The Great American Mathematical Physics Course Outline

1 Theorem Summary

1.1 Linear Algebra Theorems

1.1.1 Dimension Theorem

Let $T : V^n \to W^m$ be a linear map from an n-dimensional vector space to an m-dimensional vector space. Then the following are true

(a) $\ker T = \{x \in V^n \mid Tx = 0\}$ is a subspace of V^n.

(b) $\operatorname{ran} T = \{x \in W^m \mid \exists\, y \in V \ni x = Ty\}$ is a subspace of W^m.

(c) If we let $\operatorname{nuly} T = \dim \ker T$ and $\operatorname{rank} T = \dim \operatorname{ran} T$, then $\operatorname{nuly} T + \operatorname{rank} T = n$.

1.1.2 Invertible Operator Theorem

Let $T \in \mathcal{L}(V)$ be a linear operator on an n-dimensional vector space. The following conditions are equivalent (if any one is true, all the others must be true):

(a) T is invertible (that is $T^{-1} : V \to V$ is well-defined),

(b) T is injective,

(c) $\ker T = \{0\}$ ("the kernel is trivial"),

(d) $\operatorname{nuly} T = 0$,

(e) $\operatorname{rank} T = n$ ("T has full rank"),

(f) $\operatorname{ran} T = V$, (equivalently, T is surjective)

(g) T maps bases to bases,

(h) $\det T \neq 0$, and

(i) $0 \notin \sigma(T)$ ($\sigma(T)$ is the spectrum of T).

1.1.3 Cauchy-Schwarz Inequality

If v and w are vectors in an inner product space V, then

$$|\langle v|w\rangle| \leq \|v\|\|w\|,$$

where $\|v\| = \sqrt{\langle v|v\rangle}$. Moreover, equality holds iff v or w is a (possibly zero) multiple of the other. An immediate consequence of this theorem is the triangle inequality, which can be written in two forms: either $\|v + w\| \leq \|v\| + \|w\|$ or $d(v, w) \leq d(v, x) + d(w, x)$, where $d(v, w) = \|v - w\|$. The first shows that one side of a triangle cannot be longer the sum of lengths of the other two sides, whereas the second states that the shortest distance between two points is a straight line.

1.2 Spectral Theorem

We have encountered no less than 6 different versions of spectral theorem, which have been a unifying theme in the course. They are listed below.

1.2.1 Ultra Baby Spectral Theorem

A matrix A is orthogonally diagonalizable if and only if it is symmetric. In other words, if A is symmetric, then it can be written $A = ODO^{-1}$, where O is orthogonal and D is real and diagonal. The entries of D are the eigenvalues of A and the columns of O are the corresponding unit eigenvectors. Conversely, if $A = ODO^{-1}$ where D is real and diagonal and O is orthogonal, then A is symmetric. Moreover, any real-valued function of A can be calculated by the formula $f(A) = Of(D)O^{-1}$.

1.2.2 Super Baby Spectral Theorem

A matrix A is unitarily diagonalizable with real eigenvalues if and only if it is Hermitian. In other words, if A is Hermitian, then it can be written $A = UDU^{-1}$, where U is unitary and D is real and diagonal. The entries of D are the eigenvalues of A and the columns of U are the corresponding unit eigenvectors. Conversely, if $A = UDU^{-1}$ where D is real and diagonal and U is unitary, then A is Hermitian. Moreover, any function of A can be calculated by the formula $f(A) = Uf(D)U^{-1}$.

1.2.3 Baby Spectral Theorem

A matrix A is unitarily diagonalizable if and only if it is normal. In other words, if A is normal, then it can be written $A = UDU^{-1}$, where U is unitary and D is diagonal. The entries of D are the eigenvalues of A and the columns of U are the corresponding unit eigenvectors. Conversely, if $A = UDU^{-1}$ where D is diagonal and U is unitary, then A is normal. Moreover, any function of A can be calculated by the formula $f(A) = Uf(D)U^{-1}$.

1.2.4 Spectral Theorem (Finite Dimensions)

Let A be a Hermitian operator on an n-dimensional inner product space $\mathbb{V}^n$ over $\mathbb{C}$. Then A has n linearly independent eigenvectors, which can be chosen to be mutually orthogonal, and all its eigenvalues are real. Moreover, the action of A on any vector v can be represented by

$$Av = \sum_{\lambda \in \sigma(A)} \lambda \langle v_\lambda | v \rangle v_\lambda = \sum_{\lambda \in \sigma(A)} \lambda P_\lambda v,$$

where v_λ is a unit eigenvector corresponding to eigenvalue λ and P_λ is the projection onto the eigenspace. Finally, any function of A can be computed by the following formula:

$$f(A)v = \sum_{\lambda \in \sigma(A)} f(\lambda) \langle v_\lambda | v \rangle v_\lambda = \sum_{\lambda \in \sigma(A)} f(\lambda) P_\lambda v.$$

1.2.5 Spectral Theorem (Infinite Dimensions)

Let A be an operator on an infinite-dimensional (Hilbert) inner-product space $\mathbb{V}^\infty$ over $\mathbb{C}$. If A is self-adjoint (that is, Hermitian on a *sufficiently large subspace of* $\mathbb{V}^\infty$) **and** A has at most a countable infinity of eigenvalues, then the eigenvalues of A are real and its eigenvectors can be chosen to form an orthonormal basis for $\mathbb{V}^\infty$. Moreover, the action of A on any vector v can be represented by

$$Av = \sum_{\lambda \in \sigma(A)} \lambda \langle v_\lambda | v \rangle v_\lambda = \sum_{\lambda \in \sigma(A)} \lambda P_\lambda v,$$

where v_λ is a unit eigenvector corresponding to eigenvalue λ. Finally, any function of A can be computed by the following formula:

$$f(A)v = \sum_{\lambda \in \sigma(A)} f(\lambda) \langle v_\lambda | v \rangle v_\lambda = \sum_{\lambda \in \sigma(A)} f(\lambda) P_\lambda v.$$

1.2.6 Generalized Spectral Theorem

Let A be an operator on a (Hilbert) inner-product space $\mathbb{V}$ over $\mathbb{C}$. If A is self-adjoint (that is, Hermitian on a *sufficiently large subspace of* $\mathbb{V}$) then the eigenvalues of A are real, and the action of A on any vector v can be represented by

$$Av = \int_{\sigma(A)} \lambda dP_\lambda v.$$

The operators dP_λ are called orthogonal projection operators or projectors, and $dP_\lambda v$ is a generalization of $\langle v_\lambda | v \rangle v_\lambda$. Moreover, any function of A can be computed by the following formula:

$$f(A)v = \int_{\sigma(A)} f(\lambda) dP_\lambda v.$$

1.3 Fourier Series

Let a_n, b_n and c_n be the unnormalized Fourier coefficients of f, defined by the equations

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(nx) f(x) dx; \qquad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} \sin(nx) f(x) dx; \qquad c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-inx} f(x) dx.$$

Notice that $n \in \{0, 1, 2, 3, \ldots\}$ for a_n, $n \in \mathbb{N}$ for b_n, and $n \in \mathbb{Z}$ for c_n. Also notice, that with the exception of a_0, all the above formulas simply give the scalar projection (i.e., component) of f onto the non-unit vectors $\cos nx$, $\sin nx$, and e^{inx}. The reason that a_0 is not a projection is so that its formula looks like the formula for the other a_n, but the price we pay is that a_0 gets special treatment in the formula for the Fourier series itself:

$$f_n = \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos kx + b_k \sin kx = \sum_{k=-n}^{n} c_k e^{ikx}.$$

The subscript n denotes that this is the nth partial sum, a notation we will use in all of the following theorems. We can normalize the orthogonal functions used in defining the Fourier series, which give rise to two sets of functions:

$$\{e_n\}_{n=-\infty}^{\infty} = \left\{\frac{e^{inx}}{\sqrt{2\pi}}\right\}_{n=-\infty}^{\infty} \quad \text{and} \quad \{E_n\}_{n=1}^{\infty} = \left\{\frac{1}{\sqrt{2\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \ldots\right\}.$$

Using these, we define the normalized Fourier coefficients of f, called $\tilde{a}_n$, $\tilde{b}_n$, and $\tilde{c}_n$ via projection:

$$\tilde{a}_n = \langle E_{2n+1}\,|f\rangle\,; \quad \tilde{b}_n = \langle E_{2n}\,|f\rangle\,; \quad \tilde{c}_n = \langle e_n\,|f\rangle\,.$$

As these are the components of f relative the unit vectors, we can rewrite f_n as

$$f_n = \frac{\tilde{a}_0}{\sqrt{2\pi}} + \sum_{k=1}^{n} \frac{\tilde{a}_k \cos kx}{\sqrt{\pi}} + \frac{\tilde{b}_k \sin kx}{\sqrt{\pi}} = \sum_{k=-n}^{n} \frac{\tilde{c}_k e^{ikx}}{\sqrt{2\pi}}.$$

Since $f_n = f_n$ no matter which way we write it, these formulas give use the relationship between the normalized and unnormalized coefficients.

1.3.1 Dirichlet Theorem

Suppose f obeys the following conditions

(a) f is single valued on $[-\pi, \pi]$,

(b) f has a finite number of discontinuities on $[-\pi, \pi]$,

(c) f has a finite number of extrema on $[-\pi, \pi]$,

(d) $\displaystyle\int_{-\pi}^{\pi} |f(x)|dx < \infty$, and

(e) and f is periodic of period 2π on the real line.

Then $f_n \to f$ *pointwise* where f is continuous, and $f_n(x) \to (f(x^+) + f(x^-))/2$ where f is discontinuous. In other words, at each point x the sequence of numbers $f_n(x)$ converges to the number $f(x)$ when f is continuous at x and converges the average of the left and right limits if f is discontinuous there. If f meets conditions (a)–(d) but not (e), then $f_n \to f$ pointwise for $x \in (-\pi, \pi)$.

1.3.2 Differentiation Theorem

Here we collect a few disparate facts about Fourier series and differentiation:

(a) If f has a continuous first derivative $[-\pi, \pi]$, then f_n converges pointwise to f on the interval $(-\pi, \pi)$ (whether or not f and f' obey Dirichlet's Theorem). However, the Fourier series of f' may not converge.

(b) If f' has a (pointwise) convergent Fourier series, then f has a convergent Fourier series, and the term-by-term derivative of f's Fourier series is f''s Fourier series. In particular, if f' obeys the first four conditions of Dirichlet's theorem, then term-by-term differentiation is justified.

1.3.3 Decay Theorem

The following is perhaps the simplest of a number of results which could be called decay theorems, being essentially a combination of Dirichlet's Theorem and the Differentiation Theorem. However, note that, for all its complexity, it still does not give us a sharp iff result. Contrast this with the nice, sharp corollary to Parseval's Theorem, below. This is another reason why spaces of square-integrable functions are often more useful than spaces of continuous functions.

Consider a function f on $[-\pi, \pi]$. Let l be the largest integer for which every $f^{(i)} \in \{f, f', f'', \ldots, f^{(l-1)}\}$ satisfies

(a) $f^{(i)}$ is continuous, and

(b) $f^{(i)}$ obeys periodic boundary conditions $f^{(i)}(-\pi) = f^{(i)}(\pi)$.

Suppose further that $f^{(l)}$ obeys the first four conditions of the Dirichlet Theorem. Then there exists a positive real number q_f (which depends on the function f) for which

$$|a_n| \leq \frac{q_f}{n^{l+1}}; \quad |b_n| \leq \frac{q_f}{n^{l+1}}; \quad |c_n| \leq \frac{q_f}{n^{l+1}}.$$

In other words, if f and its first $l-1$ derivatives are continuous and periodic on $[-\pi, \pi]$, and f's lth derivative is "nice", then the coefficients of f decay like $1/n^{l+1}$ for large n. Conversely, if the Fourier coefficients of f obey a formula like the preceding for some integer l, then f and its first $l-1$ derivatives are continuous (but we can't conclude much about the lst derivative). By the differentiation theorem, the first $l-2$ derivatives (notice it is **not** $l-1$, the prime deficiency of this theorem) have pointwise convergent Fourier series.

1.3.4 Parseval's Theorem—Coefficients Version

If $f \in L^2([-\pi, \pi])$, i.e. $\int_{-\pi}^{\pi} |f(x)|^2 dx < \infty$, then the following formulas hold:

$$\langle f \rangle := \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 dx = \frac{|a_0|^2}{4} + \frac{1}{2} \sum_{n=1}^{\infty} \left(|a_n|^2 + |b_n|^2 \right) = \sum_{n=-\infty}^{\infty} |c_n|^2$$

and

$$\|f\|^2 := \int_{-\pi}^{\pi} |f(x)|^2 dx = |\tilde{a}_0|^2 + \sum_{n=1}^{\infty} \left(|\tilde{a}_n|^2 + |\tilde{b}_n|^2 \right) = \sum_{n=-\infty}^{\infty} |\tilde{c}_n|^2.$$

Corollary: A function and its first l derivatives are in $L^2([-\pi, \pi])$ iff

$$\sum_{n=0}^{\infty} |n^l c_n|^2 < \infty.$$

This is simply a result of Parseval's theorem and integration by parts. Since the integral of any ordinary function is continuous, we also have that f and its first $l-1$ derivatives are continuous if $l > 0$. The lth derivative is square integrable, but it may or may not be continuous (or any other nice property).

1.3.5 Parseval's Theorem—Basis Version

The two sets of functions $\{e_n\}$ and $\{E_n\}$ defined at the start of this section are both orthonormal bases for $L^2([-\pi,\pi])$ (there are many others, of course). If $f \in L^2([-\pi,\pi])$, then $f_n \to f$ in the L^2 sense (even if f does not meet the criteria of the Dirichlet Theorem). In other words, if $\int_{-\pi}^{\pi} |f(x)|^2 dx < \infty$, and we define the error in the n^{th} partial sum as

$$\mathcal{E}_n := \|f - f_n\| = \sqrt{\int_{-\pi}^{\pi} |f(x) - f_n(x)|^2 dx},$$

then $\mathcal{E}_n \to 0$ as $n \to \infty$. Finally, the two sets of functions

$$\left\{e^{inx}\right\}_{n=-\infty}^{\infty} \quad \text{and} \quad \{1, \sin x, \cos x, \sin 2x, \cos 2x, \ldots\}$$

are both *complete sets*—meaning each is an orthogonal basis but is not orthonormal.

1.3.6 Parseval's Theorem—High Falutin' Version (for Cocktail Parties Only)

Fourier series are an isometric isomorphism from the (Hilbert) space $L^2([-\pi,\pi])$ of square-integrable functions to the (Hilbert) space l^2 ("little el-two") of square-summable sequences.

1.3.7 Other Domains

For simplicity, the above discussion was all in terms of the domain $[-\pi,\pi]$. However, all of the above results remain valid for any symmetric domain $[-L,L]$, as long as we replace e^{ikx} but $e^{i\pi kx/L}$ and similarly for the sines and cosines. Of course, all normalizations of π, 2π, $\sqrt{\pi}$, and $\sqrt{2\pi}$ become L, $2L$, $\sqrt{L}$, and $\sqrt{2L}$.

1.4 Fourier Transform

In the following theorems, we assume that $f : \mathbb{R} \to \mathbb{C}$ (i.e, f is a complex valued function on the entire real line) and consider the symmetric Fourier pair

$$\hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} f(x) dx \qquad f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{+ikx} \hat{f}(k) dk.$$

As long as f is integrable on $\mathbb{R}$, then its Fourier transform is defined. However, the inverse transform of $\hat{f}$ may not exist or be equal to f for badly behaved functions.

1.4.1 Fourier Integral Theorem

If f satisfies

(a) conditions (a)–(c) of the Dirichlet Theorem on every finite domain $[-L,L]$, and

(b) $\displaystyle\int_{-\infty}^{\infty} |f(x)| dx < \infty,$

then $\int_{-\infty}^{\infty} |\hat{f}(k)| dk < \infty$. This means the inverse transform is well defined. In particular, $\check{\hat{f}}(x) = f(x)$ where f is continuous and the average of the left and right limits where it is discontinuous. This is the analog of the Dirichlet Theorem.

1.4.2 Multiplication and Differentiation Theorem

The relationship between differentiation and the Fourier transform is both simpler and more complicated than the relationship between differentiation and Fourier series. On the one hand, no amount of differentiability guarantees that $\hat{f}$ can be inverse transformed back to f (unlike Fourier series, where a continuos first derivative guarnteed convergence of the series), because behaviour at infinity can result in various problems. On the other hand, since $\hat{f}(k)$ is defined for every real k, there is a symmetric relationship between differentiation and multiplication. In particular, if both f and f' are integrable (i.e., $\int_{-\infty}^{\infty} |f(x)|dx < \infty$ and $\int_{-\infty}^{\infty} |f'(x)|dx < \infty$), then they both have a Fourier transforms and

$$\left(\frac{df}{dx}\right)\hat{}\,(k) = ik\hat{f}(k)$$

Conversely, if both $f(x)$ and $xf(x)$ are integrable, than $\hat{f}$ is differentiable, and

$$i\frac{d}{dk}\hat{f}(k) = (xf)\hat{}(k).$$

Turning differentiation into multiplication is what makes the Fourier transform so useful in solving PDEs. While these formulas can be proven directly (for "nice" functions the first via integration by parts, the second by differentiation under the integral sign), they can also be understood in terms the Generalized Spectral Theorem applied to the differentiation operator $\hat{p} = -i\frac{d}{dx}$. This is because $e_k = e^{ikx}/\sqrt{2\pi}$ is an almost-eigenvector of $\hat{p}$ with eigenvalue k, but it is not square-integrable. Thus, the Fourier inversion formula can be written as

$$f(x) = \int_{-\infty}^{\infty} \frac{e^{ikx}}{\sqrt{2\pi}}\,\hat{f}(k)dk = \int_{-\infty}^{\infty} e_k \langle e_k | f\rangle\, dk = \int_{-\infty}^{\infty} k^0 dP_k f = If,$$

where $I = \hat{p}^0$ is the identity operator, and dP_k is our one example of the projectors of the generalized spectral theorem. Moreover, we have that

$$\hat{p}f(x) = -i\frac{df}{dx} = \int_{-\infty}^{\infty} ke_k \langle e_k | f\rangle\, dk = \int_{-\infty}^{\infty} kdP_k f,$$

which simultaneously expresses both of the above relationships.

1.4.3 Decay Theorem

Suppose f and its first $l-1$ derivatives are continuous and satisfy the conditions of the Fourier Integral Theorem, and $f^{(l)}$ obeys the hypotheses of the Fourier Integral Theorem. The there exists a positive real number q_f such that

$$|\hat{f}(k)| \leq \frac{q_f}{k^{l+1}}.$$

In other words, $\hat{f}(k)$ decays like $1/k^{l+1}$ for large k. Conversely, if $\hat{f}(k)$ decays like $1/k^{l+1}$ for some l, then f has at least $l-1$ continuous, integrable derivatives.

1.4.4 Parseval's Theorem

If $f \in L^2(\mathbb{R})$, i.e., $\int_{-\infty}^{\infty} |f(x)|^2 dx < \infty$, then $\hat{f}(k)$ and $\check{\hat{f}}(x)$ well-defined *whether or not* f *satisfies the the hypotheses of the Fourier Integral Theorem.* Moreover, $\hat{f} \in L^2(\mathbb{R})$ and $\|f\|^2 = \|\hat{f}\|^2$. In equations,

$$\int_{-\infty}^{\infty} |f(x)|^2 dx = \int_{-\infty}^{\infty} |\hat{f}(k)|^2 dk.$$

This shows that the Fourier transform is a unitary operator on $L^2(\mathbb{R})$. Moreover, f has $l-1$ derivatives which are continuous, integrable, and square-integrable, and an lst derivative which is merely square integrable (although it **may** be continuous) iff $\int_{-\infty}^{\infty} |k^l \hat{f}(k)|^2 dk < \infty$.

1.4.5 Convolution Theorem

Let $f : \mathbb{R} \to \mathbb{C}$ and $g : \mathbb{R} \to \mathbb{C}$ both satisfy the conditions of the Fourier Integral Theorem, and put $h = fg$. It follows that h also obeys the conditions of the Fourier Integral Theorem, and the Fourier transform of h can be computed by the convolution integral:

$$\hat{h}(k) = \hat{f}(k) \star \hat{g}(k) =: \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(k-z)\hat{g}(z)dz.$$

Similarly, if H is the convolution of $f(x)$ and $g(x)$, then its Fourier transform is $\hat{f}(k)\hat{g}(k)$:

$$H(x) := f(x) \star g(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x-z)g(z)dz \quad \Rightarrow \quad \hat{H}(k) = \hat{f}(k)\hat{g}(k).$$

2 Eigenvectors

Eigenvectors are solutions to the eigenvector equation, $Av = \lambda v$, where A is a matrix or linear operator, and λ is a scalar. Eigenvectors are vectors which are **not** rotated when the operator/matrix acts on them.

2.1 Finite Dimensional Case

If the vector space has dimension $n < \infty$, we first find the eigenvalues, then the eigenvectors. Let $p_\lambda = \det(A - \lambda I)$, where I is the identity matrix. Notice that p_λ is an polynomial of degree n in λ, called the characteristic polynomial. The eigenvalues are the solutions of the equation $p_\lambda = 0$. We then find the solutions of $(A - \lambda I)v_\lambda = 0$ by row-reduction. While any solution v_λ is an eigenvector, in most applications we need to normalize it by dividing by its norm.

2.2 Infinite Dimensional Case

The infinite dimensional case is more complex. Typically we have to first find potential eigenvectors by solving a differential equation. We then apply boundary conditions, which

will restrict us to certain eigenvalues. At the end we normalize to get our eigenvectors. All potential eigenvectors which obey the boundary conditions must be included in the list of eigenvectors in order to use the spectral theorem. We studied in detail the eigenvectors of the operator $H = -\frac{d^2}{dx^2}$.

2.2.1 A-type Boundary Conditions

A-type boundary conditions, or periodic boundary conditions as they are more formally known, corresponds to a problem on a circle. For an n-th order differential operator, we must impose $\psi_\lambda(-\pi) = \psi_\lambda(\pi)$ and the same on all derivatives up to order $n-1$. Since $n = 2$ for H, we must also have $\psi'_\lambda(-\pi) = \psi'_\lambda(\pi)$. The result is $\lambda = n^2$ for $n \in \{0, 1, 2, \ldots\}$, and that non-zero eigenvalues are order two (meaning they have two linearly independent eigenvectors). The eigenvectors are nothing but the Fourier basis:

$$\psi_0 = \frac{1}{\sqrt{2\pi}} \quad \psi_{n,1}(x) = \frac{\cos(nx)}{\sqrt{\pi}} \quad \text{and} \quad \psi_{n,2}(x) = \frac{\sin(nx)}{\sqrt{\pi}}.$$

Notice that we have switched to the conventional labeling where we label according to where in the list the eigenvector appears, even though the actual eigenvalue is n^2 (so the notation really should be ψ_{n^2} in order to be consistent with ψ_λ).

2.2.2 B-type Boundary Conditions

B-type boundary conditions, or Dirichlet boundary conditions as they are more formally known, correspond to holding the variable (temperature, displacement, wavefunction) fixed at the end points. The eigenvectors can be written in two ways:

$$\psi_{n-1/2} = \frac{\cos\left((n-1/2)x\right)}{\sqrt{\pi}}; \quad \psi_n = \frac{\sin nx}{\sqrt{\pi}}$$

or

$$\psi_n = \frac{\cos(nx/2)}{\sqrt{\pi}} \quad (n \text{ odd}) \quad \psi_n = \frac{\sin(nx/2)}{\sqrt{\pi}} \quad (n \text{ even})$$

for $n \in \mathbb{N}$. In the first way of writing the eigenvalues are $(n-1/2)^2$ and n^2, while in the second notation $\lambda = n^2/4$. Notice that $\lambda = 0$ is **not** an eigenvalue.

2.2.3 C-type Boundary Conditions

C-type boundary conditions, or Neumann boundary conditions as they are more formally known, corresponds to holding the gradient of the variable (e.g., heat flow for temp, tension for oscillating string) fixed at the end points. The eigenvectors can be written in two ways:

$$\psi_0 = \frac{1}{\sqrt{2\pi}}; \quad \psi_{n-1/2} = \frac{\sin\left((n-1/2)x\right)}{\sqrt{\pi}}; \quad \psi_n = \frac{\cos nx}{\sqrt{\pi}},$$

or

$$\psi_0 = \frac{1}{\sqrt{2\pi}}; \quad \psi_n = \frac{\sin(nx/2)}{\sqrt{\pi}} \quad (n \text{ odd}) \quad \psi_n = \frac{\cos(nx/2)}{\sqrt{\pi}} \quad (n \text{ even})$$

for $n \in \mathbb{N}$. In the first way of writing the eigenvalues are 0, $(n-1/2)^2$ and n^2, while in the second notation $\lambda = n^2/4$. Notice that $\lambda = 0$ is an eigenvalue, just like for A-type eigenfunctions.

3 Useful Tools

3.1 Projection

Projection means to find the part of one vector which is parallel to another vector or the space spanned by a set of vectors. Large portions of this course can be understood as an extended exercise in projection.

3.1.1 The Projection Formula (Vector on Vector)

Given two vectors v and w in an inner product space, the projection of v onto w is given by

$$v_{\|} = \frac{\langle w\,|v\rangle}{\langle w\,|w\rangle}\, w.$$

Notice that $v_{\|}$ is parallel to w, not v. If w happens to be a unit vector, then the denominator is one and the formula simplifies to

$$v_{\|} = \langle w\,|v\rangle\; w \qquad (w \text{ a unit vector.})$$

We then have that the component of v perpendicular to w is

$$v_{\perp} = v - v_{\|}.$$

We often re-express the above two formulas by means of the projection operator $P_{|v>}$ and $P_{|v>^{\perp}}$, which eat vectors and spit out the components parallel to and perpendicular to v. In bra-ket notation, this is written

$$P_{|v>} = \frac{|v\rangle\,\langle v|}{\langle v\,|v\rangle} \qquad \text{and} \qquad P_{|v>^{\perp}} = I - P_{|v>}$$

Notice that the formula for the unnormalized Fourier coefficients a_n and b_n is simply the projection formula in action, since $\pi = \langle \cos nx\,|\cos nx\rangle = \langle \sin nx\,|\sin nx\rangle$.

3.1.2 Projection onto a Space

If we are projecting onto a space higher than dimension one, we must first find a basis for that space which is orthonormal (or at least orthogonal). We then apply the projection formula to each basis vector in turn and sum the results:

$$v_{\|} = \langle e_1\,|v\rangle\; e_1 + \langle e_2\,|v\rangle\; e_2 + \langle e_3\,|v\rangle\; e_3 + \ldots.$$

(If the basis $\{b_n\}$ is merely orthogonal, we must divide each term by $\langle b_n\,|b_n\rangle$ as in projection formula for vector on vector.) Fourier series are an example of this formula. This relationship can also be written in terms of operators, namely

$$P_{\text{space spanned by } e_1 \text{ through } e_k} = \sum_{i=1}^{k} P_{|e_i>} = \sum_{i=1}^{k} |e_i\rangle\,\langle e_i|$$

3.1.3 Gram-Schmidt

Gram-Schmidt is an iterative process to generate an orthonormal basis from any other basis. Let $\{v_i\}_{i=1}^n$ be non-orthonormal basis. Define

$$e_1 = \frac{v_1}{\|v_1\|}.$$

If we have found $e_1, \ldots e_{i-1}$, then

$$e_i = \frac{b_i}{\|b_i\|}, \qquad \text{where } b_i = v_i - \sum_{k=1}^{i-1} \langle e_k | b_i \rangle \; e_k.$$

The set $\{e_i\}_{i=1}^n$ is an orthonormal basis for the space spanned by $\{v_i\}_{i=1}^n$.

3.2 Index Notation/δ Function

It may seem a bit strange to combine these two topics, but there is a powerful analogy between the δ function and the Kronecker δ.

3.2.1 Index Notation: Definitions

In index notation, we write a_{ij} for the matrix A. We imagine i and j ranging from 1 to n to gives us all possible entries of the matrix A. The first subscript refers to the row, the second to the column. The identity matrix is represent by a special symbol, called the Kronecker δ:

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}.$$

Vectors are matrices with only one row or column, so they just get a single index.

3.2.2 Index Notation: Examples

The product equation $C = AB$ can be written:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

If the matrix A acts on a vector v, we get $w = Av$, or

$$w_i = \sum_{j=1}^{n} A_{ij} v_j.$$

If $A = I$, we get $v = Iv$, or

$$v_i = \sum_{j=1}^{n} \delta_{ij} v_j.$$

Notice that we can think of the sum over j as "forcing" i to be equal to j because of the Kronecker delta,, so we end with v_i. Transpose interchanges rows and columns, so

$$(c^T)_{ij} = c_{ji}.$$

3.2.3 δ Function

One definition of the δ function is

$$\delta(x-x_0) = \begin{cases} \infty & \text{if } x = x_0 \\ 0 & \text{if } x \neq x_0 \end{cases} \quad \text{and} \quad \int_{-\infty}^{\infty} \delta(x-x_0)dx = 1.$$

Notice the similarity to the definition of the Kronecker δ. The variables x and x_0 are playing the role the indices i and j played for the Kronecker δ. An immediate consequence of the definition is that

$$f(x_0) = \int_{-\infty}^{\infty} f(x)\delta(x-x_0)dx.$$

This is the analog of the equation $v_i = \sum_j \delta_{ij} v_j$. We also studied two other definitions of δ. One was in terms of the Fourier transform:

$$\delta(x-x_0) = \left(\frac{e^{-ikx_0}}{\sqrt{2\pi}}\right)^{\check{}} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ik(x-x_0)}dk.$$

The other was in terms of Gaussians:

$$\delta(x-x_0) = \lim_{\sigma\to 0} \frac{\exp\left(-(x-x_0)^2/\,2\sigma^2\right)}{\sqrt{2\pi}\sigma},$$

where the limit means a limit *in the sense of distributions* (see next section).

3.3 Distributions

The δ "function" is the prototype for a new mathematical object called a distribution or generalized function. Distributions, rather than assigning complex numbers to real numbers (or vectors in $\mathbb{R}^n$), assign numbers to certain classes of *functions* on the reals. This is typically written in the form of an integral

$$T(f) = \int_{-\infty}^{\infty} t(x)f(x)dx$$

where $t(x)$ is some *representation* of the distribution T. However, as in the case of δ, $t(x)$ may not be any sort of function in the ordinary sense. The functions which a distribution maps to a number are called test functions. Different spaces of test functions can be defined (leading to different types of distributions), but the two most common spaces are the "space of smooth functions of compact support"

$$\mathcal{D} = \{f \mid f \text{ is smooth and equals zero outside of some finite-length interval}\},$$

and the "Schwartz space"

$$\mathcal{S} = \left\{f \mid f \text{ is smooth and } \max_{x\in\mathbb{R},\ k\in\mathbb{N}} \left|x^k f(x)\right| < \infty\right\}$$

In other words, $\mathcal{D}$ consists of smooth (i.e., infinitely differentiable) functions which are literally zero when far enough away from the origin, whereas $\mathcal{S}$ consists of smooth functions

which do not necessarily vanish but go to zero faster than any power of x as $x \to \pm\infty$. Clearly these are both vector spaces, and $\mathcal{D} \subseteq \mathcal{S}$.

The space distributions, denoted $\mathcal{D}'$, consists of all continuous linear mapping from from $\mathcal{D}$ to $\mathbb{C}$. This is in some sense the largest space of distributions which can be defined on the reals. The space of *tempered distributions*, denoted $\mathcal{S}'$, contains all continuous linear mappings from $\mathcal{S}$ into $\mathbb{C}$. Because $\mathcal{D} \subseteq \mathcal{S}$, it is immediate that any element of $\mathcal{S}'$ defines an element of $\mathcal{D}'$ (i.e., $\mathcal{S}' \subseteq \mathcal{D}'$), but the reverse is not true. Thus $\mathcal{S}'$ is a more restricted set of distributions. However, it is more useful in most physics applications, especially quantum mechanics. The first reason is that most distributions which arise in practice are, in fact, tempered. The second reason is that the eigenstates of various Hamiltonians, most notably the simple harmonic oscillator Hamiltonian, live in $\mathcal{S}$, so it makes sense to restrict attention to distributions which live in $\mathcal{S}'$.

We defined convergence for distributions to mean that the sequence of numbers $T_n(f)$ converged to the number $T(f)$ for every test function f. Thus, statement in the previous section that the limit of normalized Gaussians is the δ function means that the equation

$$\lim_{\sigma\to 0}\int_{-\infty}^{\infty}\frac{\exp\left(-(x-x_0)^2/\,2\sigma^2\right)}{\sqrt{2\pi}\sigma}f(x)dx = f(0)$$

holds for every test function f. Since the above integral will converge and the limit will exists for any element $f \in \mathcal{S}$, we have that the δ-function is a tempered distribution. This is important for Fourier theory (see below).

3.3.1 Differentiation and Distributions

The derivative of a distribution is defined using a formula obtained from integration by parts:

$$T'(f) := \int_{-\infty}^{\infty} t'(x)f(x)dx = -\int_{-\infty}^{\infty} t(x)f'(x)dx = -T(f')$$

If the representation $t(x)$ is differentiable, then integration by parts holds and distributional derivative coincides with the ordinary derivative. If $t(x)$ is not differentiable, like the step function, the distributional derivative will not be like ordinary derivative. For example, we saw that $\frac{d}{dx}H(x-x_0) = \delta_{x_0}$, and

$$\delta'_{x_0}(f) = \int_{-\infty}^{\infty} f(x)\delta'(x-x_0)dx = -f'(x_0),$$

since f is differentiable (as all test functions are).

3.3.2 Fourier Transforms and Distributions

Not all distributions have Fourier transforms, but every tempered distribution (i.e., member of $\mathcal{S}'$) does. These are defined by the relationship $\hat{u}(\phi) = u(\hat{\phi})$. Just as the Fourier transform is a unitary operator on $L^2(\mathbb{R})$, it can be shown that it is also a unitary-like operator on

$\mathcal{S}'$. Applying the definition of Fourier transform to the δ-function gives three important relationships:

$$\hat{\delta}(k) = \frac{1}{\sqrt{2\pi}} \qquad \left(\frac{1}{\sqrt{2\pi}}\right)\hat{}\,(k) = \delta(k) \qquad \delta(x - x') = \frac{1}{2\pi}\int dk e^{ik(x-x')}$$

Of course, all three equations are to be understood as holding in the sense of distributions: the two sides are equal when integrated against a nice function. For example:

$$\int_{-\infty}^{\infty} \hat{\delta}(k) f(k) dk = \int_{-\infty}^{\infty} \delta(k)\hat{f}(k) dk = \hat{f}(0) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} f(x) dx,$$

where $f \in \mathcal{S}$. In the first step, we've used the definition of Fourier transform of a distribution, in the second the definition of δ, and in the third the definition of Fourier transform of an ordinary function. The equality of the leftmost integral with the rightmost shows that $\hat{\delta} = 1/\sqrt{2\pi}$, as claimed above.

3.4 Row Reduction

Row reduction can be used to test for linear independence and solve linear equations. It is performed using the elementary row operators:

- interchanging two rows;
- multiplying a row by a non-zero scalar;
- adding a multiple of one row to another row.

These row operations change the range of the matrix, but leave the rank, nullity, and null space unaltered. When we row-reduce, we usually work to reduced row-echelon form. Any invertible matrix is reduced to the identity. An singular matrix will have one or more zero rows.

3.4.1 Linear Independence

The non-zero rows of the row-echelon form are linearly independent. If we placed a set of vectors in the rows of a matrix, the row-reduced rows are a basis for the span those original vectors. Alternatively, we can put the vectors in columns, and then the pivot columns indicate a subset of of the original vectors which form a basis for the span.

3.4.2 Solving Equations

If we wish to solve $Mx = y$, we form the augmented matrix $A = [My]$ and row reduce it (if $y = 0$ we often omit it, since we know the last column will remain zero). If $\operatorname{rank} M < \operatorname{rank} A$, there are no solutions. If $\operatorname{rank} M = \operatorname{rank} A = \#$ of unknowns, there is a unique solution is which given by the the last column of the row-reduced augmented matrix. If $\operatorname{rank} M = \operatorname{rank} A < \#$ of unknowns, there are infinitely many solutions. Some variables remain free and the remaining ones can be expressed in terms of the free variables. Alternatively, we can say an arbitrary solution is the sum of any one solution plus an element of the kernel.

3.5 Even and Odd Functions

3.5.1 Even Functions

An even function obeys $f(-x) = f(x)$. Graphically, this means that $f(x)$ is symmetric when reflected across the $x = 0$ axis. An analytic function is even if and only if its Taylor polynomial *only* has even powers of x in it. For even function integrated on a symmetric domain, we have

$$\int_{-a}^{a} f(x)dx = 2\int_{0}^{a} f(x)dx.$$

3.5.2 Odd Functions

An odd function obeys $f(-x) = -f(x)$. Graphically, this means that $f(x)$ is symmetric when reflected across the origin. Analytic functions are odd if and only if their Taylor polynomial *only* has odd powers of x in it. The integral of an odd function on a symmetric domain vanishes:

$$\int_{-a}^{a} f(x)dx = 0.$$

Notice that this means that the Fourier coefficients a_n vanish for an odd function, and the b_n vanish for an even function.

3.5.3 General Functions

Not every function is even or odd, but every function can be split into even and odd parts:

$$f(x) = f^E(x) + f^O(x); \quad f^E(x) = \frac{f(x) + f(-x)}{2}; \quad f^O(x) = \frac{f(x) - f(-x)}{2}.$$

Notice that $f^O = 0$ if f is even, and $f^E = 0$ if f is odd. Using the result that the integral of an odd function over a symmetric domain vanishes, we have that only f^E contributes to the a_n, and only f^O contributes to the b_n. The use of this fact can sometimes result in a considerable simplification of the integrals.

3.5.4 Friends of Even and Odd Functions

(a) The product of two even or odd functions is even.

(b) The product of an even function with an odd function is odd.

(c) The derivative of an even/odd function is odd/even, respectively. Notice that this also gives the same result for the *indefinite* integral.

(d) The Fourier transform of even/odd functions are likewise even/odd. If f is real and even/odd, is Fourier transform is real/imaginary.

3.6 Complex Algebra

3.6.1 Basic Relationships

We can express a complex number in either Cartesian or polar form, which are related by

$$z = \operatorname{Re} z + i \operatorname{Im} z = x + iy = r\cos\theta + ir\sin\theta = re^{i\theta}.$$

Notice that $\operatorname{Im} z$ does not include the i, so it is a real number. The complex conjugate is the reflection along the real axis of z:

$$\bar{z} = x - iy = r\cos\theta - ir\sin\theta = re^{-i\theta}.$$

Complex conjugation commutes with addition and multiplication, i.e.,

$$\overline{zw} = \bar{z}\bar{w} \quad \text{and} \quad \overline{z+w} = \bar{z} + \bar{w}.$$

The modulus or absolute value of z is given by

$$|z| = \sqrt{z\bar{z}} = \sqrt{x^2 + y^2} = r.$$

Notice that $z^2 \neq |z|^2$ unless $\operatorname{Im} z = 0$ Multiplication is a pain in Cartesian form but a snap in in polar form:

$$z = re^{i\theta} \text{ and } w = se^{i\varphi} \quad \Rightarrow \quad zw = rse^{i(\theta+\varphi)}.$$

This immediately gives $|zw| = |z||w|$. We also see that $z^n = r^n e^{in\theta}$, which when $r = 1$ allows us to derive trigonometric identities for $\sin nx$ and $\cos nx$. Taking the n-th root of the polar form of z gives yields

$$\sqrt[n]{z} = \sqrt[n]{r}e^{i\theta + 2\pi k/n}.$$

3.6.2 Elementary Functions

Most functions of interest are "elementary functions" which can be expressed in terms of the exponential or logarithm. The most familiar are the circular and hyperbolic functions:

$$\cos x = \frac{e^{ix} + e^{-ix}}{2} \qquad \sin x = \frac{e^{ix} - e^{-ix}}{2i} \qquad \cosh x = \frac{e^{x} + e^{-x}}{2} \qquad \sinh x = \frac{e^{x} - e^{-x}}{2}.$$

From these basic formulas we can derive many useful trigonometric as well as relationships between circular and hyperbolic functions. We also used the polar form of complex numbers together the log function to compute arbitrary powers starting from the base formula

$$\log z = \log r + i(\theta + 2\pi k).$$

Copyright Notice