\int_{-1}^{1} \phi(x)dx = 1$. Then $\{j\phi(jx)\}_j$, considered as a sequence in $L^1[-1,1]$, has no weakly convergent subsequence, yet considered as a sequence in $M[-1,1]$ converges weak* to $\delta(0)$ as $j \to \infty$.*

For a weak* convergent sequence, we also have

Proposition 4. *A weak* convergent sequence is bounded in norm.*

Proposition 5. *Suppose $x_j \rightharpoonup x$ weak* in X, then $||x|| \leq \liminf_{j\to\infty} ||x_j||$.*

When X is not reflexive, Theorem 3 implies that the closed unit ball in X is not sequentially compact. However, when $X = Y^*$ and Y is a separable Banach space, we have

Theorem 4 (Helly). *Suppose $X = Y^*$ and Y is a separable Banach space, then the closed unit ball in X is weak* sequentially compact.*

This theorem applies to l^∞ (as well to $L^\infty(\Omega)$), so any bounded sequence in l^∞ (or $L^\infty(\Omega)$) has a subsequence that converges weak* in l^∞.

Proof. Let $\{y_j\}_j$ be a countable dense set of Y. Given a sequence $\{x_k\}_k$ in X with $||x_k||_X \leq 1$. Then for any j, $\{\langle y_j, x_k \rangle\}_k$ is bounded, hence has a convergent subsequence. By a diagonal selection process, we can find a subsequence $\{x_{k_l}\}_l$ of $\{x_k\}_k$ such that $\{\langle y_j, x_{k_l} \rangle\}_l$ is convergent as $l \to \infty$ for any j. We can define $x_\infty \in Y^*$ by first defining $\langle y_j, x_\infty \rangle = \lim_{l\to\infty} \langle y_j, x_{k_l} \rangle$, then extending to $\langle y, x_\infty \rangle$ for any $y \in Y$ by density. This is possible because $\langle y_j - y_{j'}, x_\infty \rangle = \lim_{l\to\infty} \langle y_j - y_{j'}, x_{k_l} \rangle$, so $|\langle y_j - y_{j'}, x_\infty \rangle| \leq ||y_j - y_{j'}||_Y$ using the bound $||x_{k_l}|| \leq 1$. We also see that $||x_\infty||_X \leq 1$. It's now easy to see that $x_{k_l} \rightharpoonup x_\infty$ weak* as $l \to \infty$. □

We summarize one of the ingredients in the proof as

Proposition 6. *Suppose that $\{y_j^*\}_j$ is a sequence in Y^* such that*

(i). $\{y_j^*\}_j$ is uniformly bounded in Y^*, namely, for some $C > 0$,

$$||y_j^*||_{Y^*} \leq C, \quad \textit{for all } j.$$

(ii). For a dense set Z of Y, $\lim_{j\to\infty}\langle y_j^*, z\rangle$ exists for all $z \in Z$.

Then the sequence $\{y_j^*\}_j$ is convergent weak.*

A similar statement does not hold true for weak convergence, namely, there exists Banach space X and a sequence $\{x_j\}_j$ in X such that

(i). $\{x_j\}_j$ is uniformly bounded in X, and

(ii). For a dense set Z of X^*, $\lim_{j\to\infty}\langle z, x_j\rangle$ exists for all $z \in Z$,

yet, $\{x_j\}_j$ does not converge weakly. (Try to understand the difference between the situation here and above. You can construct an example with $X = c_0$, $x_j = (1, \ldots, 1, 0, 0, \ldots)$, where the last entry 1 is in the jth place.) We do have the following

Proposition 7. *Suppose that $\{x_j\}_j$ is a sequence in X such that*

(i). $\{x_j\}_j$ is uniformly bounded in X, and

(ii). There exists $x_\infty \in X$ and a dense set Z of X^*, $\lim_{j\to\infty}\langle z, x_j\rangle = \langle z, x_\infty\rangle$ for all $z \in Z$. Then $x_j \rightharpoonup x_\infty$ as $j \to \infty$.

Example 5. *Let $P_r(x) = (1-r^2)/(1+r^2-2r\cos(2\pi x))$ denote the Poisson kernel on $\mathbb{T}$. Then for any $f \in L^p(\mathbb{T})$, $1 \le p \le \infty$, $f * P_r$ is a harmonic function of $z = re^{2\pi ix}$ (treating r and $2\pi x$ as the polar coordinates of z) defined inside the unit disk $|z| < 1$. Since $\{P_r\}_{0<r<1}$ is a family of good kernels, we also know that when $1 \le p < \infty$, $f * P_r \to f$ in $L^p(\mathbb{T})$ as $r \nearrow 1$. When $p = \infty$, we see that $\{f * P_r\}_{0<r<1}$ is a bounded family in L^∞, and that $f * P_r \rightharpoonup f$ weak* as $r \nearrow 1$.*

We ask the converse question: given a harmonic function $F(z)$ defined in the unit disk $|z| < 1$, under what conditions can $F(z)$ be represented as $f * P_r$ for some $f \in L^p(\mathbb{T})$ for some p? In our discussion in the paragraph above, since we have $f * P_r \to f$ in $L^p(\mathbb{T})$ as $r \nearrow 1$ in the case $1 \le p < \infty$, and $f * P_r \rightharpoonup f$ weak as $r \nearrow 1$ in the case $p = \infty$, a necessary condition is that $\{F(re^{2\pi ix})\}_r$, considered as a family of functions of $x \in \mathbb{T}$ is bounded in $L^p(\mathbb{T})$. It turns out that we have*

Proposition 8. *Suppose that $F(z)$ is a harmonic function in the unit disk $|z| < 1$, and that $\{F(re^{2\pi ix})\}_r$ is a bounded family in $L^p(\mathbb{T})$. Then when $1 < p \le \infty$, there is a unique $f \in L^p(\mathbb{T})$ such that $F(re^{2\pi ix}) = f * P_r$, and when $p = 1$, there is a unique finite Borel measure μ on $\mathbb{T}$ such that $F(re^{2\pi ix}) = P_r * \mu$ in the sense that*

$$F(re^{2\pi ix}) = \int_{\mathbb{T}} P_r(x-y)\, d\mu(y).$$

2.4 weak (weak*) compactness

The notion of weak (weak*) convergence can be described in terms of the corresponding weak (weak*) topology on X, which is useful in some context.

Definition 5. *Let X be a normed linear space. The weak topology on X is generated by subsets of X of the form $\{x \in X : a < l(x) < b\}$ where $l \in X^*$ and $a < b$ are arbitrary scalars. We denote this topology by $\sigma(X, X^*)$. In other words, any open set in $\sigma(X, X^*)$ contains a subset of the form $\cap_{j=1}^N \{x \in X : a_j < l_j(x) < b_j\}$ for some finite number of $l_j \in X^*$ and $a_j < b_j$.*

Definition 6. *Suppose $X = Y^*$ where Y is a normed linear space. The weak* topology on X is generated by subsets of X of the form $\{y^* \in X : a < y^*(y) < b\}$ where $y \in Y$ and $a < b$ are arbitrate scalars. We denote this topology by $\sigma(Y^*, Y)$.*

We can verify easily that $x_j \rightharpoonup x$ weakly in X in the sense we defined earlier iff $x_j \rightharpoonup x$ in the topology $\sigma(X, X^*)$, and $y_j^* \rightharpoonup y^*$ weak* in X in the sense we defined earlier iff $y_j^* \rightharpoonup y^*$ in the weak* topology $\sigma(Y^*, Y)$.

The main cautionary words are that these topologies are in general not metrizable, so the closure of a subset in these topologies may not be described in terms of sequences. There is, however, a theorem of Eberlin and Smulyan that says that *For a Banach space X, a subset $K \subset X$ is compact in the weak topology iff it is sequentially weak compact.* The same statement does not hold for weak* topology, namely, there exists a Banach space X and a subset $K \subset X^*$ which is compact in the weak* topology, yet it contains no sequence that converges weak*. For an example, take $X = l^\infty$, and K to be the weak* closure of $\{e_j\}_j \subset X^*$. One most useful feature of the weak* topology is the Alaoglu theorem (5.18), the proof of which depends on the Tychonoff theorem (4.42).

Review problems for the midterm. Turn in the *ed problems at the beginning of the midterm on March 25, which will be counted as 32 points out of the 72 points for the midterm. Use the rest as practice problems for the midterm.
Chapter 5: 6*, 22, 25, 47, 49(a), 50, 53, 62, 63.
Chapter 6: 19, 20(a)*, 22(compare with 9, 10),

A. For any $f \in L^p(\mathbb{T})$ with $1 \le p < \infty$, let $S_N(f) = \sum_{m=-N}^N \hat{f}(m) e^{2\pi i m x}$ denotes the Nth partial sum of the Fourier series of f and regard $S_N : L^p(\mathbb{T}) \mapsto L^p(\mathbb{T})$ as a bounded linear operator, with

$$||S_N||_{p \mapsto p} := \sup\{||S_N(f)||_p : ||f||_p = 1\}.$$

(a) Prove that the following two statements are equivalent

i. For any $f \in L^p(\mathbb{T})$, $S_N(f) \to f$ in $L^p(\mathbb{T})$ as $N \to \infty$.

ii. There is a bound $C > 0$ such that $||S_N||_{p \mapsto p} \leq C$.

(b) Prove that neither (i) nor (ii) above holds in $L^1(\mathbb{T})$.

B. For any $t > 0$ let $P_t(x) = (e^{-2\pi t|\xi|})^\vee$, where $e^{-2\pi t|\xi|}$ is treated as a function in $\mathbb{R}^n$.

(a) Use problem 26 of chapter 8 to evaluate $P_t(x)$.

(b) Verify that $(\partial_t^2 + \Delta_x)P_t(x) = 0$ for all $(x,t) \in \mathbb{R}^n \times \mathbb{R}^+$. (Note that $(\partial_t^2 + \Delta_x)e^{-2\pi t|\xi| + 2\pi i x\cdot\xi} = 0$.)

(c)* Verify that for any $f \in L^p(\mathbb{R}^n)$ with $1 \leq p < \infty$, $(\partial_t^2 + \Delta_x)P_t * f = 0$ for all $(x,t) \in \mathbb{R}^n \times \mathbb{R}^+$, and $P_t * f \to f$ in $L^p(\mathbb{R}^n)$ as $t \searrow 0$.

C. Let $1 \leq q \leq p \leq \infty$ and $a(x)$ be a measurable function on Ω. Assume that $au \in L^q(\Omega)$ for every $u \in L^p(\Omega)$. Prove that $a \in L^r(\Omega)$ with

$$r = \begin{cases} \frac{pq}{p-q} & \text{if } p < \infty, \\ q & \text{if } p = \infty. \end{cases}$$

D. Let $W^{1,p}(\mathbb{R})$ denotes the completion of $C_c^1(\mathbb{R})$ in the norm

$$||u||_{W^{1,p}(\mathbb{R})} := ||u||_{L^p(\mathbb{R})} + ||u'||_{L^p(\mathbb{R})}.$$

Prove that each element u in $W^{1,p}(\mathbb{R})$ can be identified as a Hölder continuous function and

$$|u(x) - u(y)| \leq |x-y|^{1/p'}||u'||_{L^p(\mathbb{R})} \quad \text{for all } x, y \in \mathbb{R},$$

where p' is the conjugate exponent of p.

E. Let $p < q$ and $\mathcal{F}$ be a bounded family in $L^p \cap L^q$ that is sequentially pre-compact in L^p. Prove that for any $p < r < q$, $\mathcal{F}$ is sequentially pre-compact in L^r.

F*. Let (f_j) be a sequence in $L^p(\Omega)$ with $1 < p < \infty$ and let $f \in L^p(\Omega)$. Prove that the following two properties are equivalent:

$$(A) \quad f_j \rightharpoonup f \quad \text{weakly in } L^p(\Omega).$$

$$(B) \begin{cases} \exists C > 0 \text{such that} ||f_j||_p \leq C \quad \text{for all } f_j \text{ and} \\ \int_E f_j \to \int_E f \ \forall E \subset \Omega \text{ measurable and } |E| < \infty. \end{cases}$$

2.5 Compact linear operators

Definition 7. *A bounded linear operator $L : X \mapsto Y$ is called compact if it maps bounded sets in X into pre-compact sets in Y. A bounded linear operator $L : X \mapsto Y$ is said to be of finite rank if its range $R(L)$ is a finite dimensional subspace of Y.*

Finite rank operators and their limits under operator norm are examples of compact operators. It turns out that compact perturbations of invertible linear operators behave in some aspects like linear operators between finite dimensional vector spaces.

Let's first list some important features of linear operators between finite dimensional vector spaces. Let A be an $m \times n$ matrix and $L = L_A : \mathbb{R}^n \mapsto \mathbb{R}^m$ be the linear operator associated with A from $\mathbb{R}^n$ to $\mathbb{R}^m$. Then

- The kernel $N(L)$ of L is a closed subspace of $\mathbb{R}^n$ and the range $R(L)$ of L is a closed subspace of $\mathbb{R}^m$.
- There is a well defined adjoint operator $L^* = L_{A^t} : \mathbb{R}^m \mapsto \mathbb{R}^n$, and $y \in R(L)$ iff $y \cdot y^* = 0$ for any $y^* \in N(L^*)$.
- $\dim N(L) = n - \text{rank}(A)$, and $\dim N(L^*) = m - \text{rank}(A^t) == m - \text{rank}(A)$, so $\dim N(L) - \dim N(L^*) = n - m$ is independent of L.
- When $n = m$, L is onto iff $\dim N(L) = 0$.

When $n = m$, one major part of linear algebra is to study the diagonalizability of A, for which the concept of eigenvalues and eigenvectors play crucial roles. Concerning the diagonalizability of an $n \times n$ matrix, we know that

- An $n \times n$ matrix is diagonalizable iff there is a basis of $\mathbb{R}^n$ consisting of eigenvectors of A.
- We often need to extend L_A to act on the complex vector space $\mathbb{C}^n$ to diagonalize A.
- When diagonalization of A fails, a Jordan canonical form may be a useful substitute.
- Certain classes of $n \times n$ matrices are diagonalizable, including real symmetric matrices, Hermitian matrices, orthogonal matrices and unitary matrices, all of which belong to the class of the so called *normal operators*.

In the infinite dimensional setting, the kernel $N(T)$ of a bounded linear operator $T : X \mapsto Y$ is a closed subspace of X, but it may not be finite dimensional; and its range $R(T)$ may not be a closed subspace of Y. So when $Y = X$ and we consider the eigenvalue problem $\lambda I - T : X \mapsto X$, even when $N(\lambda I - T) = \{0\}$, $\lambda I - T$ may fail to have a bounded inverse.

Definition 8. *The spectrum $\sigma(T)$ of $T : X \mapsto X$ is the set of scalars λ for which $\lambda I - T$ fails to have a bounded inverse. The complement of the $\sigma(T)$ is called the resolvent of T.*

So an eigenvalue of T lies in its spectrum; but $\sigma(T)$ may contains non-eigenvalues; and a scalar λ is in the resolvent of T iff

(i) $\lambda I - T$ is onto;

(ii) There exists a constant $C > 0$ such that $||x|| \le C||(\lambda I - T)x||$ for all $x \in X$.

We first list some basic properties on compact operators.

Theorem 5. *1. Composition of a compact operator with any bounded linear operator is compact.*

2. If $T : X \mapsto Y$ is compact, then $T^\dagger : Y^* \mapsto X^*$ is compact. The converse also holds.

3. Let $\mathcal{K}(X, Y)$ denote the set of compact linear operators from X into Y, then $\mathcal{K}(X, Y)$ is a closed linear subspace of $L(X, Y)$ in the operator norm.

4. If Y is a Hilbert space, then any operator in $\mathcal{K}(X, Y)$ can be approximated by finite rank (compact) operators from X to Y.

We will sketch some results on compact operators $K : X \mapsto X$ on a Banach space X. The first result we will discuss is the Fredholm alternative.

Theorem 6. *Let $K : X \mapsto X$ be a compact linear operator. Then*

(a) $N(I - K)$ is finite dimensional.

(b) $R(I - K)$ is closed; more precisely $R(I - K) = N(I - K^\dagger)^\perp$.

(c) $N(I - K) = \{0\}$ iff $R(I - K) = X$.

(d) $dim N(I - K) = dim N(I - K^\dagger)$.

As a consequence, we have

Corollary 7. *Let $K : X \mapsto X$ be a compact linear operator. Then*

(a) $\lambda \neq 0$ is in the spectrum of K iff λ is an eigenvalue of K.

(b) For any eigenvalue $\lambda \neq 0$ of K, its corresponding eigenspace $N(\lambda I - K)$ is finite dimensional, and $\lambda I - K$ has a closed range with codimension $dim N(\lambda I - K)$.

Remark 6. *The conclusions in the Corollary does not hold for $\lambda = 0$. Using the operator K in problem 1 of Assignment 5, with a sequence $\lambda_i \to 0$ but $\lambda_i \neq 0$, then K does not have a closed range, and $\lambda = 0$ is in $\sigma(K)$ (give a direct proof here) but not an eigenvalue of K; on the other hand, if one chooses a sequence $\lambda_i \to 0$ with infinitely many of them zero, then $N(K)$ is infinite dimensional.*

The compactness of K is used crucially in proving (a), the closedness of $R(I-K)$ in (b) and the following two facts, which are used to prove (c).

Lemma 1. *Suppose that $K : X \mapsto X$ is compact, and set $T = I - K$. Then*

$$\exists k \text{ such that } N(T^k) = N(T^{k+1}); \tag{23}$$

and

$$\exists l \text{ such that } R(T^l) = R(T^{l+1}). \tag{24}$$

$R(I-K) = N(I-K^\dagger)^\perp$ follows from $R(I-K) = \overline{R(I-K)}$ and the general fact that $\overline{R(I-K)} = N(I-K^\dagger)^\perp$. (c) follows from (23) and (24) as follows. First assume that $N(T) = \{0\}$. Then it follows directly that $N(T^n) = \{0\}$ for all $n \in \mathbb{N}$. For any $x \in X$, $T^l x \in R(T^l)$. Then $T^l x = T^{l+1} y$ for some $y \in X$ by (24). This implies that $T^l(x - Ty) = 0$, namely, $x - Ty \in N(T^l)$. But $N(T^l) = \{0\}$, therefore $x - Ty = 0$, proving that $x = Ty \in R(T)$ and $R(T) = X$.

Conversely, assume $R(T) = X$, then $R(T^n) = X$ for all $n \in \mathbb{N}$. Suppose $x \in N(T)$, then it follows that $x = T^k y$ for some $y \in X$. This implies that $T^{k+1} y = Tx = 0$, namely $y \in= N(T^{k+1})$. But due to (23), we infer that $x = T^k y = 0$, proving that $N(T) = \{0\}$.

Remark 7. *The closedness of $R(I-K)$ is needed for establishing* (24). (23) *and* (24) *would imply each other if $dimX < \infty$,as $dimN(T^n) + dimR(T^n) = dimX$, but not so if $dimX = \infty$, as the left and right shift operators on l^p demonstrate.*

(23) *and* (24) *can also used to establish the following useful results using routine linear algebra argument:*

$$N(T^k) = N(T^{k+1}) = N(T^{k+2}) = \dots \quad \text{and} \quad R(T^l) = R(T^{l+1}) = R(T^{l+2}) = \dots;$$

denoting $m = \min\{k, l\}$,

$$N(T^m) = N(T^{m+1}) \text{ and } R(T^m) = R(T^{m+1}); \tag{25}$$

$$N(T^m) \cap R(T^m) = \{0\}; \tag{26}$$

$$X = N(T^m) \oplus R(T^m). \tag{27}$$

Each of the spaces $N(T^n)$ and $R(T^n) = X$ are invariant under T and K, and the restriction of T onto $N(T^n)$ is nilpotent and easily understood, thus a decomposition such as (27) allows one to reduce the study of K to invariant subspaces. This is one approach to proving the Jordan canonical form.

One way to prove (d) is to use (27). (27) implies that $R(T) = R(T^m) \oplus (R(T) \cap N(T^m))$. Since $N(T^m)$ is finite dimensional by (a), there is a complementary subspace V of $R(T) \cap N(T^m)$ in $N(T^m)$: $N(T^m) = V \oplus (R(T) \cap N(T^m))$. Thus V is also a complementary subspace of $R(T)$ in X. So $dimN(I - K^\dagger) = dimV$. On the other hand, define the restriction of T on $N(T^m)$ as $\hat{T} : N(T^m) \mapsto N(T^m)$. Then $N(\hat{T}) = N(T)$, and $R(\hat{T}) = R(T) \cap N(T^m)$. The first equality and $R(\hat{T}) \subset R(T) \cap N(T^m)$ are obvious; now for any $x \in R(T) \cap N(T^m)$, $T^m x = 0$ and there exists some $y \in X$ such that $x = Ty$, which implies that $T^{m+1}y = 0$. Thus $y \in N(T^{m+1}) = N(T^m)$ due to (23), and $x = Ty \in R(\hat{T})$. Now in the finite dimensional space $N(T^m)$, standard linear algebra implies that $dimN(\hat{T}) = dimV$, which allows us to conclude (d).

Further properties concerning the eigenvalues of compact operators are established in

Theorem 7. *Let $K : X \mapsto X$ be a compact linear operator. Then $\sigma(K)$ is bounded and $\sigma(K) \setminus \{0\}$ has no accumulation points. When $dimX = \infty$, we also have $0 \in \sigma(K)$.*

Just as in finite dimensional settings, diagonalization can be achieved only under some additional assumptions. We will present a diagonalization result on self-adjoint compact operators.

Definition 9. *Let $\mathcal{H}$ be a Hilbert space. A bounded linear operator $T : \mathcal{H} \mapsto \mathcal{H}$ is called self-adjoint if*

$$(Tx, y) = (x, Ty) \quad \textit{for all } x \in \mathcal{H}.$$

Theorem 8. *Let $\mathcal{H}$ be a Hilbert space and $K : \mathcal{H} \mapsto \mathcal{H}$ be a self-adjoint compact operator, and let $\{\lambda_1, \lambda_2, \ldots\}$ be an enumeration of all the distinct non-zero eigenvalues of K, and P_n denote the orthogonal projection from $\mathcal{H}$ onto $N(\lambda_n I - K)$. Then*

$$K = \sum_n \lambda_n P_n, \tag{28}$$

where the series converges to K in the operator norm.

Let P_0 denote the orthogonal projection of $\mathcal{H}$ onto $N(K)$, then

$$I = \sum_{n=0}^{\infty} P_n,$$

where the series converges to the identity map strongly, namely, for any $x \in \mathcal{H}$,

$$x = \sum_{n=0}^{\infty} P_n x. \tag{29}$$

Furthermore, for $n \neq m$, $P_m P_n = P_n P_m = 0$, the λ_n's are all real, and $N(K) = R(K)^{\perp}$.

A large part of Theorem 8 can be proved using the same ingredients as dealing with self-adjoint matrices in finite dimensions. For example, eigenvectors associated with distinct eigenvalues are orthogonal to each other, which is stated here as $P_m P_n = P_n P_m = 0$ for for $n \neq m$; each eigenspace $N(\lambda_i I - K)$ is an invariant subspace of K, and so is the linear span $\mathcal{Y}$ of all the eigenspaces of K; and for any invariant subspace V of K, $V^{\perp}$ and $\overline{V}$ are also invariant for K. We first assume the following claim and complete a proof for Theorem 8.

Claim 1. *$\mathcal{Y}$ is dense in $\mathcal{H}$, namely, $\overline{\mathcal{Y}} = \mathcal{H}$.*

Now for any $x \in \mathcal{H}$ and $\epsilon > 0$, there exists J and $x_j \in N(\lambda_i I - K)$ such that $||x - \sum_{i=0}^{J} x_i|| < \epsilon$. But orthogonal projections in a Hilbert space imply that, for all $j \geq J$,

$$||x - \sum_{i=0}^{j} P_i x|| \leq ||x - \sum_{i=0}^{J} P_i x|| \leq ||x - \sum_{i=0}^{J} x_j|| < \epsilon,$$

proving (29). Noting that $KP_i = \lambda_i P_i$ for all i, we find

$$||Kx - \sum_{i=0}^{j} \lambda_i P_i x|| = ||K(x - \sum_{i=0}^{j} P_i x)|| \leq ||K|| \cdot ||x - \sum_{i=0}^{j} P_i x|| \to 0$$

as $j \to \infty$. Thus

$$Kx = \sum_{i=0}^{\infty} \lambda_i P_i x$$

converges for all x, and

$$\begin{aligned} ||Kx - \sum_{i=0}^{j} \lambda_i P_i x||^2 &= ||\sum_{i>j} \lambda_i P_i x||^2 \\ &= \sum_{i>j} \lambda_i^2 ||P_i x||^2 \quad \text{ as } P_i x \perp P_{i'} x \text{ for } i \neq i' \\ &\leq \max_{i>j}\{\lambda_i^2\} \sum_{i>j} ||P_i x||^2 \\ &\leq \max_{i>j}\{\lambda_i^2\} ||x||^2 \quad \text{by Bessel's inequality.} \end{aligned}$$

It follows now that

$$||K - \sum_{i=0}^{j} \lambda_i P_i|| \leq \max_{i>j}\{|\lambda_i|\} \to 0$$

as $j \to \infty$.

We will provide a proof for Claim 1 using the following

Lemma 2. *Let $K : X \mapsto X$ be a compact linear operator. Then*

$$\text{for any weakly convergent sequence } x_j \rightharpoonup x_\infty \text{ in } X,\ Kx_j \to Kx_\infty \text{ in norm.} \tag{30}$$

$$||K|| = \sup_{||x||=1} ||Kx|| \text{is attained by some } x_* \text{ with } ||x_*|| = 1, \text{ if } X \text{ is reflexive.} \tag{31}$$

An operator satisfying (30) is called *completely continuous.*

Proof. Let $x_j \rightharpoonup x_\infty$ in X, we will first prove that $Kx_j \rightharpoonup Kx_\infty$. By definition of $x_j \rightharpoonup x_\infty$, $\langle x_j, z^* \rangle \to \langle x_\infty, z^* \rangle$ for any $z^* \in X^*$. We will take $z^* = K^\dagger y^*$, where $y^* \in X^*$ is arbitrary. Then

$$\langle Kx_j, y^* \rangle = \langle x_j, K^\dagger y^* \rangle \to \langle x_\infty, K^\dagger y^* \rangle = \langle Kx_\infty, y^* \rangle,$$

proving $Kx_j \rightharpoonup Kx_\infty$. Now by compactness of K, any subsequence of $\{Kx_j\}$ has a convergent subsequence. We will assume $Kx_{j_k} \to y$ for some $y \in X$. Then $\langle Kx_{j_k}, y^* \rangle \to \langle Kx_\infty, y^* \rangle$ and $\langle Kx_{j_k}, y^* \rangle \to \langle y, y^* \rangle$, which implies that $\langle y, y^* \rangle = \langle Kx_\infty, y^* \rangle$ for all $y^* \in X^*$. So $y = Kx_\infty$, namely, there is a unique limit, Kx_∞, to the subsequential limits of $\{Kx_j\}$. Therefore $Kx_j \to Kx_\infty$.

For (31), we take a maximizing sequence x_j for $\sup_{||x||=1} ||Kx||$. Again we may assume $Kx_j \to y$ for some $y \in X$ and $x_j \rightharpoonup x_*$ for some x_* by the assumption that X is reflexive. Therefore $Kx_j \to Kx_*$, and $||Kx_*|| = \lim_{j\to\infty} ||Kx_j|| = ||K||$. We finally claim that $||x_*|| = 1$ when $||K|| > 0$. Otherwise, since $||x_*|| \leq \liminf_{j\to\infty} ||x_j|| = 1$, $x_* = \theta n$ for some $0 < \theta < 1$ and $||n|| = 1$, which would lead to $||Kn|| = ||Kx_*||/\theta = ||K||/\theta > ||K||$, contradicting the definition of $||K||$. □

We now supply

Proof of Claim 1. If $\overline{\mathcal{Y}} \neq \mathcal{H}$, then $\overline{\mathcal{Y}}^\perp$ is a non-trivial subspace of $\mathcal{H}$, invariant for K. Let $\hat{K} : \overline{\mathcal{Y}}^\perp \mapsto \overline{\mathcal{Y}}^\perp$ be the restriction of K to $\overline{\mathcal{Y}}^\perp$. If $||\hat{K}|| = 0$, then $\overline{\mathcal{Y}}^\perp$ is a non-trivial subspace of $N(K) \subset \mathcal{Y}$, impossible. So $||\hat{K}|| > 0$, and there exists some $x_* \in \overline{\mathcal{Y}}^\perp$ with $||x_*|| = 1$ attaining $||\hat{K}||$. Then a simple calculation of the first variation of $||\hat{K}x||^2$ on the constraint $\{x \in \overline{\mathcal{Y}}^\perp : ||x|| = 1\}$, using the self-adjoint assumption of K at this point, shows that x_* is an eigenvector of $K : Kx_* = \pm||\hat{K}||x_*$. This implies that x_* should have been included in $\mathcal{Y}$, again impossible. Therefore we conclude that $\overline{\mathcal{Y}}^\perp = \{0\}$ and $\overline{\mathcal{Y}} = \mathcal{H}$. □

We now outline an application of Theorem 8 to eigenfunction expansions. Our motivation comes from solving

$$\begin{cases} r(x)u_t(x,t) = \left(p(x)u_x(x,t)\right)_x + q(x)u(x,t) & a < x < b, t > 0, \\ \quad u(a,t) = u(b,t) = 0 & t > 0, \\ \quad u(x,0) = f(x) & a < x < b, \end{cases} \tag{32}$$

for some given $f(x)$. We first look for separable solutions to the first two of (32), of the form $u(x,t) = X(x)T(t)$. Then

$$r(x)X(x)T'(t) = T(t)\left[(p(x)X'(x))' + q(x)X(x)\right],$$

and it follows that $T'(t) = -\mu T(t)$ for some constant μ and

$$(p(x)X'(x))' + q(x)X(x) + \mu r(x)X(x) = 0.$$

The boundary conditions in (32) become $X(a)T(t) = X(b)T(t) = 0$ for all $t > 0$. In order to obtain non-trivial solutions of the form $u(x,t) = X(x)T(t)$, we must choose $X(a) = X(b) = 0$. Thus we need to find a non-trivial solution X to the Sturm-Liouville problem

$$\begin{cases} (p(x)X'(x))' + q(x)X(x) + \mu r(x)X(x) = 0, \quad a < x < b, \\ \qquad\qquad X(a) = X(b) = 0. \end{cases} \tag{33}$$

Suppose that $\{\mu_i\}$ is a collection of values for which (33) has non-trivial solution $X_i(x)$, then $e^{-\mu_i t}X_i(x)$ solves the first two of (32), and so does any (finite) linear combination of them $\sum_i c_i e^{-\mu_i t}X_i(x)$, which takes on initial data $\sum_i c_i X_i(x)$.

Question 10. *What kind of functions $f(x)$ can be expanded in terms of $\{X_i(x)\}$? More precisely, is there a way to find a "complete" set of eigenfunctions of (33) such that all functions $f(x)$ in a reasonable function space can be expanded in terms of $\{X_i(x)\}$?*

We know one case well for (33): $p(x) \equiv r(x) \equiv 1$ and $q(x) \equiv 0$. If we set $a = 0$, then all the eigenvalues for (33) in this case are given by $\mu_n = (n\pi/b)^2$ with $X_n(x) = \sin(n\pi x/b)$ as a corresponding eigenfunction. Note that $\mu_n \to \infty$ as $n \to \infty$.

It turns out we can cast the eigenfunction expansion problem as the eigenfunction expansion of a compact operator K constructed through the solution operator to (33). We will fix a μ_0 such that there is a unique solution X to

$$\begin{cases} (p(x)X'(x))' + (q(x) - \mu_0 r(x))\, X(x) = f(x), \quad a < x < b, \\ \qquad\qquad X(a) = X(b) = 0. \end{cases} \tag{34}$$

for every f in a function space to be chosen. This can be done if we assume

$$p(x), q(x) \text{and } r(x) \text{ are bounded measurable functions on } [a,b], \text{ with } p(x), r(x) \geq c > 0. \tag{35}$$

A solution $X(x)$ to (34) can be constructed by extending our variational approach: we look for a minimizer $X(x)$ to

$$\min\{I[X] := \int_a^b \left[p(x)|X'(x)|^2 + (\mu_0 r(x) - q(x))\, X^2(x) + f(x)X(x)\right] dx\}$$

in the class $X \in E := H_0^1(a,b)$. We will take $f \in L^2[a,b]$. Then because of (35), we can choose μ_0 sufficiently large to prove that $I[X]$ has a unique minimizer X in E for every $f \in L^2[a,b]$. Denote this X by $S[f]$ and *regard S as an operator on $L^2[a,b]$ through $S : L^2[a,b] \mapsto E \subset L^2[a,b]$*. Our solution process will show that S is linear and maps bounded set in $L^2[a,b]$ into bounded set in E. But bounded set in E are pre-compact in $L^2[a,b]$! So we have constructed a compact operator S on $L^2[a,b]$. Let $X = S[f]$ and $Y = S[g]$ for some $f, g \in L^2[a,b]$, then multiplying both sides of the equation for X by Y gives

$$-\int_a^b \left[p(x)X'(x)Y'(x) + (\mu_0 r(x) - q(x))\, X(x)Y(x)\right] dx = \int_a^b f(x)Y(x)dx.$$

But

$$\begin{aligned} &-\int_a^b \left[p(x)X'(x)Y'(x) + (\mu_0 r(x) - q(x))\, X(x)Y(x)\right] dx \\ &= \int_a^b X(x) \left[(p(x)Y'(x))' + (q(x) - \mu_0 r(x))\, Y(x)\right] dx = \int_a^b X(x)g(x)dx. \end{aligned}$$

Using $(\cdot,\cdot)$ to denote the L^2 inner product in $L^2[a,b]$, we have established

$$(f, S[g]) = (S[f], g) \quad \text{for all } f, g \in L^2[a,b]. \tag{36}$$

Define $K[X] = S[r(x)X]$. Then replacing f by $r(x)X(x)$ and g by $r(x)Y(x)$, (36) becomes

$$(r(x)X, K[Y]) = (K[X], r(x)Y), \tag{37}$$

namely, S is self-adjoint in the standard L^2 inner product, while K is self-adjoint in the weighted L^2 inner product $(X,Y)_r := (r(x)X, Y)$. Let us repeat the computation to show that eigenfunctions of (33) corresponding to distinct eigenvalues are orthogonal with weight $r(x)$. Suppose $\mu_1 \neq \mu_2$ are two distinct eigenvalues of (33) with corresponding eigenfunctions $X_1(x)$ and $X_2(x)$, then

$$\left(p(x)X_i'(x)\right)' + (q(x) - \mu_0 r(x))\, X_i(x) = -(\mu + \mu_0) r(x) X_i(x),$$

for each $i = 1, 2$, from which it follows that $X_i = -(\mu + \mu_0)S[r(x)X_i]$. Applying (36) for $f = -(\mu_1 + \mu_0)r(x)X_1(x)$ and $g = -(\mu_2 + \mu_0)r(x)X_2(x)$, we get $S[f] = X_1$, $S[g] = X_2$, and

$$-(\mu_1 + \mu_0)(r(x)X_1, X_2) = -(\mu_2 + \mu_0)(X_1, r(x)X_2).$$

Thus $(X_1, X_2)_r = (r(x)X_1, X_2) = 0$. Let $\mathcal{H}$ denote the L^2 inner product space with weight $r(x)$. Under our assumption (35), $K : \mathcal{H} \mapsto \mathcal{H}$ is well defined and compact, self-adjoint. Thus Theorem 8 applies to K. In addition, each eigenfunction X of (33) is an eigenfunction of K with eigenvalue $\lambda = -1/(\mu + \mu_0)$, and vice verse. In the case here $N(K) = \{0\}$. By Theorem 8, we can construct an orthonormal basis for each of the non-zero eigenspace of K, and the union of these bases will become a basis for $\mathcal{H}$, that is, denoting by $\{X_i\}$ such an orthonormal basis, then any $f \in \mathcal{H}$ can be expanded as

$$f = \sum_i c_i X_i,$$

where the convergence is in the $\mathcal{H}$ norm, and $c_i = (f, X_i)_r = \int_a^b f(x)X_i(x)r(x)dx$. $\lambda_i \to 0$ now corresponds to $\mu_i \to \infty$.

Theorem 9. *Assume* (35). *Then there exists an orthonormal basis of $\mathcal{H}$, consisting of eigenfunctions of* (33).

Remark 8. *We have outlined the eigenfunction expansion in the simple setting of homogeneous Dirichlet boundary conditions in* (32). *The same conclusion can be reached with any homogeneous boundary conditions of the type $\alpha_i X(a) + \beta_i X'(a) = 0$ and $\alpha_2 X(b) + \beta_2 X'(b) = 0$, assuming $|\alpha_i| + |\beta_i| \neq 0$.*

The eigenfunction expansion is often presented as an eigenvalue problem for the unbounded operator $L[X] := \left[(p(x)X'(x))' + q(x)X(x)\right]/r(x)$ *on $\mathcal{H}$. The domain of L can't be $\mathcal{H}$, but is taken to be some dense subspace of $\mathcal{H}$. Assuming $p(x) \in C^1[a,b]$ for instance, then $(p(x)X'(x))' = p(x)X''(x) + p'(x)X'(x)$ is in $L^2[a,b]$ if $X \in H^2[a,b]$. One could regard L as a bounded operator from $H^2[a,b]$ into $\mathcal{H}$, but it does not quite make sense to talk about eigenvalues of an operator from one space into a different space. So a convenient approach is to regard $H^2[a,b]$ as a dense subspace of $\mathcal{H}$ and treat L an an unbounded operator. This is the way many differential operators are treated.*

Assignment 5 The following problems are adapted from the new edition of Brezis' text. Turn all four problems in by 5pm, April 20.

1. Let (λ_n) be a bounded sequence in $\mathbb{R}$ and consider the operator $K : l^p \mapsto l^p$ for $1 \leq p \leq \infty$:

$$Kx = (\lambda_1 x_1, \lambda_2 x_2, \ldots, \lambda_n x_n, \ldots),$$

where

$$x = (x_1, x_2, \ldots, x_n, \ldots).$$

Prove that K is compact iff $\lambda_n \to 0$ as $n \to \infty$.

2. Let X and Y be two Banach spaces and $T \in \mathcal{K}(X, Y)$. Assume $\dim X = \infty$. Prove that there exists a sequence (u_n) in X such that $||u_n||_X = 1$ and $||Tu_n||_Y \to 0$ as $n \to \infty$.

3. Let (λ_n) be a sequence of positive numbers such that $\lim_{n\to\infty} \lambda_n = \infty$. Let V denote the space of sequences $(u_n)_{n\geq 1}$ such that

$$\sum_{n=1}^{\infty} \lambda_n |u_n|^2 < \infty.$$

V is equipped with the inner product

$$((u, v)) = \sum_{n=1}^{\infty} \lambda_n u_n v_n.$$

Prove that V is a Hilbert space and the injection $V \subset l^2$ is compact.

4. Define $Tu(x) = \int_0^x u(t)dt$.

   (a) Prove that $T : C[0, 1] \mapsto C[0, 1]$ is compact, but $T(B)$ is not closed, where B is the closed unit ball in $C[0, 1]$.
   In the following, let $X = L^p(0, 1)$ with $1 \leq p < \infty$, and consider $T \in L(X)$.

   (b) Prove that $T \in \mathcal{K}(X)$.

   (c) Determine $\sigma(T)$.

   (d) Give an explicit formula for $(\lambda I - T)^{-1}$ when λ is in the resolvent of T.

   (e) Determine $T^\dagger$.

Assignment 6 The following are recommended for chapter 3.
Chapter 3: 1, 2, 7(a), 9, 12, 20, 21, 24, 25, 28, 29, 36, 39 Here are a couple more for practice.

1. Suppose that f is monotone increasing on $[a, b]$ and $f(b) - f(a) = \int_a^b f'(x)dx$. Prove that f is absolutely continuous on $[a, b]$.

2. Suppose that $f \in BV([a, b])$. Prove that $|f| \in BV([a, b])$. Construct a function g such that $|f| \in BV([a, b])$, but $g \notin BV([a, b])$.

3. Suppose that f and g are in $BV([a,b])$. Prove that $\max\{f,g\} \in BV([a,b])$.

4. Suppose that $f \in L[a,b]$ and $F(x) = \int_a^x f(u)du$. Prove that $F \in BV([a,b])$ and $V_a^b(F) = \int_a^b |f(u)|du$.

5. Suppose that $f \in BV([a,b])$ and $V_a^b(f) = \int_a^b |f'(u)|du$. Prove that f is absolutely continuous on $[a,b]$.

3 Some comments on "signed measures and differentiation"

A large part of the material on "signed measures and differentiation" originated in extending the Fundamental Theorem of Calculus to the context of Lebesgue integral. Namely, we would like to answer

Question 11. *How to extend the statement that*

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x) \quad \text{at every } x \in [a,b], \text{ if } f \text{ is continuous on } [a,b],$$

to the case when $f \in L[a,b]$?

Question 12. *When F is differentiable on (a,b), continuous on $[a,b]$, and $F'(x)$ is Riemann integrable on $[a,b]$, we have $F(b) - F(a) = \int_a^b F'(x)dx$. What would be an appropriate extension of this theorem in the context of Lebesgue integral?*

For question 11, one approach is to study

$$\begin{aligned} &\lim_{h\to 0} \frac{\int_a^{x+h} f(t)\,dt - \int_a^x f(t)\,dt}{h} \\ =&\lim_{h\to 0} \frac{1}{h}\int_x^{x+h} f(t)\,dt, \end{aligned}$$

and try to prove that the limit is $f(x)$ almost everywhere. Denoting

$$\hat{A}_h f(x) := \frac{1}{h}\int_x^{x+h} f(t)\,dt = f * \left(\frac{\chi_{[0,h]}}{h}\right),$$

and recalling that that $\{h^{-1}\chi_{[0,h]}\}_{h>0}$ forms an approximation to identity, so $||\hat{A}_h f - f||_{L[a,b]} \to 0$ as $h \searrow 0$, therefore, there exists a subsequence $h_i \searrow 0$ such that $\hat{A}_{h_i} f(x) \to f(x)$ almost everywhere on $[a,b]$. So we would have an affirmative answer that $\left(\int_a^x f(t)\,dt\right)' = f(x)$ almost everywhere on $[a,b]$, provided we can prove that $\lim_{h\to 0}\hat{A}_h f(x)$ exists almost everywhere. For this purpose, it is natural to study the function

$$\Omega f(x) := |\limsup_{h\to 0}\hat{A}_h f(x) - \liminf_{h\to 0}\hat{A}_h f(x)|,$$

in particular $m(\{x : \Omega f(x) > \alpha\})$ for any $\alpha > 0$. It turns out to be extremely useful to study the Hardy-Littlewood maximal function

$$Hf(x) := \sup_{r>0} A_r|f|(x) = \sup_{r>0}\frac{1}{m(B(r,x))}\int_{B(r,x)}|f(y)|dy.$$

The main property of this function is given in Theorem 3.17. Using $\Omega f(x) \le 2Hf(x)$ and Theorem 3.17, Folland gives a proof in Theorem 3.18 that $\hat{A}_h f(x) \to f(x)$ almost everywhere on $[a,b]$ as $h \to 0$.

Another approach to question 11 is to use the relation

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

for $f \in L[a,b]$, where $f^+(x) = \max\{f(x),0\}$ and $f^-(x) = -\min\{f(x),0\}$. Thus $\int_a^x f(t)\,dt$ is written as the *difference of two monotone increasing functions.* We could answer questions on the differentiability of $\int_a^x f(t)\,dt$ if we could answer questions on the differentiability of monotone increasing functions. Interest in the latter was related to Weierstrass' construction of a continuous function which is nowhere differentiable: his construction used a sum of trigonometric series and exploited the oscillations of these functions. So it was natural to ask whether monotone functions enjoy better differentiability properties. Using a Vitali covering argument, Lebesgue proved that *a monotone function is differentiable almost everywhere*—this is part of Theorem 3.23 in Folland's text, although with a more modern proof using Theorem 3.22.

Folland's approach studies the differentiability of a monotone increasing $F(x)$ from the point of view of measures. When F is monotone increasing(and right continuous), there is an associated Borel measure μ_F on $(\mathbb{R},\mathcal{B})$ with $\mu_F((a,b]) = F(b) - F(a)$ for any $a < b$. Noting

$$\frac{F(x+h) - F(x)}{h} = \frac{\mu_F((x,x+h])}{m((x,x+h])}, \quad \text{for } h > 0,$$

with a similar expression when $h < 0$, the differentiability of $F(x)$ can be studied through the differentiability of the measure μ_F, for which the Lebesgue-Radon-Nikodym theorem can be applied as in Theorem 3.22 to give an affirmative answer. Here is a bit more detail. Let $\mu_F = \lambda + f(x)dx$ be the Lebesgue decomposition of μ_F with respect to the Lebesgue measure m. We know that λ is a positive measure and

$f(x) \geq 0$ m-a.e., and there exists a Borel set Z_1 with $m(Z_1) = 0$ such that $\lambda(Z_1^c) = 0$. Thus

$$\frac{\mu_F((x, x+h])}{m((x, x+h])} = \frac{\lambda((x, x+h])}{m((x, x+h])} + \frac{\int_x^{x+h} f(t)dt}{h}.$$

According to Theorem 3.18, there exists a Borel set Z_2 with $m(Z_2) = 0$ such that for any $x \in [a, b] \setminus Z_2$,

$$\frac{\int_x^{x+h} f(t)dt}{h} \to f(x), \quad \text{as } h \to 0.$$

It remains to prove that

$$\lim_{h\to 0} \frac{\lambda((x, x+h])}{m((x, x+h])} \quad \text{exists for } m\text{-a.e } x.$$

In fact, Theorem 3.22 proves that

$$\lim_{h\to 0} \frac{\lambda((x, x+h])}{m((x, x+h])} = 0 \quad \text{ for } m\text{-a.e } x.$$

Therefore, $F'(x) = f(x)$ for m-a.e x. In terms of the Lebesgue decomposition of μ_F, this result says $\mu_F = \lambda + F'(x)dx$.

Either of these approaches tells us that $\int_a^x f(t)\,dt$ is differentiable almost everywhere if $f \in L[a, b]$, but does not tell us at this point what it's derivative is. For that, one uses again, as above, that there exists a sequence of $h_i \to 0$ such that $\hat{A}_{h_i} f(x) \to f(x)$ almost everywhere on $[a, b]$. Since we now know at this point that $\lim_{h\to 0} \hat{A}_h f(x)$ exists almost everywhere on $[a, b]$, we can now conclude that this limit is $f(x)$ almost everywhere on $[a, b]$, i.e., $\left(\int_a^x f(t)\,dt\right)' = f(x)$ almost everywhere on $[a, b]$.

The study above also identifies a class of functions of interest: the difference of two monotone increasing functions. Such functions are characterized by the property of *bounded variation*. Properties of these BV functions are studied in Examples 3.25, Lemma 3.26, Theorem 3.27, and Theorem 3.29. Folland then summarized the main differentiability properties of these functions in Proposition 3.30. It is useful to expand part of Proposition 3.30 into the following more concrete form:

Proposition 9. *Let F be monotone increasing and right continuous, then $\mu_F = \lambda + F'(x)dx$, where λ is a positive measure on $(\mathbb{R}, \mathcal{B})$, singular with respect to the Lebesgue measure m: there is a Borel measurable set Z with $m(Z) = 0$, such that $\lambda(E) = \lambda(E \cap Z)$ for any Borel set E. As a consequence,*

$$F(b) - F(a) = \mu_F((a, b]) \geq \int_a^b F'(x)dx, \tag{38}$$

for any $a < b$, with equality iff $\lambda((a,b]) = 0$. In general, $\lambda((a,b]) = 0$ iff $\mu_F << m$ on $[a,b]$.

For question (12), the standard Cantor function $C(x)$ is a monotone increasing continuous function, with $C'(x) = 0$ almost everywhere, but $C(b) - C(a) > \int_a^b C'(x)dx$ for many $0 < a < b < 1$! In the case of monotone increasing function $F(x)$, the answer to question (12) is already partly provided by the decomposition $\mu_F = \lambda + F'(x)dx$ above: $F(b) - F(a) = \int_a^b F'(x)dx$ iff $\mu_F << m$. This absolute continuity condition can be expressed in terms of F directly, as done in the text. The general condition for $F(b) - F(a) = \int_a^b F'(x)dx$ is also given in terms of the absolute continuity of F, as in Propositions 3.30, 3.32, and Theorem 3.35.

Remark 9. *A function $f(x)$ may be differentiable everywhere on $[a,b]$, yet fails to be in $BV[a,b]$. Here is an example.*

$$f(x) = \begin{cases} x^2 \sin(\frac{1}{x^2}), & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Uniform limit of a sequence of $BV[a,b]$ functions may not be in $BV[a,b]$. For example,

$$f_m(x) = \begin{cases} x^2 \sin(\frac{1}{x^2}), & |x| \geq \frac{1}{m}, \\ 0, & |x| < \frac{1}{m}. \end{cases}$$

However, if $\{f_m\}$ is a sequence of $BV[a,b]$ with bounded BV semi-norm*, namely,*

there exists $M > 0$ such that $T_{f_m}(b) - T_{f_m}(a) \leq M$ for all m.

and

$$\lim_{m\to\infty} f_m(x) = f(x) \quad \textit{for all } x \in [a,b].$$

Then $f \in BV[a,b]$, and $T_f(b) - T_f(a) \leq M$.

Remark 10. *BV functions arise in defining and computing the arclength of a curve. Suppose that $\gamma : [a,b] \mapsto \mathbb{C}$ defines a continuous parametrized curve in $\mathbb{C}$. It's natural to approximate the arclenth of γ by inscribed linear segments and define*

$$\textit{arclength}(\gamma) = \sup\{\sum |\gamma(t_{i+1}) - \gamma(t_i)| : a = t_0 < t_1 < \cdots < t_N = b\}.$$

This is identified to be the total variation of the complex valued function $\gamma(t)$ on $[a,b]$, and according to (38)*, is $\geq \int_a^b T'_\gamma(t)dt$, with equality iff $T_\gamma(t)$ is AC on $[a,b]$. But*

$$\left|\frac{T_\gamma(t+h) - T_\gamma(t)}{h}\right| \geq \left|\frac{\gamma(t+h) - \gamma(t)}{h}\right|,$$

for any $h \neq 0$, so $T'_\gamma(t) \geq |\gamma'(t)|$ for almost all $t \in [a,b]$. Thus we conclude arclength(γ) $\geq \int_a^b |\gamma'(t)|dt$, with equality iff $\gamma(t)$ is AC on $[a,b]$, using the following

Fact. *$\gamma(t)$ is AC on $[a,b]$ iff $T_\gamma(t)$ is AC on $[a,b]$.*

4 Comments on the notions of weak derivatives

We have encountered several notions of weak (strong) L^p derivatives. Here are some comments to clarify their relations. The notion of weak derivatives was formally defined in Definition 15. But we first saw a need for extending the notion of derivatives through our variational approach, in which we encountered a Cauchy sequence $u_j \in C^1(\mathbb{R}^n)$ (or $C_c^\infty(\mathbb{R}^n), \mathcal{S}$) such that $u_j \to u$ in $L^2(\mathbb{R}^n)$ for some $u(x) \in L^2(\mathbb{R}^n)$, as $j \to \infty$, and $\{\partial_l u_j\}_j$ is Cauchy in $L^2(\mathbb{R}^n)$. Then for each $l = 1, \cdots, n$, there exists $w_l \in L^2(\mathbb{R}^n)$ such that $\{\partial_l u_j\}_j \to w_l$ $L^2(\mathbb{R}^n)$. We will see now that this u has weak L^2 derivatives in the sense of Definition 15. Fix any $\eta \in C_c^\infty(\mathbb{R}^n)$, for each j, we have

$$\int_{\mathbb{R}^n} u_j(x)\partial_l\eta(x)\,dx = -\int_{\mathbb{R}^n} \partial_l u_j(x)\eta(x)\,dx.$$

Since $\{u_j\}_j \to u$ and $\{\partial_l u_j(x)\}_j \to w_l$ in $L^2(\mathbb{R}^n)$, as $j \to \infty$, we know that

$$\int_{\mathbb{R}^n} u_j(x)\partial_l\eta(x)\,dx \to \int_{\mathbb{R}^n} u\partial_l\eta(x)\,dx,$$

and

$$\int_{\mathbb{R}^n} \partial_l u_j(x)\eta(x)\,dx \to \int_{\mathbb{R}^n} w_l(x)\eta(x)\,dx.$$

(Can you supply the details?) Thus

$$\int_{\mathbb{R}^n} u\partial_l\eta(x)\,dx = -\int_{\mathbb{R}^n} w_l(x)\eta(x)\,dx$$

for every $l = 1, \cdots, n$ and $\eta \in C_c^\infty(\mathbb{R}^n)$, proving that u has weak L^2 derivatives, and $\partial_l u = w_l$.

Conversely, let u have weak L^2 derivatives $\partial_l u$. Fix a family of good kernels $\{g_j\} \subset C_c^\infty(\mathbb{R}^n)$, and consider $G_j = u * g_j$. Then each $G_j \in C^\infty(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, and

$$\partial_l G_j = (\partial_l u) * g_j.$$

We know $G_j \to u$ and $\partial_l G_j \to \partial_l u$ in $L^2(\mathbb{R}^n)$. This essentially proves that u can be approximated in $L^2(\mathbb{R}^n)$ by smooth (decaying) $L^2(\mathbb{R}^n)$ functions whose derivatives are Cauchy in $L^2(\mathbb{R}^n)$. The G_j may not have compact support, but one can further construct smooth cut-offs that have compact support and approximate u as well.

In problems # 8 and 9 of chapter 8, two more notions of derivatives are introduced for the one dimensional case. If u has strong L^p derivative h on $\mathbb{R}$, then it is easy to see that it has h as weak L^p derivative, which can be defined almost identically as in Definition 15.

Here is an outline to prove that if u has weak L^p derivative h, then it has strong L^p derivative. We will first prove the property in #9, namely, u is absolutely continuous on any finite interval. Consider the same regularized G_j, then, for any $a < x < y < b$,

$$G_j(y) - G_j(x) = \int_x^y G_j^{'}(t)dt = \int_x^y u^{'} * g_j(t)dt. \tag{39}$$

It follows that

$$|G_j(y) - G_j(x)| \leq \int_x^y |u^{'} * g_j(t)|dt \leq \left(\int_x^y |u^{'} * g_j(t)|^p dt\right)^{1/p} |y-x|^{1/q}, \tag{40}$$

for $1 < p < \infty$, where q is the conjugate exponent of p, so $\{G_j\}$ is equicontinuous, as $u^{'} * g_j(t) \to u^{'}$ in $L^p(\mathbb{R})$, so there is an upper bound for $(\int_x^y |u^{'} * g_j(t)|^p dt)^{1/p}$. In the case $p = 1$, (40) still gives equicontinuity of $\{G_j\}$, as $u^{'} * g_j(t) \to u^{'}$ in $L^1(\mathbb{R})$. Since $G_j \to u$ in $L^p(\mathbb{R})$, there is also a bound for $\int_a^b G_j(x)dx$, which together with (40) implies that there exists a subsequence of $\{G_j\}$ that converges to a continuous limit w uniformly on any compact interval of $\mathbb{R}$. Furthermore, since $G_j^{'}(t) \to u^{'}$ in $L^p(\mathbb{R})$, we can pass to the limit in (39) to conclude that

$$w(y) - w(x) = \int_x^y u^{'}(t)dt \quad \text{for any } x < y.. \tag{41}$$

But $G_j \to u$ in $L^p(\mathbb{R})$, we deduce that $w = u$ a.e., namely after modifying u on a Lebesgue null set, u is identical to a continuous function w that satisfies (41). We can now conclude that u is absolutely continuous on every bounded interval.