Lecture Notes For Mathematics 556

Jeffrey RAUCH

Fall 1994

Contents

Mathematics 556 Syllabus

Text. G. Folland, *Fourier Analysis and its Applications*, Wadworth & Brooks/Cole

Prerequisites. Advanced Calculus and Linear Algebra.

1. Finite dimensional scalar product spaces.

(a) Scalar product spaces. [H]
  (1) Orthonormal bases .
  (2) Least squares and orthogonal projection
  (3) Diagonalizability of symmetric matrices.
(b) Diffusion among a finite number of containers. [H]
  (1) Two containers.
  (2) General Diffusion Matrix, $\dot{u} = Au + h$.
    (a) Nullspace.
    (b) Physical spectral heuristics.
  (3) Diffusion on a circular array.
    (a) Eigenfunctions and eigenvectors.
    (b) Second difference operator.
(c) Perturbation Theory for $\dot{u} = A(\epsilon)u + h$.
  (a) Neglect quadratic terms approach.
  (b) Formal series expansion approach.
  (c) Justification via differentiable and analytic dependence on parameters.
(d) Variational properties of eigenvalues.
  (a) Application to discrete diffusion.
(e) Oscillatory heat sources and steady states, $(i\omega - A)w = h$.
  (a) Perturbation Theory for $(i\omega - A(\epsilon))w = h$. Homework.
    (a) Neglect quadratic terms approach.
    (b) Formal series expansion approach.
    (c) Justification by inplicit function theorem.
(f) Perturbation of spectrum.
  (a) Neglect quadratic terms approach.
  (b) Formal series expansion approach.
  (c) Justification via implicit function theorem.
  (d) Application to discrete diffusion.

2. Fourier analysis in $\mathbf{C}^d$ and $C^r(S^1)$. [F Ch. 1,2,3,4], [Strang 290-308], Handout, [Dym-Mckean 60-72].

(a) Trigonometric interpolation. [H]

(1) The formal continuum limit.

(2) Fourier Series.

(b) Fourier series on $C^r(S^1) \quad r = 1, 2, 3...$

(c) L^2 as a completion and Riesz-Fischer Theorems. [H]

(d) Dirichlet kernel: $D_N * f \to f$.

(e) Heat equation with periodic data.

(a) Dissipation of L^2 norm and uniqueness.

(b) Convergence to equilibrium.

(c) Smoothing.

(d) Irreversibility.

(e) Convergence to steady states and the Dirichlet problem.

(f) Wave equation with periodic data.

(a) Travelling waves.

(b) Energy of wave motion on S^1 as sum of oscillator energies.

(g) Exploiting symmetry.

(a) Method of Images for Dirichlet and Neumann problem to reduce to periodic case.

(b) Fourier sine and cosine series.

3. The Fourier Transform F. Ch. 7], [Dym-Mckean pp. 86-91], Handout.

(a) Derivation of the Fourier Transform on Schwartz space, $\mathcal{S}$, from Fourier Series. [H]

(b) The L^2 theory by completion.

4. Fundamental solutions and Green's functions [F, ch 7,9,10]

(a) Solution for the heat equation with data in $\mathcal{S}$: $u(x,t) = K_t * f$ [DM 107-108, Stak 180].

(a) Summability kernels.

(a) Prove that $\lim_{t\to 0} K_t * f = f$ uniformly.

(b) State validity for general f.

(b) Heat and wave equations on $[0, \infty[$ [DM 107] using images.

(b) Introduction to distributions. [Stakgold 86-106, Folland §9.1-9.2, H].

(a) Definition

(b) Examples. functions, Dirac delta, layer potentials.

(c) Multiplication by a smooth and differentiation.

(d) Convergence.

(d) Convolution [H]. $K_t * f \to f$, as approximations to Dirac delta.

(c) Green's Functions for ODE's. [Folland §10.1].

(a) Delta function sources.

(b) Initial value problems for d^j/dt^j.

(a) Taylor's Theorem.

(c) Boundary Value Problems.

(a) Explicit examples.

(a) Analogy between GF and inverse matrix.

(b) Sign of GF and physical interpretation.

(b) Green's Fn. exists *iff* 0 not eigenvalue.

(c) Eigenvalues of regular differential operator are discrete.

(d) Introduce Laplace equation.

(a) Via the large time asymptotics of heat flow.

(b) Via electrostatic potential.

(c) Via velocity potential in fluid mechanics.

(e) The Newtonian potential.

(a) $\Gamma(x)$ in R^n for $n = 1, 2, 3$ from spherically symmetric harmonic functions.

(b) Compute $\Delta\,\Gamma$ in $\mathcal{D}'$. by cutting out the ε- ball [Stakgold 73-76, 102-103].

(c) $\Delta\Gamma * f = f$, for $f \in \mathcal{S}$.

(d) Poisson's equation *iff* inverse square law.

(f) Method of images.

(a) Poisson kernels for the upper half plane.

(a) Comparison for Dirichlet and Neumann conditions.

(b) Poisson kernels for the disk and the ball. Homework. [CH II 264-268, 377],

(g) Integral operator **K** defined by Green's Kernel.

(a) Reduction of boundary value problems for ODE to integral equations.

(a) Eigenvalue problems for ODE $\rightarrow$ eigenvalue problem for integral operator.

(c) Smoothing, $\mathbf{K} : C([a, b]) \rightarrow C^2([a, b])$.

4. Hilbert space. [F. Ch. 3], [Stakgold 275-333]]

(a) Bounded Linear Operators.

(a) Matrices in finite dimension.

(b) Multiplication operators.

(c) Orthogonal projections on finite diml. subspace.

(d) **K**. Hilbert-Schmidt estimate for the norm.

(b) Adjoint of a bounded operator.

(a) Adjoints of above.

(c) Neumann Series and Perturbation Theory.

(d) Differential operators are unbounded.

(e) Formal adjoint of a differential operator.

(a) Formally selfajoint boundary value problems.

(a) Symmetry of Green's Function.

(f) The spectrum of an bounded operator.

(a) Symmetric operators without eigenfunctions.

(b) Noninvertible operator with trivial nullspace.

5. Self-adjoint compact operators in Hilbert space. [Stakgold 335-346]

(a) Compact operators. Definition and examples.

(b) State spectral theorem for compact selfadjoint opoerators.

(c) Variational properties of eigenfunctions.

(j) Perturbation of eigenvalues from the implicit function theorem.

(d) Fredholm alternative

(e) Fourier series revisited: Eigenfunctions of D_θ^2.

(f) Regular Sturm Liouville theory: work it out for Dirichlet and/or Neumann B.C.'s but state it generally. [Stak 411-438]

(g) Application of Fredholm Alternative: perturbation theory of a simple eigenvalue

6. Asymptotic Methods.

(a) Stationary and non-stationary phase: the formal idea. [Whitham 371-374]

(b) Applications: contrast long time behavior of Klein-Gordon and wave equations. [Handout]

References.

R. Courant and D. Hilbert, Methods of Mathematical Physics vol.I and II, John Wiley, 1963 and 1966.

H. Dym and H.P. McKean - Fourier Series and Integrals, Academic Press 1972.

G. Hellwig, Differential Operators of Mathematica Physics, Addison Wesley, 1964

T. Kato: A Short Introduction to Perturbation Theory of Linear Operators, Springer -Verlag

M. Reed and B. Simon, Modern Methods in Mathematical Physics Volumes I and II,

A. Sommerfeld, Partial Differential Equations of Mathematical Physics,

I. Stakgold, Green's Functions and Boundary Value Problems, John Wiley and Sons, 1979.

G. Strang, Introduction to Applied Mathematics, Wellesley Cambridge Press 1986.

G. Whitham, Linear and Nonlinear Waves, John Wiley 1974.

H. Widom - Integral Equations,

Additional Topics.

The following topics were not listed for lack of time. They may be considered as recommendations for supplementary material and also as topics suitable for Math 557.

a) Gibbs Phenomenon.

b) Examples of with discrete and continuous spectrum (e.g. $\sqrt{-1}D_x$, square well)

c) Van der Corput's lemma (statement only) and the idea of chopping up k-space into regions of stationary and nonstationary phase

d) dispersion for the KdV, and Schrödinger equations

e) The method of steepest descent and Laplace integrals.

f) Analysis of the stability of nonlinear motion by linearization followed by spectral analysis.

§1. Scalar Products and Symmetric Matrices

The vector space $\mathbf{R}^d$ with the Euclidean scalar product

$$\langle x, y \rangle := \sum x_j \, y_j \tag{1}$$

is the basic example of a scalar product space. The Euclidean geometries $d = 1, 2, 3$ are the foundation for our geometric intuition. The complex analogue is $\mathbf{C}^d$ with scalar product

$$\langle x, y \rangle := \sum x_j \, \overline{y}_j \, . \tag{2}$$

Definition. A *scalar product space* is a vector space V together with a mapping $v, w \in V \times V \to \langle v_1, v_2 \rangle \in \mathbf{C}$ such that

$$\langle v, v \rangle \geq 0 \, , \tag{3}$$

$$\langle v, v \rangle = 0 \ \text{ if and only if } \ v = 0 \, , \tag{4}$$

$$\langle v, w \rangle = \overline{\langle w, v \rangle} \, , \tag{5}$$

$$\langle \alpha_1 v_1 + \alpha_2 v_2, w \rangle = \alpha_1 \langle v_1, w \rangle + \alpha_2 \langle v_2, w \rangle \, . \tag{6}$$

In a scalar product space the norm or length of a vector is defined to be $||v|| := \langle v, v \rangle^{1/2}$. This is possible thanks to 1.4 and the only vector of length equal to zero is the zero vector. More subtle is that fact that the triangle inequality

$$||v + w|| \leq ||v|| + ||w|| \tag{7}$$

is satisfied. It is a consequence of the important *Cauchy-Schwarz Inequality*

$$|\langle v, w \rangle| \leq ||v|| \, ||w||. \tag{8}$$

The elegant proof of inequality 1.8, begins with the observation that

$$0 \leq ||v + tw||^2 = \langle v + tw, v + tw \rangle = ||v||^2 + 2t \Re \langle v, w \rangle + t^2 ||w||^2 \, . \tag{9}$$

There is nothing to prove if either v or w is equal to zero. When they are nonzero, the quadratic polynomial in t is minimized when $t = -\Re \langle v, w \rangle / ||w||^2$. Plugging in this value of t and multiplying the resulting expression by $||w||^2$ yields

$$(\Re \langle v, w \rangle)^2 \leq ||v||^2 ||w||^2 \, . \tag{10}$$

If $\arg \langle v, w \rangle = \theta$, apply this inequality to $\tilde{v} := e^{-i\theta} v$ and w to prove 1.8.

A one line proof of 1.7 comes from folding the entire computation above into the identity

$$||v||^2 ||w||^2 - |\langle v, w \rangle| ||v|| \, ||w|| = \langle ||w|| e^{-i\theta} v - ||v|| w, ||w|| e^{-i\theta} v - ||v|| w \rangle \geq 0 \, .$$

The first proof of the Cauchy-Schwarz inequality gives a recipe for computing scalar products from distances

$$\Re \langle v, w \rangle = \frac{1}{2} \frac{d}{dt} \bigg|_{t=0} ||v + tw||^2 \, . \tag{11}$$

The imaginary part is then found from

$$\Im\langle v, w\rangle = \Re\langle -iv, w\rangle = \frac{1}{2}\frac{d}{dt}\bigg|_{t=0} ||-iv+tw||^2 . \tag{12}$$

Another way to compute scalar products from norms are the *polarization identities*. The first of these is

$$2\Re\langle v, w\rangle = ||v+w||^2 - ||v-w||^2 \tag{13}$$

which is sufficient in itself for real scalar product spaces. In the complex case one adds a similar expression for the imaginary part to get

$$4\langle v, w\rangle = ||v+w||^2 - ||v-w||^2 + i||v+iw||^2 - i||v-iw||^2 . \tag{14}$$

The derivation of 1.7 is given on page 64 of Folland. A consequence is that defining

$$\text{dist}\,(v, w) := ||v-w|| \tag{15}$$

yields a metric on the scalar product space. This allows one to talk of balls, distances, and continuity in any scalar product space.

Here are two more subtle examples of scalar product spaces, both of infinite dimension. The first is the space of continuous functions $C([a,b])$ with scalar product

$$\langle f, g\rangle := \int_a^b f(x)\overline{g(x)}\,dx . \tag{16}$$

Second is the space of infinite sequences $(x_1, x_2, \ldots)$ with at most a finite number of nonzero entries with scalar product

$$\langle x, y\rangle := \sum x_j\,\overline{y}_j . \tag{17}$$

Many applied problems come with natural scalar products which need have little to do with the Euclidean scalar product and are sometimes not obvious on first view. Finding the appropriate algebraic structure is an important step in solving many problems.

Definition. Two vectors v and w in a scalar product space are *orthogonal* if and only if $\langle v, w\rangle = 0$. A set of vectors is orthogonal if any two vectors in the set are orthogonal. Orthogonal families of vectors of unit length are called *orthonormal.*

Proposition 1. *A family of nonzero othogonal vectors is linearly independent.*

Proof. If v_j is a finite set of nonzero orthogonal vectors and $\sum a_j\,v_j = 0$, then taking the scalar product with v_k yields $a_k||v_k||^2 = 0$. Since the v_j are nonzero, this shows that $a_k = 0$. $\diamondsuit$

Proposition 2. Pythagorean Theorem. *If v_j is a finite set of othogonal vectors then*

$$||\sum v_j\,||^2 = \sum_j ||v_j||^2 . \tag{18}$$

Proof. Folland page 65. $\diamondsuit$

There is a basic computation with orthonormal families which is crucially important.

Problem. Given a finite orthonormal family $\{e_k\}$ and a vector v, find a vector w in the $W :=$ Span $\{e_k\}$ such that $v - w$ is orthogonal to W. **Discussion.** The vector w is called the *projection* of v on W.

Solution. Seek $w = \sum a_k\, e_k$ such that $\langle w, e_k\rangle = 0$ for all k. Take the scalar product of the expression for w with e_m. All the summands vanish except for $k = m$. Thus a necessary condition is $\langle w, e_m\rangle = a_m$. Conversely, the vector $\sum\langle w, e_k\rangle e_k$ has the dersired properties. $\diamondsuit$

A first application of this is the *Gramm-Schmidt process.* Given is a finite linearly independent $\{w_j\}$ set, an orthonormal family of $\{e_j\}$ with the same span is constructed as follws. Set $e_1 = w_1/||w_1||$. If $j > 1$ and one has constructed e_m for $m \leq j-1$,

$$e_j := \frac{w_j - \sum_{k=1}^{j-1}\langle w_j, e_k\rangle\, e_k}{||w_j - \sum_{k=1}^{j-1}\langle w_j, e_k\rangle\, e_k||}\,. \tag{19}$$

The numerator in 19 is the unique vector in Span$(\{e_j\})$ which has the same scalar product as w with each of the vectors e_j.

Any finite dimensional subspace of a scalar product space has a basis and therefore, by the Gramm-Schmidt process, an orthonormal basis.

Prosposition 3. Projection on Finite Dimensional Subspaces. *If W is a finite dimensional subspace of a scalar product space V and $v \in V$, then there is a unique vector $w \in W$ which is closest to v. This vector is characterized by the fact that $w \in W$ and $v - w$ is orthogonal to all vectors in W. If e_j is an othonormal basis for W, then $w = \sum\langle v, e_j\rangle e_j$.*

Proof. Set $\underline{w} := \sum\langle v, e_j\rangle e_j$ and $u := v - \underline{w}$. Then u is orthogonal to each vector e_j and therefore to any vector in W. For any $w \in W$, $v - w = (v - \underline{w}) + (\underline{w} - w)$. These are othogonal so by the Pythagorean Theorem

$$||v - w||^2 = ||v - \underline{w}||^2 + ||w - \underline{w}||^2\,.$$

The first term on the right is independent of w and the second is is nonnegative and vanishes only when $w = \underline{w}$. Thus the left hand side has a unique minimum when $w = \underline{w}$.

On the other hand, if $w \in W$ is such that $v - w$ is orthogonal to W then

$$v = u + \underline{w} = (v - w) + w$$

so $w - \underline{w} = u - (v - w)$. The right hand side is orthogonal to W so taking the scalar product with $w - \underline{w} \in W$ yields $||w - \underline{w}||^2 = 0$ so $w = \underline{w}$ proving the characterization in the Proposition. $\diamondsuit$

This proposition is one of the few simple results yielding closest elements or equivalently best approximations to an arbitrary vector by an element of a given subspace. It is the basis of many *least squares approximations* in applications.

A particularly useful case of the last identity in the proposition occurs when $v \in W$ in which case one has $v = \sum \langle v, e_j \rangle e_j$. When $V = W$ this holds for all $v \in V$.

Definition. For a linear transformation $A : V \to V$ of a finite dimensional vector space to itself, the *spectrum* of A is the set of scalars λ such that $A - \lambda I$ is not invertible. The complement of the spectrum is called the *resolvent set.* The function $(A - \lambda I)^{-1}$ defined on the resolvent set, is called the *resolvent.*

A point λ belongs to the spectrum if and only if there is a nonzero vector v satisfying $Av = \lambda v$. Such a vector is called an *eigenvector* and λ is called the corresponding *eigenvalue.* Alternatively, the points in the spectrum are the solutions of $\det(A - \lambda I) = 0$. This is a polynomial equation

$$0 = \det(\lambda I - A) = \lambda^d + \text{ lower order terms} \tag{20}$$

of degree exactly equal to $d := \dim V$. Thus there are no more than d points in the spectrum. As a function on the resolvent set, Cramer's rule expresses $(A - \lambda I)^{-1}$ as a quotient of a matrix valued polynomial of degree $d - 1$ divided by $\det(A - \lambda I)$. Thus the resolvent is a rational function of λ

Definition. A linear transformation $A : V \to V$ from a scalar product space to itself is called *symmetric* if and only if for all v, w in V, $\langle Av, w \rangle = \langle v, Aw \rangle$.

If V is finite dimensional these transformations are exactly those whose matrices A_{ij} with respect to an orthonormal basis satisfy the condition of hermitian symmetry, $A_{ij} = \overline{A}_{ji}$ for all i, j.
If A is symmetric then

$$\langle Av, v \rangle = \langle v, Av \rangle = \overline{\langle Av, v \rangle} \tag{21}$$

so is real.

Proposition 4. *Suppose that $A : V \to V$ is a symmetric linear transformation. All eigenvalues of A are real. Eigenvectors corresponding to distinct eigenvalues are orthogonal. If W is a linear subspace invariant under A then the set of vectors orthogonal to W is also a subspace invariant under A.*

Proof. If v is an eigenvector with eigenvalue λ then

$$\lambda = \langle Av, v \rangle / \langle v, v \rangle \tag{22}$$

is real by the observation in the last paragraph.
If v_1 and v_2 are eigenvectors with distinct eigenvalues λ_j then

$$\lambda_1 \langle v_1, v_2 \rangle = \langle Av_1, v_2 \rangle = \langle v_1, Av_2 \rangle = \lambda_2 \langle v_1, v_2 \rangle \,.$$

Since the eigenvalues are distinct, this implies that $\langle v_1, v_2 \rangle = 0$.
If v is orthogonal to W, then for any $w \in W$

$$\langle Av, w \rangle = \langle v, Aw \rangle = 0$$

since $Aw \in W$ by the invariance of W and $v \perp W$. Therefore Av is perpendicular to W. $\diamondsuit$

Theorem 5. *If A is a symmetric linear transformation on a finite dimensional complex scalar product space V then V is the orthogonal direct sum of the eigenspaces $E_\lambda := \text{kernel}\,(A - \lambda I)$ with λ running over the spectrum of A. In particular, V has an othonormal basis of eigenvectors.*

Proof. The proof is by induction on the dimension $1 \le d$ of the space. If $d = 1$ choose a nonzero $v \in V$. Then $V = \mathbf{C}v$ so $Av = \lambda_1 v$ for some λ_1. Proposition 4 shows that λ_1 is real. Then $A = \lambda_1 I$, $\text{spec}\, A = \{\lambda_1\}$, $E_{\lambda_1} = V$, and the result is true.

Suppose that $d > 1$ and the theorem has been proven for dimensions less than or equal to $d - 1$.

The key step is to show that a symmetric transformation on a d dimensional space has an eigenvalue. The Fundamental Theorem of Algebra implies that the characteristic equation 20 has a root λ_1 which is therefore an eigenvalue of A.

If $E_{\lambda_1} = V$ then $A = \lambda_1 I$ and the result is true as for the case $d = 1$.

Otherwise let $V_1 := E_{\lambda_1}$. Then $V_1^\perp$, be the space of vectors in V orthogonal to E_{λ_1}, is a scalar product space of dimension strictly smaller than d which is invariant under A by Proposition 4. In addition λ_1 is not in the spectrum of the restriction of A to $V_1^\perp$.

The restriction of A to $V_1^\perp$ is symmetric so the inductive hypothesis implies that $V_1^\perp$ is the orthogonal direct sum of eigenspaces of the restriction of A to $V_1^\perp$. Therefore V is the orthogonal direct sum of V_1 and those spaces.

Thus, there is a finite set of distinct eigenvalues $\{\lambda_1, \dots \lambda_k\}$ and an orthogonal direct sum decomposition $V = \oplus V_j$ and on V_j, $A = \lambda_j I$.

In this case $\text{spec}\, A = \{\lambda_1, \dots \lambda_k\}$ and $E_{\lambda_j} = V_j$ and the proof is complete. $\diamondsuit$

The use of the Fundamental Theorem of Algebra to produce an eigenvector can be avoided, see problem 8 of assignment 1.

The treatment above is all algebraic and geometric. The next two results concern continuity so have a topological flavor.

Proposition. *The scalar product is continuous in the sense that if $v_n \to v$ and $w_n \to w$ then*

$$\langle v_n, w_n \rangle \to \langle v, w \rangle\,.$$

Proof. Write

$$\langle v_n, w_n \rangle - \langle v, w \rangle = \langle v_n - v, w_n \rangle + \langle v, w_n - w \rangle\,.$$

Two applications of the Schwarz inequality yield

$$|\langle v_n, w_n \rangle - \langle v, w \rangle| \le ||v_n - v||\, ||w_n|| + ||v||\, ||w - w_n||\,.$$

Use the triangle inequality $||w_n|| \le ||w|| + ||w_n - w||$ to give

$$|\langle v_n, w_n \rangle - \langle v, w \rangle| \le ||v_n - v||\, (||w|| + ||w_n - w||) + ||v||\, ||w - w_n||\,.$$

As $n \to \infty$ the right hand side tends to zero. $\diamondsuit$

Proposition. *If A is a linear transformation on a finite dimensional scalar product space, then A is continuous, that is if $v_n \to v$ then $Av_n \to Av$.*

Proof. For any vector $e \in V$ the previous Proposition implies that

$$\langle Av_n - Av, e\rangle = \langle A(v_n - v), e\rangle = \langle v_n - v, A^* e\rangle \to 0\,. \tag{23}$$

If e_j is an othonormal basis, one has

$$||Av_n - Av||^2 = \sum |\langle Av_n - Av, e_j\rangle|^2 \to 0$$

where (23) is used in the last step. $\diamondsuit$

§2. Discrete Diffusion

The material in these notes serve several purposes. A nontrivial application is modeled and a study begun. This introduces some of the manners of thinking about applied problems, and serves as a source of interesting questions on which to test our mathematical muscle.

Next it will lead to the discrete Fourier transform and from there to the amazing suggestion that any 2π periodic function is a sum of the functions e^{in} with $n \in \mathbf{Z}$. This is the usual starting point of a course such as this and I for one have always found that extremely unmotivated.

Consider two containers of solution of some substance. The containers have volumes V_j, concentrations C_j, and N_j molecules of the solute where in all cases $j = 1, 2$. The containers are supposed to be in contact separated so that the rate of flow of molecules of the solute is proportional to the difference of concentrations. Typically the proportionality constant is the product of the area of contact between the containers and the permeability of the material serparating them. The differential equations describing the diffusion of molecules is then

$$\begin{aligned} \frac{dN_1}{dt} &= -k(C_1 - C_2) \\ \frac{dN_2}{dt} &= -k(C_2 - C_1)\,. \end{aligned} \tag{1}$$

The concentration of solute in conatiner j, measured in molecules per unit volume is given by $C_j = N_j/V_j$. Dividing the j$^{\text{th}}$ equation by V_j yields the system of equations for $C(t) := (C_1(t), C_2(t))$

$$\frac{dC}{dt} = \begin{pmatrix} -k/V_1 & k/V_1 \\ k/V_2 & -k/V_2 \end{pmatrix} C := A\,C\,, \tag{2}$$

defining the 2×2 matrix A. Intuition suggests that solute will flow from the container with higher concentration to the other. When the concentrations are equal there should be no flow. One expects that any initial condition will lead to an evolution tending to that equilibrium.

It is worth remarking that exactly the same equation $dN/dt = AN$ describes the evolution of the particle numbers. One would also find an equation of exactly the same form for the dynamics of the temperatures of containers in contact. Here the heat flow is proportional to temperature difference, and temperature change is proportional to heat flow in with proportionality constant the heat capacity of the material in the container. The capacity is the product of volume and heat capacity per unit volume. Briefly, one finds exactly the same form for the dynamics.

Recall that to solve a constant coefficient linear system of ordinary differential equation, $dC/dt = AC$, one looks for eigenvectors v of A, that is nonzero vectors such that $Av = \lambda v$ for some complex λ. Then $C(t) = e^{\lambda t} v$ is a solution. If there are a basis of eigenvectors then the general solution is a linear combination, with real or complex coefficients, of such exponential solutions.

The system 2 can be solved without setting pen to paper. The equilibrium solution corresponds to the fact that $(1,1)$ is an eigenvector of the matrix A with eigenvalue equal to zero. Since the sum of the diagonal elements (a.k.a. the trace of A) is equal to the sum of the eigenvalues, and, one eigenvalue is zero, it follows that a second eigenvalue is equal to the trace, namely $-k(1/V_1 + 1/V_2)$. The range of A is then spanned by the second eigenvector so reading directly the columns of A gives the eigenvector $(-1/V_1, 1/V_2)$. Thus the general solution of (2) is

$$C(t) = a\,(1,1) + b\,e^{-kt(1/V_1 + 1/V_2)}\,(-1/V_1, 1/V_2) \tag{3}$$

with arbitrary real constants a, b. All solutions decay to equilibrium with charateristic time $V_1V_2/k(V_1+V_2)$. The quantity $V_1C_1 + V_2C_2$ is independent of time, expresssing the conservation of solute.

It is not obvious from the formula, but is not difficult to check, that initially positive concentrations remain positive for all time. To show this it suffices to check that for vectors C on the boundary of the quadrant $\{C : C_1 \geq 0 \text{ and } C_2 \geq 0\}$ the velocity vector AC points toward the interior of the quadrant. For example on the boundary where $C_1 = 0$ and $C_2 > 0$ one has $AC = kC_2(1/V_1, -1/V_2)$ which points inward since its first component is positive. A similar argument works at the boundary points $(C_1, 0)$.

The equations for diffusion among m containers is similar. The flow of solute from container i to j is $k_{ij}(C_i - C_j)$ where $k_{ij} = k_{ji} \geq 0$. The dynamics is then given by

$$\frac{dC_i}{dt} = \sum_{j\neq i} -\frac{k_{ij}}{V_i}(C_i - C_j), \quad 1 \leq i \leq m. \tag{4}$$

In matrix form this is $dC/dt = AC$ with

$$A = \begin{pmatrix} -\sum_{j\neq 1} k_{1j}/V_1 & k_{12}/V_1 & \cdots & k_{1m}/V_1 \\ k_{21}/V_2 & -\sum_{j\neq 2} k_{2j}/V_2 & \cdots & k_{2m}/V_2 \\ \vdots & \vdots & \vdots & \vdots \\ k_{m1}/V_m & k_{m2}/V_m & \cdots & -\sum_{j\neq m} k_{mj}/V_m \end{pmatrix} \tag{5}$$

Note that the sum of the elements in each row is equal to zero. This shows that $(1, 1, \ldots, 1)$ belongs to the kernel of A and equivalently that equal concentrations give stationary solutions. The diagonal elements are negative and the off diagonal elements are positive. The biggest element in each row is the diagonal element.

Since concentrations *must* be positive one must show that arbitrary positive initial data should stay positive. Otherwise the model is unacceptable. Positivity can be proven as in the two by two case, showing that the velocity vector never points outward at points of the boundary of the set of vectors all of whose components are nonnegative.

Suppose that enough of the coefficients k_{ij} are positive so that it is possible for heat to flow from any container to any other by passing through a finite number of others across passages with positive permeability. Then, physical intuition suggests that the only equilibria are those of equal concentrations and that arbitrary initial states converge to equilibrium. Conservation of concentrate suggests that $\sum V_jC_j$ is independent of time.

If these assertions are correct then the kernel of A is equal to $\mathbf{R}(1, 1, \ldots, 1)$, and, all nonzero eigenvalues have strictly negative real part. An interesting question is whether increasing one of the k_{ij} automatically increases the rate of decay to equilibrium, equivalently moves the nonzero eigenvalues to the left. Another interesting question is whether the eigenvalues are all real. The answer to the latter question is yes in case of equal volumes for then the matrix A is symmetric. Studying these questions serves as an interesting source of test problems for us, and a good exposure to the sort of thinking that is common in applied mathematics. Physically reasonable assertions are translated to mathematical assertions whose validity must be studied. True mathematical assertions which were not suggested by intuition lead to new intuition.

Next turn to a particularly simple configuration of containers. Suppose that there are N containers of equal volume arranged in a circular array with connectivity to nearest neighbors only. All connections are supposed equally permeable. This is like a pearl necklace with touching pearls. The common value of k_{ij}/V_j is denoted by k. The matrix A has the form

$$A = k \begin{pmatrix} -2 & 1 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 1 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & & & & & & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 1 & -2 \end{pmatrix} . \tag{6}$$

To find the eigenvalues and eigenvectors of the symmetric matrix A, It suffices to treat the case $k = 1$ which simplifies the notation. We present two distinct strategies for the computations. In these notes we will exploit directly the rotational symmetry of the problem, and in the homework an approach using difference equations is presented. The matrix A plays a role in many different problems so that a thorough analysis is worthwhile. For example, it is the discretization of the one dimensional Laplacian d^2/dx^2 by centered second differences or by the finite element method using piecewise linear continuous elements.

Since A is a symmetric matrix it defines a symmetric linear transformation if we adopt the standard Euclidean scalar product in the m dimensional real space of concentrations $(C_1, \ldots, C_m)$.

The key to analysing the matrix A is to use the following symmetry argument. If you twist the ring of containers one notch counterclockwise the concentration C_j is replaced by C_{j+1}. If $C(t)$ is a solution of the dynamic equation $dC/dt = AC$ then it is quite reasonable to expect that the new rotated concentrations also yield a solution. Introduce some notation to help in the verification.

The rotation operator τ is given by

$$\tau\big((C_1, C_2, \ldots, C_{m-1}, C_m) \big) := (C_2, C_3, \ldots, C_m, C_1) \tag{7}$$

The inverse transformation τ^{-1} twists one notch in the other direction replacing C_j with C_{j-1} and C_1 with C_m. The matrix of τ and τ^{-1} with respect to the standard basis in $\mathbf{R}^N$ are equal to

$$\tau = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \qquad \tau^{-1} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix} . \tag{8}$$

The symmetry argument above suggests that $\tau(C(t))$ satisfies $(d/dt - A)(\tau C) = 0$. Compute

$$(\frac{d}{dt} - A)(\tau C) = \tau \frac{d}{dt} C - A(\tau C) = (\tau A - A\tau)C \,. \tag{9}$$

Since $C(0)$ is an arbitrary vector, the left hand side vanishes for arbitrary solutions if and only if A commutes with τ that is $\tau A - A\tau = 0$. This relation can be checked by direct computation but here is another way to see it. Using formula (8) shows that

$$A = \tau - 2I + \tau^{-1} \tag{10}$$

and therefore commutes with τ since each of the summands in 10 commutes with τ.

Equation (10) shows that if v is an eigenvector of τ with eigenvalue λ then v is an eigenvector of A with eigenvalue $\lambda - 2 + \lambda^{-1}$ However τ is not symmetric at first it is not clear whether to expect that it has many nice eigenvectors. There is a class of matrices more general than symmetric matrices which also have orthonormal bases of eigenvectors. The next section describes this class and applies the results to finish the calculation of the eigenvalues and eigenvectors of A.

Adjoints and Normal Matrices

Proposition 1. *If V is a finite dimensional scalar product space and $B : V \to V$ is a linear transformation, then there is a unique linear transformation, denoted B^* and called the adjoint of B, satisfying*

$$\langle Bv, w\rangle = \langle v, B^* w\rangle \quad \text{for all } v, w \text{ in } V. \tag{1}$$

The matrix of B^* with respect to any orthonormal basis is equal to the conjugate transpose of the matrix of B, that is

$$(B^*)_{ij} = \overline{B_{ji}}. \tag{2}$$

Proof. If $\{e_j\}$ is an orthonormal basis then if (1) holds one has

$$(B^*)_{ij} = \langle B^* e_i, e_j\rangle = \overline{\langle e_j, B^* e_i\rangle} = \overline{\langle Be_j, e_i\rangle} = \overline{B_{ji}}.$$

which proves uniqueness and formula (2). It is left as an exercise for the reader to show that the linear transformation with matrix given by (2) satisfies (1) . ♢

From (1), it follows that the adjoint satisfies

$$(B_1 + B_2)^* = B_1^* + B_2^* \quad \text{and} \quad (B_1 B_2)^* = B_2^* B_1^* \tag{3}$$

for any pair B_1, B_2 of linear transformations. For example to prove the second identity use (1) twice to compute

$$\langle B_1 B_2 v, w\rangle = \langle B_2 v, B_1^* w\rangle = \langle v, B_2^* B_1^* w\rangle \,.$$

The characterization (1) of the adjoint completes the verification of the product rule in (3). An important special case of the product rule for adjoints is $(aB)^* = \overline{a} B^*$. Note the complex conjugate on the scalar a.

The symmetric transformations are exactly the transformations which are equal to their own adjoints. They are sometimes called *selfadjoint*. A transformation B is called *antisymmetric* if $B^* = -B$. They are sometimes called *antiselfadjoint*. A transformation is antisymmetric if and only if it is of the form iA with A symmetric. An arbitrary transformation can be written in exactly one way (assignment 2, problem 1) as the sum of a symmetric plus an antisymmetric transformation, and the expresssion is explicit

$$B = (B + B^*)/2 + (B - B^*)/2\,. \tag{4}$$

This is analogous to writing an arbitrary real number as a sum $x + iy$ with real numbers x, y. Equation (4) writes a matrix as $X + iY$ with symmetric matrices X, Y. The first summand in (4) is called the *symmetric part* and the second the *antisymmetric part* of B.

Definition. A linear transformation B on a finite dimensional scalar product space is called *normal* if it commutes with its adjoint, that is $B^*B = BB^*$.

A transformation is normal if and only if its symmetric and antisymmetric parts commute (assignment 2, problem 2).

Examples. (i.) A symmetric matrix is normal. **(ii.)** If A_1 and A_2 are commuting symmetric matrices then $B := A_1 + iA_2$ is normal. There is an othonormal basis e_j each of whose elements is an eigenvector of both A_1 and A_2. Denote the eigenvalues by λ_{1j} and λ_{2j}. Then e_j is an eigenvector of B with eigenvalue $\lambda_{1j} + i\lambda_{2j}$. In this basis the matrix of B is diagonal

$$B = \begin{pmatrix} \lambda_{11} + i\lambda_{12} & 0 & 0 & \cdots & 0 \\ 0 & \lambda_{12} + i\lambda_{22} & 0 & \cdots & 0 \\ 0 & 0 & \lambda_{13} + i\lambda_{23} & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_{1m} + i\lambda_{2m} \end{pmatrix}$$

(iii.) Since the adjoint of τ is τ^{-1} which commutes with τ, τ is normal. The argument in (ii.) shows that τ has an eigenbasis.

Beware, most transformations are not normal.

Two normal transformation commute if and only if the the family of four symmetric and antisymmetric parts commutes (assignment 2, problem 2). Therefore, the result of assignment 1, problem 6 implies the following important theorem.

Theorem 2. *If* $\mathbf{A}$ *is a finite family of commuting normal transformations, there is an orthonormal basis consisting of vectors which are simultaneously eigenvectors of all the transformations in* $\mathbf{A}$.

Example. This result is used to compute the eigenvalues of A. We have already observed that eigenvectors of τ are automatically eigenvectors of A and, that the normal transformation τ has an orthonormal basis of eigenvectors.

To compute the eigenvectors of τ note that a vector $C \in \mathbf{C}^m$ satisfies $\tau C = \lambda C$ if and only if

$$C_2 = \lambda C_1, \quad C_3 = \lambda C_2, \quad \ldots, \quad C_m = \lambda C_{m-1}, \quad C_m = \lambda C_1 \,. \tag{5}$$

If $\lambda = 0$ it follows that $C = 0$ so the eigenvalues are nonzero. If $C \neq 0$, then running through the cycle of equations (2) starting at a nonzero component C_μ yields $C_\mu = \lambda^m C_\mu$. Thus the eigenvalues are m^{th} roots of unity. Conversely, if $\lambda^m = 1$ then the unit vector $(1, \lambda, \lambda^2, \ldots, \lambda^{(m-1)})/\sqrt{m}$ is an eigenvector. Since the scalar multiples of these m vectors are the only eigenvectors of τ Theorem 4 implies that they form an orthonormal basis. These evaluations prove the following theorem.

Theorem 3. *For* $j = 0, 1, 2, \ldots, m-1$ *let*

$$\lambda_j := e^{i2\pi j/m} \tag{6}$$

and

$$e_j := (1, \lambda_j^1, \lambda_j^2, \ldots, \lambda_j^{(m-1)})/\sqrt{m} \,. \tag{7}$$

Then, the e_j form an orthonormal basis of $\mathbf{C}^m$ and satisfy the eigenvalue equations

$$\tau e_j = \lambda_j e_j \,, \qquad A\, e_j = 2(\cos \frac{2\pi j}{m} - 1\,) e_j \,. \tag{8}$$

Proof. It remains to prove the last equation in (8). Equation (10) shows that e_j is an eigenvalue of A with eigenvalue $\lambda_j - 2 + \lambda_j^{-1}$ which (8). $\diamondsuit$

That e_j is an orthonormal basis is verified by direct computation in Lemma 7.1 on page 251 of Folland.

The eigenvalues of A lie between -2 and 0. For the matrix in equation (6) of the last section, the eigenvalues are bewteen $-2k$ and 0. The eigenvalue 0, has one dimensional eigenspace spanned by the vector $e_0 = (1, 1, \ldots, 1)/\sqrt{m}$. It follows that arbitrary solutions decay to a state with equal concentrations. The nonequilibrium part of the solution is a sum of terms decaying like $e^{k\lambda_j t}$ with $j > 0$. The slowest to decay are those corresponding to $j = 1, m-1$ and the rate of decay is given by

$$e^{2kt(\cos \frac{2\pi}{m} - 1\,)} \,.$$

Note that the rate of decay is proportional to $k = k_{ij}/V$. Increasing the permeability or decreasing the size of the cells increases the speed at which equilibrium is approached. The more subtle question of whether increasing one of the permeabilities speeds equilibration, is not answered by this calculation.

§3. Lines and Circles, Discrete and Otherwise

The flow of solute or heat among a circular array of equal containers has produced for us an eigenbasis for $\mathbf{C}^m$. Heat flow in a circular thin wire can be approximated by replacing the wire by m equal arcs, and heat transfer from one arc to the next yields exactly the problem solved by the basis (17). For m large one expects a good approximation to heat flow in the ring. That is the theme of the current section.

Consider the circle as the set of complex numbers of modulus equal to 1,

$$S^1 := \{z \in \mathbf{C} \,:\, |z| = 1\} \tag{1}$$

This is a commutative group with complex multiplication as the operation. The exponential function

$$\theta \in \mathbf{R} \to e^{i\theta} \tag{2}$$

is a group homomorphism from the additive group of reals to the circle group. The kernel of this homomorphism is equal to subgroup $2\pi\mathbf{Z}$, so as groups, $S^1 = \mathbf{R}/2\pi\mathbf{Z}$.[1] Each point in S^1 is the image of an infinite set of points in $\mathbf{R}$, the points differing by integer multiples of 2π.

If Y is any set and $g : S^1 \to Y$, then composition with the exponential function yields a function $f(\theta) := g(e^{i\theta})$ from $\mathbf{R}$ to Y. The function f is 2π periodic, that is

$$f(\theta + 2\pi) = f(\theta) \quad \text{for all} \quad \theta \in \mathbf{R}\,. \tag{3}$$

Conversely, any such periodic function f is induced by a unique g. In this way one can view periodic functions on the reals as functions on S^1 and *vice versa.*

In the circle group S^1 the elementary m^{th} root of unity $\omega = e^{2\pi i/m}$ generates a subgroup isomorphic to the cyclic group with m elements. It consists of the points $\omega^\nu = e^{2\pi i\nu/m}$ with $0 \le \nu \le m-1$. This copy of the cyclic group will be called the *discrete circle.* Arcs on the discrete circle between adjacent points have length given by $h = 2\pi/m$. This discrete circle with m elements is denoted S^1_h and h is called the *meshwidth.* It is the image under the mapping (2) of the *discrete line* $h\mathbf{Z}$. The discrete line is an additive group. The spacing between adjacent points is equal to h. The discrete line is denoted $\mathbf{R}_h$.

By composition with the exponential function, a function $S^1_h \to Y$ induces a 2π periodic function $\mathbf{R}_h \to Y$. Conversely a 2π periodic function from the discrete line to Y is induced by exactly one function on the discrete circle. In this way there is a one to one correspondence between 2π periodic complex functions on $\mathbf{R}_h$ and the complex functions on S^1_h. This correspondence is an invertible linear transformation identifying the two spaces.

A function on S^1_h is determined by its values at the m points of the discrete circle. A 2π periodic function on the discrete reals is determined by its values at m consecutive points, for example the points jh_m for $0 \le j \le m-1$. These are m equally spaced points between 0 and 2π. The space of complex functions on S^1_h and the space of 2π periodic functions on $\mathbf{R}_h$ are vector spaces with dimension equal to m.

There is a natural scalar product on the space of functions on S^1_h. It is the product which weights all points equally. Giving them each weight equal to $1/m$ has the advantage that the norm is then exactly equal to the root mean square value of the components.

[1] This example illustrates a lot of first year graduate course vocabulary. The relation $S^1 = \mathbf{R}/2\pi\mathbf{Z}$ is also valid as topological spaces, 1-manifolds, and Lie Groups. The exponential map is a covering map which exhibits $\mathbf{R}$ as the simply connected cover of S^1.

Definition. Thr space of complex valued functions on S^1_h is given the scalar product

$$\langle f, g\rangle_{S^1_h} := \frac{1}{m} \sum_{z\in S^1_h} f(z)\,\overline{g(z)}\,.$$

This induces the following scalar product on the 2π periodic functions on $\mathbf{R}_h$

$$\langle f, g\rangle_{\mathbf{R}_h^{per}} := \frac{1}{m} \sum f(\theta)\,\overline{g(\theta)}\,.$$

where the sum is over m consecutive points in $\mathbf{R}_h$.

There are some familiar and important 2π periodic functions. The simplest are $\sin n\theta$ and $\cos n\theta$ with integer n. More generally there are the complex exponentials given by De Moivre's relation $e^{in\theta} = \cos n\theta + i \sin n\theta$. The function $e^{i\theta}$ from $\mathbf{R} \to S^1$ is induced by the identity map on S^1. Equivalently, by the function z on S^1. Taking powers positive and negative shows that $e^{in\theta}$ is induced by the function z^n on the circle for every $n \in \mathbf{Z}$.

Evaluating the function $e^{in\theta}$ at the first m points of the discrete line yields the sequence of values

$$(\,1,\, e^{2\pi i n/m},\, e^{2\pi i 2n/m},\, e^{2\pi i 3n/m},\, \ldots,\, e^{2\pi i(m-1)n/m}\,)\,. \tag{4}$$

Comparing with Theorem 3 of the last section yields the following interesting relation.

Theorem 1. Discrete Fourier Transform. *If m is a positive integer and $h = 2\pi/m$, then the functions $e^{in\theta}$ for $0 \le n \le m-1$ form an orthonormal basis for the 2π periodic functions on the discrete line $\mathbf{R}_h$. In particular, if f is such a function then there is a unique choice of a_n such that*

$$f(\theta) = \sum_{n=0}^{m-1} a_n\, e^{in\theta} \quad \text{for all} \quad \theta \in \mathbf{R}_h\,. \tag{5}$$

The coefficients are given by the explicit formula

$$a_n = \frac{1}{m} \sum_{j=0}^{m-1} f(jh)\, e^{-injh}, \quad h := 2\pi/m\,. \tag{6}$$

If $p(z) := \sum_{n=0}^{m-1} a_n z^n$ and $g : S^1_h \to \mathbf{C}$ is the function such that $f(\theta) = g(e^{i\theta})$, then $p(z) = g(z)$ for all $z \in S^1_h$.

Proof. The relation (6) is exactly the expression $f = \sum \langle f, e_n\rangle e_n$ for the orthonormal basis $e^{in\theta}$. $\diamondsuit$

The relations (5-6) describe the *Discrete Fourier Transform* and its inverse. Though simple to state and to prove the identities have a remarkable number of deep mathematical consequences and are astonishingly useful in Applied Mathematics, in particular in conjunction with the *Fast Fourier Transform* which shows that the computation of the Discrete Transform with $m = 2^N$ can be performed in $O(m \log m)$ floating point operations (see Folland pages 252-253).

Note that $p(z)$ is the polynomial of degree $m-1$ interpolating the values of $g(z)$ at the m points of S^1_h. For m large these points tend to fill out the cirlce in a regular manner. Thus if one is given a differentiable function on the circle and interpolates at S^1_h it is not unreasonable to hope that this yields a good approximation over the whole circle. This expectation is correct. However it is quite remarkable. Usually polynomial interpolation at many regularly spaced points is very risky indeed. For example, Runge proved that interpolation of the function $1/(1+x^2)$ by a polynomial of degree m at equally spaced points in $[-5,5]$ yields a sequence of functions which is badly divergent (see, for example, E. Isaacson and J.B. Keller *Analysis of Numerical Methods*, ch.6 §3.4). This makes the convergence for interpolation at equally spaced points on the circle surprising.

To analyse convergence it is import to rewrite (5). The reason is that the terms $e^{in\theta}$ with n close to m look more rapidly oscillating than they really are. The key observation is that, since $mh = 2\pi$, the function $e^{in_1\theta}$ and $e^{in_2\theta}$ define the same function on the discrete line $\mathbf{R}_h$ whenever the n_j differ by a multiple of m. For example the exponentials $e^{i(m-1)\theta}$ and $e^{-i\theta}$ agree on $\mathbf{R}_h$. The first expression *looks* like it is rapidly varying. One would expect to find difference quotients of size $O(m)$. However, the second expression shows that the difference quotients are all bounded by 1. The motto is that it is best to choose the representation $e^{in\theta}$ with $|n|$ as small as possible. For odd $m = 2M+1$ this means that the sum is from $-M$ to M. For even $m = 2M$ if we break the tie in favor of positive integers, the sum is from $1-M$ to M. For the odd case the representation formula is then

$$f\big|_{\mathbf{R}_h} = \sum_{n=-M}^{M} c_{n,h}\, e^{in\theta} \tag{7}$$

$$c_{n,h} = \frac{1}{m} \sum_{j=-M}^{M} f(jh)\, e^{-injh}, \quad h := 2\pi/m, \quad m := 2M+1\,. \tag{8}$$

In this case, the right hand side is a trigonometric polynomial of degree M according to the following definition.

Definition. A *trigonometric polynomial* of degree M is a complex linear combination of the functions $e^{in\theta}$ with $|n| \le M$.

A trigonometric polynomial is a function $\sum_{n=-M}^{M} a_n e^{in\theta}$. It is the function on $\mathbf{R}$ induced by the function $\sum_{n=-M}^{M} a_n z^n$ on S^1. Writing

$$\cos(\theta) = \frac{e^{i\theta} + e^{-i\theta}}{2} \quad \text{and} \quad \sin(\theta) = \frac{e^{i\theta} - e^{-i\theta}}{2i}$$

shows that $\sin^n(\theta)$ and $\cos^n(\theta)$ are trigonometric polynomials of degree n. Conversely

$$e^{i\theta} = \cos\,\theta + i \sin\,\theta\,, \quad \text{and} \quad e^{-i\theta} = \cos\,\theta - i\sin\,\theta$$

show that every trigonometric polynomial of degree M is a complex linear combination of the functions $\sin^n(\theta)$ and $\cos^n(\theta)$ for $|n| \le M$. Thus trigonometric polynomials are exactly such linear comginations.

Equations (7-8) give a reciple for computing a trigonometric polynomial which agrees with (a.k.a. *interpolates)* a 2π periodic function on the discrete line. Here are two ways that the formula can be used. If one is given discrete periodic data values on $\mathbf{R}_h$, then one can use (7-8) to construct a trigonometric polynomial which gives a function on the whole line assuming those values.

Alternatively, given a nice function on $\mathbf{R}$, one can replace it with a trigonometric polynomial which aggrees with the given function on $\mathbf{R}_h$. This replaces a function by a simpler one which one hopes is close to the given function. This idea gives the following approximation algorithm.

Approximation Algorithm. A nice 2π periodic function $f : \mathbf{R} \to \mathbf{C}$ is approximated by the trigonometric polynomial P_M of degree M

$$f \approx P_M(\theta) = \sum_{n=-M}^{M} c_{n,h}\, e^{in\theta} \tag{9}$$

where the coefficients are given by equation (8).

Next consider the limit as $h \to 0$ of these approximations. Since the interval between adjacent points has length $h = 2\pi/m$ the sum on the right hand side of (8) are Riemann sums approximating the integral $\int_0^{2\pi} f\, d\theta/2\pi$. Therefore

$$c_{n,h} \to \frac{1}{2\pi} \int_0^{2\pi} f(\theta)\, e^{-in\theta}\, d\theta \quad \text{as} \quad h \to 0\,.$$

Since the trigonometric interpolation is on an increasingly dense set of points on the real line, and the expression on the right hand side of (7) have clear limit this strongly suggests the following fundamental result.

Fourier's Representation Theorem. *If f is a nice 2π periodic function on $\mathbf{R}$ then*

$$f(\theta) = \sum_{n=-\infty}^{\infty} c_n\, e^{in\theta} \tag{9}$$

with

$$c_n := \frac{1}{2\pi} \int_0^{2\pi} f(\theta)\, e^{-in\theta}\, d\theta\,. \tag{10}$$

Definition. The expression (9) is called the **Fourier Series** of f and the coefficients c_n, defined in (10), are called the *Fourier coefficients* of f.

To make this into a useful result we need to provide precise statements specifying both what is meant by nice, and how well the series converges. For example one must investigate whether the series can be integrated or differentiated term by term and whether the convergence is uniform or only pointwise or even weaker than that.

One of the beauties of the motivation for (9-10) which we have just finished is that it is not hard to turn it into a proof.

Summary. Using the eigenvectors of the specific symmetric matrix (5) from discrete diffusion, we were lead to conjecture one of the basic, and seemingly nonintuitive, theorems of mathematics. At the same time, this leaves us with at least two hard problems. The first is to study the validity of the representation formula, and the second is to prove the convergence of the approximating trigonometric polynomials P_M.

§4. Properties of Fourier Series

In the last section the representation formula for well behaved periodic functions

$$f(\theta) = \sum_{n=-\infty}^{\infty} c_n\, e^{in\theta} \quad \text{where} \quad c_n := \frac{1}{2\pi} \int_0^{2\pi} f(\theta)\, e^{-in\theta}\, d\theta\,, \tag{1}$$

was motivated by passing to the limit of many interpolation points in the formulas for interpolating trigonometric polynomials. The formulas were derived using the eigenvectors of the symmetric matrix A describing the diffusion among a circularly symmetric array of containers.

In this section properties of the resulting Fourier series are investigated. Once this is done, it will not be difficult to prove that (1) is valid, and the properties are of great independent importance. It is important to grasp the logic here. Given a nice periodic f one computes the c_n as in (1) and thereby define a series $\sum c_n e^{in\theta}$ whose properties we now study. Whether the series sums to f or not is neither used nor addressed at this time.

The material in Folland §2.1,2.3,2.4 correspond in large part to what is done here. We will then return to give a second proof of the representation formula, that given in §2.2.

Notice that the Fourier coefficients c_n are well defined as soon as f is integrable. For example if f is a 2π periodic continuous function on $\mathbf{R}$. More generally one could take f to be piecewise continuous as in Folland, or even more generally to be Riemann or even Lebesgue integrable. All of the latter choices suffer a minor inconvenience. A nonzero function may have all Fourier coefficients equal to zero. For example let f be the function which is equal to zero except at $2\pi\mathbf{Z}$ where f is equal to 1. This shows that for such functions a little care needs to be taken with (1). The first propositions are valid for such general functions and follow immediately from the defintion (1). For definiteness, you may suppose that f is piecewise continuous or continuous, but the option of more general integrable functions is left open. The $n^{\rm th}$ Fourier coefficient is denoted

$$c_n = c_n(f) = \hat{f}(n) := \frac{1}{2\pi} \int_0^{2\pi} f(\theta)\, e^{-in\theta}\, d\theta\,. \tag{2}$$

Proposition 1. *The Fourier coefficients of the sum of two functions are the sum of the Fourier coefficients of the functions,*

$$c_n(f+g) = c_n(f) + c_n(g)\,. \tag{3}$$

The Fourier coefficient of a scalar multiple of a function is that scalar times the Fourier coefficients of the original function,

$$c_n(af) = a c_n(f)\,. \tag{4}$$

Define the reflection and translation operators by

$$(Rf)(\theta) := f(-\theta) \quad \text{and} \quad (\tau_\phi f)(\theta) := f(\theta - \phi)\,. \tag{5}$$

Then

$$c_n(Rf) = c_{-n}(f) \quad \text{and} \quad c_n(\tau_\phi f) = e^{-in\phi}\, c_n(f)\,. \tag{6}$$

The Fourier coefficients of a periodic function satisfy the following estimates

$$|c_n(f)| \le \frac{1}{2\pi} \int_0^{2\pi} |f(\theta)|\, d\theta \le \sup_{\theta\in\mathbf{R}} |f(\theta)|\,. \tag{7}$$

The next group of facts about Fourier series depend on the basic results about scalar product spaces. To make the connection, we first derive a family of orthogonality relations satisfied by the exponential functions. Theorem 1 of *Lines and Circles, Discrete and Otherwise* implies that if $n - n' \notin m\mathbf{Z}$ then

$$h \sum_{j=0}^{m-1} e^{injh}\, e^{-in'jh} = 0\,, \quad \text{where} \quad h = h_m = \frac{2\pi}{m}\,. \tag{8}$$

Fix n and n' and let $m \to \infty$. Then the sum in (8) is a Riemann sum which converges to $\int_0^{2\pi} e^{in\theta} e^{-in'\theta}$. This proves that

$$0 = \int_0^{2\pi} e^{in\theta}\, e^{-in'\theta}\, d\theta \quad \text{whenever} \quad n \neq n'\,. \tag{9}$$

These identities can be proved by direct computation as on the bottom of page 19 in Folland.

Theorem 2. i. *Let V denote the scalar product space of continuous 2π periodic functions on $\mathbf{R}$ endowed with the scalar product*

$$\langle f, g \rangle := \frac{1}{2\pi} \int_0^{2\pi} f(\theta)\, \overline{g}(\theta)\, d\theta\,. \tag{10}$$

Then the family of functions $\{e^{in\theta} \;:\; n \in \mathbf{Z}\}$ is an orthonormal family.
ii. *If $J \in \mathbf{Z}$ is a finite set and $f \in V$ then $\sum_{n\in J} c_n(f)\, e^{in\theta}$ is the closest element to v in the span of $\{e^{in\theta} \;:\; n \in J\}$. This best approximant satisfies*

$$||f - \sum_{n\in J} c_n(f)\, e^{in\theta}||^2 + \sum_{n\in J} |c_n(f)|^2 = ||f||^2 \tag{11}$$

where the first term is the mean square error and the other terms are easy to compute.
iii. *The Fourier coefficients are square summable and satisfy Bessel's inequality*

$$\sum_{-\infty}^{\infty} |c_n|^2 \leq ||f||^2\,. \tag{12}$$

Proof. Assertion i. follows from (9).
The first assertion of ii. is part of the Projection Theorem.
To prove (11), use the fact that $f - \sum_J c_n e^{in\theta}$ and the projection $\sum_J c_n e^{in\theta}$ are orthogonal together with the Pythagorean Theorem to show that

$$||f - \sum_{n\in J} c_n(f)\, e^{in\theta}||^2 + ||\sum_{n\in J} c_n(f)\, e^{in\theta}||^2 = ||f||^2\,. \tag{13}$$

Since the exponentials are orthogonal, the Pythagorean Theorem applies again to give

$$||\sum_{n\in J} c_n(f)\, e^{in\theta}||^2 = \sum_{n\in J} |c_n|^2\,. \tag{14}$$

Together these yield (11).

Finally, the left hand side of (12) is the supremeum over J of $\sum_{n\in J}|c_n|^2$ which is bounded by the right hand side of (12) thanks to (11). $\diamondsuit$

As remarked on page 21 of Folland, this result implies that the Fourier coefficients tend to zero as $|n|$ tends to infinity.

Formula (11) applied when J is the set of integers between $-N$ and N yields

$$||f-\sum_{n=-N}^{N} c_n(f)\, e^{in\theta}||^2 = ||f||^2 - \sum_{n=-N}^{N} |c_n(f)|^2\,. \tag{15}$$

Convergence of the Fourier series to f in the mean square sense is equivalent the convergence to zero of the left hand side, which is then equivalent to *Parseval's identity*

$$||f||^2 = \sum_{n\in\mathbf{Z}} |c_n|^2\,. \tag{16}$$

For the moment this remains unproved.

The excellent examples on pages 22-25 of Folland illustrate the fact that the smoother is the graph of f the more rapidly its Fourier coefficients decay the more rapid is the convergence of the partial sums of the series. For the examples considered the partial sums appear to converge to the original function suggesting that the relation (1) holds at least in these cases. The next material corresponds closely to Folland 2.3, with the exception that we do not yet know what he has proved in Theorem 2.1.

Theorem 3. *If f is a coniuously differerntiable 2π periodic function then the* n^{th} *Fourier coefficient of $f'(\theta)$ is equal to in time the n^{th} Fourier coefficient of $f(\theta)$,*

$$c_n(f') = inc_n(f)\,. \tag{17}$$

Remarks. i. Folland Theorem 2.2 is slightly more general.

ii. One can guess and also remember formula (17) by simply differentiating the identity (1) to find,

$$f' = \frac{d}{d\theta}\sum_{n=-\infty}^{\infty} c_n\, e^{in\theta} = \sum_{n=-\infty}^{\infty} inc_n\, e^{in\theta}$$

which is a Fourier representation for f'. Of course that is **not** a proof as we do not know yet that (1) is true and also the interchange of differentiation and infinite sum might not be permissible.

Proof. The proof is by an integration by parts in the definition,

$$2\pi\, c_n(f') := \int_0^{2\pi} f'(\theta)\, e^{-in\theta}\, d\theta = -\int_0^{2\pi} f(\theta)\,(-in)e^{-in\theta}\, d\theta + f(\theta)e^{-in\theta}\Big|_0^{2\pi}\,.$$

The relation (17) follows upon noticing that last term vanishes because $fe^{-in\theta}$ is 2π periodic. $\diamondsuit$

By induction one has that for k times continuously differentiable periodic functions and $n \neq 0$

$$c_n(f) = \frac{c_n(f^{(k)})}{(in)^k}. \tag{18}$$

Together with the estimate (7) this yields the important decay estimate

$$|c_n(f)| \leq |n|^{-k} \max_{\theta \in \mathbf{R}} |f^{(k)}(\theta)| \tag{19}$$

An immediate consequence is that $\sum |c_n| < \infty$ for the Fourier coefficients of a C^2 periodic function since $\sum_{n \neq 0} 1/n^2 < \infty$. Then the Weirstrass M-Test (Folland page 40) implies that the Fourier series is uniformly convergent to a continuous function. Of course we believe that the limit is f and will prove it shortly. Actually one derivative is sufficient for the same conclusion. Folland Theorem 2.5 is slightly more general.

Theorem 4. *The Fourier coefficients of a continuously differentialbe function are absolutely summable. Therefore, the Fourier series is absolutely and uniformly convergent on* $\mathbf{R}$.

Proof. For any finite subset $J \subset \mathbf{Z}$, equation (17) yields

$$\sum_{n \in J} |c_n(f)| \leq \sum_{n \in J} \frac{1}{|n|} \; |c_n(f')|$$

The Cauchy-Shwarz inequality implies

$$\sum_{n \in J} |c_n(f)| \leq \Big(\sum_{n \in J} \frac{1}{n^2} \Big)^{1/2} \Big(\sum_{n \in J} |c_n(f')|^2 \Big)^{1/2}.$$

Finally, Bessel's inequality yields the basic estimate

$$\sum_{n \in J} |c_n(f)| \leq \Big(\sum_{n \in J} \frac{1}{n^2} \Big)^{1/2} ||f'||. \tag{21}$$

Apply (21) with $J = \{0 < |n| < N\}$ to get

$$\sum_{0<|n|<N} |c_n(f)| \leq \Big(\sum_{0<|n|<\infty} \frac{1}{n^2} \Big)^{1/2} ||f'||.$$

Since the right hand side is finite and independent of N, the $c_n(f)$ are absolutely summable and

$$\sum_{n \neq 0} |c_n| \leq \Big(\sum_{0<|n|<\infty} \frac{1}{n^2} \Big)^{1/2} ||f'||.$$

The Weierstrass M-Test implies the absolute and uniform convergence. $\diamondsuit$

Corollary 5. *If $k \geq 1$ and f is a C^k periodic function and $j \leq k-1$ then, the Fourier series of $f^{(j)}$ converges absolutely and uniformly on* $\mathbf{R}$. *As $N \to \infty$, the sum of the tail of the series is estimated from above by*

$$\Big|\sum_{N<|n|<\infty} c_n(f)\, e^{in\theta}\Big| \leq \frac{\sqrt{2/(2k-1)}}{N^{k-1/2}}\, ||f^{(k)}||\,. \tag{22}$$

Proof. The first assertion is an immediate consequence of Theorem 4. To prove (22) note that the proof of (21) can be repeated using (18) for arbitrary k and with $J = \{N < |n| \leq M\}$ to give

$$\sum_{N<|n|\leq M} |c_n(f)| \leq \Big(\sum_{N<|n|\leq M} \frac{1}{n^{2k}}\Big)^{1/2} ||f^{(k)}||\,. \tag{23}$$

The sum on the right is bounded above by

$$2\int_N^\infty \frac{1}{x^{2k}}\, dx = \frac{2}{(2k-1)\, N^{2k-1}}\,.$$

Then passing to the limit $M \to \infty$ in (23) yields

$$\sum_{N<|n|<\infty} |c_n(f)| \leq \Big(\frac{2}{(2k-1)\, N^{2k-1}}\Big)^{1/2} ||f^{(k)}||\,,$$

and (22) follows. $\diamondsuit$

§5. Properties of Trigonometric Interpolation

Some of the results of *Properties of Fourier Series* have useful analogues for interpolating polynomials. Let $h := 2\pi/(2M+1)$ and

$$P_M(\theta) = \sum_{-M}^{M} c_{n,h}\, e^{in\theta} \tag{1}$$

denote the unique trigonometric polynomial of degree M interpolating f on the discrete line $\mathbf{R}_h$, that is

$$f(\theta) = P_M(\theta) \quad \text{whenever} \quad \theta \in \mathbf{R}_h\,. \tag{2}$$

Formula (8) of *Circles* reads

$$c_{n,h}(f) = \langle f, e^{in\theta}\rangle_{\mathbf{R}_h^{per}} = \frac{h}{2\pi} \sum_{j=-N}^{N} f(jh)\, e^{-injh}\,. \tag{3}$$

The Pythagorean Theorem gives

$$\sum_{n=-M}^{M} |c_{n,h}|^2 = ||f||^2_{\mathbf{R}_h^{per}} = \frac{1}{2M+1} \sum_{n=-M}^{M} |f(nh)|^2 \le \max_{\theta\in\mathbf{R}_h} |f(\theta)|\,. \tag{4}$$

For smooth f, the coefficients $c_{n,h}$ decay as n grows and this decay is *uniform in* h. To show this, imitate the integration by parts trick used in *Properties of Fourier Series.* The role of the derivative is played by difference quotients with $\Delta\theta = h$. Introduce translation and difference operators acting on function on the discrete line by

$$(\tau_+ f)(\theta) := f(\theta - h)\,, \qquad (\tau_- f)(\theta) := f(\theta + h)\,, \tag{5}$$

$$D_+ f := (\tau_- - I)f/h = \frac{f(\theta+h) - f(\theta)}{h}\,, \qquad D_- f := (I - \tau_+)f/h = \frac{f(\theta) - f(\theta-h)}{h}\,. \tag{6}$$

All four of these operators map the m dimensional scalar product space of 2π periodic functions on the discrete line itself.

Proposition 1. i. *The functions* $\{e^{in\theta} : n \in \mathbf{Z}\}$ *are eigenfunctions of the operators* $\tau_\pm$ *and* $D_\pm$ *for all* h. *They satisfy*

$$\tau_\pm e^{in\theta} = e^{\mp inh}\, e^{in\theta}, \quad \text{and} \quad D_\pm e^{in\theta} = \frac{\pm(e^{\pm inh} - 1)}{h}\, e^{in\theta}\,. \tag{7}$$

Choosing m *consecutive values of* n *gives an orthonormal basis of eigenfunctions.*

ii. *The adjoint of* τ_+ *is* τ_-. *The adjoint of* D_+ *is* $-D_-$.

iii. *The coefficients of* f *and* $D_\pm f$ *are related by*

$$c_{n,h}(D_+ f) = \frac{(e^{inh} - 1)}{h}\, c_{n,h}(f)\,, \qquad c_{n,h}(D_- f) = \frac{(1 - e^{-inh})}{h}\, c_{n,h}(f)\,. \tag{8}$$

Proof. i. Compute

$$\tau_+ e^{in\theta} = e^{in(\theta-h)} = e^{-inh}\, e^{in\theta}$$

proving the first of the relations (7). The others are similar.

ii. Choosing the eigenbasis $e^{in\theta}$ the transformation τ_+ is diagonal with diagonal entries e^{-inh}. The adjoint is then diagonal with entries e^{inh}. That is the matrix of τ_-. This shows that $\tau_+^* = \tau_-$. The relation for the adjoint of D_+ follows from the sum and product laws for adjoints,

$$D_+^* = \left(\frac{\tau_+ - I}{h}\right)^* = \frac{\tau_+^* - I^*}{h} = \frac{\tau_- - I}{h} = -D_- \,.$$

iii. To prove the result for D_+, use the adjoint of D_+ in the definition of $c_{n,h}$ to get

$$c_{n,h}(D_+ f) = \langle D_+ f, e^{in\theta}\rangle_{\mathbf{R}_h^{per}} = \langle f, -D_- e^{in\theta}\rangle_{\mathbf{R}_h^{per}} \,.$$

Then the eigenvalue relation for D_- yields

$$= \langle f, \frac{e^{-inh}-1}{h} e^{in\theta}\,\rangle_{\mathbf{R}_h^{per}} = \overline{\left(\frac{e^{-inh}-1}{h}\right)}\,\langle f, e^{in\theta}\rangle_{\mathbf{R}_h^{per}} = \frac{e^{inh}-1}{h}\; c_{n,h}(f)\,. \tag{9}$$

$\diamondsuit$

The identity (8) is analogous to the case $k = 1$ of the formula (18) for Fourier series. In fact, for $nh << 1$, Taylor's Theorem shows that $e^{inh} - 1 \approx inh$ so (9) shows that $c_{n,h}(D_+ f) \approx in\, c_{n,h}(f)$ and the analogy is even clearer. To derive an estimate analogous to (19) requires a lower bound for $|e^{inh} - 1|$. Note that if nh is an integer multiple of 2π then this quantity vanishes. It is crucial that nh avoids these values.

Proposition 2. *If f is a 2π periodic function on $\mathbf{R}_h$, then for all $M > 0$ and $n \neq 0$ the coefficients of the interpolation polynomial satisfy*

$$|c_{n,h}(f)| \;\le\; \frac{\pi}{2n}\; |c_{n,h}(D_+ f(\theta))| \;\le\; \frac{\pi}{2n}\; \max_{\theta\in\mathbf{R}_h} |D_+ f(\theta)|\,. \tag{10}$$

Proof. For $h = 2\pi/(2M+1)$ and $|n| \le M$, one has $|nh| \le \pi$.

The quantity $|e^{i\theta} - 1|$ is the length of the line segment from $z = 1$ to $z = e^{i\theta}$, and $|\theta|$ is the arclength on the unit circle between these two points. For $|\theta|$ small these two lengths are about equal, with the circle being slightly larger. As θ increases so does the discrepancy. It is plausible, from geometric intuition, that $|e^{i\theta} - 1|/|\theta|$ decreases as $|\theta|$ increases from 0 to π, and therefore, the minmum $2/\pi$ is achieved at $\theta = \pm\pi$. To verify this, write

$$|e^{i\theta} - 1|^2 = (e^{i\theta} - 1)\overline{(e^{-i\theta} - 1)} = (e^{i\theta} - 1)(e^{-i\theta} - 1) = 2 - 2\cos\theta = 4\sin^2(\theta/2)\,. \tag{11}$$

The function $(\sin x)/x$ is a smooth even function whose derivative, $(-x\,\cos x - \sin x)/x^2$, is negative for $0 < x < \pi/2$. Therefore, that for $|\phi| \le \pi/2$, $|\sin\phi| \ge |2\phi/\pi|$. Using this in (11) yields

$$|e^{i\theta} - 1|^2 \ge 4\theta^2/\pi^2 \quad \text{for} \quad |\theta| \le \pi\,. \tag{12}$$

Using (12) in (8) yields the desired result. $\diamondsuit$

The next result answers the basic question about the Approximation Algorithm and the Fourier Representation Theorem.

Theorem 3. *If f is a continuously differentiable 2π periodic function on $\mathbf{R}$ and P_M denotes the trigonometric polynomial of degree M interpolating f on $\mathbf{R}_h$ with $h = 2\pi/(2M+1)$, then as $M \to \infty$, P_M converges uniformly to f on $\mathbf{R}$. The Fourier Series of f also converges uniformly to f.*

Proof. Theorem 4 in *Properties of Fourier Series* shows that the Fourier series is uniformly to a limit which we call g,

$$g := \sum_{-\infty}^{\infty} c_n(f)\, e^{in\theta}\,.$$

Inequality (22) from that section with $k = 1$ yields an estimate for the tail of the series

$$\Big|\sum_{N<|n|} c_n\, e^{in\theta}\Big| \le \sqrt{\frac{2}{N}}\,\left(\frac{1}{2\pi}\int_0^{2\pi} |f'(\theta)|^2\, d\theta\right)^{1/2}. \tag{13}$$

The first and key step is to prove that $P_M \to g$. Since the coefficients $c_{n,h}$ are Riemann sums and $f(\theta)e^{-in\theta}$ is a continuous function of θ, the fundamental existence theorem of integral calculus asserts that for each n

$$c_{n,h}(f) \to c_n(f) \quad \text{as} \quad h \to 0\,. \tag{14}$$

Splitting both P_M and g into low and high frequency parts yields

$$P_M(\theta) - g = \sum_{|n|\le N} (c_{n,h} - c_n)\, e^{in\theta} + \sum_{N<|n|\le M} c_{n,h}\, e^{in\theta} - \sum_{N<|n|} c_n\, e^{in\theta}\,. \tag{15}$$

The second summand is the high frequency part of P_M and the last term is the high frequency part of g. To bound the second sum, use the decay of the coefficients proved in the last proposition to find

$$\Big|\sum_{N<|n|\le M} c_{n,h}(f)\, e^{in\theta}\Big| \le \sum_{N<|n|\le M} |c_{n,h}(f)| \le \sum_{N<|n|\le M} \frac{\pi\, |c_{n,h}(D_+f)|}{2|n|}\,. \tag{16}$$

The Cauchy-Schwarz inequality yields

$$\le \frac{\pi}{2}\Big(\sum_{N<|n|\le M} \frac{1}{n^2}\Big)^{1/2} \Big(\sum_{N<|n|\le M} |c_{n,h}(D_+f)|^2\Big)^{1/2}.$$

The sum inside the first parentheses is bounded above by $2\int_N^\infty x^{-2}dx = 2/N$. The sum inside the second parentheses is bounded by the sum over all n, which is equal to the square of the norm of D_+f. This yields the first of the estimates

$$\Big|\sum_{N<|n|\le M} c_{n,h}(f)\, e^{in\theta}\Big| \le \frac{\pi}{2}\,\sqrt{\frac{2}{N}}\; ||D_+f||_{\mathbf{R}_h^{per}} \le \frac{\pi}{\sqrt{2N}}\,\max_{\theta\in\mathbf{R}} |f'(\theta)|\,. \tag{17}$$

For the second,

$$||D_+f|| \le \max_{\mathbf{R}_h} |D_+f|$$

and for all θ,

$$|D_+f(\theta)| = |f(\theta-h) - f(\theta)| = \Big|\int_{\theta-h}^{\theta} f'(\theta)\,d\theta\Big| \le \max_{\theta\in\mathbf{R}} |f'(\theta)|\,.$$

The uniform convergence $P_M \to g$ is then proved as follows. Given a challenge number $\epsilon > 0$ we must show that there is an $M_0 > 0$ so that whenever $M > M_0$, $|P_M(\theta) - g(\theta)| < \epsilon$ for all θ.

Estimates (13) and (17) show that there is an N_0 so that for all M

$$\Big|\sum_{N_0<|n|\le M} c_{n,h}\,e^{in\theta} - \sum_{N_0<|n|} c_n\,e^{in\theta}\Big| < \epsilon/2\,. \tag{18}$$

In the first sum in (15) there are $2N_0+1$ summands, and each tends to zero by (13). Thus there is an $M_0 > 0$ depending on ϵ such that for $M > M_0$

$$\Big|\sum_{|n|\le N_0} (c_{n,h} - c_n)\,e^{in\theta}\Big| < \epsilon/2\,. \tag{19}$$

Using (18) and (19) in (15) proves the uniform convergence $P_M \to g$.

It remains to show that $g(\theta) = f(\theta)$. Consider the subsequence of $M(k) \to \infty$ mesh widths $h_k \to 0$ defined by

$$2M(k) + 1 = 3^k\,,\quad \text{and,}\quad h(k) = 2\pi/3^k\,.$$

The picture would be simpler $M = 2^k$ but our M must be odd! The discrete line $\mathbf{R}_{h(k+1)}$ arises from $\mathbf{R}_{h(k)}$ by trisecting each of the subintervals. In particular the points of the discrete line $\mathbf{R}_{h(j)}$ belong to all of the discrete lines $\mathbf{R}_{h(k)}$ with $k \ge j$.

Equation (2) shows that

$$P_{M(k)}(\theta) = f(\theta)\,,\quad \text{for all}\quad \theta \in \mathbf{R}_{h(k)}\,. \tag{21}$$

For any point in $\mathbf{R}_{h(j)}$, passing to the limit $k \to \infty$ implies that

$$g(\theta) = f(\theta)\,,\quad \text{for all}\quad \theta \in \mathbf{R}_{h(j)}\,. \tag{22}$$

Since this is true for all j, it follows that f and g are equal on the dense set of points $\bigcup_j \mathbf{R}_{h(j)}$. By continuity, they are equal everywhere.

Since both P_M and the Fourier series of f converge uniformly to g and $g = f$, this completes the proof of the Theorem. $\diamondsuit$

§6. Corollaries of the Convergence Theorem

The uniform convergence of the Fourier series of continuously differentiable functions implies several important corollaries. The first is that Bessel's inequality is an equality.

Theorem 1. Parseval's Identity. *If f is a continuously differenitalbe 2π periodic function, then*

$$\sum_{n=-\infty}^{\infty} |c_n(f)|^2 = ||f||^2 := \frac{1}{2\pi}\int_0^{2\pi} |f(\theta)|^2\, d\theta\,. \tag{1}$$

Proof. The Projection Theorem implies that

$$\frac{1}{2\pi}\int_0^{2\pi} \Big|f(\theta) - \sum_{-N}^{N} c_n(f)\, e^{in\theta}\Big|^2\, d\theta = ||f||^2 - \sum_{-N}^{N} |c_n(f)|^2\,.$$

As $N \to \infty$, the left hand side tends to zero since the integrand converges uniformly to zero. Thus the rigth hand side tends to zero which is the desired result. $\diamondsuit$

The next result shows that our results are not too far from capturing exactly the differentiability properties of

Theorem 2. *(**i.**) If f is a k times continuously differentiable 2π periodic function then*

$$\sum_{n=-\infty}^{\infty} |n^k c_n(f)|^2 \ < \infty. \tag{2}$$

*(**ii.**) If f is a continuously differentiable 2π periodic function and*

$$\sum_{n=-\infty}^{\infty} |n^k c_n(f)| \ < \infty\,, \tag{3}$$

then, f is a k times continuously differentiable 2π periodic function.
*(**iii.**) A continuously differentiable 2π periodic function f is an infinitely differentiable if and only if*

$$(\forall k)\ \lim_{|n|\to\infty} n^k\, c_n(f) = 0\,. \tag{4}$$

Proof. The first result follows from Parseval's identity together with the formula for the Fourier coeffcients of derivatives.

$$\infty > \frac{1}{2\pi}\int_0^{2\pi} \left|\frac{d^k f(\theta)}{d\theta^k}\right|^2 d\theta \ = ||f^{(k)}||^2 = \sum_{n=-\infty}^{\infty} |c_n(f^{(k)})|^2 = \sum_{n=-\infty}^{\infty} |n^k c_n(f)|^2\,. \tag{5}$$

To prove (**ii.**), consider the $k+1$ series

$$\sum_{-\infty}^{\infty}(in)^j\, c_n(f)\, e^{in\theta} \quad \text{for} \quad 0 \le j \le k\,.$$

The j^{th} is obtained from the $(j-1)^{st}$ by differentiating term by term. The terms with $n \neq 0$ are dominated by the convergent sequence of positive reals $\sum |nc_n|^k$.

It follows by the Weirstrass M-test that the series are uniformly convergent.

A basic theorem in advanced calculus asserts that if a uniformly convergent series of C^k functions and the series obtained by termwise differentiating k times both coverge uniformly, then the sum of the first is k times continously differentiable and its derivatives up to order k are equal to the sum of the series of differentiated functions.

This shows that $\sum_{-\infty}^{\infty} c_n(f)\, e^{in\theta}$ is a C^k function. Since the sum is equal to f, the proof of **ii.** is complete.

The assertion **iii.** is a consequence of the previous two. ♢

Condition (3) is stronger than condition (2). That is

$$\sum_{n=-\infty}^{\infty} |n^k c_n(f)| < \infty \Rightarrow \sum_{n=-\infty}^{\infty} |n^k c_n(f)|^2 < \infty. \tag{6}$$

In the opposite direction,

$$\sum_{n=-\infty}^{\infty} |n^k c_n(f)|^2 < \infty \Rightarrow \sum_{n=-\infty}^{\infty} |n^{k-1} c_n(f)| < \infty\,. \tag{7}$$

This follows from the Cauchy-Schwartz inequality as follows. With sums omitting the value $n=0$,

$$\sum_{n=-N}^{N} |n^{k-1}\, c_n(f)| = \sum_{n=-N}^{N} |\frac{1}{n}\, n^{k-1}\, c_n(f)| \le \Big(\sum_{n=-\infty}^{\infty} \frac{1}{n^2}\Big)^{1/2} \Big(\sum_{n=-\infty}^{\infty} |n^{k-1}\, c_n(f)|^2\Big)^{1/2}. \tag{8}$$

This can be strengthened. For ϵ a positve real number as small as you like,

$$\sum_{n=-\infty}^{\infty} |n^k c_n(f)|^2 < \infty \Rightarrow \sum_{n=-\infty}^{\infty} |n^{k-1/2-\epsilon}\, c_n(f)| < \infty\,. \tag{9}$$

For proof, repeat (8) beginning with

$$\sum_{n=-N}^{N} |n^{k-1/2-\epsilon}\, c_n(f)| = \sum_{n=-N}^{N} |\frac{1}{n^{1/2+\epsilon}}\, n^k\, c_n(f)|\,.$$

The fact that $\sum 1/n^{1+2\epsilon} < \infty$ is used at the next stage.

Thus (7), (9), and the first two parts of the previous theorem show that by testing the convergence of the series with terms $|n^k c_n|$ and $|n^k c_n|^2$ one can determine the differentiability of f with an uncertainty of at most 1 derivative. If one accepts fractional derivative, which can be done using the notion of Hölder continuity, the uncertainty is reduced to 1/2.

Definition. If $0 < \alpha < 1$, a 2π periodic function f is Hölder continuous of order α denoted $f \in C^\alpha$ if and only if there is a constant C such that for all θ and ϕ

$$|f(\theta) - f(\phi)| \leq C\, |\theta - \phi|^\alpha\,. \tag{11}$$

If $k = m + \alpha$ with nonegative integer m and $0 < \alpha < 1$, so k is between the integers m and $m+1$, then f belongs to the Hölder class C^k if and only if $f \in C^m$ and $d^m f/dx^m \in C^\alpha$.

The condition (11) is only interesting for θ near ϕ. If it is true for $|\theta - \phi| \leq \delta$ then (11) is satisfied for all θ , ϕ, with perhaps a larger value of C.

Examples. The function $|\sin\theta|^{1/2}$ belongs to the Hölder space $C^{1/2}$ but not the the space C^α for any $\alpha > 1/2$. This is a precise assertion which corresponds to the idea that this function has exactly 1/2 a derivative. More generally, if $k > 0$ is not an integer then $|\sin\theta|^k$ belongs to the Hölder space C^k and to no C^l with $l > k$.

The assertions of the Theorem are valid if the phrase "is a k times differentiable function" is replaced by "is a member of the Hölder class C^k". The interested reader is referred to the standard texts on Fourier analysis for the proofs.

Examples. 1. Consider the function f whose Fourier series is $\sum_{n\neq 0} |n|^{-7.1}\, e^{in\theta}$. Part (i.) of the Theorem shows that if $f \in C^k$ then $\sum |n^{k-7.1}|^2 \leq \infty$. This holds if and only if $k < 6.6$. Thus f cannot have more than 6.6 derivatives in the sense that f does not belong to the Hölder spaces C^k with $k \geq 6.6$.

In the other direction, $\sum |n^k\, c_n| < \infty$ if and only if $k < 6.1$ Thus f is guaranteed to belong to the Hölder space C^k for all $k < 6.1$.

If you are only interested in integer values this shows that $f \in C^6$ and $f \notin C^7$.

2. The function $\sum e^{-|n|^{1/2}}\, e^{in\theta}$ is infinitely differentiable by part (iii.) since the Fourier coefficients decay more rapidly than any power of n. This example is interesting as it can be shown that the Taylor expansion about any point $\underline{\theta}$ diverges at all points other than $\underline{\theta}$. This is an infinitely differentiable function which never has a convergent Taylor series. The proof of divergence is by showing that the terms of the Taylor series do not tend to zero. Thus the series cannot be convergent. The proof that the terms do not tend to zero is tricky.

§7. Complete Scalar Product Spaces

One of the key properties of the real numbers is that they are complete. This is defined as follows.

Definition. A sequence r_n of real numbers is a *Cauchy sequence* if for any $\epsilon > 0$ there is an N so that if n and m are greater than or equal to N then $|r_n - r_m| < \epsilon$.

It is easy to see that any convergent sequence is Cauchy. On the other hand there are Cauchy sequnces of rational numbers which do not converge to rational numbers. For example, the seqence 1, 1.4, 1.41, 1.414, 1.4142, ... of successive decimal approximations of the square root of 2. The key inspiration in the construction of the real numbers is to fill in these gaps.

Theorem 1. *There is an ordered field* $\mathbf{R}$ *containing the rational numbers as a dense subset, that is for any* $r \in \mathbf{R}$ *there is a sequence of rationals* $q_n \to r$ *and which is complete in each of the following equivalent senses.*

i. Every set which is bound above has a least upper bound.

ii. Every bounded sequence has a convergent sequence.

iii. Every Cauchy sequence is convergent.

Any two such fields $\mathbf{R}_1$, *and* $\mathbf{R}_2$ *are isomorphic in the sense that there is an order preserving bijection* $\phi : \mathbf{R}_1 \to \mathbf{R}_2$ *which is an isomorphism of fields, that is for allgoto-line* $a, b \in \mathbf{R}_1$,

$$\phi(a + b) = \phi(z) + \phi(b) \quad \text{and} \quad \phi(ab) = \phi(a)\phi(b)$$

and, such that $\phi(q) = q$ *for all rationals* q.

The set $\mathbf{R}$ is the real numbers. The beauty of the condition iii. is that it allows one to construct a real number without giving an exact formul and it lets you prove convergence without knowing the limit. The definition of $r_n \to r$ is that for any ϵ there is an n_0 so that for $n > n_0$, $|r_n - r| < \epsilon$. For example, if one needs to know the limit to a precision of six decimal places one takes $\epsilon = 0.5 \times 10^{-6}$ and then for $n > n_0$, r_n gives the limit to six places. Howver, this definition requires r while in practice one almost never knows r. The criterion iii. allows one to verify convergence without knowing the limit. Furthermore if one wants the limit of a Cauchy sequence to an accuracy ϵ it suffices to take r_n with $n > N$ where N is as in the definition of Cauchy sequence.

You often use the criteria of the Theorem. For example, when you use the comparison test for an infinite series, you are showing that the partial sums are Cauchy and that they converge to the sum at least as fast as do the partial sums of the comparison series.

A second example is the integral of a continuous function which is defined as the limit of Riemann sums. The Riemann sums are shown to converge by showing that they form a Cauchy sequence.

Definition. A sequence v_n of vectors in a scalar product space is a *Cauchy sequence* if for any $\epsilon > 0$ there is an N so that if n and m are greater than or equal to N then $||v_n - v_m|| < \epsilon$.

A convergent sequence $v_n \to v$ is Cauchy, since given $\epsilon > 0$ one can choose n_0 so that $||v_n - v|| < \epsilon/2$ for $n > n_0$. Then for n and m larger than n_0,

$$||v_n - v_m|| \leq ||v_n - v|| + ||v - v_m|| < \epsilon \,.$$

Definition. A scalar product space V is *complete* if every Cauchy sequence is convergent. A complete scalar product space is called a *Hilbert space.*

Theorem 2. *Every finite dimensional scalar product space is complete.*

Proof. Choose an orthormal basis $e_1, e_2, \ldots, e_d$. For any vector $v = \sum \langle v_j, e_j \rangle e_j$, $||v||^2 = \sum |\langle v_j, e_j \rangle|^2$. It follows that if v_n is a Cauchy sequence of vectors then the Fourier coefficients $\langle v_n, e_j \rangle$ are Cauchy sequences of complex numbers for $1 \le j \le d$.

Thus the real and imaginary parts of $\langle v_j, e_j \rangle$ are Cauchy sequences of reals. It follows by completeness of the real numbers, that the real and imaginary parts converge, and therefore, that $\langle v_j, e_j \rangle$ converges.

Thus, for each j the complex sequence $\langle v_j, e_j \rangle$ is convergent as $n \to \infty$ to limits c_j.

This in turn implies that $v_n \to (c_1, c_2, \ldots, c_n)$ since

$$||v_n - (c_1, c_2, \ldots, c_n)||^2 = \sum |\langle v_j, e_j \rangle - c_j|^2 \to 0\,.$$

$\diamondsuit$

The conclusion here is that the completeness of the real numbers implies the completeness of finite dimensional scalar product spaces.

The bad news is that infinite dimensional spaces are not as well behaved. Worse still, the spaces we have been dealing with are not complete. To show that it suffices to exhibit Cauchy sequences which are not convergent. As with the case of the decimal approximations to the square root of 2, these series apprear to converge, but to elements which are not in the underlying spaces, suggesting that these new elements be added. The goal, which can eventually be achieved is to add enough new elements so that the resulting space is complete. This is the same strategy that leads from the rationals to the reals.

In each of the following spaces of 2π periodic functions the scalar product is equal to $\int_0^{2\pi} f\overline{g}\, d\theta/2\pi$.

Examples 1. The spaces of infinitely differentialbe 2π periodic functions is not a complete scalar product spaces.

Proof. Denote by $f_N := \sum_{n=1}^N (1/n^2) e^{in\theta}$ the partial sum of the Fourier series $\sum_{n=1}^\infty (1/n^2) e^{in\theta}$. Then f_N is We know that Cauchy. In fact f_N converges uniformly to $f := \sum_{n=1}^\infty (1/n^2) e^{in\theta}$ and therefore in norm. However f is not infinitely differentiable since its Fourier coefficients do not decay faster than any power. If f_N converged in norm to a smooth function g it would follow that $f = g$, but f is not smooth. $\diamondsuit$

The limit function f in the above example is continuous and piecewise smooth. You might hope that by throwing in all such funcions you would get a complete space.

Example 2. The space of continuous piecewise smooth 2π periodic functions is not a complete scalar product spaces.

Proof. Let f be the periodic function which is equal to 1 for $0 \le \theta \le \pi$ and is equal to 0 for $-\pi < \theta < 0$.

Let f_n be the continuous piecewise smooth function which is equal to f in $1/n \le |\theta| \le \pi - 1/n$. In $[-1/n, 1/n]$, f_n is the linear function whose graph rises from 0 at the left hand endpoint to 1 at the right hand endpoint. In $[\pi - 1/n, \pi + 1/n]$, f_n decreases linearly from 1 to 0.

If $N < n < m$,

$$||f_n - f_m||^2 \leq 2/\pi N \,,$$

proving that f_n is a Cauchy sequence. If there were a continuous piecewise smooth function g such that $f_n \to g$ in norm, then the triangle inequality shows that

$$||f - g|| \leq ||f_n - g|| + ||f - f_n|| \,.$$

Both terms tend to zero as $n \to \infty$ so one must have $f = g$ showing that g is not continuous. $\diamondsuit$

This example shows that you must also add the the discontinuous piecewise smooth functions. That too, is insufficient.

Example 3. The piecewise smooth 2π periodic functions are not complete. This is demonstrated by the example on the middle of page 73 of Folland. The sequence of piecewise smooth periodic functions f_n which are defined by

$$f_n(\theta) = \frac{1}{\theta^{1/4}} \quad \text{for} \quad \frac{1}{n} \leq \theta \leq 2\pi, \quad \text{and} \quad f_n(\theta) = 0 \quad \text{for} \quad O < \theta < \frac{1}{n}$$

is Cauchy and has the limit $f(\theta) := 1/\theta^{1/4}$. This limit has square which is integrable as an improper Riemann integral.

Example 4. The 2π periodic functions whose squares have a well defined improper Riemann integral do not define a complete space.

Proof. Let $q_1, q_2, \ldots$ be an enumeration of the rational numbers between 0 and 2π. Let f denote the 2π periodic function equal to $1/\theta^{1/4}$ on $]0, 2\pi]$. Then for any q, $f(\theta - q)$ has an infinite spike at q. Let

$$F_n := \sum_{k=1}^{n} f(\theta - q_n)/2^n \,.$$

Then F_n has a good improper Riemann integral and infinite discontinuities at each of the first n rational. In addition,

$$||F_m - F_n|| \leq \sum_{k=m+1}^{n} ||f(\theta - q_n)/2^n|| \leq \sum_{k=m+1}^{\infty} ||f||/2^n = ||f||/2^m \,,$$

showing that F_n is Cauchy. The limit is clearly

$$\sum_{k=1}^{\infty} f(\theta - q_n)/2^n$$

which is infinite at every rational number, so can't possibly be square integrable in any Riemann sense. Somewhat surprisingly the function f is finite almost everywhere and is square integrable when the latter is interpreted in the sense of Lebesgue. $\diamondsuit$

Pursuing this process systematically would lead us to the general theory of integration due to Lebesgue. However, you might have the uneasy feeling that the process may never end and in fact that no matter what we do we will never be lead to a complete space. Happily such pessimism is uncalled for. There is a complete space that arises from filling in the gaps. In fact, the construction is exactly the same as the construction of the reals from the rationals, and follows the strategy indicated above. Whenever there is a Cauchy sequence which does not coverge we put in a point to be the limit. The key is that two different Cauchy sequences v_n and w_n should correspond to the same new point when $||v_n - w_n|| \to 0$. Thus many sequences correspond to the same point.

Theorem 3. *If V is a scalar product space which is not complete then there is a complete scalar product space $\mathbf{V}$ such that V is a dense subset of $\mathbf{V}$, that is, for every $\mathbf{v} \in \mathbf{V}$ there is a sequence $v_n \in V$ such that $v_n \to \mathbf{v}$. If $\mathbf{V}_1$ and $\mathbf{V}_2$ are two such completions, then they are isomorphic in the sense that there is a linear bijection $U : \mathbf{V}_1 \to \mathbf{V}_2$ which preserves scalar products and is equal to the identity on V.*

Definition. $\mathbf{V}$ is called the *completion* of V. It is determined up to isomorphic copies.

Outline of Proof. Let Γ denote the set of all Cauchy sequences of vectors in V. Define an equivalence relation $\sim$ on Γ by

$$v_n \sim w_n \quad \text{iff} \quad \lim_{n\to\infty} ||v_n - w_n|| = 0\,.$$

Let $\mathbf{W}$ be defined to be the set of equivalence classes of Cauchy sequences. Denote the equivalence class of the Cauchy sequence v_n by $[v_n]$. It is not difficult to show that if v_n, v_n^1 and v_n^2 are Cauchy sequences then so are av_n, $v_n^1 + v_n^2$, and $\langle v_n^1, v_n^2\rangle$. The last being a Cauchy sequence of complex numbers. In addition if $v_n \sim w_n$, $v_n^1 \sim w_n^1$ and $v_n^2 \sim w_n^2$ then

$$av_n \sim aw_n, \quad v_n^1 + v_n^2 \sim w_n^1 + w_n^2, \quad \text{and} \quad \langle v_n^1, v_n^2\rangle \sim \langle w_n^1, w_n^2\rangle\,.$$

This allows us to define operations of $\mathbf{W}$ by

$$[v_n^1] + [v_n^2] := [v_n^1 + v_n^2], \quad a[v_n] := [av_n], \quad \text{and} \quad \langle [v_n^1], [v_n^2]\rangle := [\langle v_n^1, v_n^2\rangle]\,.$$

It is not difficult to show that $\mathbf{W}$ is then a complete scalar product space. In fact the proof is virtually identical to the proof that the reals constructed as equivalence classes of Cauchy sequences of rationals is a complete ordered field.

Finally, there is a copy if V sitting inside $\mathbf{W}$. For each $v \in V$ the constant sequence $v, v, \ldots$ is Cauchy. Its equivalence class is the set of sequences in V which converge to v. Simply rename that equivalence class in $\mathbf{W}$ as v to give a new set $\mathbf{V}$ whose elements are either elements of V or Cauchy sequences of V which are not convergent to elements of V. It inherits the structure of complete scalar product space from $\mathbf{W}$.

For the uniqueness statement, define $U : \mathbf{V}_1 \to \mathbf{V}_2$ by setting $U(v) = v$ for $v \in V \subset \mathbf{V}_1$. Then U is linear, densely defined, and is an isometry since for all $v, w \in V$,

$$||U(v) - U(w)||_{\mathbf{V}_2} = ||v - w||_{\mathbf{V}_2} = ||v - w||_V = ||v - w||_{\mathbf{V}_1}\,. \tag{1}$$

It follows that U extends uniquely to a linear isometry $\mathbf{V}_1 \to \mathbf{V}_2$.

To show that U is onto suppose that $w \in \mathbf{V}_2$. Choose a sequence w_n in V converging to w in $\mathbf{V}_2$. Then w_n is Cauchy in V hence in $\mathbf{V}_1$ so converges to an element $z \in \mathbf{V}_1$. Then by continuity,

$$U(z) = \lim U(w_n) = \lim w_n = w$$

so w is in the image of U.

Finally for $v, w \in V$

$$\langle v, w\rangle_{\mathbf{V}_1} := \langle v, w\rangle_V := \langle v, w\rangle_{\mathbf{V}_2} := \langle Uv, Uw\rangle_{\mathbf{V}_2}. \tag{2}$$

Since the set of such v, w is sense it follows by continuity that $\langle v, w\rangle_{\mathbf{V}_1} = \langle Uv, Uw\rangle_{\mathbf{V}_2}$ for all $v, w \in \mathbf{V}_1$, proving that U preserves the scalar product. $\diamondsuit$

From the projection theorem it follows that if $\phi_1, \phi_2, \ldots, \phi_n, \ldots$ is a sequence of orthonormal vectors in a scalar product space V and $f \in \mathbf{V}$, then, for all N

$$\sum_{n=1}^{N} |\langle f, \phi_n\rangle|^2 = ||\sum_{n=1}^{N}\langle f, \phi_n\rangle\phi_n||^2 = ||f||^2 - ||f - \sum_{n=1}^{N}\langle f, \phi_n\rangle\phi_n||^2 \le ||f||^2. \tag{3}$$

In particular we have Bessel's inequality

$$\sum_{n=1}^{\infty} |\langle f, \phi_n\rangle|^2 \le ||f||^2. \tag{4}$$

Proposition. Folland Lemma 3.2. *Suppose that* $\mathbf{V}$ *is a scalar product space and that* $\phi_1, \phi_2, \ldots, \phi_n, \ldots$ *is a sequence of orthonormal vectors. Then the partial sums of* $\sum a_n\phi_n$ *form a Cauchy sequence if and only if* $\sum |a_n|^2 < \infty$. *If, in addition,* $\mathbf{V}$ *is complete then these series converge and*

$$||\sum a_n\phi_n||^2 = \sum |a_n|^2. \tag{5}$$

In particular, the Fourier series $\sum_{n=1}^{\infty}\langle f, \phi_n\rangle\phi_n$ *of any* $f \in \mathbf{V}$ *is convergent in* $\mathbf{V}$.

Notice how much better is the conclusion in the case of a complete space. This proposition is typical of how Cauchy sequences and completeness arise naturally in analysis.

Proof. Only the first assertion needs to be proved and it follows from the Pythagorean identity

$$||\sum_{M}^{N} a_n\,\phi_n||^2 = \sum_{M}^{N} |a_n|^2.$$

If $\sum |a_n|^2 < \infty$, the right hand side is bounded above by $\sum_{n>N} |a_n|^2$ which tends to zero as $N \to \infty$. This proves that the partial sums are Cauchy.

Conversely, if the partial sums are Cauchy, take $\epsilon = 1$ in the defintion of Cauchy sequence to show that there is an N_0 so that for $N_0 < M < N$, $\sum_M^N |a_n|^2 < 1$. Letting $M = N_0$ and $N \to \infty$ yields $\sum_{N_0}^{\infty} |a_n|^2 \le 1$. Therefor the $|a_n|^2$ are square summable. $\diamondsuit$

The next result is fundamental and very important even though the proof is easy.

Theorem 4. Folland 3.4. *If* $\phi_1, \phi_2, \ldots, \phi_n, \ldots$ *is a sequence of orthonormal vectors in a Hilbert space* $\mathbf{V}$ *then the following conditions are equivalent.*

i. $f \in \mathbf{V}$ and $\langle f, \phi_n\rangle = 0$ *for all* n *only if* $f = 0$.

ii. For every $f \in \mathbf{V}$, $f = \sum \langle f, \phi_n \rangle \, \phi_n$.

iii. The set of finite linear combinations of the ϕ_n is dense in $\mathbf{V}$.

iv. For every $f \in \mathbf{V}$, we have Parseval's equation

$$\sum |\langle f, \phi_n \rangle|^2 = ||f||^2 \,. \tag{6}$$

Proof. The statement i, ii, and iv correspond to Folland's a, b, and c. Thus the implications i implies ii and iv implies i are in Folland. We must show ii implies iii implies iv. It is the implication i implies ii which uses completeness.

That ii implies iii is automatic since the partial sums of the series $\sum_1^N \langle f, \phi_n \rangle \, \phi_n$ are finite linear combinations converging to f.

To show that iii implies iv suppose given a challenge number ϵ and $f \in \mathbf{V}$. Choose a finite linear combination $\sum_1^N a_n \phi_n$ such that $||f - \sum_1^N a_n \phi_n|| < \epsilon$. Bessel's inequality and the Pythagorean Theorem imply that

$$0 \le ||f||^2 - \sum_{n=1}^{\infty} |\langle f, \phi_n \rangle|^2 \le ||f||^2 - \sum_{n=1}^{N} |\langle f, \phi_n \rangle|^2 = ||f - \sum_{n=1}^{N} \langle f, \phi_n \rangle \, \phi_n||^2 \,.$$

The best approximation property of Fourier series shows that this last expression is

$$\le ||f - \sum_1^N a_n \phi_n||^2 < \epsilon^2 \,.$$

Thus $0 \le ||f||^2 - \sum_{n=1}^{\infty} |\langle f, \phi_n \rangle|^2 \le \epsilon^2$ for any ϵ. Choosing ϵ tending to 0 proves Parseval's identity. $\diamondsuit$

Definition. An orthonormal family satisfying the properties of the previous theorem is called an *orthonormal basis.* It is also called *complete.*

Both of these names are confusing. Recall that a basis of a vector space has the property that any element of the space is a finite linear combination of elements of the basis. An *orthonormal basis* is **NOT** a basis. It takes infinite combinations.

The second name is confusing since there are two utterly unrelated use of the adjective complete. The first says that Cauchy sequences converge. It has nothing to do with orthonormal sets.

Combining the above results for the orthonormal family of exponentials in the Hilbert space of square integrable functions yields the following.

Theorem 5. Riesz-Fischer I. *Let V denote the scalar product space of smooth 2π periodic functions and $\mathbf{V}$ its completion.*

i. *The orthonormal family $e^{in\theta}$ is complete in $\mathbf{V}$.*

ii. *The doubly infinite sequence c_n, $-\infty < n < \infty$, are the Fourier coefficients of an element of $\mathbf{V}$ if and only if $\sum |c_n|^2 < \infty$. In that case, the Fourier series converges in norm, that is*

$$||f - \sum_{n=-M}^{N} c_n(f) \, e^{in\theta}||^2 \;\to\; 0 \tag{7}$$

as N, M tend to infinity.

Proof. To prove completeness we verify condition iii. of the previous Theorem. Given $f \in \mathbf{V}$ and a challenge number $\epsilon > 0$ there is a $g \in V$ such $||f - g|| < \epsilon/2$.

Expand $g = \sum c_n(g)e^{in\theta}$ in its uniformly convergent Fourier series. Choose $N > 0$ so that $|g(\theta) - \sum_{-N}^{N} c_n(g)e^{in\theta}| < \epsilon/2$ for all θ. Then $||g(\theta) - \sum_{-N}^{N} c_n(g)e^{in\theta}|| < \epsilon/2$.

The triangle inequality yields

$$||f - \sum_{-N}^{N} c_n(g)e^{in\theta}|| \leq ||f - g|| + ||g(\theta) - \sum_{-N}^{N} c_n(g)e^{in\theta}|| < \epsilon/2|| < \epsilon\,.$$

This shows that f can be arbitrarily well approximated by finite linear combinations of the $e^{in\theta}$ which proves iii. $\diamondsuit$

The abstract completion works for any scalar product space. Most of the time, as in the above proof, all that one needs to know about the completion are the two facts that it is complete and that $\mathbf{V}$ is dense. Sometimes, as for the real numbers, there are concrete realizations of the completion. For the reals, the realization is as infinite decimal expansions. This can aid in thinking at times.

In the case of the scalar product space V of continuously differentiable 2π periodic functions there is also a concrete realization of the completion using the ideas of Lebesgue integral. Lebesgue integration assigns an integral to functions appreciably less regular than the Riemann integrable functions of calculus courses. It assigns an integral (possibly equal to ∞) to every nonnegative measureable function. It is unlikely that you will ever encounter a nonmeasureable function. On the other hand you are quite likely to meet functions which are not Riemann integrable. In fact the function f of example 4 is a fine example. Another example is the function which takes value 1 on the rational numbers and value 0 on the irrational numbers.

A measureable function defined on $[a, b]$ is *square integrable* if $\int_a^b |f|^2 d\theta < \infty$. Define an equivalence relation on the square integrable functions by

$$f \sim g \quad \text{iff} \quad \int_a^b |f(\theta) - g(\theta)|^2\, d\theta = 0\,. \tag{8}$$

Two functions are equivalent if and only if $f(\theta) = g(\theta)$ almost everywhere in the sense that the set of θ where they are unequal is a set of Lebesgue measure zero. Such sets are small but may be very complicated.

Similarly, two 2π periodic measureable functions are defined to be equivalent if and only iff they are equal almost everywhere. The equivalence class is determined by the values of the function on any period. This space is given a scalar product structure via

$$\langle [f], [g] \rangle := \frac{1}{2\pi} \int_0^{2\pi} f(\theta)\, \overline{g}(\theta)\, d\theta\,. \tag{9}$$

The resulting scalar product space is denoted L^2_{per}.

Theorem 6. Riesz-Fischer II. *The set L^2_{per} of equivalence classes of 2π periodic measureable functions which are square integrable on $[0, 2\pi]$ is a complete scalar product space, and, the set of equivalence classes $[f]$ for infinitely differentiable 2π periodic f is dense.*

This identifies the completion of smooth periodic functions as the set of equivalence classes of square integrable functions on $[0, 2\pi]$. This set is denoted $L^2([0, 2\pi])$. This shows how natural and important is the Lebesgue integral. The proof and the theory of Lebesgue integral is given in Math 596.

Combining the two Riesz-Fischer Theorems yields a third.

Theorem 6. Riesz-Fischer III. *The mapping $f \to c_n(f)$ which associates to a Lebesgue square integrable 2π periodic function its Fourier coefficients, establishes a one to one correspondence between L^2_{per} and the set of sequences such that $\sum |c_n|^2 < \infty$. For such f the Fourier series converges in norm, that is*

$$\lim_{M,N\to\infty} \frac{1}{2\pi} \int_0^{2\pi} |f(\theta) - \sum_{-M}^{N} c_n(f)\, e^{in\theta}|^2 \, d\theta = 0\,, \tag{10}$$

and one has Parseval's identity,

$$\int_0^{2\pi} |f(\theta)|^2 \, d\theta = \sum |c_n(f)|^2 \,. \tag{11}$$

The assertion (10) has many names; root mean square convergence, convergence in the mean, convergence in norm, and L^2 convergence. It is the most natural notion of convergence for Fourier series. Note that it does not imply convergence at a single point. Nor does pointwise convergence at all points necessarily imply convergence in the mean as shown by examples like the g_n on Folland page 83. On the other hand, uniform convergence does imply convergence as in (10).

§8. The Fourier Transform

The formula

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\xi)\, e^{ix\xi}\, d\xi \tag{1}$$

with

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

is motivated by a passage to the limit in Fourier series as on Pages 204-205 of Folland. The mapping from $\hat{f}$ to f in (1) is denoted by an inverted hat, and *temporarily*, as $\mathcal{G}$

$$\check{g}(x) := \mathcal{G} g(x) := \int_{-\infty}^{\infty} g(\xi)\, e^{ix\xi}\, d\xi \tag{3}$$

The motivation can be made into a proof. To do so, one first studies the properties of the Fourier transform $\hat{f}$ defined in (2). This is entirely analogous to our approach to Fourier series. There, the fundamental relation was motivated by trigonometric interpolation. That motivation became a proof after properties of the Fourier coefficients were understood.

The class of rapidly decreasing functions are defined on pages 87-88 of the Dym and McKean handout. We adopt the more standard notation $\mathcal{S}(\mathbf{R})$ for this space. To the material in Dym-Mckean we add the following definition.

Definition. As sequence $\{f_n\} \in \mathcal{S}(\mathbf{R})$ converges to f in $\mathcal{S}(\mathbf{R})$ if and only if for all integers m, k

$$x^m \frac{d^k}{dx^k} f_n \to x^m \frac{d^k}{dx^k} f$$

uniformly on $\mathbf{R}$.

For example, $e^{-x^2/n} f \to f$ for any $f \in \mathcal{S}(\mathbf{R})$.

Theorem 1. ≈ Folland Theorem 7.5. *If* $f \in \mathcal{S}(\mathbf{R})$ *then*

$$\widehat{xf} = i \frac{d\hat{f}}{d\xi} \tag{4}$$

$$\widehat{\frac{df}{dx}} = i\xi \hat{f} \tag{5}$$

$$\mathcal{F}(f(x-a))(\xi) = e^{ix\xi}\, \hat{f}(\xi) \tag{6}$$

$$\mathcal{F}(f(ax))(\xi) = \frac{1}{a}\, \hat{f}(\xi/a) \quad \text{for} \quad a > 0\,. \tag{7}$$

Theorem 2. *The Fourier Transform* $\mathcal{F}$ *is a continuous one to one and onto mapping form* $\mathcal{S}(\mathbf{R})$ *to itself. Its inverse is equal to* $\mathcal{G}$*, that is for all* $f \in \mathcal{S}(\mathbf{R})$

$$\mathcal{G}\mathcal{F}\, f = f\,, \tag{8}$$

One also has Plancherel's identity,

$$2\pi \int_{-\infty}^{\infty} |f(x)|^2 \, dx = \int_{-\infty}^{\infty} |\hat{f}(\xi)|^2 \, d\xi \,. \tag{9}$$

This important result is proved following pages 88-90 of Dym-McKean. The exercises on those pages are solved in the handout on the Riemannian Dominated Convergence Theorem.

Now that $\mathcal{G}$ is identified as the inverse of the mapping $\mathcal{F}$ we adopt the more standard notation $\mathcal{F}^{-1}$ for the mapping $\mathcal{G}$.

The space $\mathcal{S}(\mathbf{R})$ is a scalar product space with the scalar product

$$\langle f, g \rangle := \int_{-\infty}^{\infty} f(x)\overline{g}(x) \, dx \,. \tag{10}$$

The Plancherel identity (9) shows that $(2\pi)^{-1}\mathcal{F}$ is an isometry of $\mathcal{S}(\mathbf{R})$. In particular, it is continuos with respect to norm convergence. This is different from the continuity with respect to $\mathcal{S}(\mathbf{R})$ convergence, and, is at the heart of the following fundamental result.

Theorem 3. L^2 Theory of the Fourier Transform. *Denote by $\mathbf{V}$ the completion the scalar product space $\mathcal{S}(\mathbf{R})$. Then both the Fourier Transform, $\mathcal{F}$, and the inverse Fourier Transform, $\mathcal{F}^{-1}$, extend uniquely to continuous linear mappings of $\mathbf{V}$ to itself. The extensions, still denoted $\mathcal{F}$ and $\mathcal{F}^{-1}$, satisfy (8) and (9) for all $f \in \mathbf{V}$. In addition*

$$2\pi\langle \mathcal{F}f, \mathcal{F}g \rangle = \langle f, g \rangle \,, \quad \text{and} \quad 2\pi\langle \mathcal{F}f, g \rangle = \langle f, \mathcal{F}^{-1}g \rangle \,. \tag{11}$$

for all $f, g \in \mathbf{V}$.

Proof. All but the second identity in (11) were proved in class, 26 October. The last step is to let $h := \mathcal{F}^{-1}g$ so $g = \mathcal{F}h$. Then using the first identity yields

$$2\pi\langle \mathcal{F}f, g \rangle = 2\pi\langle \mathcal{F}f, \mathcal{F}h \rangle = \langle f, h \rangle = \langle f, \mathcal{F}^{-1}g \rangle \,.$$

$\diamondsuit$

The second identity in (11) shows that $\mathcal{F}^{-1}$ is naturally interpreted as the adjoint of $2\pi\mathcal{F}$.

It is sometimes useful to know that there is a realization of $\mathbf{V}$ as equivalence classes of Lebesgue square integrable functions on $\mathbf{R}$, that is, as $L^2(\mathbf{R})$. Note however that this interpretation causes more problems than one might think. For a square integrable function $f(x)$ it would be tempting to try to interpret Fourier Transforms as integrals as in (1) and (2). However the intgrands are typically not integrable. For example for the square integrable function $f(x) = (1 + |x|)^{-2/3}$, the integrand in (2) is not integrable. Correct interpretations require approximating f in norm by functions which are integrable. For example one has the symmetric formulas

$$f = \lim_{\epsilon \to 0+} \int_{-\infty}^{\infty} e^{-\epsilon\xi^2} \, \hat{f}(\xi) \, e^{ix\xi} \, d\xi \tag{12}$$

and

$$\hat{f} := \lim_{\epsilon \to 0+} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\epsilon x^2} \, f(x) \, e^{-ix\xi} \, dx \,. \tag{13}$$

This means that for each $\epsilon > 0$, the right hand sides are integrals which define functions of x and ξ respectively which converge in $L^2(\mathbf{R})$ to f and $\hat{f}$ respectively.

§9. A Riemannian Dominated Convergence Theorem

The Lebesgue Dominated Convergence Theorem is the basic convergence theorem of integration theory. In the treatment of Folland and in that of Dym and McKean it is used many times. However, in the context of Fourier transforms of functions in the Schwartz class $\mathcal{S}(\mathbf{R})$, all the integrals are integrals of smooth rapidly decreasing functions so can be interpreted as improper Riemann integrals, that is $\lim_{M,N\to\infty}\int_{-M}^{N} f(x)dx$. This is the same as taking improper Rieman integrals from $-\infty$ to 0 and from 0 to ∞ and adding the results. With exactly this interpretation, we proved the crucial formulas for the Fourier transforms of xf and df/dx. That the Fourier Transform is a continuous mapping of $\mathcal{S}(\mathbf{R})$ to itself follows.

In the motivation of the Fourier Inversion Formula,

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

we used Fourier series to show that if f is an infinitely differentiable function which vanish for $|x|\ge l$, and if $|x|\le l$, then

$$f(x) = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} \hat f(\xi_n)\, e^{ix\xi_n}\, \Delta\xi_n \qquad \xi_n := \frac{n\pi}{l}\,.$$

At the same time, Bessel's identity showed

$$\int_{-l}^{l} |f(x)|^2 dx = 2\pi \sum_{n=-\infty}^{\infty} |\hat f(\xi_n)|^2\, \Delta\xi_n\,.$$

One needs to pass to the limit in the Riemann sums on the right hand sides to get integrals over $]-\infty,\infty[$. Both Dym-McKean and Folland rely on Lebesgue's Dominated Convergence Theorem. This is not necessary. One can stay completely within the Riamannian context by using the following elementary analogue.

Riemannian Dominated Convergence Theorem. *Suppose that*

1. f_n, f, g are Riemann integrable on each finite interval $[0,R]$ with $R>0$,

2. For all n and $x\in[0.\infty[$, $|f_n(x)|\le g(x)$, and $\int_0^\infty g(x)\,dx<\infty$.

3. For all $R>0$, $f_n\to f$ uniformly on $[0,R]$.

Then, the improper integrals $\int_0^\infty f_n(x)\,dx$ are all finite and as $n\to\infty$,

$$\int_0^\infty f_n(x)\,dx \to \int_0^\infty f(x)\,dx\,.$$

Proof. To show that the improper integrals of the f_n exist it suffices to show that for each n the sequence $I_N := \int_0^N f_n dx$ is a Cauchy sequence. Since the improper integral of g exists we know that the sequence $\int_0^N g dx$ is Cauchy. Thus for each $\epsilon>0$ there is an N_0 so that if $N_0\le M\le N$, $\int_M^N g dx<\epsilon$. Then for the same M, N,

$$|I_N - I_M| = |\int_M^N f_n dx| \le \int_M^N |f_n| dx \le \int_M^N g dx < \epsilon\,,$$

proving that the I_N are Cauchy. At the same time this estimate shows that

$$|\int_{N_0}^{N} f_n(x)\,dx| \le \int_{N_0}^{N} g dx < \epsilon$$

for all N. Letting $N \to \infty$ shows that for all n

$$|\int_{N_0}^{\infty} f_n(x)\,dx| \le \int_{N_0}^{\infty} g(x)\,dx < \epsilon\,. \tag{1}$$

Thanks to the uniform convergence, one can choose an n_0 so that for all $n \ge n_0$,

$$\int_0^{N_0} |f_n(x) - f(x)|\,dx \le \epsilon\,. \tag{2}$$

Then, for $n \ge n_0$,

$$|\int_0^{\infty} f_n\,dx - \int_0^{\infty} f\,dx| \le |\int_0^{N_0} f_n\,dx - \int_0^{N_0} f\,dx| + \int_{N_0}^{\infty} |f_n|\,dx + \int_{N_0}^{\infty} |f|\,dx < 3\epsilon\,.$$

$\diamondsuit$

The pattern of this proof is worth remembering. The domination is used to control the tail as in (1). Then a separate argument is used for the part of the integral in the bounded interval that remains.

Example 1. If $f \in \mathcal{S}$ then

$$\lim_{\epsilon \to 0} \int_{-\infty}^{\infty} e^{-\epsilon x^2} f(x)\,dx = \int_{-\infty}^{\infty} f(x)\,dx\,. \tag{3}$$

Proof. It suffices to show that for any sequence $\epsilon_n \to 0$

$$\int_{-\infty}^{\infty} e^{-\epsilon_n x^2} f(x)\,dx \to \int_{-\infty}^{\infty} f(x)\,dx\,.$$

It suffices to apply the Riemannian Dominated Convergence Theorem with $f_n(x) := e^{-\epsilon_n x^2} f(x)$ and $g(x) = |f(x)|$. $\diamondsuit$

The next example shows that infinite Riemann sums converge to integrals. This is needed to justify our derivation of the Fourier Inversion Formula.

Example 2. If $f \in \mathcal{S}$ then

$$\lim_{h \to 0+} \sum_{n=-\infty}^{\infty} f(nh)\,h = \int_{-\infty}^{\infty} f(x)\,dx\,. \tag{4}$$

Proof. It suffices to show that

$$\sum_{n=1}^{\infty} f(nh)\, h \to \int_0^{\infty} f(x)\, dx \quad \text{and} \quad \sum_{n=-\infty}^{-1} f(nh)\, h \to \int_{-\infty}^{0} f(x)\, dx\,. \tag{5}$$

We prove the first. The second is analogous.
If $h_k \to 0+$ define $f_k(x)$ by

$$f_k(x) := f(nh_k) \quad \text{for} \quad x \in [(n-1)h_k, nh_k[\,.$$

Then

$$\sum_{n=1}^{\infty} f(nh_k)\, h_k = \int_0^{\infty} f_k(x) dx\,.$$

The strategy is to prove that $\int_0^\infty f_k\, dx \to \int_0^\infty f\, dx$ by applying the dominated convergence theorem. Since $|f_k(x) - f(x)| \le h_k \max_{\mathbf{R}} |f'(x)|$ it follows that $f_k \to f$ uniformly on $[0, \infty[$ as $k \to \infty$. Choose $c > 0$ so that $|f(x)| \le c/(1+x^2)$ and set $g := c/(1+x^2)$. Then $\int_0^\infty g\, dx < \infty$, and, for any k and $x \in [0, \infty[$ there is a unique n so that $x \in [(n-1)h_k, nh_k[$ and then

$$|f_k(x)| = |f(nh_k)| \le g(nh_k) \le g(x)\,, \tag{6}$$

the last inequality because g is monotone decreasing and $nh_k > x$. The dominated convergence theorem finishes the proof. $\diamondsuit$

The next result is needed in the justification of the Fourier Inversion Formula and Plancherel's Theorem, when one passes from the case of functions which vanish outside a bounded interval to the case of arbitrary functions by inserting a cutoff function $f_n := \chi(x/n) f$ and then passing to the limit. For $f \in \mathcal{S}$ one has $f_n \to f$ in $\mathcal{S}$ and more generally,

$$f_n\, e^{-ix\xi} \to f\, e^{-ix\xi}\,, \quad \hat{f}_n\, e^{ix\xi} \to \hat{f}\, e^{ix\xi}\,, \quad \text{and,} \quad |f_n|^2 \to |f|^2$$

in $\mathcal{S}$.

Example 3. If $f_n \to f \in \mathcal{S}$ then

$$\int_{-\infty}^{\infty} f_n(x)\, dx \to \int_{-\infty}^{\infty} f(x)\, dx\,. \tag{7}$$

Proof. It suffices to show that

$$\int_0^{\infty} f_n(x)\, dx \to \int_0^{\infty} f(x)\, dx \quad \text{and} \quad \int_{-\infty}^{0} f_n(x)\, dx \to \int_{-\infty}^{0} f(x)\, dx\,.$$

We prove the first. The second is analogous.
Since f_n and f belong to $\mathcal{S}$, they are continuous so the first hypothesis of the dominated convergence theorem is satisfied. Since $f_n \to f$ in $\mathcal{S}$ it follows that f_n converges uniformly to f on $\mathbf{R}$ verifying the second hypothesis.
It remains to find a dominating function g. Convergence in $\mathcal{S}$ implies the uniform convergence $(1+x^2)f_n \to (1+x^2)f$ on $\mathbf{R}$. Thus, there is an n_0 so that for all $n > n_0$,

$$|(1+x^2)f_n - (1+x^2)f| \le 1 \quad \text{for all} \quad x \in \mathbf{R}\,.$$

Let $M := \max_{\mathbf{R}} |f(x)|$. The triangle inequality for real numbers shows that for $n > n_0$,

$$|(1+x^2)f_n(x)| \le 1 + |(1+x^2)f| \le 1 + M \quad \text{for all} \quad x \in \mathbf{R}\,. \tag{8}$$

This proves the desired domination with $g := (1+M)/(1+x^2)$. $\diamondsuit$