Advanced Analysis: Skeleton notes

1. Fourier Theory

The *Fourier series* of a function $f(\theta)$ on $[-\pi, \pi]$ is

$$\sum_{-\infty}^{\infty} a_n e^{in\theta},$$

where

$$a_n = a_n(f) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)e^{-in\theta}d\theta.$$

The *Fourier transform* of a function $f(x)$ on $\mathbf{R}$ is the function

$$\hat{f}(\xi) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-ix\xi}dx,$$

and the Fourier inversion formula is

$$f(x) = \int_{-\infty}^{\infty} f(\xi)e^{ix\xi}d\xi.$$

One circle of basic questions is to ask under what conditions, and in what sense the Fourier series represents the original function/the Fourier transform is defined and the inversion formula is true.

Proposition 1. *(a) If $|f|$ is integrable on $[0, 2\pi]$ the Fourier co-efficients a_n tend to zero as $n \to \pm\infty$. (b) If $|f|$ is integrable on $\mathbf{R}$ the FT $\hat{f}$ is defined as a continuous function of ξ and $\hat{f}(\xi) \to 0$ as $|\xi| \to \infty$.*

This leads easily to a simple criterion for pointwise convergence

Proposition 2. *If $|f|$ and $|\frac{f(x)-f(0)}{x}|$ are integrable on $[-\pi, \pi]$ (respectively $\mathbf{R}$) then the Fourier series at 0 converges (to $f(0)$) (respectively the Fourier Inversion Formula holds at $x = 0$).*

For example the hypotheses are satisfied for *Holder continuous*, integrable functions.

It is not enough merely to assume the function is continuous: we have the famous

Proposition 3. *There is a continuous function whose Fourier series diverges at* 0.

(With a similar statement for Fourier integrals.)

One can prove this either by constructing an example explicitly, or by appeal to the Banach-Steinhaus Theorem. Either way the crucial ingredient involes the Dirichlet Kernel. The partial sum $S_N = \sum_{-N}^{N} a_n e^{in\theta}$ of the Fourier Series of a function f is represented by a convolution

$$S_N(\phi) = \int_{-\pi}^{\pi} D_N(\phi - \theta)f(\theta)d\theta,$$

where

$$D_N(x) = \frac{1}{2\pi} \frac{\sin(N+1/2)x}{\sin x/2}$$

and the point is that

$$\int_{-\pi}^{\pi} |D_N| \to \infty$$

as $N \to \infty$.

A substitute is

Proposition 4. *The Fourier series of a continuous function converges to the function at each point in the sense of Cesaro means.*

The key here is that the Cesaro sum $\frac{s_1+s_2+\dots s_N}{N}$ is given by convolution with a function F_N and the sequence F_N is an "approximate identity".

Here are some other frameworks in which one can do Fourier theory.

(1) **Complex variables** For Fourier series, put $z = e^{i\theta}$. We are conidering periodic functions as functions on the boundary of the unit disc in $\mathbf{C}$. Suppose such a function extends to a holomorphic function on an annulus. Then the Fourier representation coincides with the Laurent series of complex function theory.

For suitable functions f on $\mathbf{R}$ we may allow the variable ξ in the Fourier transform to be complex. Thus the Laplce transform may be regarded as a special case of the Fourier transform (for functions supported on a half-line).

(2) L^2 **theory** For Fourier series this amounts to the ordinary discussion of $e^{in\theta}$ as a complete orthonormal system in the Hilbert space $L^2(-\pi,\pi)$. The Fourier series of an L^2 function converges in L^2. For Fourier transforms we prove, for well-behaved functions f, the Parseval formula

$$\int_{-\infty}^{\infty} |\hat{f}(\xi)|^2 d\xi = 2\pi \int_{-\infty}^{\infty} |f(x)|^2 dx.$$

Using this, and a density argument, we extend the FT to an isomorphism from $L^2(\mathbf{R})$ to itself.

(3) *Distributions* Also called "generalised functions". The Schwartz space $\mathcal{S}$ consists of C^∞ functions on $\mathbf{R}$ all of whose derivatives decay faster than any $|x|^{-k}$ at infinity. A distibution is a linear map $L : \mathcal{S} \to \mathbf{C}$ such that there is a sequence f_i in $\mathcal{S}$ with

$$L(F) = \lim \int_{-\infty}^{\infty} f_i(x) F(x) dx.$$

For example the delta function

$$\delta_0(F) = F(0),$$

is obtained via a sequence forming an "approximate identity". We can define differentiation of distributions via integration by parts: in PDE theory this is the idea of a "weak solution". The Fourier transform of a distribution is defined by

$$\hat{L}(\hat{f}) = L(f(-x)).$$

For example consider the function $g_\alpha(x) = x^\alpha$ if $x > 0$, $g_\alpha(x) = 0$ if $x < 0$. Suppose α is real and not an integer. The function defines a distribution whose Fourier transform can be interpreted as the function

$$\hat{g}_\alpha(\xi) = \frac{\Gamma(\alpha+1)}{2\pi} e^{-\pi/2i(\alpha+1)\mathrm{sgn}(\xi)} |\xi|^{-(\alpha+1)}.$$

The Poisson Summation formula
This asserts that, for suitable functions F on $\mathbf{R}$,

$$\sum_{k=-\infty}^{\infty} F(2\pi k) = \sum_{n=-\infty}^{\infty} \widehat{F}(n).$$

It can be viewed as the equation of distributions

$$\left(\sum \delta_{2\pi k}\right) = \widehat{\sum \delta_n}.$$

Applied to the Gaussian functions e^{-ax^2} one deduces that the Theta function

$$\theta(\tau) = \sum_{k=-\infty}^{\infty} e^{\pi i \tau k^2},$$

where τ lies in the upper half-plane, satisfies

$$\theta(\tau) = \frac{1}{\sqrt{-i\tau}} \theta(-1/\tau).$$

It follows that θ^8 is a "modular form of weight 2" for the group of Mobius maps generated by $\tau \mapsto \tau + 2, \tau \mapsto -1/\tau$.

2. L^p.

For any measure space X and $p \geq 1$ the L^p norm on functions on X is defined by

$$\|f\|_{L^p} = \left(\int_X |f|^p\right)^{1/p}.$$

Holder's inequlaity states that

$$\int_X fg \leq \|f\|_{L^p} \|g\|_{L^{p'}},$$

where p' is the conjugate exponent $\frac{1}{p} + \frac{1}{p'} = 1$. For $p > 1$ this sets up a duality $(L^p)^* = L^{p'}$. It is often useful to consider the supremum norm

$$\|f\|_{L^\infty} = \sup |f|,$$

as a limiting case of the L^p norms, but one must be careful in going to $p = \infty$–some natural statements one might hope to be true are actually false. Likewise for the other extreme case $p = 1$ as we shall see.

Why should one bother with the L^p norms, as opposed to sticking with L^2 say?

(1) One obvious answer is that the family of L^p norms gives a more accurate picture of the whole function. The distibution function of f is defined to be

$$\mu_f(\alpha) = |\{x : |f(x)| \geq \alpha\}|.$$

(We use the notation $|A|$ to denote the measure of a set A.) Then

$$\|f\|_{L^p}^p = p \int_0^\infty \mu_f(\alpha)\alpha^{p-1} d\alpha.$$

Knowing all the L^p norms of f is roughly equivalent to knowing the distribution function.

(2) In nonlinear problems (e.g. PDE) one is forced to consider L^p.

(3) Related to the above: the Sobolev inequalities involve particular L^p norms, depending on dimension.

Proposition 5.

Let f be a smooth function of compact support on $\mathbf{R}^n$.

(a) If f is supported in the ball of radius R and $p > n$

$$|f(0)| \leq C_{n,p} R^{1-n/p} \|\nabla f\|_{L^p},$$

(b) If $p < n$ and q is defined by

$$1 - \frac{n}{p} = -\frac{n}{q}$$

then

$$\|f\|_{L^q} \leq C_{n,p} \|\nabla f\|_{L^p}.$$

Here the constants $C_{n,p}$ depend only on n and p.

The numerical relations between p, q, n are dictated by the scaling behaviour of the norms.

These Sobolev inequalities extend in a routine way to larger classes of functions by a density argument, and can be viewed equivalently as embedding theorems for various function spaces. The inequalities are intimately related with *isoperimetric inequalities.*

Young's inequality asserts that for a function f on $\mathbf{R}$ with Fourier transform $\hat{f}$ one has

Proposition 6.

$$\|\hat{f}\|_{L^{p'}} \leq \frac{1}{(2\pi)^{1/p}} \|f\|_{L^p},$$

where $1 < p \leq 2$ *and* p, p' *are conjugate exponents.*

This can be proved as an application of the Riesz-Thorin Interpolation Theorem, using the fact that the result is true when $p = 1, p' = \infty$(obvious) and when $p = p' = 2$(Parseval). The idea is to extend the real variable $p^{-1} \in (1/2, 1)$ to a strip in the complex plane.

3. The Hilbert transform.

The *Hilbert transform* on the circle can be defined via Fourier series.

$$H(\sum a_n e^{in\theta}) = -i \sum \mathrm{sgn}(n) a_n e^{in\theta}.$$

Clearly H defines a bounded map on L^2: indeed on the subspace of functions of integral zero H is an isometry and $H^2 = -1$. Note also that H takes real-valued functions to real-valued functions. If u is a real valued function on the circle then $v = H(u)$ is the unique function of integral 0 such that $u + iv$ extends to a holomorphic function over the disc.

Proposition 7. *For each* $p > 1$*, the Hilbert transform defines a bounded map* $H : L^p \to L^p$.

In concrete terms this says that there are constants c_p such that

$$\|H(f)\|_{L^p} \leq c_p \|f\|_{L^p},$$

and as usual it suffices to prove this for some dense set of functions f.

Corollary. *For* $p > 1$ *the Fourier series of an* L^p *function converges in* L^p

The first proof of Proposition 7 (Riesz) uses complex function theory. For simplicity consider the case when the exponent p is an even integer $2k$. (Actually the general case can be deduced from this using interpolation and duality.) It suffices to consider the case when $f = u$ is real and of integral 0. Then $u + iH(u) = u + iv$ is the restriction of a holomorphic function $f(z)$ on the disc, vanishing at the origin. Cauchy's Theorem gives

$$\int_{-\pi}^{\pi} (u + iv)^{2k} d\theta = 0,$$

and the proof follows easily when one expands using the binomial theorem.

The Hilbert transform on the line is defined in a similar fashion:

$$H(\hat{f})(\xi) = -i\mathrm{sgn}(\xi) \hat{f}(\xi).$$

There is a version of Proposition 7 in this case too.

The two Hilbert transforms can be represented by *singular integral operators.* On the circle

$$H(f)(\theta) = \int \cot(\frac{\theta - \phi}{2}) f(\phi) d\phi :$$

and on the line

$$H(f)(x) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{f(y)}{y-x} dy.$$

These integrals are interpreted via their "principal values". (Which are defined at each point if f is smooth.)

There is a contour-integral point of view on these formulae. In the case of the line, suppose that f is a function (with suitable decay) on $\mathbf{R}$ which extends to a holomorphic function on a thin strip $\{|\Im z| < \epsilon\}$ in $\mathbf{C}$. Let C_+, C_- be contours parrallel to the real axis and slightly below and above respectively. Then the Cauchy integrals

$$F_+(w) = \frac{1}{2\pi} \int_{C_+} \frac{f(z)}{z-w} dz \, F_-(w) = -\frac{1}{2\pi} \int_{C_-} \frac{f(z)}{z-w} dz,$$

yield holomorphic functions F_+, F_- on the (slightly extended) upper and lower half-planes respectively. Cauchy's formula gives

$$f(z) = F_+(z) + F_-(z),$$

on the strip about $\mathbf{R}$. The Hilbert transform is given by $Hf = i(F_+ - F_-)$ i.e. the principal value is the average of the values obtained by deforming the path of integration either side of the pole. There is a similar discussion in the case of circle, with holomorphic functions on the regions $|z| < 1 + \epsilon$, $|z| > 1 - \epsilon$.

4. Higher dimensions.

Singular integral operators occur naturally in analysis on $\mathbf{R}^n$. For example consider the Laplace operator $\Delta = -\sum_{i=1}^{3} \nabla_i^2$ on $\mathbf{R}^3$. In potential theory one shows that if ρ has, say, compact support then the solution to the equation $\Delta\phi = \rho$ with $\phi \to 0$ at infinity is given by the Green's function

$$\phi(x) = G(\rho) = \int_{\mathbf{R}^3} g(x-y)\rho(y)dy,$$

where

$$g(x) = \frac{1}{4\pi|x|}.$$

The operator G is a straightforward integral operator–the kernel function g has a singularity but is integrable near zero. Now let $i \neq j$ and consider the partial derivative $\nabla_i \nabla_j \phi$. Manipulating formally this is given by the integral

$$\int_{\mathbf{R}^3} q_{ij}(x-y)\rho(y)dy,$$

where

$$q_{ij}(x) = \frac{x_i x_j}{2\pi|x|^5},$$

and this is correct if the integral is interpreted as a principal value. This defines an operator $Q_{ij} = \nabla_i \nabla_j \Delta^{-1}$. (Beware that the story needs to be modified a bit when $i = j$. The operator $\nabla_i \nabla_i \Delta^{-1}$ is given by $-1/3$ times the identity map plus a singular integral operator.)

There are (at least) three ways of approaching the relation between $\rho = \Delta\phi$ and $\nabla_i\nabla_j\phi$. One is to work directly with the differential operators. For example, integrating by parts one proves the identity

$$\int_{\mathbf{R}^3} |\Delta\phi|^2 = \sum_{i,j} \int_{\mathbf{R}^3} |\nabla_i\nabla_j\phi|^2,$$

(assuming sufficient decay at infinity). This shows that the Q_{ij} are bounded on L^2. The second is to use the Fourier transform. We have

$$\widehat{Q_{ij}(\rho)}(\xi) = m_{ij}(\xi)\hat{\rho}(\xi),$$

where $m_{ij}(\xi)$ is the homogeneous function

$$m_{ij}(\xi) = c\frac{\xi_i\xi_j}{|\xi|^2}.$$

Thus the operator Q_{ij} goes over under Fourier transform to a multiplier operator with homogeneous, hence bounded, multiplier m_{ij}. The fact that the multiplier is bounded shows again, by Parseval, that Q_{ij} is bounded on L^2. The third way is to use the singular integral description, with kernel q_{ij}, as above.

In general a classical (translation-invariant) singular differential operator on $\mathbf{R}^n$ is defined by a kernel function $k(x)$ which is homogeneous of order $-n$ so

$$k(x) = \frac{1}{|x|^n}\Omega(\frac{x}{|x|}),$$

where Ω is a function on the unit sphere S^{n-1} of integral 0. This means that, for smooth f of compact support say, one can define the integrals

$$(Tf)(x) = \int_{\mathbf{R}^n} k(x-y)f(y)dy$$

as principle values. The operator T goes over under Fourier transform to a multiplier operator

$$\widehat{Tf}(\xi) = M(\xi)\hat{f}(\xi),$$

where M is homogeneous of degree 0, and again with integral 0 on the unit sphere. Hence the operator T extends to L^2. This can all be expressed by saying that the Fourier transform of k, viewed as a distribution, is M.

The Hilbert transform (on $\mathbf{R}$) and the Q_{ij} are examples of this general set-up. Other important examples arise from Boundary Value problems. The Riesz transforms R_i on $\mathbf{R}^n$ are defined by kernel function $r_i(x) = \frac{x_i}{|x|^{n+1}}$ and multipliers $m_i(\xi) = c\frac{\xi_i}{|\xi|}$. They have the property that if Ψ is a harmonic function on the half-space $\mathbf{R}^n \times \mathbf{R}^+$ with suitable decay at infinity then the tangential derivatives of Ψ on the boundary are obtained from the normal derivative by applying the Riesz transforms. In dimension $n = 1$ we get the Hilbert transform.

The central result in this course is

Proposition 8. *Any classical singular integral operator of the above kind defines a bounded operator $T : L^p \to L^p$ for each $p > 1$.*

It is possible to prove this for most (perhaps all) such operators that occur in nature by a trick to reduce to the case of the Hilbert transform on $\mathbf{R}$–separating into radial and angular variables. This applies initially to the case when the kernel k is an odd function; for example the operators R_i. But then we can write $Q_{ij} = R_i \circ R_j$ to deduce the result for the Q_{ij}. However the n-dimensional proof in the next section has the virtue that it extends to more general situations.

5. The Calderon-Zygmund Theory.

The strategy of this prrof of Proposition 8 is to interpolate, using the transparent behaviour on L^2. The first obstacle is that the result is not even true for the other extremes $p = 1, \infty$. For example the real and imaginary parts of the function which maps the disc conformally to an infinite strip are related by the Hilbert transform on the circle, but one is bounded and the other not. It is even more obvious that a singular integral operator T of the kind considered on $\mathbf{R}^n$ cannot map L^1 to L^1: for any function f, say of compact support, with

$$\int_{\mathbf{R}^n} f \neq 0$$

$|Tf|$ decays like $|x|^{-n}$ at infinity and so is not integrable. The way around this is to substitute "weak-type" bounds when $p = 1$. A map T is said to be of weak type $(1, 1)$ if

$$|\{x : |Tf(x)| > \alpha\}| \leq \frac{C}{\alpha} \|f\|_{L^1}.$$

Proposition 9. *(Marcinkiewicz) If T is of weak type (1,1) and bounded as a map from L^2 to L^2 then T defines a bounded map from L^p to L^p for $1 < p \leq 2$*

The proof gives explicit bounds on the L^p operator norms.

Now consider an operator T defined by a kernel $k(x, y)$

$$(Tf)(x) = \int_{\mathbf{R}^n} k(x, y) f(y) dy.$$

We may even suppose k is smooth and of compact support initially if we like.

Theorem 10. *Suppose the operator T above with kernel k satisfies*

(1) $\|Tf\|_{L^2} \leq c_1 \|f\|_{L^2}$
(2) $|\nabla_x k(x, y)| + |\nabla_y k(x, y)| \leq \frac{c_2}{|x-y|^{n+1}}$;

then

$$|\{x : |Tf(x)| < \alpha\}| \leq \frac{C}{\alpha} \|f\|_{L^1},$$

where the constant C depends only on c_1, c_2.

Corollary. *Such an operator defines a bounded map on L^p for all $p > 1$. Moreover the L^p operator norm can bounded above by an explicit expression involving only c_1, c_2.*

To obtain the the corollary one uses the interpolation theorem to handle the range $1 < p \leq 2$ and a duality argument to handle the range $2 \leq p < \infty$.

From this Corollary one easily deduces Proposition 9, for example by approximating the singular kernel by smooth ones. (In fact the possible presence of a singularity on the diagonal is irrelevant in the proof of Theorem 10.)

A vital ingredient in the proof is a simple

Lemma. *Let $\Box$ be a cube in $\mathbf{R}^n$ of side-length r, and let $\Box^*$ be the cube with the same centre and side length $2r$. Then if β is a function supported in $\Box$ with $\int_\Box \beta = 0$ we have*

$$\|T\beta\|_{L^1(\mathbf{R}^n\setminus\Box^*)} \leq c\|\beta|_{L^1},$$

where the constant C depends only on c_2 (and not on the scale r).

The proof of Theorem 10 uses the famous Calderon-Zygmund cube decomposition of a function. Suppose initially that f is continuous and has compact support. Given α we find a large cube containing the support of f such that the average value of $|f|$ on the cube is less than α. Then we subdivide this cube into 2^n smaller ones. The rule is that we stop subdividing any cube once the average value of $|f|$ exceeds α. In this way we get a collection of disjoint cubes Q_i such that the average value of f on each Q_i lies between α and $2^n\alpha$ and $|f| \leq \alpha$ outside $\bigcup Q_i$. Now write $f = g + b$ where $b = \sum b_i$, each b_i is supported in Q_i and on Q_i,

$$b_i = f - \frac{1}{|Q_i|}\int_{Q_i} f.$$

The proof now follows by applying the Lemma to each b_i, the L^2 hypothesis to g, and piecing together what we know.

6. Wavelet analysis:preliminaries.

We know that a Fourier series $\sum a_n e^{in\theta}$ defines an L^2 function if and only if $\sum |a_n|^2 < \infty$ and a smooth function if and only if $|a_n|$ is rapidly decreasing (faster than any inverse power). It is however difficult to extract detailed information about the behaviour of the function—both the degree of smoothness (differentiablity, Holder continuity....) and the growth (L^p-norms)—from the Fourier coefficients. This is illustated by the following fact: if $\sum |a_n|^2 < \infty$ then for almost all choices of signs $\sum \pm a_n e^{in\theta}$ defines an L^p function. To state this precisely we introduce the Rademacher fuctions $R_n(t), t \in [0,1], n \geq 0$ defined by

$$R_n(t) = \rho(2^n t)$$

where $\rho(x) = (-1)^{[x]}$. Then R_n are *independent random variables* on $[0,1]$, each taking values ± 1 with equal probablity $1/2$. Choose an identification of $\mathbf{Z}$ with $\mathbf{N}$, and hence define functions r_n for $n \in \mathbf{Z}$. Then we have

Proposition 11. *If $\sum |a_n|^2 < \infty$ and*

$$F(t,\theta) = \sum r_n(t) a_n e^{in\theta}$$

then for any $p > 1$ the integral

$$\int\int |F(t,\theta)|^p dt d\theta$$

is finite. Hence (by Fubini's Theorem) $F(t,\ )$ is in L^p for almost all $t \in [0,1]$.

The proof is a straightforward application of Kinchine's inequality, which states that all L^p norms are equivalent on the subspace of $L^2[0,1]$ spanned by the R_n.

Another way of viewing this result is in terms of multiplier transformations. If λ_n is any sequence of complex numbers with $|\lambda_n| = 1$ (say), the map which takes $\sum a_n e^{in\theta}$ to $\sum \lambda_n a_n e^{in\theta}$ is bounded (in fact an isometry) on L^2, but cannot be bounded on L^p for all multipliers (λ_n). Precisely which sequences (λ_n) *do* define bounded maps on L^p is a delicate matter. Similarly for Fourier transforms. Now consider more generally a measure space X (e.g. the line or the circle) and an orthonormal basis ψ_n of $L^2(X)$—where the ψ_n lie in $L^1 \cap L^\infty$ say. We may consider the class of linear maps given by multipliers in this basis

$$T_{\underline{\lambda}}(\sum a_n \psi_n) = \sum \lambda_n a_n \psi_n.$$

Proposition 12. *For $p > 1$ the following are equivalent:*
(1) all $T_{\underline{\lambda}}$ are bounded on L^p:

$$\|T_{\underline{\lambda}} f\|_{L^p} \leq C\|f\|_{L^p}.$$

(2) the L^p norm is equivalent to the norm

$$\|f\|_{\underline{\psi}} = \| \left(\sum |a_n(f)|^2 |\psi_n|^2\right)^{1/2} \|_{L^p(X)}.$$

Here $a_n(f) = \langle f, \psi_n \rangle$.

The proof uses the same trick as in Proposition 11. The background to this is the Littlewood-Paley theory: if we decompose the Fourier series of a function into "dyadic blocks"

$$f = \sum \Delta_j(f),$$

then any multiplier which is constant on the Δ_j is bounded on L^p, or in turn the L^p norm is equivalent to an norm

$$\| \left(\sum |\Delta_j(f)|^2\right)^{1/2} \|_{L^p}.$$

Another (equivalent) way of expressing the condition of Proposition 12 is that the ψ_n form an "unconditional basis" for L^p. In particular for $f \in L^p$ the sum $\sum a_n(f)\psi_n$ converges to f in L^p norm, and this is true whatever choice of ordering is used for the basis elements.

Definition 13. *An orthonormal wavelet basis for $L^2(\mathbf{R})$ is an orthonormal basis of the form*

$$\psi_{j,k}(x) = 2^{-j/2}\psi(2^{-j}x - k) \quad j,k \in \mathbf{Z},$$

for some $\psi \in L^2(\mathbf{R})$.

It is not at all obvious that such bases exist, but leaving that aside for the moment we have:

Proposition 14. *If $\psi_{j,k}$ is an orthonormal wavelet basis, where ψ is piecewise differentiable and $|\psi|, |\nabla\psi|$ decay exponentially, then the basis $\psi_{j,k}$ satisfies the conditions of Proposition 12.*

(Here we have to make the obvious change of notation to accomodate the different indexing of the basis elements.)

The proof uses the Calderon-Zygmund theory to show that the multipliers with kernel

$$K(x,y) = \sum \lambda_{j,k}\psi_{j,k}(x)\overline{\psi_{j,k}(y)},$$

are bounded on L^p.

7. Construction of orthonormal wavelets.

If f is a function on $\mathbf{R}$ we define $\tau(f)(x) = f(x-1)$ and $E(f)(x) = 1/\sqrt{2}f(x/2)$, so τ, E are isometries of $L^2(\mathbf{R})$. We write ℓ^2 for the space of square summable sequences (a_n) indexed by the integers, and ℓ^2_0 for the subset of sequences which are zero for all but finitely many indices.

Definition 15. *A multiresolution analysis of $L^2(\mathbf{R})$ is a chain of closed subspaces*

$$\dots V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \dots,$$

such that

(1) $\bigcap_j V_j = 0$;
(2) $\overline{\bigcup_j V_j} = L^2(\mathbf{R})$;
(3) $V_n = E^n(V_0)$;
(4) τ *maps* V_0 *onto itself.*

We also assume that there is a $\rho \in V_0$ such that the map $A_\rho : \ell^2_0(\mathbf{Z}) \to V_0$ defined by

$$A_\rho((a_i)) = \sum a_i \tau^i(\rho),$$

extends to a topological isomorphism from ℓ^2 to V_0 (i.e. is bounded and has a two-sided, bounded, inverse).

Theorem 16. *Given a multiresolution analysis as above there is an associated orthonormal wavelet basis $\psi_{j,k}$ such that V_0 is the closure of the span of the $\psi_{j,k}$ with $j > 0$.*

Example 1. Let V_0 consist of L^2 step functions, constant on all integer intervals $[n, n+1)$. Here we can take ρ to be the characteristic function of the interval $[0,1)$. It is clear that A_ρ is an isomorphism, in fact an isometry. In this case the wavelet function ψ is supported in $[0,1]$; equal to 1 on the interval $[0,1/2)$ and to -1 on $[1/2,1)$. The system $\psi_{j,k}$ is the *Haar basis* of $L^2(\mathbf{R})$.

Example 2. Let V_0 consist of L^2 piecewise linear functions, smooth in each interval $(n, n+1)$. We can take ρ to be the function supported in $[0,2]$ with

$\rho(1) = 1$: we shall see later that A_ρ is an isomorphism. In this case the function ψ is piecewise linear and with exponential decay, so satisfies the conditions of Theorem 14.

Remark. Let $\chi_{j,k}$ be 2^{-j} times the characteristic function of the interval $[2^{-j}k, 2^{-j}(k+1)]$. These are the squares of the norm of the Haar basis elements from the first example above. Using the techniques of the previous section one can show that for any wavelet system $\psi_{j,k}$ satisfying the conditions of Proposition 14 the L^p norm is equivalent to the norm

$$|||f||| = \| \left(\sum_{j,k} \langle f, \psi_{j,k} \rangle |^2 \chi_{j,k} \right)^{1/2} \|_{L^p}.$$

This gives an "almost explicit" description of the L^p norm of a function in terms of its wavelet co-efficients $\langle f, \psi_{j,k} \rangle$.

One can extend these examples to consider, for general r, the space of C^{r-1} functions given by polynomial of degree at most r on each interval $[n, n+1)$. Then one gets C^{r-1} wavelets with exponential decay. There are also many other examples, which you will find in the literature.

Theorem 16: Abstract Proof

We assume V_j, ρ is a multiresolution analysis as above. We begin with

Lemma 17. *There exists $\Phi \in V_0$ such that the map $A_\Phi : \ell^2 \to V_0$ is an isometry.*

Proof. Let $A = A_\rho$. The fact that this is a topological isomorphism means that A^*A is a positive self-adjoint operator on ℓ^2, bounded above and below: $C^{-1} \leq A^*A \leq C$. Hence, by the spectral theorem, we can define $(A^*A)^{-1/2}$ and so $B = A(A^*A)^{-1/2} : \ell^2 \to V_0$, which is an isometry. The shift map on ℓ^2 commutes with A^*A and hence with $(A^*A)^{-1/2}$: this means that B interwtines the shift map on ℓ^2 with the translations on V_0, so $B = A_\Phi$ for some (unique) Φ. (We shall see later how to do this explicitly, so you don't need to know the spectral theorem.)

Now let W_1 be the orthogonal complement of V_1 in V_0, so $V_0 = V_1 \oplus W_1$. The translation τ acts on V_0 and τ^2 maps V_1 isomorphcally to itself, so τ^2 also maps W_1 isomorphically to itself.

Proposition 18. *There is a $\psi_1 \in W_1$ such that the translates $\tau^{2k}(\psi_1)$, for $k \in \mathbf{Z}$, form an orthonormal basis of W_1.*

Given this Proposition, the proof of Theorem 16 is immediate. We let $W_j = E^{j-1}(W_1)$ so there is an orthogonal direct sum decomposition

$$L^2(\mathbf{R}) = \bigoplus_{j=-\infty}^{\infty} W_j \qquad (**)$$

Applying E^{-1} we see that $\psi = E^{-1}(\psi_1)$ has the property that the translates $\tau^k(\psi)$ form an orthonormal basis for W_0, and the wavelet property is a restatement of this and (**) above.

Proof of Proposition 18

Consider the Hilbert space $H = L^2(S^1) \oplus L^2(S^1)$. An element of H is a pair of functions (g_1, g_2) on the circle. We define a complex antilinear map $J : H \to H$ by

$$J(g_1, g_2) = (-\overline{g_2}, \overline{g_1}).$$

(This will be familiar if you have encountered the quaternions.) Given an element $\underline{g} = (g_1, g_2)$ of H we define a map

$$m_{\underline{g}} : C^0(S^1) \to H$$

by $m_{\underline{g}}(f) = f\underline{g} = (fg_1, fg_1)$.

Lemma 19. *Suppose $\underline{g} \in H$ has the property that $\|m_{\underline{g}}(f)\|_{L^2} = \|f\|_{L^2}$ for all $f \in C^0$. Then $|g_1(\theta)|^2 + |g_2(\theta)|^2 = 1$ for almost all $\theta \in S^1$ and $m_{\underline{g}}$ extends to an isometry from $L^2(S^1$ to a closed subspace $U_{\underline{g}}$ of H. The orthogonal complement $U_{\underline{g}}^{\perp}$ is the image $U_{J(\underline{g})}$ of the similar map defined by $J(\underline{g})$.*

Notice that the hypothesis of Lemma 19 is equivalent to saying that the multiples $m_{\underline{g}}(e^{in\theta})$ form an orthnormal system. The proof of the Lemma is immediate if one knows that an L^1 function on the circle all of whose Fourier co-efficients are zero is zero almost everywhere.

Now to complete the proof of Proposition 18, consider the isometry R from ℓ^2 to H which maps a sequence (a_N) to

$$R((a_n)) = \frac{1}{\sqrt{2\pi}}(\sum_m a_{2m}e^{im\theta}, \sum_m a_{2m+1}e^{im\theta}).$$

The composite $R_\Phi = R \circ A_\Phi^{-1}$ is an isometry from V_0 to H which intertwines the translation τ^2 on V_0 and the multiplication map $\sigma : H \to H$ given by $\sigma(\underline{g}) = e^{i\theta}\underline{g}$. Let $U \subset H$ be the closed subspace $R_\Phi(V_1)$. Now we know that the translates $\tau^{2k}(E(\Phi))$ form an orthonormal basis of V_1. So if we set $\underline{g} = R_\Phi(E(\phi))$ the images $\sigma^k(\underline{g})$ form an orthonormal basis of U. Thus $\underline{g}$ satisfies the hypothesis of Lemma 19, and $U = U_{\underline{g}}$. So we know that $U^\perp \subset H$ is $U_{J(\underline{g})}$. Going backwards, the images $\sigma^k(J(\underline{g}))$ form an orthnormal basis for $U^\perp$, hence the translates $\tau^{2k}(R_\Phi^{-1}(J(\underline{g}))$ form an orthonormal basis of W_1: thus we can take $\psi_1 = R_\Phi^{-1}(J(\underline{g}))$, and we have completed the proof of Theorem 16.

Theorem 16:explicit construction

It is easier to work in the Fourier transformation representation. So we define $\hat{V}_j$ to be the FT of V_j. The map τ goes over to $\hat{\tau}(f) = e^{i\xi}f$, and E to $\hat{E}$, which is just E^{-1}. The composite $\hat{A}_\rho$ of A_ρ with $(2\pi)^{-1/2}$ times the Fourier transfrom maps a sequence (a_n) in ℓ^2 to the function

$$\hat{A}_\rho((a_n))(\xi) = (2\pi)^{-1/2} \sum a_n e^{in\xi} \hat{\rho}(\xi).$$

Let $S : \ell^2 \to L^2(S^1)$ be the standard isometry

$$S((a_n)) = (2\pi)^{-1/2} \sum a_n e^{in\theta}.$$

Then the map $\alpha_\rho = \hat{A}_\rho \circ S^{-1}$ sends a function $f \in L^2(S^1)$ to

$$\alpha_\rho(f)(\xi) = f(\xi)\hat{\rho}(\xi).$$

Here we are regarding a function f on the circle S^1 as a 2π-periodic function on $\mathbf{R}$. The adjoint α_ρ^* maps a function $h(\xi)$ on $\mathbf{R}$ to

$$\sum_{k=-\infty}^{\infty} \overline{\hat{\rho}(\xi + 2k\pi)}.h(\xi + 2k\pi).$$

Thus $\alpha_\rho^* \alpha_\rho$ is the map which mutiplies a function f on the circle by

$$\sum_{k=-\infty}^{\infty} |\hat{\rho}(\xi + 2k\pi)|^2.$$

This gives

Proposition 20. *For $\rho \in V_0$ the map $A_\rho : \ell_0^2 \to V_0$ extends to a topological isomorphism if and only if the translates $\tau^k(\rho)$ span a dense subspace of V_0 and there is a $C > 0$ such that*

$$C^{-1} \leq \sum_{k=-\infty}^{\infty} |\hat{\rho}(\xi + 2k\pi)|^2 \leq C$$

for all $\xi \in \mathbf{R}$.

When the above condition holds the operator $(\alpha_\rho^* \alpha_\rho)^{-1/2}$ is multiplication by

$$\left(\sum_{k=\infty}^{\infty} |\hat{\rho}(\xi + 2k\pi)|^2 \right)^{-1/2}.$$

It follows that the Fourier transform of Φ is

$$(*) \qquad \hat{\Phi}(\xi) = \left(\sum_{k=-\infty}^{\infty} |\hat{\rho}(\xi + 2k\pi)|^2 \right)^{-1/2} \hat{\rho}(\xi).$$

Now the general element of $\hat{V}_0$ can be written as

$$m(\xi)\hat{\Phi}(\xi),$$

where m is a function in $L^2(S^1)$, regarded as a 2π-periodic function on $\mathbf{R}$. Let $\hat{R}_\Phi$ be the composite of the inverse Fourier transform with the isometry R_Φ above, so $\hat{R}_\Phi$ is an isometry from $\hat{V}_0$ to the Hilbert space H. Let $J_\Phi : \hat{V} \to \hat{V}$ be the antilinear map $J_\rho = \hat{R}_\Phi^{-1} J \hat{R}_\Phi$. Thus J_Φ coresponds to the map J on H under the isometry $\hat{R}_\Phi$.

Lemma 21. *The map* J_Φ *takes an element* $m(\xi)\hat{\Phi}(\xi)$ *of* $\hat{V}_0$ *to* $\overline{m(\xi+\pi)}\hat{\Phi}(\xi)e^{i\xi}$.

This is a matter of checking on the generators $e^{in\xi}\hat{\Phi}(\xi)$.

We are now able to write down the FT of the wavelet. By hypothesis $\hat{E}(\hat{\Phi}) = 2^{1/2}\hat{\Phi}(2\xi)$ lies in $\hat{V}_0$, so

$$M(\xi) = 2^{1/2}\frac{\hat{\Phi}(2\xi)}{\hat{\Phi}(\xi)}$$

is a 2π-periodic function. The FT of ψ_1 is

$$J_\rho(M)\hat{\Phi} = \overline{M(\xi+\pi)}e^{i\xi}\hat{\Phi}(\xi).$$

Thus

$$\hat{\psi}(\xi) = \overline{\left(\frac{\hat{\Phi}(\xi+2\pi)}{\hat{\Phi}(\xi/2+\pi)}\right)}e^{i\xi/2}\ \hat{\Phi}(\xi/2),$$

where $\hat{\Phi}$ is given by (*).

It is also useful to have the formulae written out making less use of the Fourier transform. We have

(***) $$\sum_{k=-\infty}^{\infty} |\hat{\rho}(\xi+2k\pi)|^2 = \sum_n b_n e^{in\xi},$$

where the b_n are the L^2-inner products

$$b_n = \langle \tau^n \rho, \rho \rangle.$$

The function ψ_1 is

$$\psi_1 = \sum (-1)^n h_{1-n} \tau^n(\Phi),$$

where $h_n = \langle E(\Phi), \tau^n \Phi \rangle$.

Example

We return to the second example above, starting with the multiresolution analysis by piecewise linear functions. First it is simple linear algebra to show that any compactly supported function in V_0 can be written as a finite linear combination of the translates $\tau^k(\rho)$, so these translates span a dense subspace of V_0. The Fourier transform of ρ is

$$\hat{\rho}(\xi) = \frac{(1-\cos\xi)}{\pi\xi^2}e^{i\xi}.$$

The formula (***) above gives

$$\sum_{k=-\infty}^{\infty} |\hat{\rho}(\xi+2k\pi)|^2 = \frac{1}{3}(2+\cos\xi),$$

which lies between 1 and 1/3, so the criteria of Prop.20 are satisfied. The function Φ has Fourier transform

$$\hat{\Phi}(\xi) = \frac{1-\cos\xi}{\pi\xi^2} e^{i\xi} \frac{\sqrt{3}}{\sqrt{2+\cos\xi}},$$

and the Fourier transform of the wavelet ψ is

$$\hat{\psi}(\xi) = \frac{\sqrt{3}}{\pi} \frac{(1-\cos\xi/2)^2}{\xi^2} e^{i\xi/2} \sqrt{\frac{2-\cos\xi/2}{(2+\cos\xi)(2+\cos\xi/2)}}.$$

Notes

(1) The FT of ψ vanishes at $\xi = 0$, so the integral of ψ is zero. In fact this is true for any wavelet system. It may seem paradoxical that any function $f \in L^1 \cap L^2$ can be written as a superposition

$$f = \sum_{j,k} a_{j,k} \psi_{j,k},$$

of functions of integral zero. The point is that the expansion does not converge in L^l, although it does converge in L^p for any $1 < p \leq 2$, as we have seen.

(2) The FT $\hat{\psi}$ extends to a holomorphic function on a strip $|\mathrm{Im}(\xi)| \leq \epsilon$. It follows that ψ does indeed have exponential decay—like $e^{-\epsilon|x|}$—as $|x| \to \infty$.

(3) We can use a Taylor expansion of the square roots to write

$$\hat{\psi}(\xi) = \xi^{-2}(\sum \alpha_n e^{in\xi/2}),$$

where the co-efficients α_n are given by explicit convergent series. From this it is standard Fourier theory to obtain the values of the function ψ explicitly, in terms of the co-efficients α_n.