FOURIER ANALYSIS & METHODS

JULIE ROWLETT

ABSTRACT. Caveat Emptor! These are just informal lecture notes. Errors are inevitable! Read at your own risk! Also, this is by no means a substitute for the textbook, which is warmly recommended: *Fourier Analysis and Its Applications,* by Gerald B. Folland. He was the first math teacher I had at university, and he is awesome. A brilliant writer. So, why am I even doing this? Good question...

1. 2018.02.27

1.1. **Integral representation of the Bessel functions.**

Let $z = e^{i\theta}$ for $\theta \in \mathbb{R}$. Then the theorem on the generating function for the Bessel functions says

$$\sum_{n\in\mathbb{Z}} J_n(x) z^n = e^{\frac{xz}{2} - \frac{x}{2z}}.$$

So, we use the fact that

$$\frac{1}{e^{i\theta}} = e^{-i\theta},$$

together with this formula to see that

$$\sum_{n\in\mathbb{Z}} J_n(x) e^{in\theta} = e^{\frac{x}{2}(e^{i\theta} - e^{-i\theta})}.$$

By Euler's formula,

$$\sum_{n\in\mathbb{Z}} J_n(x) e^{in\theta} = e^{ix\sin\theta} = \cos(x\sin\theta) + i\sin(x\sin\theta).$$

Therefore, the left side is the Fourier expansion of the function on the right. OMG!!! Hence, the Bessel functions are actually *Fourier coefficients* of this function! So,

$$J_n(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ix\sin\theta} e^{-in\theta} d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos(x\sin\theta - n\theta) + i\sin(x\sin\theta - n\theta) d\theta.$$

Note that

$$\sin(x\sin(-\theta) - n(-\theta)) = \sin(-x\sin\theta - n(-\theta)) = -\sin(x\sin\theta - n\theta).$$

So the sine part is odd and integrates to zero. We therefore have

$$J_n(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos(x\sin\theta - n\theta) d\theta.$$

This formula can be super useful. For example, we see that the Bessel functions have yet another property similar to their straight-laced Swedish ancestors, the sine and cosine. They satisfy $|J_n(\theta)| \le 1 \forall x$.

1.2. **Solving PDEs and special functions.** We have seen how the process of solving PDEs like the heat and wave equation often leads to a set of functions which comprise an orthogonal basis for $\mathcal{L}^2$ or a weighted $\mathcal{L}^2$ space. These basis functions generally come from separation of variables. When we solve the "space" part of the PDE, we very often end up solving a type of SLP. The easiest examples are:

$$f'' = \lambda f, \quad f(a) = 0 = f(b), \text{ for f defined on the interval, } [a,b]$$
$$f'' = \lambda f, \quad f'(a) = 0 = f'(b), \text{ for f defined on the interval, } [a,b]$$
$$f'' = \lambda f, \quad f(a) = 0 = f'(b), \text{ for f defined on the interval, } [a,b]$$
$$f'' = \lambda f, \quad f'(a) = 0 = f(b), \text{ for f defined on the interval, } [a,b].$$

A more challenging example comes from solving the heat and wave equations on a circular sector. There, when we did separation of variables, we got the nice type of SLP above for the angular variable (θ), and we got a more complicated SLP for the radial variable. Just in case you don't remember exactly how this worked, we wrote the Laplace operator in polar coordinates. The μ_m number comes from solving the SLP for the $\Theta(\theta)$ function, and the equation for the function $R(r)$ which depends on the radial variable, r was

$$r^2 R'' + rR' + (\lambda r^2 + \mu_m)R = 0.$$

We made a small clever change of variables. Let

$$x = \sqrt{\lambda} r, \quad f(x) := R(r), \quad r = \frac{x}{\sqrt{\lambda}}.$$

Then by the chain rule

$$R'(r) = \sqrt{\lambda} f'(x), \quad R''(r) = \lambda f''(x).$$

So, the equation is

$$\left(\frac{x^2}{\lambda}\right) \lambda f''(x) + \frac{x}{\sqrt{\lambda}} \sqrt{\lambda} f'(x) + (x^2 + \mu_m) f(x) = 0.$$

This simplifies, recalling that $\mu_m = -m^2\pi^2/\alpha^2$,

besseleq

$$x^2 f''(x) + x f'(x) + (x^2 - m^2\pi^2/\alpha^2) f(x) = 0. \tag{1.1}$$

This is the definition of Bessel's equation of order $\frac{m\pi}{\alpha}$. So, depending on the boundary condition, the function $R_{m,k}(r)$ is $J_{m\pi/\alpha}(\sqrt{\lambda_{m,k}} r)$, where the $\lambda_{m,k}$ ensures the boundary condition.

In *other* geometric settings, this same process will lead to *other* special functions. In the last part of this course, based on chapter 6 in Folland, we will look at *the French polynomials,*

(1) Legendre polynomials
(2) Hermite polynomials
(3) Laguerre polynomials

1.3. **Origin of the French polynomials.**

1.3.1. *Legendre polynomials.* These French polynomials arise from using spherical coordinates to solve the wave and heat equations on a three-dimensional sphere.

1.3.2. *Hermite polynomials.* These French polynomials arise from using parabolic coordinates to solve the wave and heat equations in a parabolic shaped region.

1.3.3. *Laguerre polynomials.* These French polynomials arise from the quantum mechanics of the hydrogem atom.

1.4. **Orthogonal polynomials general theory.** For the purpose of this course, it is most important that you learn how to *use* the orthogonal polynomials. Depending on how much time we have, we may go into the details of the origins of the French polynomials, but these details are rather complicated, and we will not be examined. So, we prioritize that which shall be examined. The following proposition will be useful.

Proposition 1. *Assume that $\{p_n\}_{n\in\mathbb{N}}$ is a sequence of polynomials such that p_n is of degree n for each n. Assume that $p_0 \neq 0$. Then for each $k \in \mathbb{N}$, any polynomial of degree k is a linear combination of $\{p_j\}_{j=0}^k$.*

Proof: The proof is by induction of course! If q_0 is a polynomial of degree 0, then we may simply write

$$q_0 = \frac{q_0}{p_0} p_0.$$

This is okay because p_0 is degree zero, so it is a constant, and $p_0 \neq 0$, so the coefficient q_0/p_0 is also a constant. Assume that we have verified the proposition for all $0, 1, \ldots k$. We wish to show that it holds for $k+1$. So, let q be a polynomial of degree $k+1$. This means that

$$q(x) = ax^{k+1} + l.o.t. \quad l.o.t. \text{ means lower order terms}$$

has

$$a \neq 0.$$

Moreover, since p_{k+1} is of degree $k+1$ (not of a lower degree), it is of the form

$$p_{k+1} = bx^{k+1} + l.o.t., \quad b \neq 0.$$

So, let us consider

$$q(x) - \frac{a}{b} p_{k+1}(x) = p(x) \text{ which is degree } k.$$

By induction, p is a linear combination of $p_0, \ldots, p_k$. Therefore

$$q(x) = \frac{a}{b} p_{k+1} + \sum_{j=0}^{k} c_j p_j,$$

for some constants $\{c_j\}_{j=0}^k$.

Proposition 2. *Let $\{p_k\}_{k=0}^\infty$ be a set of polynomials such that each p_k is of degree k, and $p_0 \neq 0$. Moreover, assume that they are $\mathcal{L}^2$ orthogonal on a finite bounded interval $[a,b]$. Then these polynomials comprise an orthogonal basis of $\mathcal{L}^2$ on the interval $[a,b]$.*

Proof: Assume that some $f \in \mathcal{L}^2$ on the interval is orthogonal to all of these polynomials. Therefore by the preceding proposition, f is orthogonal to *all* polynomials. To see this, note that if p is a polynomial of degree n, then there exist numbers $c_0, \ldots, c_n$ such that

$$p = \sum_{j=0}^{n} c_j p_j \implies \langle f, p \rangle = \sum_{j=0}^{n} \overline{c_j} \langle f, p_j \rangle = 0.$$

We shall use the fact that continuous functions are dense in $\mathcal{L}^2$. Therefore given $\varepsilon > 0$, there exists a continuous function, g, such that

$$||f-g|| < \frac{\varepsilon}{2(||f||+1)}.$$

Next, we use the Stone-Weierstrass Theorem which says that all continuous functions on bounded intervals can be approximated by polynomials. Therefore, there exists a polynomial p such that

$$||g-p|| < \frac{\varepsilon}{2(||f||+1)}.$$

Finally, we compute

$$||f||^2 = \langle f, f\rangle = \langle f-g+g-p+p, f\rangle = \langle f-g, f\rangle + \langle g-p, f\rangle + \langle p, f\rangle$$
$$= \langle f-g, f\rangle + \langle g-p, f\rangle.$$

By the Cauchy-Schwarz inequality,

$$||f||^2 \le ||f-g||||f|| + ||g-p||||f|| < \frac{||f||\varepsilon}{2(||f||+1)} + \frac{||f||\varepsilon}{2(||f||+1)} < \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, this shows that $||f|| = 0$. Hence by the three equivalent conditions to be an orthogonal basis, we have that the polynomials are an orthogonal basis of $\mathcal{L}^2$ on the interval.

1.5. **Best approximations.** We recall a slight variation of the best approximation theorem:

Theorem 3. *Let $\{\phi_n\}_{n\in\mathbb{N}}$ be an orthonormal set set in a Hilbert space, H. If $f \in H$,*

$$||f - \sum_{n\in\mathbb{N}} \langle f, \phi_n\rangle \phi_n|| \le ||f - \sum_{n\in\mathbb{N}} c_n\phi_n||, \quad \forall \{c_n\}_{n\in\mathbb{N}} \in \ell^2,$$

and $=$ holds $\iff$ $c_n = \langle f, \phi_n\rangle$ holds $\forall n \in \mathbb{N}$. More generally, let $\{\phi_n\}_{n=0}^N$ be an orthogonal, non-zero set in a Hilbert space H. Then,

$$||f - \sum_{n=0}^N \frac{\langle f, \phi_n\rangle}{||\phi_n||^2}\phi_n|| \le ||f - \sum_{n=0}^N c_n\phi_n||, \quad \forall \{c_n\}_{n=0}^N \in \mathbb{C}^{N+1}.$$

Equality holds if and only if

$$c_n = \frac{\langle f, \phi_n\rangle}{||\phi_n||^2}, \quad n = 0, \dots, N.$$

How to prove it? The only difference is the last part, but we can use the proof of the first part. Define $\psi_n = 0$ for $n > N$. Next define

$$\psi_n = \frac{\phi_n}{||\phi_n||}, \quad n = 0, \dots, N.$$

Repeat the argument in the proof of the best approximation theorem using $\{\psi_n\}_{n\in\mathbb{N}}$ instead of ϕ_n.

$$||f - \sum_{n\in\mathbb{N}} c_n\psi_n||^2 = ||f - \sum_{n\in\mathbb{N}} \hat{f}_n\psi_n + \sum_{n\in\mathbb{N}} \hat{f}_n\psi_n - \sum_{n\in\mathbb{N}} c_n\psi_n||^2$$

$$= ||f - \sum_{n\in\mathbb{N}} \hat{f}_n \psi_n||^2 + ||\sum_{n\in\mathbb{N}} \hat{f}_n \psi_n - \sum_{n\in\mathbb{N}} c_n \psi_n||^2 + 2\Re\langle f - \sum_{n\in\mathbb{N}} \hat{f}_n \psi_n, \sum_{n\in\mathbb{N}} \hat{f}_n \psi_n - \sum_{n\in\mathbb{N}} c_n \psi_n \rangle.$$

The scalar product

$$\langle f - \sum_{n\in\mathbb{N}} \hat{f}_n \psi_n, \sum_{n\in\mathbb{N}} \hat{f}_n \psi_n - \sum_{n\in\mathbb{N}} c_n \psi_n \rangle = \langle f, \sum_{n\in\mathbb{N}} (\hat{f}_n - c_n) \Psi_n \rangle - \sum_{n\in\mathbb{N}} \hat{f}_n \langle \psi_n, \sum_{m\in\mathbb{N}} (\hat{f}_m - c_m) \Psi_n \rangle.$$

By the orthogonality and definition of Ψ_n, and the definition of $\hat{f}_n$,

$$= \sum_{n\in\mathbb{N}} \hat{f}_n \overline{(\hat{f}_n - c_n)} - \sum_{n\in\mathbb{N}} \hat{f}_n \sum_{m\in\mathbb{N}} \overline{(\hat{f}_m - c_m)} \langle \psi_n, \psi_m \rangle$$

$$= \sum_{n\in\mathbb{N}} \hat{f}_n \overline{(\hat{f}_n - c_n)} - \sum_{n\in\mathbb{N}} \hat{f}_n \overline{(\hat{f}_n - c_n)} = 0.$$

Therefore

$$||f - \sum_{n\in\mathbb{N}} c_n \psi_n||^2 = ||f - \sum_{n\in\mathbb{N}} \hat{f}_n \psi_n||^2 + ||\sum_{n\in\mathbb{N}} \hat{f}_n \psi_n - \sum_{n\in\mathbb{N}} c_n \psi_n||^2$$

$$= ||f - \sum_{n=0}^{N} \hat{f}_n \psi_n||^2 + \sum_{n=0}^{N} |\hat{f}_n - c_n|^2 \le ||f - \sum_{n=0}^{N} \hat{f}_n \psi_n||^2,$$

with equality if and only if $c_n = \hat{f}_n$ for all n. Since

$$\sum_{n=0}^{N} \hat{f}_n \psi_n = \sum_{n=0}^{N} \frac{\langle f, \phi_n \rangle}{||\phi_n||^2} \phi_n,$$

this completes the proof.

1.5.1. *Applications: best approximation problems.* This shows us that if we have a finite orthogonal set of non-zero vectors in a Hilbert space, then for any element of that Hilbert space, the best approximation of f in terms of those vectors is given by

$$\sum_{n=0}^{N} \frac{\langle f, \phi_n \rangle}{||\phi_n||^2} \phi_n.$$

Here is the setup of questions which can be solved using this theory. Either:

(1) You are given functions defined on an interval which are $\mathcal{L}^2$ orthogonal on that interval (possibly with respect to a weight function which is also specified). Either you recognize that they orthogonal because you've seen them before (like sines, cosines, from problems you have solved previously) *or* you *compute* that they are $\mathcal{L}^2$ orthogonal on the interval. Then, you are asked to find the numbers $c_0, c_1, \ldots c_N$ so that the $\mathcal{L}^2$ norm, or the weighted $\mathcal{L}^2$ norm of $f - \sum_{k=0}^{N} c_k \phi_k$ is minimized, where the function f is also specified.

(2) You are asked to find the polyonomial of at most degree N such that the $\mathcal{L}^2$ norm (or weighted $\mathcal{L}^2$ norm) of $f - p$ where p is a polynomial is minimized.

In the first case, you compute

$$c_k = \frac{\langle f, \phi_k \rangle}{||\phi_k||^2}.$$

In the second case you need to *build up a set of orthogonal or orthonormal polynomials.* Then, you let ϕ_k be defined to be the polynomial of degree k you have built. Proceed the same as in the first case, and your answer shall be

$$\sum_{k=0}^{N} c_k \phi_k.$$

If you don't like the thought of building up a set of orthogonal polynomials, if you are lucky, then it may be possible to suitably modify some of the French polynomials to be orthogonal on the interval under investigation, with respect to the (possibly weighted) $\mathcal{L}^2$ norm. So, we shall proceed to study the French polynomials. Depending on how much time we have, we may also be able to get into their "origin stories."

1.6. **The Legendre polynomials.** The *Legendre polynomials,* are defined to be

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} \left((x^2 - 1)^n \right).$$

OMG like why on earth are they defined in such a bizarre way, right? What did you expect, they are French polynomials! Of course they are not defined in some simple way, mais non, they must be all fancy and shrouded in mystery and intrigue. Actually though, the reason comes from the PDE in which they arise as solving one part of the separation of variables for the heat and wave equations in three dimensions using spherical coordinates. First, let us do a small calculation involving these polynomials:

$$(x^2 - 1)^n = \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} (x^2)^k = \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} x^{2k}.$$

Therefore, if we differentiate n times, only the terms with $k \geq n/2$ survive. Differentiating a term x^{2k} once we get $2kx^{2k-1}$. Differentiating n times gives

$$\frac{d^n}{dx^n}(x^{2k}) = x^{2k-n} \prod_{j=0}^{n-1} (2k - j).$$

If we want to be really persnickety, we prove this by induction. For $n = 1$, we get that

$$(x^{2k})' = 2kx^{2k-1}.$$

Which is correct. If we assume the formula is true for n, then differentiating $n+1$ times using the formula for n we get

$$(2k - n)x^{2k-(n+1)} \prod_{j=0}^{n-1} (2k - j) = x^{2k-(n+1)} \prod_{j=0}^{n} (2k - j).$$

See, it is correct. As a result,

$$P_n(x) = \frac{1}{2^n n!} \sum_{k \geq n/2}^{n} (-1)^{n-k} \binom{n}{k} x^{2k-n} \prod_{j=0}^{n-1} (2k - j).$$

So, we see that this is indeed a polynomial of degree n.