Physics 172 - Mathematical Methods in Physics II

0.1 Lecture 1:

In this section I discuss infinite dimensional vectors spaces. Most of the concepts that I introduced last semester on vector spaces apply to both the finite and infinite dimensional cases.

Infinite dimensional vector spaces are important because they can be used to reduce the solution of linear differential, integral, or partial differential equations to problems in linear algebra. While the resulting infinite dimensional systems cannot always be solved exactly, the exact solutions can often be expressed as limits of solutions of approximate finite-dimensional linear equations, of the type studied previously. On a computer these approximate problems involve the inversion of a large matrix or the solution of an eigenvalue problem.

The solutions of linear differential, integral, or partial differential equations are functions. The infinite dimensional vectors spaces that contain the solutions of these problems are vector spaces of functions.

There are different vector spaces of functions that have properties that are needed to solve a given problem. If I want to construct a function that vanishes at $x = a$ and $x = b$, it is useful to express it as a linear combination of functions that vanish at $x = a$ and $x = b$.

For example, consider the equation

$$f(x) = g(x) + \int_a^b K(x, y) f(y) dy \tag{1}$$

where $g(x)$ and $K(x, y)$ are known functions, and $f(x)$ is an unknown function. If we can write

$$f(x) = \sum_{n=0}^{\infty} \phi_n(x) f_n \tag{2}$$

where the $\phi_n(x)$ are known functions and f_n are unknown complex coefficients, then the above equation becomes

$$\sum_{n=0}^{\infty} \phi_n(x) f_n = g(x) + \int_a^b K(x, y) \sum_{n=0}^{\infty} \phi_n(y) dy f_n \tag{3}$$

If the functions $\phi_n(x)$ are chosen to satisfy the orthogonality condition

$$\int_a^b \phi_m^*(x)\phi_n(x)dx = \delta_{mn} \tag{4}$$

then multiplying the equation by $\phi_m^*(x)$ and integrating from a to b, assuming that we can change the order of the sum and integral, gives the infinite algebraic system

$$f_m = g_m + \sum_{n=0}^{\infty} K_{mn} f_n \tag{5}$$

where

$$g_m = \int_a^b \phi_m^*(x)g(x)dx \qquad K_{mn} = \int_a^b \phi_m^*(x)K(x,y)\phi_n(y)dxdy. \tag{6}$$

I can attempt to solve this by approximating the sum by the first N terms. This leads to a linear algebraic equation for the first N coefficients $f_0, \cdots, f_{N-1}$.

This example will be studied in more detail later in the semester. It provides the motivation for why it is interesting to study infinite dimensional vector spaces.

As a first example of an infinite-dimensional vector space I define the space $C[a,b]$ to be the space of continuous complex valued functions of a real variable x on the interval $[a,b]$. These functions are defined to be continuous on the entire closed interval, including the endpoints.

It is obvious that if $f(x), g(x) \in C[a,b]$ then $h(x) = \alpha f(x) + g(x) \in C[a,b]$, for any complex constant α. It is a simple exercise to show that $C[a,b]$ satisfies all of the requirements needed to be a vector space.

Many infinite-dimensional vector spaces are metric spaces, normed linear spaces or inner product spaces.

The norm on the space $C[a,b]$ is

$$\|f\| = \sup_{x\in[a,b]} |f(x)| \tag{7}$$

This is defined for all continuous functions on $[a,b]$. To show that (7) is a norm note

$$\|f\| = \sup_{x\in[a,b]} |f(x)| \geq 0 \tag{8}$$

$$\|\alpha f\| = \sup_{x\in[a,b]} |\alpha f(x)| = |\alpha|\|f\| \tag{9}$$

for any complex α,

$$\|f+g\| = \sup_{x\in[a,b]} |f(x)+g(x)| \leq$$

$$\sup_{x\in[a,b]} |f(x)| + \sup_{x\in[a,b]} |g(x)| = \|f\| + \|g\| \tag{10}$$

and

$$0 = \|f\| = \sup_{x\in[a,b]} |f(x)| \Rightarrow f(x) = 0 \qquad x \in [a,b]. \tag{11}$$

The norm (7) is not derived from an inner product. It is not the only norm that can be used as a “distance function” for the continuous complex valued functions on $[a,b]$.

An inner product that is defined on the space of continuous complex valued functions on the interval $[a,b]$ is given by

$$\langle g|f\rangle := \int_a^b g(x)^* f(x) dx \tag{12}$$

The definition (12) implies $\langle f|g\rangle = \langle g|f\rangle^*$, $\langle f|g+\alpha h\rangle = \langle f|g\rangle + \alpha\langle f|h\rangle$, and $\langle f|f\rangle \geq 0$.

If $\langle f|f\rangle = 0$ and $f(x_0) = c \neq 0$ is not zero for some $x_0 \in [a,b]$ then by continuity there is a small interval $I \subset [a,b]$ of finite width δ containing x_0 where $|f(x)| > |c|/2$. It follows that

$$0 = \langle f|f\rangle = \int_a^b |f(x)|^2 dx \geq \int_I |f(x)|^2 dx > \frac{\delta|c|}{2} > 0 \tag{13}$$

This leads to the contradiction $0 > 0$ which implies that $f(x)$ must be identically zero on the interval $[a,b]$.

The function $\langle f|f\rangle^{1/2}$ is a norm on the space of continuous functions on $[a,b]$. This norm is not identical to the norm (7).

One property of finite dimensional vector spaces is that Cauchy sequences of vectors converge to vectors. In the finite dimensional case the Cauchy sequence was used to construct a sum which was shown to converge to a vector. This is not automatic in infinite dimensional vector spaces. Previously we defined a vector space as complete if every Cauchy sequence of vectors converged to a vector in the space.

Consider the space of continuous complex valued functions on $[a,b]$ with inner product $\langle f|g\rangle$. It is possible to construct a sequence of continuous functions $\{f_n\}$ with the property $\langle f_m - f_n | f_m - f_n \rangle \to 0$ and $m, n \to \infty$ (i.e. they are a Cauchy sequence) with the property that the sequences converges to a function that is discontinuous. Since the limiting function is discontinuous, it is not a element of the inner product space. This means that the space of continuous functions on $[a, b]$ with the inner product (12) is not complete.

A set of functions with this property are

$$f_n(x) = \begin{cases} 0 & : \quad a \leq x \leq c_{n-} \\ \frac{x - c_{n-}}{c_{n+} - c_{n-}} & : \quad c_{n-} < x < c_{n+} \\ 1 & : \quad c_{n+} \leq x < b \end{cases} \tag{14}$$

where $c_{n-} := \frac{(b-a)(n-1)}{2n}$, $c_{n+} := \frac{(b-a)(n+1)}{2n}$. This converges to the step function that jumps discontinuously from 0 to 1 at the midpoint of the interval $[a, b]$.

This shows that the space of continuous complex valued functions on $[a, b]$ is not complete in the inner product $\langle f | g \rangle$ defined by equation (12).

This is a problem if I want to construct solutions of differential or integral equations by taking limits of solutions of finite dimensional approximations to a given equation. I need to be sure that the limiting functions are also vectors in the desired infinite dimensional vector space. This will be true if the vector space is complete in the inner product.

For example, if the function is a solution to a second order differential equation, and a sequence of approximate solutions have second derivatives, we would like the limit to have second derivatives so it can be used in the differential equation.

0.2 Lecture 2:

If I would like to use continuous functions with the inner product (12) there are two options. The first is to extend the class of functions in the space by defining new vectors in the space as Cauchy sequences of vectors already in the space. One complication is that a given Cauchy sequence may not converge to a single function. In the above example, the limiting step function could be defined as having the value 1 or 0 at the midpoint of (a, b). This defines two different functions that differ at the point $x = a + \frac{b-a}{2}$, but both of them are limit points of the above Cauchy sequence (14).

Since the difference of any two of the limiting functions is a non-zero function with a 0 norm, we have a new problem. This can be avoided by defining abstract vectors as "equivalence classes" of functions or Cauchy sequences that differ by zero norm vectors. An equivalence relation $\sim$ is a relation on the set that divides a set into mutually distinct equivalence classes, where elements in the same class are related by

$$a \sim a \tag{15}$$

$$a \sim b \quad \rightarrow \quad b \sim a \tag{16}$$

$$a \sim b \quad b \sim c \quad \rightarrow \quad a \sim b \tag{17}$$

In this case $f(x) \sim g(x)$, or $f(x)$ and $g(x)$ are in the same equivalence class if and only if

$$\langle f - g | f - g \rangle = 0 \tag{18}$$

This method of making a vector space complete by (1) adding all Cauchy sequences and (2) replacing vectors by equivalence classes of vectors whose difference has zero norm can always be used to make infinite dimensional vector spaces complete. In this case vectors are equivalence classes of Cauchy sequences of functions.

The inner product of a Cauchy sequence $\{f_n\}$ with a vector g is defined by

$$\langle g | f \rangle = \lim_{n \to \infty} \langle g | f_n \rangle \tag{19}$$

This becomes a Cauchy sequence of Complex numbers, which does converge to a complex number. The convergence is independent of g because

$$|\langle g | f - f_n \rangle| \leq \|g\| \|f - f_n\| \tag{20}$$

which vanishes independent of g if $\|f - f_n\| \to 0$.

This method is perfectly good, but it not popular because a space of functions has been replaced by a space of equivalence classes of functions. This means that every operation on a vector has to be checked to make sure that if maps ever element of one equivalence classes to the same equivalence class.

There is a second method that is more commonly used to make spaces complete. In this case the class of functions is enlarged. In order to accomodate the larger class of functions it is necessary to use a more general definition of the integral. The functions have the property that (1) the inner product $\langle \cdot | \cdot \rangle$ is defined for all functions in the enlarged class of functions and it agrees with the previously defined scalar poduct, $\langle \cdot | \cdot \rangle$, for all continuous functions.

The first step in this second process is to redefine what is meant by an integral. The Riemann Integral is defined as a limit of finite sums

$$\int_a^b f(x)dx = \lim_{N\to\infty} \sum_{n=1}^N f(a + (b-a)\frac{2n+1}{2N})(\frac{b-a}{N}). \tag{21}$$

In this definition the interval $[a, b]$ is divided up into N smaller subintervals. The integral is approximated by summing products of the value of the function at a point in each subinterval multiplied by the width of the subinterval. With this definition all continuous functions on $[a, b]$ are integrable.

There is an alternate way to define the above integral. Instead of dividing up the interval $[a, b]$ into small subintervals, it is also possible to divide values (range) of the function into subintervals. Consider the following definition

$$\int_a^b f(x)dx = \lim_{N\to\infty} \sum_{n=-\infty}^{\infty} \frac{n}{N} \text{Volume}\{x \in [a, b] \,|\, \frac{n}{N} \leq f(x) < \frac{n+1}{N}\} \tag{22}$$

While this sum is formally infinite, if the function is continuous its value will range over a finite minimum and maximum value on the interval $[a, b]$, resulting in a finite sum. The number of terms in the sum will increase with N.

To make sense of this definition it is necessary to define the "volume" of a set. It turns out that this is difficult to do in a useful way for arbitrary sets. The allowed sets should be large enough so the integral defined by (22) can be used to define the integral of any continuous function.

The problem is that it can be shown that there is no volume function m that

a. Is defined for all sets

b. Satisfies $m(\cup_n A_n) = \sum_n m(A_n) \qquad A_m \cap A_n = \emptyset$.

c. Is translationally invariant

d. Satisfies $m([a,b]) = b - a$

To motivate the limited choice of sets note that continuous functions have the property that the inverse image of any open interval

$$f^{-1}((a,b)) = \{x | a < f(x) < b\} \tag{23}$$

can be expressed as a union of disjoint open intervals. This is actually an alternative definition of a continuous function that is equivalent to the standard definition with ϵ-δ definition. I state this problem without proof. This property will be used to ensure that the modified integral is defined for all continuous functions.

To limit the class of sets to a meaningful sub class with properties that permit the formulation of an integral, it is useful to introduce the concept of a σ-algebra of sets. A collection of subsets $\mathcal{M}$ of a set X is called a σ-algebra if it has the following three properties

1. $X \in \mathcal{M}$

2. $A \in \mathcal{M} \qquad \to \qquad A^c = X - A \in \mathcal{M}$

3. $A_n \in \mathcal{M}$, $n \in \mathbb{Z}$ then $\cup_{n\in\mathbb{Z}} A_n \in \mathcal{M}$

The identity $A \cap B = (A^c \cup B^c)^c$ can be used to express intersections of sets in terms of complements unions. As a consequence of this relation it follows that countable intersections of elements of a σ algebra are also members of the algebra.

A measure m on a σ-algebra $\mathcal{M}$ is real function with values in $[0,\infty]$ that is countably additive. This means that for disjoint $A_n \in \mathcal{M}$, $n \in \mathbb{Z}$

$$m(\cup_n A_n) = \sum_n m(A_n) \tag{24}$$

The σ-algebra of Borel sets of the real line is smallest σ-algebra of the real line $\mathbb{R}$ that contains all open subsets of $\mathbb{R}$. This definition implies that the σ-algebra of Borel sets is the smallest σ algebra that contains the inverse image of all open sets with respect to any continuous function.

For homework you will show, using properties of σ-algebras that closed sets and single points are also Borel sets.

Lebesgue measure is the measure defined on the Borel sets of the real line satisfying $\mathrm{Vol}((a, b)) = b - a$.

The countable additivity can be used to show that Lebesgue measure satisfies $\mathrm{Vol}((a, b)) = \mathrm{Vol}([a, b]) = b - a$ and $\mathrm{Vol}([a, a]) = a - a = 0$, so points have 0 Lebesgue measure.

A function $f(x)$ on the real line is Lebesgue measurable if the inverse image of open subset of the real line is a Borel set. By this definition all continuous functions are Lebesgue measurable. Since the class of borel sets are bigger than the the class of open sets, the class of measurable functions is larger than the class of continuous functions.

The Lebsegue integral is the integral given by (22) where the volume function is Lebesgue measure.

0.3 Lecture 3:

The class of Lebesgue measurable functions contains the continuous functions, but is larger than the class of continuous functions. It contains functions that are piecewise continuous, the function that is 1 on the rationals and 1/2 on the irrationals, and many other highly discontinuous functions. The important property of measurable functions is that they are functions and Cauchy sequences of measurable functions with respect to the inner product (6) converge to measurable functions in a sense mad precise by the theorems below. While the limit may still not be unique, the condition that $\langle f|f\rangle = 0$ means that $f(x)$ vanishes except possibly on a set of Lebesgue measure zero. This is sometimes written $f(x) = 0_{[a.e.]}$ which is stated that $f(x)$ is zero almost everywhere with respect to Lebesgue measure.

What has been gained is that the space of equivalence classes of functions that differ at most on a set of Lebesgue measure zero and are square integrable with respect to Lebesgue measure are complete with respect to the inner product (6). I state the three key important properties of the Lebesgue integral:

Monotone Convergence Theorem

a. $f_n(x)$ measurable

b. $f_n(x) \to f(x)$ each x

c. $f_n(x) \leq f_{n+1}(x)$ each x

d. $\int |f_n(x)| dx < C$ all n

$$\Downarrow$$

$f(x)$ is measurable

$\int |f(x)| dx < \infty$

$\lim_{n\to\infty} \int |f(x) - f_n(x)| dx = 0$

Dominated Convergence Theorem

a. $f_n(x)$ measurable

b. $f_n(x) \to f(x)$ each x

c. $f_n(x) \leq g(x)$, all x; $\int |g(x)|dx < \infty$

$$\Downarrow$$

a. $f(x)$ measurable

$\int |f(x)|dx < \infty$

$\lim_{n\to\infty} \int |f(x) - f_n(x)|dx = 0$

Define the $L^p(\mathbb{R})$ spaces as normed spaces of Lebesgue measurable functions with the L^p norm given by

$$\|f\|_p := [\int |f(x)|^p dx]^{1/p} \tag{25}$$

The proof that this actually defines a norm is not obvious for $p \neq 2$. It uses a generalization of the Schwartz inequality called Jensen's inequality. The argument can be found in Real and Complex Analysis by W. Rudin in the reference list. The theorem also applies if the interval is restricted to a subinterval of the real line or if the integral includes a positive weight function $dx \to w(x)dx$.

Riesz-Fischer Theorem: The L^p spaces for $(1 \leq p \leq \infty)$ are complete normed linear spaces. The Riesz-Fischer theorem can be proved from the monotone convergence theorem which follows the dominated convergence theorem. The proof of these theorems can be found in any graduate level text om mathematical analysis. They are consequences of the definitions above, and also hold for more abstract measure spaces. The important thing is that the Riesz-Fischer theorem provides a large class of complete functions spaces.

It is useful to include a positive weight function in the definition of these spaces of Borel measurable functions

$$L^p(\mathbb{R}, w) := \{f(x) | \|_p f\| := [\int |f(x)|^p w(x) dx]^{1/p} < \infty \tag{26}$$

When $p = 2$ the space $L^2(\mathbb{R}, w)$ is a complete inner product space or Hilbert space with inner product

$$\langle f|g\rangle = \int f^*(x) g(x) w(x) dx \tag{27}$$

A final important result, which I state without proof, is that measurable functions that are bounded except on sets of measure zero can be approximated by continuous functions in the sense that $|c(x) - m(x)| < \epsilon$ except on sets of measure $< \epsilon$, where $c(x)$ is continuous and $m(x)$ is measurable.

In this section I discuss some properties of complete inner product spaces or Hilbert spaces. In what follows the integrals will always be assumed to be Lebesgue integrals and functions will be equivalence classes of Lebesgue measurable functions that differ only on sets of zero measure.

A set of complex valued functions $\{\phi_n(x)\}$ on $[a, b]$ are orthogonal with respect to a positive weight $w(x)$ if they satisfy the orthogonality relations

$$\langle \phi_n | \phi_m \rangle = \int_a^b w(x) \phi_n^*(x) \phi_m(x) dx = \delta_{mn}. \tag{28}$$

Here a and b do not have to be finite.

We want to investigate when a set of orthonormal functions is a basis or spans a Hilbert space, $\mathcal{H}$.

Let $|\phi_n\rangle$ be an orthonormal set of functions and let $|v\rangle$ be an arbitrary vector. Define

$$|v_N\rangle = \sum_{n=1}^{N} |\phi_n\rangle c_n \tag{29}$$

with

$$\langle \phi_m | v \rangle = c_m \tag{30}$$

It follows from the definitions that

$$\langle v | v_N \rangle = \sum_{n=1}^{N} |\langle \phi_n | v \rangle|^2 \tag{31}$$

Applying the Cauchy-Schwartz inequality to (31) gives

$$|\langle v | v_N \rangle|^2 \leq \langle v | v \rangle \langle v_N | v_N \rangle = \langle v | v \rangle \sum_{n=1}^{N} |\langle \phi_n | v \rangle|^2 \tag{32}$$

Combining (32) with (31) gives

$$|\langle v | v_N \rangle| \leq \langle v | v \rangle \tag{33}$$

or equivalently

$$\sum_{n=1}^{N} |c_n|^2 \leq \langle v|v\rangle \tag{34}$$

This inequality is called Bessel's inequality. It says that the sum of the squares of the first N expansion coefficients is alway less than or equal to the square of the norm of the vector.

Next I give a definition of a basis for infinite dimensional vector spaces. The problem in infinite dimensional spaces is that if one basis vector is removed from an infinite set, the remaining set of vectors is still infinite, but not necessarily a basis. This means that if we keep adding orthogonal vector to a set, the limiting set is not necessarily a basis.

Definition: A set of orthonormal vectors, $\{|\phi_n\rangle\}$, is a basis for a Hilbert space $\mathcal{H}$ if and only if the only vector orthogonal to all of the $|\phi_n\rangle$ is the zero vector.

Let $|v\rangle \in \mathcal{H}$ be arbitrary and consider the partial sums

$$|v_N\rangle := \sum_{n=1}^{N} |\phi_n\rangle\langle\phi_n|v\rangle \tag{35}$$

Consider the norm of the difference of the partial sums for $M > N$ terms.

$$\||v_M\rangle - |v_N\rangle\|^2 = \langle v_M|v_M\rangle - \langle v_N|v_M\rangle - \langle v_M|v_N\rangle + \langle v_N|v_N\rangle \tag{36}$$

Note that for $M > N$

$$\langle v_N|v_M\rangle = \langle v_M|v_N\rangle = \langle v_N|v_N\rangle \tag{37}$$

so (36) becomes

$$\langle v_M|v_M\rangle - \langle v_N|v_N\rangle = \sum_{n=N+1}^{M} |\langle\phi_n|v\rangle|^2 \tag{38}$$

Since the right hand side of Bessel's inequality is independent of N it follows that

$$\sum_{n=1}^{\infty} |\langle\phi_n|v\rangle|^2 < \infty \tag{39}$$

which means that the coefficients $|\langle \phi_n | v \rangle|^2$ are a convergent sequence of numbers. This implies that the difference in (36) converges to zero and the sequence of vectors $\{|v_N\rangle\}$ is a Cauchy sequence of vectors. Since by definition, the Hilbert space is complete, the infinite sum

$$|w\rangle := \lim_{N\to\infty} |v_N\rangle = \sum_{n=1}^{\infty} |\phi_n\rangle\langle\phi_n|v\rangle \tag{40}$$

converges to a vector $|w\rangle \in \mathcal{H}$.

Next consider the difference

$$|u\rangle = |v\rangle - |w\rangle \tag{41}$$

There are two possibilities; either $|u\rangle$ is the zero vector or it is not the zero vector.

If $|u\rangle$ is not zero then

$$\langle\phi_m|u\rangle = \langle\phi_m|v\rangle - \langle\phi_m|w\rangle = \langle\phi_m|v\rangle - \langle\phi_m|v\rangle = 0 \tag{42}$$

which means that there is a non-zero vector orthogonal to every element of the orthonormal set. When this happens $\{|\phi_n\rangle\}$ cannot be a basis.

Conversely, if this set is a basis, the vector $|u\rangle$ must vanish for any choice of $|v\rangle \in \mathcal{H}$ which implies

$$|v\rangle = \sum_{n=1}^{\infty} |\phi_n\rangle\langle\phi_n|v\rangle \tag{43}$$

It follows that

$$\langle v|v\rangle = \sum_{n=1}^{\infty} |\langle\phi_n|v\rangle^2 \tag{44}$$

in which Bessel's inequality becomes an equality for every vector in $\mathcal{H}$. Equation (44) is called Parseval's relation.

This shows that a condition for completeness is for Bessel's inequality to be an equality for all vectors. If is fails to be an equality for 1 vector the orthogonal set is not a basis.

When $\{|\phi_n\rangle\}$ an orthonormal basis then the what we have shown that any vector in the Hilbert space has the form

$$|v\rangle = \sum_{n=1}^{\infty} |\phi_n\rangle c_n \tag{45}$$

where

$$\sum_{n=1}^{\infty} |c_n|^2 < \infty \tag{46}$$

The infinite collection of numbers c_n are the coordinates of the vector $|v\rangle$ in the basis $|\phi_n\rangle$

In this representation the components of the sum two vectors or the product of a vector and a complex scalar are the sums of the individual components or the product of each component by the scalar. In both cases if the original coefficients satisfy (46) then the new vectors will also satisfy this condition.

The abstract requirements for an infinite set of orthogonal vectors to be a basis are interesting, but they are not of real practical value. In applications it is alway important to work with an explicit set of functions that are known to be a basis.

It turns out that for the Hilbert space of measurable functions on a finite interval $[a, b]$ that the polynomials

$$\phi_n(x) = x^n \qquad n = 0, 1, 2 \cdots \tag{47}$$

are a basis. While these are not orthonormal functions, they can be used to construct an infinite orthonormal set using the Gram Schmidt method.

The theorem that we will prove is even stronger. It is called the Weierstrass approximation theorem. The statement of the theorem is

Theorem: (Weierstrass Approximation Theorem) Let $f(x)$ be a continuous function on the finite interval $[a, b]$. Then for every $\epsilon > 0$ there exists an integer $n > 0$ and a polynomial $p_n(x)$ of degree n with the property that

$$\sum_{x \in [a,b]} |f(x) - p_n(x)| < \epsilon \tag{48}$$

This theorem shows that the approximation holds on the space $C[a, b]$ with its natural norm. Formally this means that the approximation is pointwise.

Pointwise convergence implies $L^2(\mathbb{R})$ convergence and more generally $L^p(\mathbb{R})$ convergence. Because Cauchy sequences of continuous functions can be used to construct the measurable functions in $L^2([a, b])$, if we let $f(x)$ be measurable, for every $\epsilon > 0$ we can find a continuous function $c_\epsilon(x)$ satisfying

$$\|c_\epsilon - f\| < \frac{\epsilon}{2} \tag{49}$$

Since c_ϵ is continuous we can find a polynomial of sufficiently hight degree n so

$$\|c_\epsilon - p_n\| < \frac{\epsilon}{2} \tag{50}$$

Using the triangle inequality

$$\|p_n - f\| = \|p_n - c_\epsilon + c_\epsilon - f\| <$$

$$\|p_n - c_\epsilon\| + \|c_\epsilon - f\| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \tag{51}$$

Since ϵ is arbitrary, we see that the convergence also applies to measurable functions with respect to a Hilbert space norm. The means that the polynomial are also a basis for $L^2(\mathbb{R})$ (and $L^p(\mathbb{R})$).

Note however with measurable functions the sup norm of the difference between the polynomial and the measurable function does not have to be small. For example the function could have very different values on sets of measure zero or sets of very small measure.

I also note that the Weierstrass theorem the key tool in the infinite-dimensional generalization of the proof that normal linear operators have complete sets of eigenfunctions. This is called the spectral theorem. It is used all the time in quantum mechanics,

0.4 Lecture 4:

The trick for proving the Weierstrass theorem is to construct very sharply peaked polynomials that integrate to 1 on $[a, b]$.

I start by transforming the interval $[a, b]$ to an interior subinterval of the standard interval $[0, 1]$ using

$$x' = \frac{1}{b-a}[(x-a)(1-\sigma) + (b-x)\sigma] \tag{52}$$

where $\frac{1}{2} > \sigma > 0$. In this case $x = a \to x' = \sigma$; $x = b \to x' = (1-\sigma)$. The problem reduces to finding a polynomial approximation that is accurate on the interior subinterval.

Next we extend the function $g(x') := f(x(x'))$, which is defined and continuous on the subinterval $[\sigma, 1-\sigma]$, to a continuous function on all of $[0, 1]$. This can be done by setting $f(x) := f(\sigma)$ for $x < \sigma$ and $f(x) = f(1-\sigma)$ for $x > (1-\sigma)$.

It is sufficient to show that $g(x)$ can be pointwise approximated by a polynomial on the interior interval $[\sigma, 1-\sigma]$.

Let $1 > \delta > 0$ and define the functions

$$A_n(\delta) := \int_\delta^1 (1-y^2)^n dy \tag{53}$$

To prove the Weierstrass theorem I will need the result

$$\lim_{n\to\infty} \frac{A_n(\delta)}{A_n(0)} = 0 \tag{54}$$

To show this note

$$A_n(\delta) := \int_\delta^1 (1-y^2)^n dy \leq (1-\delta^2)^n(1-\delta) \leq (1-\delta^2)^n \tag{55}$$

and

$$A_n(0) := \int_0^1 (1-y^2)^n dy \geq \int_0^1 (1-y)^n dy = \frac{1}{n+1} \tag{56}$$

Using these together I get

$$\frac{A_n(\delta)}{A_n(0)} \leq (n+1)(1-\delta^2)^n \tag{57}$$

which vanishes as $n \to \infty$ for $0 < \delta < 1$.

The desired Weierstrass polynomial approximation to $g(x)$ is

$$P_{2n}(x) := \frac{1}{2A_n(0)} \int_0^1 g(z)[1-(z-x)^2]^n dz = \frac{1}{2A_n(0)} \int_{-x}^{1-x} g(x+y)[1-y^2]^n dy \tag{58}$$

By inspection it is clear that this is a polynomial of degree $2n$ in the variable x.

Pick any $\epsilon > 0$ and choose δ so for $|x-y| < \delta$

$$|g(x+y) - g(x)| < \epsilon \tag{59}$$

This can be done because the interval is closed and bounded. The proof this statement is non-trivial and is based on a compactness argument.

Consider

$$P_{2n}(x) = \frac{1}{aA_n(0)}[\int_{-x}^{-\delta} + \int_{-\delta}^{\delta} + \int_{\delta}^{1-x}]g(y+x)(1-y^2)^n dy \tag{60}$$

where without loss of generality I can always choose δ small enough so that $\delta < x$.

We estimate all three integrals

$$P_{2n}(x) = I_1(n) + I_2(n) + I_3(n) \tag{61}$$

where

$$|I_1(n)| := \frac{1}{2A_n(0)}|\int_{-x}^{-\delta} g(y+x)(1-y^2)^n dy| \leq$$

$$\frac{1}{2A_n(0)}|g_{max}|\int_{-x}^{-\delta}(1-y^2)^n dy| =$$

$$\frac{1}{2A_n(0)}|g_{max}|\int_{\delta}^{x}(1-y^2)^n dy| \leq$$

$$|g_{max}|\frac{A_n(\delta)}{2A_n(0)}| \to 0 \tag{62}$$

and

$$|I_3(n)| := \frac{1}{2A_n(0)}|\int_{\delta}^{1-x} g(y+x)(1-y^2)^n dy| \leq$$

$$\frac{1}{2A_n(0)}|g_{max}|\int_{\delta}^{1-x}(1-y^2)^n dy| \leq$$

$$|g_{max}|\frac{A_n(\delta)}{2A_n(0)}| \to 0 \tag{63}$$

and

$$|I_2(n)| := \frac{1}{2A_n(0)}|\int_{-\delta}^{\delta} g(y+x)(1-y^2)^n dy| \leq$$

$$\frac{1}{2A_n(0)}|\int_{-\delta}^{\delta}(g(x)+g(y+x)-g(x))(1-y^2)^n dy| \leq$$

$$g(x)\frac{1}{2A_n(0)}|\int_{-\delta}^{\delta}(1-y^2)^n dy| + \frac{1}{2A_n(0)}|\int_{-\delta}^{\delta}|g(y+x)-g(x)|(1-y^2)^n dy| \leq=$$

$$\frac{1}{2A_n(0)}2|[\int_0^1 - \int_{\delta}^1](1-y^2)^n dy + \frac{1}{2A_n(0)}|\int_{-\delta}^{\delta}|g(y+x)-g(x)|(1-y^2)^n dy| \leq=$$

$$g(x)\frac{2A_n(0)-2A_n(\delta)}{2A_n(0)} + \frac{\epsilon}{2A_n(0)}\frac{2A_n(0)-2A_n(\delta)}{2A_n(0)} \tag{64}$$

This has the form $g(x)+\epsilon+$(terms that vanish as as $n \to \infty$). This means that

$$|P_{2n}(x) - g(x)| \tag{65}$$

can be made as small as desired by choosing ϵ small enough and n large enough. This completes the proof of the Weierstrass theorem

What we get from the Weierstrass theorem is an explicit basis for all of the L^p spaces of functions on a finite interval $[a,b]$. These are well understood functions that can be used in explicit computations. On $L^2([a,b])$ these basis functions can be replaced by orthogonal polynomials using the Gram Schmid method. The orthogonal polynomials defined this way on the interval $[-1,1]$ are proportional to the Legendre polynomials (to be discussed later).

0.5 Lecture 5:

Now that we know that the orthogonal polynomials are a basis for a number of function spaces, it is useful to discuss properties of orthogonal polynomials with different choices of weight functions.

While in principle it is possible to start with any positive weight function $w(x)$ on the interval $[a, b]$ and construct orthogonal Polynomials using the Gram Schmidt method, there is a special set of orthogonal polynomials, called the classical orthogonal polynomials, that arise in many physics applications. These polynomials are all constructed using similar methods.

These polynomials are constructed using the following formula, called a generalized Rodrigues formula:

$$C_n(x) = \frac{1}{w(x)} \frac{d^n}{dx^n} (w(x) s^n(x)) \tag{66}$$

where

(1.) $s(x)$ is a polynomial of degree ≤ 2 in x with real roots.

(2.) $C_1(x)$ is a polynomial of degree 1 in x.

(3.) $w(x)$ is a positive weight function satisfying $w(a)s(a) = w(b)s(b) = 0$.

We will show that the functions $C_n(x)$ are orthogonal polynomials on the interval $[a, b]$ with weight $w(x)$, i.e.:

$$\int_a^b w(x) C_n(x) C_m(x) dx = K_n \delta_{nm} \tag{67}$$

where K_n is a positive normalization constant, These can be made orthonormal by replacing $C_n(x)$ by $C_n(x)/\sqrt{K}_n$

In addition to naturally appearing in applications, classical polynomials of high orders are easily generated by computer.

In what follows all of these properties will be established. I begin by considering consequences of the assumptions above.

(1.) The Rodriguez formula applied to $C_1(x)$ implies

$$C_1(x) = \frac{1}{w(x)} \frac{d}{dx} (w(x) s(x)) = \frac{1}{w(x)} \frac{dw(x)}{dx} s(x) + \frac{ds(x)}{dx} \tag{68}$$

This equation can be written as

$$\frac{dw(x)}{dx}s(x) = w(x)(c_1(x) - \frac{ds(x)}{dx}) = w(x)p_{k\le 1}(x) \tag{69}$$

where $p_{k\le 1}(x)$ is a polynomial of degree one or less. This is because $C_1(x)$ has degree 1 and $s(x)$ is of degree two or less.

$$\boxed{\frac{dw(x)}{dx}s(x) = w(x)p_{k\le 1}(x)} \tag{70}$$

(2.) Next consider

$$\frac{d^m}{dx^m}\left(w(x)s^n(x)p_{\le k}(x)\right) =$$

$$\frac{d^{m-1}}{dx^{m-1}}\left(\frac{dw(x)}{dx}s^n(x)p_{\le k}(x) + nw(x)s^{n-1}(x)\frac{ds(x)}{dx}p_{\le k}(x) + w(x)s^n(x)\frac{dp_{\le k}(x)}{dx}\right) \tag{71}$$

Next factor $w(x)s^{n-1}(x)$ out of the () using the result of (1.) to get

$$= \frac{d^{m-1}}{dx^{m-1}}\left(w(x)s^{n-1}(x)(p_{\le 1}(x)p_{\le k}(x) + n\frac{ds(x)}{dx}p_{\le k}(x) + s(x)\frac{dp_{\le k}(x)}{dx})\right) \tag{72}$$

which has the general form

$$\boxed{\frac{d^m}{dx^m}\left(w(x)s^n(x)p_{\le k}(x)\right) = \frac{d^{m-1}}{dx^{m-1}}\left(w(x)s^{n-1}(x)p_{\le k+1}(x)\right)} \tag{73}$$

Continuing, using mathematical induction it follows that if this is repeated $m \le n$ times gives

$$\boxed{\frac{d^m}{dx^m}\left(w(x)s^n(x)p_{\le k}(x)\right) = w(x)s^{n-m}(x)p_{\le k+m}(x)} \tag{74}$$

(3.) Next note that for $m < n$ equation (2.) implies

$$\frac{d^m}{dx^m}\left(w(x)s^n(x)p_{\le k}(x)\right) = 0 \qquad x = a \quad \text{or} \quad b \tag{75}$$

This follows because a factor of $w(x)s(x)$ can be factored out of the right side of (**??**) when $m < n$.

I use these three results to show that the $C_n(x)$ are (a.) polynomials and (b.) mutually orthogonal.

First note that equation (74) with $k = 0$ and $m = n$ gives

$$C_n(x) = \frac{1}{w(x)} \frac{d^n}{dx^n}(w(x)s^n(x)) = \frac{1}{w(x)} w(x)s^0(x)p_{\leq n}(x) = p_{\leq n}(x) \tag{76}$$

which establishes that $C_n(x)$ is a polynomial of degree no higher than n.

Next I observe that for $n > m$

$$\int_a^b C_n(x)P_m(x)w(x)dx = \int_a^b \frac{1}{w(x)} P_m(x)w(x) \frac{d^n}{dx^n}(w(x)s^n(x))dx \tag{77}$$

Integrating by parts, using $w(a)s(a) = w(b)s(b) = 0$ gives

$$\int_a^b C_n(x)P_m(x)w(x)dx = (-) \int_a^b \frac{dP_m(x)}{dx} \frac{d^{n-1}}{dx^{n-1}}(w(x)s^n(x))dx \tag{78}$$

Repeating this $n - 1$ times gives

$$\int_a^b C_n(x)P_m(x)w(x)dx = (-)^n \int_a^b \frac{d^n P_m(x)}{dx^n} w(x)s^n(x)dx = 0 \tag{79}$$

because $n > m$. This shows that $C_n(x)$ is orthogonal to all polynomials of degree $m < n$. Since $C_n(x)$ has degree no more than n it must be a polynomial of degree n.

This shows that $C_n(x)$ are indeed orthogonal polynomials on $[a, b]$ with weight $w(x)$.

The requirements on the classical orthogonal polynomials are somewhat restrictive. In fact they are so restrictive that it is possible to classify all of the classical orthogonal polynomials.

I start with the case that $s = k$ where k is a constant. Then the Rodrigues formula for $C_1(x)$ becomes

$$c_1 + c_2 x = C_1(x) = \frac{1}{w(x)} \frac{d}{dx}(kw(x)) \tag{80}$$

This gives

$$\frac{d}{dx} \ln(w(x)) = \frac{c_1}{k} + \frac{c_2}{k} x \tag{81}$$

and

$$w(x) = \exp(\frac{c_2}{2k}x^2 + \frac{c_1}{k}x + c_0) \tag{82}$$

This function will vanish at $x = \pm\infty$ provided that $\frac{c_2}{2k} < 0$

By completing the square in the exponent and rescaling the resulting variable I reduce the general case to

$$s = 1 \qquad w(x) = e^{-x^2} \qquad a = -\infty \qquad b = +\infty \tag{83}$$

The corresponding orthogonal polynomials are called Hermite polynomials. They appear in solutions to the three dimensional quantum harmonic oscillator. The polynomials satisfy

$$\int_{-\infty}^{\infty} H_n(x)H_m(x)dx = 0 \qquad m \neq n \tag{84}$$

Next I consider the case that $s(x) = k_1 + k_2 x$. In this case the Rodrigues formula for $C_1(x)$ gives

$$c_1 + c_2 x = C_1(x) = \frac{1}{w(x)}\frac{d}{dx}((k_1 + k_2 x)w(x))) \tag{85}$$

This can be expressed as

$$\frac{c_1 - k2 + c_2 x}{k_1 + k_2 x} = \frac{d\ln(w(x))}{dx} \tag{86}$$

The left side of this equation can be written as

$$\frac{c_1 - k2 + c_2/k_2(k_2 x + k_1) - c_2 k_1/k_2}{k_1 + k_2 x} = \frac{c_2}{k_2} + \frac{c_1 - k_2 - c_2 k_1/k_2}{k_1 + k_2 x} = \frac{d\ln(w(x))}{dx} \tag{87}$$

Integrating both side of this equation gives

$$\frac{c_2}{k_2}x + \frac{c_1 - k_2 - c_2 k_1/k_2}{k_2}\ln(k_1 + k_2 x) = \ln(w(x)) + c_0' \tag{88}$$

Evaluating the weight function gives

$$w(x) = (k_1 + k_2 x)^{\frac{c_1 - k_2 - c_2 k_2/k_2}{k_2}} e^{c_0 + \frac{c_2}{k_2}x} \tag{89}$$

In this case $s(x)w(x) = 0$ requires $x = -k_1/k_2$ with $\frac{c_1 - k_2 - c_2 k_2/k_2}{k_2} > -1$ and $x = \infty$ if $c_2/k_2 < 0$ and $x = -\infty$ if $c_2/k_2 > 0$.

By suitable transformations I can put this in the form

$$w(x) = e^{-x} x^{\nu} \qquad \nu > -1 \qquad s(x) = x \quad a = 0, b = \infty \tag{90}$$

The corresponding classical polynomials are called the Associated Laguerre polynomials. These polynomials appear in the wave functions for the Hydrogen atom.

The last case is when $s(x) = k_0(x - k_1)(x - k_2)$. The Rodrigues formula for $C_1(x)$ gives

$$c_1 + c_2 x = \frac{1}{w(x)} \frac{d}{dx} (w(x) k_0 (x - k_1)(x - k_2)) =$$

$$\frac{d \ln(w(x))}{dx} k_0 (x - k_1)(x - k_2)) + k_0(x - k_2) + k_0(x - k_1) \tag{91}$$

This can be written in the form

$$\frac{(c_1 + k_0 k_1 + k_0 k_2) + x(c_2 - 2k_0)}{k_0(x - k_1)(x - k_2)} = \frac{d \ln(w(x))}{dx} \tag{92}$$

To compute this we use the rational function expansion that we proved last semester:

$$\frac{(c_1 + k_0 k_1 + k_0 k_2) + x(c_2 - 2k_0)}{k_0(x - k_1)(x - k_2)} =$$

$$\frac{(c_1 + k_0 k_1 + k_0 k_2) + k_1(c_2 - 2k_0)}{k_0(k_1 - k_2)} \frac{1}{x - k_1} +$$

$$\frac{(c_1 + k_0 k_1 + k_0 k_2) + k_2(c_2 - 2k_0)}{k_0(k_2 - k_1)} \frac{1}{(x - k_2)} \tag{93}$$

Integrating both sides of this equation gives

$$\frac{(c_1 + k_0 k_1 + k_0 k_2) + k_1(c_2 - 2k_0)}{k_0(k_1 - k_2)} \ln(x - k_1) +$$

$$\frac{(c_1 + k_0 k_1 + k_0 k_2) + k_2(c_2 - 2k_0)}{k_0(k_2 - k_1)} \ln(x - k_2) = \ln(w(x)) + d_1 \tag{94}$$

If we define

$$\nu_1 = \frac{(c_1 + k_0 k_1 + k_0 k_2) + k_1(c_2 - 2k_0)}{k_0(k_1 - k_2)} \tag{95}$$

$$\nu_2 = \frac{(c_1 + k_0 k_1 + k_0 k_2) + k_2(c_2 - 2k_0)}{k_0(k_2 - k_1)} \tag{96}$$

then

$$w(x) = e^{-d_1}(x - k_1)^{\nu_1}(x - k_2)^{\nu_2} \tag{97}$$

Note that

$$w(x)s(x) = e^{-d_1}(x - k_1)^{\nu_1+1}(x - k_2)^{\nu_2+1} \tag{98}$$

vanishes at k_1 and k_2 provided $\nu_1, \nu_2 > -1$

In this case the corresponding orthogonal polynomials, $P_n^{\nu_1\nu_2}$ are related to the Jacobi Polynomials, which are associated with $k_1 = -1$ and $k_2 = 1$.

There are several special cases of the Jacobi polynomials:

$\mu_1 = \nu_2 = \lambda - \frac{1}{2}$	Gegenbauer	$C_n^\lambda(x)$
$\mu_1 = \nu_2 = 0$	Legendre	$P_n(x)$
$\mu_1 = \nu_2 = -\frac{1}{2}$	Tchebichef 1st kind	$T_n(x)$
$\mu_1 = \nu_2 = \frac{1}{2}$	Tchebichef 2nd kind	$U_n(x)$

0.6 Lecture 6:

The classical orthogonal polynomials often arise as solutions to ordinary differential equation. Because of this it is useful to recognize the form of the differential equation satisfies by each of the Classical orthogonal polynomials.

To derive the differential equation start with

$$\frac{d}{dx}\left(w(x)s(x)\underbrace{\frac{dC_n(x)}{dx}}_{P_{n-1}(x)}\right) = w(x)s(x)^{1-1}P_{n-1+1}(x) \tag{99}$$

The polynomial on the right hand side of this equation can be expanded as a linear combination of the $C_m(x)$ for $m \leq n$. I write this as

$$P_{n-1+1}(x) = -\sum_{m=0}^{n} \lambda_{nm} C_m(x) \tag{100}$$

First I show that $\lambda_{mn} = 0$ unless $m = n$. To do this multiply the above expression by $C_l(x)$ with $l \leq n$ and integrate from a to b:

$$\int_a^b C_l(x)\frac{d}{dx}\left(w(x)s(x)\frac{dC_n(x)}{dx}\right)dx$$

$$= -\sum_{m=0}^{n}\int_a^b w(x)C_l(x)\lambda_{nm}C_m(x) = -\lambda_{nl}K_l \tag{101}$$

where

$$K_l := \int_a^b w(x)C_l^2(x)dx \tag{102}$$

Integrate the left side of the above equation twice by parts, using $w(x)s(x)$ vanishes at the endpoints, to get

$$\int_a^b \frac{d}{dx}\left(\frac{dC_l(x)}{dx}w(x)s(x)\right)C_n(x) = dx$$

$$\int_a^b w(x)P_l(x)C_n(x) = dx \tag{103}$$

which vanishes unless $l = n$. This leads to the differential equation

$$\frac{1}{w(x)}\frac{d}{dx}\left[s(x)w(x)\frac{dC_n(x)}{dx}\right] + \lambda_{nn}C_n(x) = 0 \tag{104}$$

In order to obtain the explicit form of this equation I need to determine the coefficient λ_{nn}.

First note

$$C_1(x) = \frac{1}{w(x)}\frac{d}{dx}(s(x)w(x)). \tag{105}$$

Using this in the differential equation gives

$$0 = \frac{1}{w(x)}\frac{d}{dx}\left[s(x)w(x)\frac{dC_n(x)}{dx}\right] + \lambda_{nn}C_n(x) =$$

$$s(x)\frac{d^2C_n(x)}{dx^2} + C_1(x)\frac{dC_n(x)}{dx} + \lambda_{nn}C_n(x) = \tag{106}$$

If

$$C_n(x) = k_{nn}x^n + k_{nn-1}x^{n-1} + \cdots \tag{107}$$

then comparing coefficients of x^n gives the relation

$$\frac{1}{2!}\frac{d^2s(x)}{dx^2}n(n-1)k_{nn} + nk_{11}k_{nn} + \lambda_{nn}k_{nn} \tag{108}$$

which gives

$$-\lambda_{nn} = nk_{11} + \frac{n(n-1)}{2}\frac{d^2s(x)}{dx^2} \tag{109}$$

where because s is at most second degree, $\frac{d^2s(x)}{dx^2}$ is a constant independent of x.

This gives an explicit expression for all coefficients in the differential equation in terms of s and leading coefficient of $C_1(x)$.

0.7 Lecture 7:

While the classical orthogonal polynomials can be computed using the Rodrigues formula, they can be computed very efficiently using recursion relations, which are satisfied by all orthogonal polynomials.

To formulate the recursion relations we assume the normalization

$$\int_a^b w(x)C_n(x)C_m(x)dx = \delta_{mn} \tag{110}$$

and

$$C_n(x) = k_{nn}x^n + k_{nn-1}x^{n-1} + p_{\leq n-2}(x) \tag{111}$$

Using these relations it follows that

$$C_{n+1}(x) = k_{n+1n+1}x^{n+1} + k_{n+1n}x^n + p_{\leq n-1}(x) \tag{112}$$

I can cancel off the coefficient of x^{n+1} by taking the difference

$$C_{n+1}(x) - \frac{k_{n+1n+1}}{k_{nn}}xC_n(x) = p_{\leq n} = \sum_{l=0}^{n} a_{n+1,l}C_l(x) \tag{113}$$

Multiplying by $C_{n-1}(x)w(x)$ and integrating from $[a,b]$ using the orthogonality relations (110) gives

$$-\frac{k_{n+1n+1}}{k_{nn}}\int_a^n w(x)xC_n(x)C_{n-1}(x)dx = a_{n+1n-1} \tag{114}$$

Since

$$xC_{n-1}(x) = \frac{k_{n-1n-1}}{k_{nn}}C_n(x) + p_{\leq n-1} \tag{115}$$

the above integral becomes

$$a_{n+1n-1} = -\frac{k_{n+1n+1}}{k_{nn}}\frac{k_{n-1n-1}}{k_{nn}} \tag{116}$$

The coefficient of x^n in Equation (113) gives the relation

$$k_{n+1n}x^n - \frac{k_{n+1n+1}}{k_{nn}}k_{nn-1}x^n = a_{n+1n}k_{nn}x^n \tag{117}$$

or

$$a_{n+1n} = \frac{1}{k_{nn}}\left(k_{n+1n} - \frac{k_{n+1n+1}k_{nn-1}}{k_{nn}}\right) \tag{118}$$

This gives the relation

$$C_{n+1}(x) = (A_n x + B_n)C_n(x) - D_n C_{n-1}(x) \tag{119}$$

where

$$A_n = \frac{k_{n+1n+1}}{k_{nn}} \tag{120}$$

$$B_n = \frac{1}{k_{nn}} \left(k_{n+1n} - \frac{k_{n+1n+1}k_{nn-1}}{k_{nn}^2} \right) \tag{121}$$

$$D_n = -\frac{k_{n+1n+1}}{k_{nn}} \frac{k_{n-1n-1}}{k_{nn}} \tag{122}$$

where in these recursions relations the normalization is assumed to be 1. Similar relations hold with more general normalizations.

Equation (4.54) can be used to generate all of the orthogonal polynomials in terms of the first two.

The differential equations are uesful for recognizing when certain calssical orthogonal polynomials arise a solution of dynamical equations. The recursion relations are useful for generating the different orthogonal polynomials. Both the recursion relation and diferential equation need explicit expressions for the coefficients of the leading two powers of x:

$$C_n(x) = k_{nn}x^n + k_{nn-1}x^{n-1} + \cdots \tag{123}$$

These can be determined directly from the Rodrigues equations. They are given on page 212-214 of the text, along with the standard normalizations.

As discussed previously the orthogonal polynomials are a complete set of functions on the Hilbert space of square integrable functions with weight w

$$\langle f|g\rangle = \int_a^b w(x)f^*(c)g(x)dx \tag{124}$$

(Technically the Weierstrass theorem does not apply to the Legendre or Hermite polynomials, however they can also be shown to be complete using a different proof.)

To use these bases I need to know how to express arbitrary vectors in this basis (in what follows I assume that the polynomials are orthonormal)

$$f(x) = \sum_{n=0}^{\infty} P_n(x)\langle P_n|f\rangle \tag{125}$$

where

$$\langle P_n|f\rangle = \int_a^b w(x)P_n^*(x)f(x)dx \tag{126}$$

In practice the infinite sum in truncated. This is usually reasonable because of the Parseval relation, which ensures that sum of the squares of the discarded coefficients gets small as more coefficients are retained. There is still the matter of computing the inner products (126) involving the retained basis functions.

In practice these integrals cannot be done analytically for an arbitrary function. The need to be approximated. There is a class of approximations that are closely related to orthogonal polynomials with a given weight.

To motivate the problem of Gaussian Quadrature let $w(x)$ be positive weight function on the intervale $[a,b]$

Problem: Find N points $\{x_i\}$ and weights $\{w_i\}$ such that for any polynomial of degree $2N-1$

$$\int_a^b w(x)P(x)dx = \sum_{n=1}^N P(x_n)w_n \tag{127}$$

Before I discuss the solution of this problem I discuss how these quadrature rules can be used. Assume that $f(x)$ is a function representing a vector in the Hilbert space. Assume that $f(x)$ can be accurately approximated by a polynomial of degree N. Let $\{x_n\}$ and $\{w_n\}$ be the points and weights for an N point quadrature with weight $w(x)$. Then

$$\langle f|P_n\rangle = \int_a^b w(x)P_n^*(x)f(x)dx \approx \sum_{m=1}^N w_m f(x_m)P_n(x_m) \tag{128}$$

This formula is exact if $f(x)$ exactly a polynomial of degree N.

To construct the points and weight associated with a given weight function $w(x)$ begin by assuming that the $\{x_n\}_{n=1}^N$ are known and define the polynomial

$$P_N(x) = \prod_{n=1}^N (x-x_n) = \sum_{n=0}^N p_n x^n \tag{129}$$

By construction the points $\{x_n\}$ are roots of $P_N(x)$:

$$P_N(x_n) = 0 \qquad n = 1, \cdots, N \tag{130}$$

Define

$$Q_m := P_N(x)x^m \qquad m = 0, \cdots, N-1 \tag{131}$$

Since $P_N(x)$ can be factored out of each $Q_m(x)$ the points $\{x_n\}$ are also roots of each $Q_m(x)$:

$$Q_m(x_n) = 0 \qquad n = 1, \cdots, N, \qquad m = 0, \cdots, N-1 \tag{132}$$

To find the desired quadrature points require

$$\int_a^b Q_m(x)w(x)dx = \sum_{n=1}^{N} Q_m(x_n)w_n = 0 \tag{133}$$

or equivalently

$$\int_a^b w(x)P_N(x)x^m dx = 0 \tag{134}$$

Using the expansion for $P_N(x) = \sum p_n x^n$ I can write

$$0 = \sum_{n=0}^{N} p_n \int_a^b x^{m+n} w(x)dx \tag{135}$$

It follows from (129) that $p_N = 1$ so this becomes a linear system for the coefficients of the polynomial $P_N(x)$:

$$\sum_{n=1}^{N-1} I_{mn} p_n = q_m \tag{136}$$

where

$$I_{mn} := \int_a^b x^{m+n} w(x)dx \tag{137}$$

$$q_m := -\int_a^b x^{N+m} w(x)dx \tag{138}$$

The coefficients p_n can be determined by solving the linear system (136). The zeros of this polynomial are the quadrature points.

To connect this with the discussion of orthogonal polynomials note that the equation

$$\int_a^b w(x)P_N(x)x^m dx = 0 \qquad m = 0 \cdots N-1 \tag{139}$$

means that up to an overall constant multiplier, $P_N(x)$ is the degree N orthogonal polynomial on the interval $[a,b]$ with weight w. Thus the points x_n are the roots of the degree N orthogonal polynomial on $[a,b]$ with weight $w(x)$.

Given the quadrature points is it still necessary to calculate the weights. To do this define the polynomials

$$L_n(x) := \frac{\prod_{m\neq n}(x-x_m)}{\prod_{m\neq n}(x_n-x_m)} \qquad L_n(x_m)=\delta_{mn} \tag{140}$$

Using these polynomials

$$\int_a^b w(x)L_m(x)dx = \sum_n w_n L_m(x_n) = \sum \delta_{mn} w_n = w_m \tag{141}$$

which gives the integral expressions for the weights

$$w_n = \int_a^b w(x)L_n(x)dx \tag{142}$$

By accurately performing the integrals (142) once , and computing the roots of $P_N(x)$ is it possible to accurately integrate a large class of function on $[a,b]$

While roots and weights still have to be evaluated, there exist standard subroutines and tables that give the points and weights for Gauss quadratures associated with all of the Classical orthogonal polynomials.

The most common of the quadratures is the Gauss-Legendre quadrature. In this case the quadrature points are the zeros on the N-ht Legendre polynomial. It is associated with the weight $w(x)=1$ on the interval $[-1,1]$ It can be linearly transformed of any other interval. To integrate functions that are not well-approximated by polynomials, it is possible to break the domain of integration into subintervals, where Guass quadrature methods are applied to each subinterval.

0.8 Lecture 8:

The Müntz-Szasz Theorem: Let $0 < \lambda_1 < \lambda_2 \cdots < \infty$ and consider $\{1, t^{\lambda_1}, t^{\lambda_2}, \cdots,\}$ is dense in the continuous functions on $[0,1]$ if an only if

$$\sum_{n=1}^{\infty} \frac{1}{\lambda_n} = \infty \tag{143}$$

Obviously $\sum_n \frac{1}{n} = \infty$ which is consistent with the Weierstrass theorem. What is interesting is that this inequality is still satisfied if we remove any finite subset of the x^m's from the basis set. For example the odd or even polynomial are separately dense on $[a,b]$.

This result does not apply to orthogonal polynomials, which is clear from the definition of basis. It is special to the infinite dimensional case.

I will outline the proof of the result, only because is uses properties of analytic functions derived last semester.

To prove the Müntz-Szasz theorem I observe that if the set of functions $\{x^{\lambda_n}\}$ are not a basis, then there is a continuous function $\phi(x)$ such that

$$\int_0^1 x^{\lambda_n} \phi(x) dx = 0 \tag{144}$$

for all n.

By contradiction I assume that a non-zero $\phi(x)$ satisfying (144) exists. For that function I define

$$f_\phi(z) := \int_0^1 t^z \phi(t) dt \qquad z = x + iy \tag{145}$$

Note that

$$t^z = e^{z\ln(t)} = t^x t^{iy} = e^{x\ln(t)} e^{iy\ln(t)} = t^x e^{iy\ln(t)} \tag{146}$$

For $x >$ and $t \in (0,1]$

$$\lim_{t\to 0} |e^{x\ln(t)} e^{iy\ln(t)}| = e^{-x|\ln(t)|} = 0 \tag{147}$$

It follows that t^z is analytic in the right half plane and

$$|t^z| = t^x = e^{x\ln(t)} \leq 1 \tag{148}$$

To demonstrate the analyticity note t^z is continuous for $x > 0$ and satisfies the Cauchy Riemann equations:

$$t^z = u + iv = e^{x\ln(t)}\cos(y\ln(t)) + ie^{x\ln(t)}\sin(y\ln(t)) \tag{149}$$

$$\frac{\partial u}{\partial x} = \ln(t)u(x) = \frac{\partial v}{\partial y} \tag{150}$$

$$\frac{\partial u}{\partial y} = -\ln(t)v(x) = -\frac{\partial v}{\partial x} \tag{151}$$

In integral representation

$$f_\phi(z) := \int_0^1 t^z\phi(t)dt \tag{152}$$

For z in the right half plane t^z is analytic for any $t \in [0,1]$ and $t^z\phi(t)$ continuous in t for $t \in [0,1]$. Theorem 2 on page 40 of the text(or page 56 of last semesters notes) implies that $f(z)$ is analytic in the right half plane.

Next note that

$$|f(z)| \leq \int_0^1 |t^z||\phi(t)|dt = \int_0^1 t^x|\phi(t)| \leq \int_0^1 |\phi(t)| = \max_{t\in[0,1]}|\phi(t)| = C < \infty \tag{153}$$

The last step follows because a continuous function on a closed interval is necessarily bounded.

The condition that

$$f_\phi(\lambda_n) := \int_0^1 t^{\lambda_n}\phi(t)dt = 0 \tag{154}$$

means that $f(z)$ has a zero for each of real exponents λ_n.

What we have established is that $f(z)$ is a bounded analytic function in the right half plane with an infinite number of zeros on the positive real line.

I use the conformal mapping

$$z = \frac{1+z'}{1-z'} \quad z' = \frac{1-z}{1+z} \tag{155}$$

to map the right half plane to the interior of the unit disc

$$z = \frac{1+re^{i\phi}}{1-re^{i\phi}} = \frac{1-r^2+2ir\sin(\phi)}{1+r^2-2r\cos(\phi)} \tag{156}$$

It is not hard to check that z takes on all values in the right half complex as z' varies over the interior of the unit disk.

Thus

$$g(z') = f(\frac{1+z'}{1-z'}) \tag{157}$$

is a bounded analytic function on the disk with and infinite number of zeros at

$$\alpha_n = \frac{\lambda_n - 1}{\lambda_n + 1} \tag{158}$$

These zeros accumulate at the point $z' = 1$

Consider the contour integral

$$\frac{1}{2\pi i} \int \frac{1-z}{1+z} \frac{1}{g(z)} \frac{dg}{dz}(z) dz \tag{159}$$

where the contour is in the unit circle enclosing the interval $[-1+\epsilon, 1-\delta]$, and we take the limit that $\delta \to 0$. The function

$$\frac{1}{g(z)} \frac{dg}{dz}(z) dz \tag{160}$$

has isolated poles at α_n. The integrand vanishes at 1 and the point -1 is avoided by the choice of contour. The theorem on page 85 implies that

$$\frac{1}{2\pi i} \int \frac{1-z}{1+z} \frac{1}{g(z)} \frac{dg}{dz}(z) dz \sum_n \frac{1-\alpha_n}{1+\alpha_n} = \sum_n \frac{1}{\lambda_n} + \text{other terms} \tag{161}$$

The other terms come from (a) additional roots in the right half plane or higher multiplicities of the roots λ_n. In all cases they have positive real parts.

Since the integrand is bounded on this contour it follows that

$$\sum_n \frac{1}{\lambda_n} < |\sum_n \frac{1}{r_n}| = C < \infty \tag{162}$$

where the second sum is over all roots times multiplicites in the right half plane.

It the sum on the left is finite, then is it possible that there is a non-zero continuous function $\phi(x)$ orthogonal to all of the x_n^λ, if on the other hand the sum is infinite then there are no continuous functions $\phi(t)$ that are orthogonal to all of the t^{λ_n}'s. This means that the set $\{z^{\lambda_n}\}$ is complete.

Clearly the powers x_n satisfy $\sum_{n=0}^{\infty} \frac{1}{n} = \infty$ which is consistent with what we learned from the Weierstrass theorem. What is interesting is that we can remove a finite and in some cases infinite number of terms from this basis and still get a basis.

This theorem does not apply if we orthogonalize our basis functions using the Gram Schmid method. It we remove one function from an orthonormal basis we have a non-trivial function that is orthogonal to the rest of the basis functions, which means that the remaining functions are not a basis.

0.9 Lecture 9:

In this section the Weierstrass approximation theorem is used to introduce Fourier series.

Let $f(\theta)$, $\theta \in [-\pi, \pi]$ be continuous and periodic, $f(\pi) = f(-\pi)$. Since this function is periodic it follows that $\tilde{f}(r, \theta) = rf(\theta)$ is a continuous function in the plane, where r and θ are considered as polar coordinates. Let

$$x = r\cos(\theta) \qquad y = r\sin(\theta) \tag{163}$$

$$g(x, y) = g(r\cos(\theta), r\sin(\theta)) = \tilde{f}(r, \theta) \tag{164}$$

Since $g(x, y)$ is continuous, it is also continuous on $x \in [-1, 1]$, $y \in [-1, 1]$. We can apply the Weierstrass theorem separately to each variable. The proof is a straightforward extension of the one-variable proof. This means that there is a polynomial

$$G_{mn} = \sum_{k,l=0}^{m,n} g_{kl} x^k y^l \tag{165}$$

It is useful to choose the polynomial in x and y to have the same order ($m = n$). This can be done by choosing maximum order to be the larger of m or n. The Weierstrass theorem means that for any $\epsilon > 0$ there is a sufficiently large N such that

$$|g(x, y) - \sum_{m,n=0}^{N} g_{mn} x^m y^n| < \epsilon \tag{166}$$

This result still holds if it is expressed in terms of polar coordinates:

$$|\tilde{f}(r, \theta) - \sum_{m,n=0}^{N} g_{mn} r^{m+n} \cos^m(\theta) \sin^n(\theta)| < \epsilon \tag{167}$$

Since this holds uniformly for all x, y in the square of side length 2 centered about the origin, we can set $r = 1$ and the inequality is still valid, which gives

$$|f(\theta) - \sum_{m,n=0}^{N} g_{mn} \cos^m(\theta) \sin^n(\theta)| < \epsilon \tag{168}$$

This series converges uniformly and pointwise for all $\theta \in [-\pi, \pi]$.

It looks more symmetric to replace

$$\cos(\theta) = \frac{e^{i\theta} + e^{-i\theta}}{2} \qquad \sin(\theta) = \frac{e^{i\theta} + e^{-i\theta}}{2i} \tag{169}$$

Using the binomial theorem we have

$$\cos^m(\theta) = \frac{1}{2^m} \sum_{k=0}^{m} \frac{m!}{k!(m-k)!} e^{i(2k-m)\theta} \tag{170}$$

$$\sin^m(\theta) = \frac{(i)^n}{2^m} \sum_{k=0}^{m} \frac{m!(i^n)}{k!(m-k)!} e^{-i(2k-m)\theta} \tag{171}$$

For both sums $-m \leq 2k - m \leq m$

$$\frac{e^{i\theta} + e^{-i\theta}}{2i} \tag{172}$$

This leads to an alternative expression for the series

$$|f(\theta) - \sum_{m=-2N}^{2N} c_m e^{im\theta}| < \epsilon \tag{173}$$

This shows that any continuous periodic function can be uniformly pointwise approximated by a trigonometric polynomial of sufficiently high order.

The basis functions $e^{im\phi}$ have the nice property that they are orthogonal for different values of m:

$$\int_{-\pi}^{\pi} e^{-im\theta} e^{in\theta} d\theta = 2\pi\delta_{mn} \tag{174}$$

We define the orthonormal basis functions

$$|e_m\rangle \qquad \langle\theta|e_m\rangle = \frac{1}{\sqrt{2\pi}} e^{im\theta} \tag{175}$$

Equation (174) is equivalent to

$$\langle e_m|e_n\rangle = \int_{-\pi}^{\pi} \langle e_m|\theta\rangle d\theta \langle\theta|e_n\rangle = \delta_{mn} \tag{176}$$

We can use the Weierstrass theorem to calculate expansion coefficients for any continuous periodic functions. This does not use the inner product.

An alternative is to choose the coefficients to minimize the $L^2[-\pi, \pi]$ norm of the difference

$$\| |f\rangle - \sum_{m=-N}^{N} f_m |e_m\rangle \|^2 =$$

$$\langle f|f\rangle + \sum_n |f_n|^2 + \sum_m (f_m^* \langle e_m|f\rangle - f_m \langle f|e_m\rangle) =$$

$$\langle f|f\rangle - \sum_n |\langle f|e_n\rangle|^2 + \sum_n |\langle f|e_n\rangle|^2 +$$

$$\sum_n |f_n|^2 + \sum_m (f_m^* \langle e_m|f\rangle - f_m \langle f|e_m\rangle =$$

$$\langle f|f\rangle - \sum_n |\langle f|e_n\rangle|^2 + \sum_n |f_m - \langle e_m|f\rangle|^2$$

Bessel's inequality tells us that

$$\sum_n |\langle f|e_n\rangle|^2 \leq \langle f|f\rangle \tag{177}$$

while the terms that depends on f_m is non-negative. A minimum will be achieved only when

$$f_m = \langle e_m|f\rangle \tag{178}$$

This leads to the Fourier expansion of the periodic function $f(\phi)$

$$|f\rangle = \sum_{m=-\infty}^{\infty} |e_m\rangle\langle e_m|f\rangle \tag{179}$$

This converges for continuous functions because the error is smaller than the Weierstrass bound with goes to zero.

0.10 Lecture 10:

There various generalization of the convergence result that extend to some non-continuous functions. For example if $f(\phi)$ is piecewise differentiable on finite subintervals and periodic then

$$\frac{1}{2}(f(\theta^+) + f(\theta^-)) = \sum_{n=-\infty}^{+\infty} \langle\theta|e_n\rangle\langle e_n vert f\rangle \tag{180}$$

where

$$f(\theta^\pm) = \lim_{\epsilon\to 0^+} f(\theta \pm \epsilon) \tag{181}$$

If θ is a point of continuity then $f(\theta^\pm) = f(\theta)$.

The proof of this is similar to the proof of the next theorem, where we let the finite interval approach the real line. I will prove the following important result:

Theorem: Let $f(x)$ be absolutely integrable and piecewise differentiable on the real line. Then

$$\frac{1}{2}(f(x^+) - f(x^-)) = \frac{1}{\pi}\int_0^\infty d\lambda \int_{-\infty}^\infty f(y)\cos(\lambda(x-y))dy \tag{182}$$

To prove this I replace the integral by

$$\frac{1}{\pi}\int_0^\infty d\lambda \int_{-\infty}^\infty f(y)\cos(\lambda(x-y))dy =$$

$$\lim_{\Lambda\to\infty}\frac{1}{\pi}\int_0^\Lambda d\lambda \int_{-\infty}^\infty f(y)\cos(\lambda(x-y))dy \tag{183}$$

Since

$$f(y)\cos(\lambda(x-y)) \tag{184}$$

is absolutely integrable on $[0,\Lambda]$ I can change the order of the integration to get

$$\frac{1}{\pi}\int_0^\infty d\lambda \int_{-\infty}^\infty f(y)\cos(\lambda(x-y))dy = \lim_{\Lambda\to\infty}\frac{1}{\pi}\int_{-\infty}^\infty dy f(y)\int_0^\Lambda d\lambda\cos(\lambda(x-y)) =$$

$$\lim_{\Lambda\to\infty}\frac{1}{\pi}\int_{-\infty}^\infty dy f(y)\frac{\sin\Lambda(x-y)}{x-y} \tag{185}$$

Next let $y' = y - x; \qquad y = y' + x$ (note that $\sin(ax)/x$ is even, to get

$$\lim_{\Lambda \to \infty} \frac{1}{\pi} \int_{-\infty}^{\infty} dy f(x + y') \frac{\sin(\Lambda y')}{y'} \tag{186}$$

From last semester we used contour integrals to calculate

$$\int_0^{\infty} \frac{\sin(\Lambda y)}{y} dy = \frac{\pi}{2} \qquad \int_{-\infty}^{0} \frac{\sin(\Lambda y)}{y} dy = \frac{\pi}{2} \tag{187}$$

Using these integrals we can write our integral as

$$I - \frac{1}{2}(f(x^+) - f(x^-)) = \lim_{\Lambda \to \infty} \left[\int_0^{\infty} (f(y + x) - f(x^+)) \frac{\sin(\Lambda y)}{y} dy + \right.$$

$$\left. \int_{-\infty}^{0} (f(y + x) - f(x^-)) \frac{\sin(\Lambda y)}{y} dy \right] \tag{188}$$

The problem is to show that the two integrals on the right side of this equation vanish as $\Lambda \to \infty$. I break each integral up into four parts

$$\int_0^{\epsilon} (f(y + x) - f(x^+)) \frac{\sin(\Lambda y)}{y} dy \tag{189}$$

$$\int_{\epsilon}^{A} (f(y + x) - f(x^+)) \frac{\sin(\Lambda y)}{y} dy \tag{190}$$

$$\int_{A}^{\infty} f(y + x) \frac{\sin(\Lambda y)}{y} dy \tag{191}$$

$$-f(x^+)) \int_{A}^{\infty} \frac{\sin(\Lambda y)}{y} dy \tag{192}$$

0.11 Lecture 11:

The integrand in (189) is uniformly bounded. This is because the function has a right handed derivative and $f(x+y)$ approached $f(x^+)$ as $y \to zero$ from above. The integral can be made as small as desired by choosing ϵ small enough.

The difference in (190) is piecewise differentiable. It can be expressed as a finite sum of integrals of differentiable functions on disjoint subintervals. This means that we can write

$$\sin(\Lambda y) = -\frac{1}{\Lambda}\frac{d}{dy}\cos(\Lambda y) \tag{193}$$

This can be integrated by parts. Since everything is finite, after integrating by parts this behaves like a constant times $\frac{1}{\Lambda}$. For any fixed ϵ, and A this vanishes $\Lambda \to \infty$.

The integral in (191) can be made as small as desired, using the absolute integrability of f/y, by choosing A sufficiently large. In practice one first choose A large enough so this is small, then choose Λ large enough to make the other integrals small.

The integral in (192) can be transformed to

$$\int_{\Lambda A}^{\infty} \frac{\sin(u)}{u} du \to 0 \tag{194}$$

as $\Lambda \to \infty$. Because this oscillates it is useful to write the integral as a sum of positive terms

$$\sum_{n=m}^{\infty} \int_{2n\pi}^{2(n+1)\pi} \frac{\sin(u)}{u} du \tag{195}$$

Since the sum converges, and the terms are all positive, this shows that the sum vanishes as $m \to \infty$.

The other integral can be made to vanish in the same way. The final result is

$$I = \frac{1}{2}(f(x^+) - f(x^-)) = \frac{1}{\pi}\int_0^{\infty} d\lambda \int_{-\infty}^{\infty} f(y)\cos(\lambda(x-y))dy \tag{196}$$

The more familiar form of this result is obtained by replacing the $\cos(\lambda(x-y))$ by its exponential form.

$$\frac{1}{2}(f(x^+) - f(x^-)) =$$

$$
\frac{1}{2\pi}\int_0^\infty d\lambda \int_{-\infty}^\infty f(y)(e^{i\lambda(x-y)} + e^{-i\lambda(x-y)}))dy =
$$

$$
\frac{1}{2\pi}\int_{-\infty}^\infty d\lambda \int_{-\infty}^\infty e^{i\lambda(x-y)} f(y))dy \tag{197}
$$

When $f(x)$ is continuous this gives

$$
f(x) =
$$

$$
\frac{1}{2\pi}\int_{-\infty}^\infty d\lambda \int_{-\infty}^\infty e^{i\lambda(x-y)} f(y))dy \tag{198}
$$

We define the Fourier transform of $f(x)$ by

$$
\tilde{f}(k) := \frac{1}{\sqrt{2\pi}} \int e^{-iky} f(y)dy \tag{199}
$$

The inverse Fourier transform is given by

$$
f(x) := \frac{1}{\sqrt{2\pi}} \int e^{iky} \tilde{f}(k)dk \tag{200}
$$

While this relation looks symmetric, except for the sign in the exponent, we started with a piecewise differentiable, absolutely integral function, $f(x)$. We have not established that $\tilde{f}(k)$ has any of these properties.

While there are a number of distinct relations between properties of functions and their Fourier transforms, one of the most useful relations involves the space $\mathcal{S}(\mathbb{R})$ of Schwartz functions of a real variable.

The useful property is that the Fourier transform of a Schwartz function is a Schwartz function.

A function $f(x)$ is a Schwartz function, written $f(x) \in \mathcal{S}(\mathbb{R})$ if and only if $f(x)$ has an infinite number of derivatives and $f(x)$ and each of its derivatives fall off faster than any inverse polynomial in the sense:

$$
\lim_{|x|\to\infty} |x^m \frac{d^m f}{dx_m}(x)| = 0 \tag{201}
$$

While this condition sounds restrictive, the functions

$$
f_n(x) = \frac{1}{\sqrt{h_n}} H_n(x) e^{-x^2/2}, \tag{202}
$$

where $H_n(x)$ are the Hermite polynomials, and the h_n are the normalization integrals, are all Schwartz functions.

0.12 Lecture 12:

While finite linear combinations of Schwartz functions are Schwartz functions, infinite linear combinations are not necessarily Schwartz functions. This follows because the functions $f_n(x)$ in (202) above are known to be an orthonormal basis for the Hilbert space of square integrable functions on the real line, but we know that there are many square integrable functions that do not have as many derivatives or fall off as fast as Schwartz functions.

It turns out that the Schwartz functions are members of a complete metric space defined using an infinite collection of norms:

$$\|f\|_{mn} := \sup_{x\in\mathbb{R}} |x^n \frac{d^m f}{dx}(x)| \tag{203}$$

$$\rho(f-g) := \sum_{mn} \frac{1}{2^{m+n}} \|(f-g)\|_{mn} \tag{204}$$

This is the only place in this course where inner product spaces and normed linear spaces are insufficient. It is worth noting that there are different equivalent ways to write this metric that appear in the literature. Two metrics are equivalent if every Cauchy sequence in one metric is Cauchy in the other metric.

Note that

$$\frac{1}{\sqrt{2\pi}} \int e^{-iky} y_n \frac{d^m f}{dy^m}(y) dy = (i)^m \frac{d^n}{dk^n}(k^m \tilde{f}(k)) \tag{205}$$

which means that the meaning of m and n get interchanged under Fourier transforms. Note that

$$\frac{1}{\sqrt{2\pi}} \int e^{-iky} f(y) dy = \frac{-i}{k} \frac{1}{\sqrt{2\pi}} \int e^{-iky} \frac{df}{dy}(y) dy = \tag{206}$$

Since $\frac{df}{dy}(y)dy$ is also a Schwartz function the integral has a bound independent of k, so the Fourier transform of a Schwartz function vanishes as $|k| \to \infty$. Because derivatives of Schwartz functions are Schwartz functions, and polynomials time Schwartz functions are Schwartz functions, the Fourier transforms of these functions also vanish in the same limit. Since polynomial time derivatives of Schwartz function get Fourier transformed into powers and derivatives, these Fourier transforms also vanish as $|k| \to \infty$. This shows that Schwartz functions are closed under Fourier transform.

A useful characterization of Schwartz functions can be given in terms of their expansion coefficients in the orthonormal basis of Hermite functions

$$\langle x|n\rangle = \frac{1}{\sqrt{h_n}} H_n(x) e^{-\frac{1}{2}x^2}. \tag{207}$$

Assume that

$$f(x) = \sum_{n=0}^{\infty} \langle x|n\rangle f_n then \tag{208}$$

$$f(x) \in L^2(\mathbb{R}) \qquad \Longleftrightarrow \qquad \sum_{n=0}^{\infty} |f_n|^2 < \infty \tag{209}$$

and it can be shown that

$$f(x) \in \mathcal{S}(\mathbb{R}) \qquad \Longleftrightarrow \qquad \sum_{n=0}^{\infty} |f_n|^2 (1+n)^m < \infty \qquad m \geq 0 \tag{210}$$

Thus we see that Schwartz functions have restrictive growth conditions in their expansion coefficients in the Hermite basis.

Later we will extend Fourier transforms to a larger class of functions. Fourier transforms are especially useful for treating linear partial differential equations. For example consider the wave equation

$$[\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2}] u(x,t) = 0 \tag{211}$$

If we express $u(x,t)$ in terms of its Fourier transform

$$u(x,t) = \frac{1}{2\pi} \int \tilde{u}(k,\omega) e^{i\omega t + ikx} dk d\omega \tag{212}$$

then equation (169) can be written as

$$\frac{1}{2\pi} \int (-\frac{\omega^2}{c^2} + k^2) \tilde{u}(k,\omega) e^{i\omega t + ikx} dk d\omega = 0 \tag{213}$$

This will vanish provided

$$\omega = \pm ck \tag{214}$$

which leads to a general solution of the form

$$u(x,t) = \frac{1}{2\pi} [\int \tilde{f}_l(k) e^{ik(x+ct)} dk + \int \tilde{f}_r(k) e^{ik(x-ct)} dk] \tag{215}$$

$$\frac{1}{\sqrt{2\pi}}(f_l(x+ct)+f_r(x-ct)) \tag{216}$$

In this case $f_l(x)$ and $f_r(x)$ are arbitrary functions. $f_l(x+ct)$ represents a disturbance moving to the left while $f_r(x-ct)$ represents a disturbance moving to the right. The function $\hat{f}(\pm\omega/c)$ is the angular frequency of this function.

It is easy to check that functions of this general form are solutions of the wave equation. This application is typical of how the Fourier transforms are used. The important property is that derivatives get converted into multiplication operators. In this simple case the partial differential equation was reduced to the algebraic equation, $c^2k^2-\omega^2=0$.

Returning to the original Fourier transformation, for Schwartz functions we have shown that

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}d\lambda\int_{-\infty}^{\infty}e^{i\lambda(x-y)}\tilde{f}(y))dy \tag{217}$$

If I change the order of the integrals without first cutting off the limit of integration in the λ integral I get the quantity

$$\delta(x-y)=\frac{1}{2\pi}\int_{-\infty}^{\infty}d\lambda e^{i\lambda(x-y)}, \tag{218}$$

which does not define a convergent integral. However $\delta(x-y)$ formally satisfies

$$f(x)=\int\delta(x-y)f(y)dy \tag{219}$$

In terms of the Lebesgue integral, if $\delta(x-y)$ represented a function it would have to be zero except on a set of measure zero. This would make the Lebesgue integral zero; so we cannot think of $\delta(x-y)$ as a function. It is usually called a Dirac delta function. Quantities of like the Dirac delta functions are called generalized functions or distributions.

We can represent the integral of a Dirac delta function multiplied by a function as a limit of integrals over a sequence of functions multiplied by the same function. For example

$$\delta(x-y)=\lim_{\Lambda\to\infty}\frac{1}{\pi}\frac{\sin(\Lambda(x-y))}{x-y} \tag{220}$$

has the property

$$\lim_{\Lambda\to\infty} \int f(y) \frac{1}{\pi} \frac{\sin(\Lambda(x-y))}{x-y} dy = f(x) = \int f(y)\delta(x-y)dy. \qquad (221)$$

This makes sense provided the y integral is performed before taking the limit. It can also be written as

$$\delta_n(x) = \begin{cases} n & -\frac{1}{2n} < x < \frac{1}{2n} \\ 0 & |x| \geq \frac{1}{2n} \end{cases} \qquad (222)$$

which also has the property

$$\lim_{n\to\infty} \int f(y)\delta_n(x-y)dy = f(x) = \int f(y)\delta(x-y)dy. \qquad (223)$$

The integrals on the left exist for each n and the limit exists if f is continuous. There are many other sequences of functions that can represent the "dirac delta function" $\delta(x-y)$. All that matters is that (1) they integrate to 1 and (2) the area that contributes to most of the integral eventually contains only the point x.

While there are many ways to represent delta functions, the basic rule

$$\int f(y)\delta(x-y)dy = f(x) \qquad (224)$$

is simple, and can be implemented without the use of sequences of auxiliary functions or the computation of complicated integrals

What we can observe about the rule (225) is that it is (1) linear in f and its value is a complex number. This is a property the it shares with the integral of a fixed function multiplied by f.

Definition: A tempered distribution is a continuous linear functional on the space of Schwartz functions.

The space of tempered distributions is denoted by $\mathcal{S}'(\mathbb{R})$.

Thus if L is a tempered distribution, α is complex, and $f(x)$ and $g(x)$ are Schwartz functions, then

$$L[f+\alpha g] = L[f] + \alpha L[g] \qquad (225)$$

and

$$\lim_{l\to\infty} \rho(f_l - g) := \lim_{l\to\infty} \sum_{mn} \frac{1}{2^{m+n}} \|(f_l - g)\|_{mn} \to 0 \qquad (226)$$

then

$$\lim_{l\to\infty} L[f_l] = L[g] \tag{227}$$

In the case of the delta function

$$L[f] = f(x) \tag{228}$$

defines a tempered distribution since

$$L[f+\alpha g] = f(x) + \alpha g(x) \tag{229}$$

$$\rho(f_n - g) \to 0 \tag{230}$$

$$\lim_{n\to\infty} L[f_n] = L[\lim_{n\to\infty}] = L[g] = g(x) \tag{231}$$

because the

$$\lim_{n\to\infty} \rho(f_n - g) \to 0 \Rightarrow sup_{y\in\mathbb{R}}|f_n(y) - g(y)| \to 0 \Rightarrow$$

$$|f_n(x) - g(x)| \to 0 \tag{232}$$

Many tempered distributions can be written as integrals. In a Hilbert space of square integrable functions, each element $|f\rangle \in \mathcal{H}$ defines a linear functional on $\mathcal{H}$ by the inner product

$$L_f[g] = \int f^*(x)g(x)dx \tag{233}$$

These linear functional are also tempered distribution because they are alco continuous on the subspace of Schwartz functions. If I expand $f(x)$ and $g(x)$ in terms of the basis functions

$$\langle x|n\rangle = \frac{1}{\sqrt{h_n}} H_n(x) e^{-x^2/2} \tag{234}$$

$$|f\rangle = \sum |n\rangle f_n \tag{235}$$

$$|g\rangle = \sum |n\rangle g_n \tag{236}$$

Then the Schwartz inequality

$$\sum |f_n^* g_n|^2 = |\langle f|g\rangle|^2 \le [\sum_n |f_n^2|][\sum_m |g_m^2|] \tag{237}$$

ensures that this scalar product is well defined.

When $g(x)$ is a Schwartz function the coefficients g_n fall off faster than $[\sum_m |g_m^2|]$. This means that it is possible to relax the fall off conditions on f so it is no longer square integrable, but gives a finite scalar product.

Specifically, if I require

$$\sum_n |g_n|^2 (1+n)^m < \infty \qquad \forall m \tag{238}$$

and

$$\sum_n |f_n|^2 \frac{1}{(1+n)^m} < \infty \qquad \text{for some } m \tag{239}$$

the Schwartz inequality

$$\sum |f_n g_n| = \sum |\frac{f_n}{\sqrt{(1+n)^m}} \sqrt{(1+n)^m} g_n|$$

$$\leq [\sum_n \frac{1}{(1+n)^m} |f_n^2|]^{1/2} [\sum_m (1+n)^m |g_m^2|]^{1/2} \tag{240}$$

This means if g is a Schwartz function and f is a tempered distribution this "inner product" is still finite.

Condition (239) is one way to characterize tempered distributions. We see that the tempered distributions have less restrictive growth conditions than square integrable functions. One consequence of the Hermite representation of tempered distributions is that any tempered distribution, including the delta function, can be approximated by an expansion in the Hermite basis functions

$$L = \sum_n c_n^* \langle n| \tag{241}$$

The only new property is that the c_n are restricted so

$$\sum_n \frac{|c_n|^2}{(1+n)^m} < \infty \tag{242}$$

for some m so

$$L[f] = \sum_n c_n^* \int \langle n|x\rangle dx \langle x|f\rangle \tag{243}$$

If I consider

$$\sum_n c_n^* \int \frac{d\langle n|x\rangle}{dx} \tag{244}$$

and apply this to a Schwartz function $\langle x|f\rangle$

$$L'[f] = \sum_n c_n^* \int \frac{d\langle n|x\rangle}{dx} dx \langle x|f\rangle = \tag{245}$$

$$L[-\frac{df}{dx}] = -\sum_n c_n^* \int \langle n|x\rangle dx \frac{d\langle x|f\rangle}{dx} \tag{246}$$

Interchanging the order of the sum an integral is valid provided $\langle x|f\rangle$ is a Schwartz function and the coefficients c_n satisfy the growth condition for a tempered distribution. Note that there are no contributions when integrating by parts because of the Schwartz functions and all of their derivatives vanish as $x \to \pm\infty$. Since derivatives of Schwartz functions are Schwartz functions the right hand side of this expression is always defined if L is a tempered distribution.

This leads us to define the derivative of the tempered distribution $L[f]$ by

$$L'[f] = -L[\frac{df}{dx}] \tag{247}$$

This can be repeated n times to get the n-th derivative of a tempered distribution

$$L^{(n)}[f] := (-)^n L[\frac{d^n f}{dx^n}] \tag{248}$$

Clear the tempered distribution have an infinite number of derivatives because the Schwartz functions have an infinite number of derivatives. These derivative are only well defined on Schwartz functions, they may not be defined on an ordinary square integrable function.

As an example consider the Heaviside function

$$\theta(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases} \tag{249}$$

$$L[f] = \int \theta(x) f(x) dx = \int_0^\infty f(x) dx \tag{250}$$

$$L'[f] = -L[f'] = -\int \theta(x) \frac{df(x)}{dx} dx =$$

$$-\int_0^\infty \frac{df(x)}{dx}dx = -f(\infty) + f(0) = f(0) \tag{251}$$

But this have the same value as the distribution $\delta(x)$. This gives the distributional identity

$$\frac{d}{dx}\theta(x) = \delta(x) \tag{252}$$

We will used this identity when we discuss Green functions. It is useful to establish some additional properties of tempered distributions.

We introduced Schwartz functions because Schwartz functions get mapped into Schwartz functions under Fourier transform. We can use this property to define the Fourier transform of a tempered distribution:

First let $f(x)$ and $g(x)$ both be Schwartz functions. Then the Fourier transform of these functions are

$$f(x) = \frac{1}{\sqrt{2\pi}} \int dk e^{ikx} \hat{f}(k) \tag{253}$$

$$g(x) = \frac{1}{\sqrt{2\pi}} \int dk e^{ikx} \hat{g}(k). \tag{254}$$

Since Schwartz functions are also square integrable, the inner product of f with g is defined

$$\langle f|g\rangle = \int f(x)^* g(x) dx = \frac{1}{2\pi} \int dx dk_1 dk_2 e^{ix(k_2-k_1)} \tilde{f}^*(k_1)\tilde{g}(k_2) = \tag{255}$$

$$\int dk_1 \tilde{f}^*(k_1)\tilde{g}(k_1) = \langle \tilde{f}|\tilde{g}\rangle \tag{256}$$

This shows that Fourier transforms preserve the Hilbert inner product.

This property can be used to define the Fourier transform of a tempered distribution. If f is a tempered distribution we can write

$$f = \sum_{n=0}^\infty f_n \phi_n(x) \tag{257}$$

$$\phi_n(x) := \langle x|n\rangle \tag{258}$$

We define finite partial sums

$$f_N(x) := \sum_{n=0}^N f_n \phi_n(x). \tag{259}$$

The f_N are all Schwartz functions.

It follows that if g is a Schwartz function

$$\langle f|g\rangle = \lim_{N\to\infty}\langle f_N|g\rangle =$$

$$\lim_{N\to\infty}\langle \tilde{f}_N|\tilde{g}\rangle := \langle \tilde{f}|\tilde{g}\rangle \tag{260}$$

This defines the Fourier transform of any tempered distribution.

Let's use this to compute the Fourier transform of a delta function. Start by letting $g(x)$ be a Schwartz function. We can write

$$\langle \tilde{\delta}_y|\tilde{g}\rangle = \int \delta(x-y)g(x)dx = g(y) = \frac{1}{\sqrt{2\pi}}\int e^{iky}\tilde{g}(k)dk \tag{261}$$

one can immediately read off the linear operator that act on $\tilde{g}(k)$:

$$\tilde{\delta}_y(k) = \frac{1}{\sqrt{2\pi}}e^{-iky} \tag{262}$$

We can also arrive at the same result using the formal rule

$$\tilde{\delta}_y(k) = \frac{1}{\sqrt{2\pi}}\int \delta(x-y)e^{-ikx}dx = \frac{1}{\sqrt{2\pi}}e^{-iky} \tag{263}$$

It often happens that while an equation may have no solutions in the space of square integrable functions, if the space is enlarged they might have solutions in the larger space. The space of tempered distributions is much larger than the space of square integrable functions.

The Schrödinger equation for a free particle with energy k^2 is

$$-\frac{d^2}{dx^2}\psi(x) = k^2\psi(x) \tag{264}$$

has solutions of the form

$$e^{\pm ikx} \tag{265}$$

These solutions are not square integrable functions on the $\mathbb{R}$, but they are immediately recognized as Fourier transforms of delta functions. Thus, they can be interpreted as tempered distributions.

These distributional solutions, e^{ikx} are complete in the sense that I can write any square integrable function in the form

$$f(y) = \frac{1}{\sqrt{2\pi}}\int e^{iky}\tilde{f}(k)dk \tag{266}$$

(technically it may have to be written as a limit of Schwartz functions which can be put in this form).

There is a lot more that can be said about distribution theory. Most of the important results can be derived by expressing the distributions as limits of finite sums of Hermite basis functions.

As a practical matter, elementary results can be obtained by acting on the distributions as if they were functions.

The tempered distributions are one a much larger class of distributions. Ordinary distributions replace Schwartz functions with infinitely differentiable functions with bounded support. In this the corresponding space of distribution is even larger. The delta function is a well defined distribution in both of these spaces.

0.13 Lecture 13:

In this section I begin the discussion of linear operators on infinite dimensional vector spaces. A Linear Operator L on an infinite dimensional vector space $\mathcal{V}$ is a function that maps vectors from a subset of $\mathcal{V}$ called the domain of L to a vectors in another subset of $\mathcal{V}$, called the range of L satisfying the linearity condition

$$L(|v\rangle + \alpha|w\rangle) = L(|v\rangle) + \alpha L(|w\rangle) \tag{267}$$

Linear operators fall into several useful classes. We will primarily be concerned with bounded operators, unbounded operators, and compact operators.

As in the finite dimensional case, if $\mathcal{V}$ is a normed linear vector space then the norm of the operator L is defined by

$$|||L||| = \sup_{|||v\rangle||=1} ||L|v\rangle|| \tag{268}$$

The operator L is bounded if is L defined for all vectors in $\mathcal{V}$ and $|||L||| = C < \infty$. Operators that are defined on subsets of $\mathcal{V}$ that are not bounded are called unbounded operators. If L is unbounded then it is always possible to find an infinite sequence of unit normalized vectors $|v_n\rangle$ such that

$$||K|v_n\rangle|| > n \tag{269}$$

Bounded operators have the property that they are continuous, because if $|||v_n\rangle - |v\rangle|| \to 0$ then

$$|||L|v_n\rangle - L|v\rangle||| = |||L(|v_n\rangle - |v\rangle)||| \leq |||L||| \cdot |||v_n\rangle - |v\rangle|| \to 0 \qquad (270)$$

Bounded sets in infinite dimensional spaces have some properties that are not shared by bounded sets in finite dimensional spaces. For example, all orthonormal basis vectors fit inside a bounded sphere of radius $1 + \epsilon$. The distance between any pair of orthonormal basis vector is $\sqrt{2}$. This means that it is possible to put an infinite collection of vectors into this bounded set and they never have to get close to each other.

In a finite dimensional ball, if I have an infinite collection of vectors of norm 1, there has to be a subsequence that are getting closer and closer to each other. This result is called the Bolzano Weierstrass Theorem. The step that was glossed over in the proof of the Weierstrass theorem, where we stated that $g(x + y) - g(x)$ could be made less than some ϵ for all x in a closed bounded set, for small enough y, is based on a variant of the Bolzano Weierstrass theorem. Our interest in this theorem is to identify subsets on infinite dimensional vector spaces that look like finite dimensional vector spaces.

0.14 Lecture 14:

Theorem: (Bolzano Weierstrass Theorem): Let $\{x_n\}$ be an infinite collection of points in the unit square $[0,1] \times [0,1]$. Then there is an infinite subsequence of the points that converge to a point in the square.

Points in a square can be thought of as vectors. Divide each square into a 10×10 array of 100 smaller squares. Label them $[m,n]$ where m,n are integers that go from 0 to 9. As least one of these squares, say $[m_1, n_1]$ has an infinite number of points. Divide that square into a 10 array of 100 smaller squares. Label them m_2, n_2. At least one of these contains an infinite number of points. Label this square by $[m_1, m_2; n_1 n_2]$. This process can be continued giving successively smaller nested squares $m_1, m_2, \cdots; n_1, n_2, \cdots]$, each containing an infinite number of points. Each of the sequences can be consider as a decimal representation of a number between 0 and 1

$$x = .m_1 m_2 m_3 \cdots \qquad y = .n_1 n_2 n_3 \cdots \tag{271}$$

This number also identifies a point in the intersection of all of the larger squares. The truncated sequences can be associated with Cauchy sequences of rational numbers. Because of the completeness of the real numbers, these Cauchy sequences converge to a pair of real number between zero and one, This pair of numbers defines a vector in the plane.

The limit point is in the intersection of an infinite number of squares, each containing an infinite number of points. Picking one vector from each of the intersecting squares gives a convergent subsequence of points. This argument can be extended to N dimensions, but it breaks down when N is infinite because then each subdivision results in an infinite number of cubes of a fixed size; each of which could contain a point.

In infinite dimensional vector spaces bounded sets do not have the Bolzano Weierstrass property. Subsets of points in infinite dimensional vector spaces that have the Bolzano Weierstrass property are called Compact sets.

Definition: A subset of a vector space is Compact if every bounded infinite sequence has a convergent subsequence that converges to a vector in the set.

Compactness depends on the definition of convergence. We will normally be concerned with compact subsets of normed linear spaces. In these spaces convergence is given in terms of the norm.

What does a compact subset of an infinite dimensional space look like? Consider an infinite sequence of vectors $|m\rangle$. I can assume that at an infinite number of them are linearly independent, otherwise all of the vectors live in a finite dimensional subspace, which is compact if it is closed (includes its limit points) and bounded.

If this set contains an infinite number of independent vectors with

$$\Vert |n\rangle - |m\rangle \Vert > \epsilon, \tag{272}$$

it cannot be compact because this subset does not have a convergent subsequence.

For the set to be compact is must be true that for all $m, n > N_\epsilon$

$$\Vert |n\rangle - |m\rangle \Vert < \epsilon \tag{273}$$

This means that for every $\epsilon > 0$, every vector in the set can expressed as the sum of a vector in a finite dimensional subspace plus an additional vector with norm less than ϵ. As ϵ gets smaller the dimension of the finite dimensional space can get larger.

What is means in words is that compact sets are very thin in all but a finite number of dimensions.

0.15 Lecture 15

Definition C1: A compact linear operator is a bounded linear operator that maps bounded sets to sets with compact closure.

In the Hilbert space case compact operators have a simple characterization.

Theorem: A linear operator on a Hilbert space is compact if and only if for every $\epsilon > 0$ it can be expressed as the sum of a finite rank operator and an operator with operator norm less than ϵ.

Note that in a Hilbert space with an orthonormal basis $\{|n\rangle\}$ we can write a general operator A in the form

$$A = \sum_{mn} |n\rangle A_{nm} \langle m|. \tag{274}$$

If I define

$$|a_n\rangle := \sum_m A_{nm} \langle m| \tag{275}$$

then

$$A = \sum_n |n\rangle\langle a_n| \tag{276}$$

where $|a_n\rangle$ is not generally a unit vector. A is a rank N operator if the n sum contains N terms.

I prove this theorem by contradiction. Pick an $\epsilon > 0$ and assume that there is an infinite collection of orthogonal vectors $\{|n\rangle\}$ with norm greater than ϵ such that there are vectors $|w_n\rangle$ with norm 1 satisfying

$$C|w_n\rangle = |n\rangle \tag{277}$$

for all n. It this is the case the image of the set of unit normed vectors is not a compact set and consequently the operator C cannot be compact. This means that for any $\epsilon > 0$ I can find a projection operator P_N on an N dimensional subspace such that

$$C = P_N C + (I - P_N)C \tag{278}$$

where $\||C(I-P)_N|v\rangle\| < \epsilon$. $P_N C$ is the desired finite rank operator.

Conversely, if for every $\epsilon > 0$ I can find F_N finite rank with

$$|||C - F_N||| < \epsilon \tag{279}$$

then every vector that gets mapped into the orthogonal complement of range of F_N necessarily gets mapped into a vector of norm less than ϵ, which means that the range of C has compact closure. They can be understood as the closure of the space of finite rank operators in the operator norm.

This shows that compact operators on Hilbert spaces can be uniformly approximated by finite dimensional matrices

Theorem: Let B be bounded and C be compact on a Hilbert space. Then $A = BC$ and $D = CB$ are compact.

This theorem means that the compact operators are a 2 sided ideal in the space of bounded operators.

To prove this let $\epsilon > 0$. Since C is compact it is possible to find a finite dimensional operator C_N such that

$$|||C - C_N||| < \epsilon. \tag{280}$$

It follows that

$$|||BC - BC_N||| \leq |||B||| \cdot |||C - C_N||| < \epsilon |||B||| \tag{281}$$

and

$$|||CB - C_N B||| \leq \cdot |||C - C_N||| \cdot |||B||| < \epsilon |||B||| \tag{282}$$

Since BC_N and $C_N B$ are finite dimensional and $|||B|||$ is a fixed finite number, it follows that A and D are compact.

An important result that illustrates the utility of compact operators is the following result, which is one of the central results in the theory of integral equations.

Theorem:(Fredholm Alternative) If A is a compact operator on a Hilbert space $\mathcal{H}$ then either

$$A|\psi\rangle = |\psi\rangle \tag{283}$$

has a non-trivial solution or $(I - A)^{-1}$ exists.

Proof: Since A is compact there exists a finite rank A_N such that

$$|||A - A_N||| < \frac{1}{2}. \tag{284}$$

It follows that

$$(1 - (A - A_N))^{-1} = \sum_{n=0}^{\infty} (A - A_N)^n \tag{285}$$

where this series converges in norm because

$$|||\sum_{n=0}^{\infty}(A-A_N)^n||| \le \sum_{n=0}^{\infty}|||(A-A_N)^n||| \le$$

$$\sum_{n=0}^{\infty}|||(A-A_N)|||^n \le \sum_{n=0}^{\infty}\frac{1}{2^n} = 2. \tag{286}$$

Next note that

$$\begin{gathered}(I-A) = I-(A-A_N)-A_N = \\ I-(A-A_N)-A_N(I-(A-A_N))^{-1}(I-(A-A_N)) = \\ (I-A_N(I-(A-A_N))^{-1})(I-(A-A_N)) = \\ (I-F_N)(I-(A-A_N))\end{gathered} \tag{287}$$

where

$$F_N := A_N(I-(A-A_N))^{-1} \tag{288}$$

is finite rank. It follows that

$$(I-A)^{-1} = (I-(A-A_N))^{-1}(I-F_N)^{-1}. \tag{289}$$

This will exist provided $(I-F_N)$ has a non-zero determinant. If $(I-F_N)$ has zero determinant $(I-F_N)|\psi\rangle$ has at least one solution.

Recall that in the finite dimensional case 0 might correspond a generalized eigenvector of some order. If the order is $k>1$, then $(I-F_N)^{k-1}$ applied to the generalized eigenvector gives the desired eigenvector.

Note that the Fredholm Alternative does not hold for bounded operators in general. To see this consider the operator

$$A\phi(x) = x\phi(x) \tag{290}$$

on $L^2([0,2])$. Then

$$A\phi(x) = \phi(x) \tag{291}$$

requires $(x-1)\phi(x)=0$, which has no solutions and $(1-A)^{-1}$ is not a bounded operator because of the singularity at $x=1$.

The Fredholm Alternative illustrates how one can extend results from finite dimensional matrices to compact operators.

0.16 Lecture 16

One of the advantages of compact operators is that it is possible to use the error estimates on the approximated operator to compute bounds on the error of the approximate solution. To show this assume that A is compact and $|||A - A_N||| < \epsilon$ for some rank N approximation A_N. Consider the equation

$$|f\rangle = |g\rangle + A|f\rangle \tag{292}$$

The solution to this equation, if it exists is

$$|f\rangle = (I - A)^{-1}|g\rangle \tag{293}$$

which can be written as

$$|f\rangle = (I - (A - A_N))^{-1}(I - F_N)^{-1}|g\rangle. \tag{294}$$

I can approximate

$$(I - (A - A_N))^{-1} \approx \sum_{m=0}^{M}(A - A_N)^m. \tag{295}$$

The error is

$$|||(I - (A - A_N))^{-1} - \sum_{m=0}^{M}(A - A_N)^m||| =$$

$$|||\sum_{m=M+1}^{\infty}(A - A_N)^m||| \leq \frac{\epsilon^{M+1}}{1-\epsilon} \tag{296}$$

This means that the error in the approximation is bounded by

$$\||f\rangle - \sum_{m=0}^{M}(A - A_N)^m(I - F_N)^{-1}|g\rangle\| \leq$$

$$\frac{\epsilon^{M+1}}{1-\epsilon}|||(I - F_N)^{-1}||| \cdot \||g\rangle\| \tag{297}$$

where

$$F_N := A_N(I - (A - A_N))^{-1}) \tag{298}$$

is a rank N operator. In principle, because F_N is finite rank, $|||(I - F_N)^{-1}|||$ can be computed by matrix methods, but in practice the consistent approximation to F_N is

$$F_N \approx B_N := A_N \sum_{m=0}^{M} (A - A_N)^m \tag{299}$$

with error

$$|||F_N - B_N||| \leq |||A_N||| \frac{\epsilon^{M+1}}{1-\epsilon} \tag{300}$$

Using this we can estimate the error in $(I - F_N)^{-1}$ by writing $F_N = B_N + \Delta$ which gives

$$(1 - F_N)^{-1} = (1 - B_N)^{-1} + (1 - B_N)^{-1}\Delta(1 - F_N)^{-1} \tag{301}$$

$$(1 - F_N)^{-1} = \sum_{l=0}^{\infty} ((1 - B_N)^{-1}\Delta)^l (1 - B_N)^{-1} \tag{302}$$

If

$$|||((1 - B_N)^{-1}||| < \frac{(1-\epsilon)}{\epsilon^m} \tag{303}$$

then

$$|||(1 - B_N)^{-1}\Delta)||| = \delta < 1 \tag{304}$$

$$|||(1 - F_N)^{-1} - \sum_{l=0}^{K} ((1 - B_N)^{-1}\Delta)^l (1 - B_N)^{-1}|||||| \leq \frac{\delta^{K+1}}{1-\delta} |||(1 - B_N)^{-1}||| \tag{305}$$

Putting all of the estimates and approximations together

$$|f\rangle \approx \sum_{m=0}^{M} (A - A_N)^m \sum_{l=0}^{K} ((1 - B_N)^{-1}\Delta)^l (1 - B_N)^{-1} |g\rangle \tag{306}$$

which has the desirable property that there is a computable upper bound on the error than can be made as small as desired. This assumes that 1 is not near an eigenvalue of B_N.

Theorem: (Riesz Schauder) Let A be a compact Hermitian operator on a Hilbert space. Then for any $\epsilon > 0$, A can have only a finite number of eigenvalues λ_n with $|\lambda_n| > \epsilon$.

Because A is Hermitian I can assume that the eigenvectors are orthonormal. Let $\{|n\rangle\}$ denote the collection of orthonormal eigenvectors with eigenvalues having magnitude larger than ϵ

This set of vectors defines an infinite sequence with the property that

$$||A|n\rangle - A|m\rangle|| = ||\lambda_n|n\rangle - \lambda_m|m\rangle|| = \sqrt{\lambda_n^2 + \lambda_m^2} > \sqrt{2}\epsilon \tag{307}$$

which contradicts the compactness of A It follows that the subspace of linear combinations of eigenvectors with eigenvalue $\lambda_n > \epsilon$ has a finite dimension.
Theorem: (Completeness) Let A be a compact Hermitian operator. Then A has a complete set of eigenvectors with eigenvalues λ_n satisfying

$$\lim_{n\to\infty} |\lambda_n| \to 0 \tag{308}$$

To prove this let $|||A|||$ be the operator norm of A. By definition

$$|||A||| = \sup_{|||v\rangle||=1} ||A|v\rangle|| \tag{309}$$

This definition means that either there is a vector $|v\rangle$ with

$$A|v\rangle = \pm|||A||| \cdot |v\rangle \tag{310}$$

or there is a sequence of unit normed vectors $|v_n\rangle$ with

$$||A|v_n\rangle|| \to |||A||| \tag{311}$$

I will show that this second possibility implies the first.

Since the sequence $A|v_n\rangle = |w_n\rangle$ satisfies the Bolzano Weierstrass property, it has a subsequence that converges to a vector $|w\rangle$ in the Hilbert space (Hilbert spaces are complete).

I claim

$$A|w\rangle = \pm|||A||| \cdot |w\rangle \tag{312}$$

To verify this note

$$|||w\rangle|| = |||A||| \tag{313}$$

by the definition of the operator norm and

$$||A|w\rangle|| \leq |||A||| \cdot |||w\rangle|| = |||A|||^2 \tag{314}$$

Pick $\epsilon > 0$ and $|v_n\rangle$ with n large enough so

$$|||w_n\rangle - |w\rangle|| < \epsilon \tag{315}$$

so

$$||A|w_n\rangle|| \leq ||A|w\rangle|| + |||A||| \cdot |||w_n\rangle - |w\rangle|| \leq ||A|w\rangle|| + |||A|||\epsilon \qquad (316)$$

Let $\epsilon' = \epsilon|||A|||$ so

$$||A|w\rangle|| \geq ||A|w_n\rangle|| - \epsilon' \qquad (317)$$

for all $n > N_\epsilon$ Next note

$$|||w_n\rangle||^2 = \langle v_n|AA|v_n\rangle \leq$$

$$\langle v_n|v_n\rangle^{\frac{1}{2}}\langle AAv_n|AAv_n\rangle^{\frac{1}{2}} = \|A|w_n\rangle\| \qquad (318)$$

Using (318) with (317) gives

$$||A|w\rangle|| \geq |||w_n\rangle||^2 - \epsilon'. \qquad (319)$$

Since ϵ is independent of n for large enough n

$$\lim_{n\to\infty} |||w_n\rangle|| \to |||w\rangle|| = |||A||| \qquad (320)$$

$$||A|w\rangle|| \geq |||w\rangle||^2 \qquad (321)$$

Combining (321) with the inequality

$$|||A||| \cdot |||w\rangle|| \geq ||A|w\rangle|| \qquad (322)$$

gives

$$||A|w\rangle|| = |||A|||^2 \qquad (323)$$

Next note

$$\langle w|AA|w\rangle \leq \langle w|w\rangle^{1/2}\langle w|A^4|w\rangle^{\frac{1}{2}} \qquad (324)$$

The left side of this expression is $|||A|||^4$ by (323). The right side is bounded by $|||A||| \cdot |||A||| \cdot |||A|||^2$ which gives the inequalities

$$|||A|||^4 \leq \langle w|w\rangle^{1/2}\langle w|A^4|w\rangle^{1/2} \leq |||A|||^4 \qquad (325)$$

leading to

$$|||A|||^3 = \langle w|A^4|w\rangle^{1/2} \qquad (326)$$

From these relations I get

$$\langle w|(A^2 - |||A|||^2)(A^2 - |||A|||^2)|w\rangle =$$

$$\langle w|A^4|w\rangle - 2|||A|||^2\langle w|A^2|w\rangle + |||A|||^6 = 0 \tag{327}$$

This gives

$$0 = (A^2 - |||A|||^2)|w\rangle = (A + |||A|||)(A - |||A|||)|w\rangle \tag{328}$$

which proves that $|w\rangle$ is an eigenvector with eigenvalue $\pm|||A|||$.

This establishes the existence of one eigenvector $|w_1\rangle$ with eigenvalue λ_1 equal to $\pm|||A|||$.

I establish completeness by induction. Define

$$A_1 = A - |w_1\rangle\lambda_1\langle w_1| \tag{329}$$

where $|w_1\rangle$ is normalized to unity. This is a compact Hermitian operator where

$$A_1|w_1\rangle = (\lambda_1 - \lambda_1)|w_1\rangle = 0 \tag{330}$$

Any other eigenvector of A is also an eigenvector of A_1 with the same eigenvalue. Repeating our analysis, we find there is an eigenvector with eigenvalue equal (in magnitude) to the largest eigenvalue of A_1 which is the second largest eigenvalue of A. The associate eigenvector is orthogonal to $|w_1\rangle$

This process can be repeated N times giving

$$A_N = \sum_{n=1}^{N} |w_n\rangle\lambda_n\langle w_n| \tag{331}$$

By the Reisz-Schauder theorem there are only finite number of eigenvalues with magnitude larger than ϵ. This means

$$|||A - A_N||| < \epsilon \tag{332}$$

$$\lim_{N\to\infty} A_N = A \tag{333}$$

To complete the proof let $|v\rangle$ be any Hilbert space vector. Then

$$A|v\rangle - \sum_{n=1}^{\infty} |w_n\rangle\lambda_n\langle w_n|v\rangle = \sum_{n=1}^{\infty} |w_n\rangle(\lambda_n - \lambda_n)\langle w_n|v\rangle \tag{334}$$

It follows that

$$|v\rangle = \sum_{n=0}^{\infty} |w_n\rangle\langle w_n|v\rangle + |v_\perp\rangle \tag{335}$$

where $|v_\perp\rangle$ is the difference. It follows by direct computation that

$$A|v_\perp\rangle = 0 \tag{336}$$

which means that $|v\rangle$ is an eigenvector of A with eigenvalue zero. (335) shows that the eigenvectors of A are complete.

0.17 Lecture 17

Next I show the general structure of a compact operator on a Hilbert space.
Theorem: Let A be compact on a Hilbert space. Then there are two orthonormal bases $\{|v_n\rangle\}$ and $\{|w_n\rangle\}$ and $\lambda_n > 0$ satisfying

$$\lim_{n\to\infty} \lambda_n \to 0 \tag{337}$$

and

$$A = \sum_{n=1}^{\infty} |v_n\rangle \lambda_n \langle w_n| \tag{338}$$

To prove this note that if A is compact then $A^\dagger A$ is compact (see homework) , Hermitian, and positive. It has a complete set of eigenvectors with positive eigenvalues

$$A^\dagger A|w_n\rangle = \lambda_n^2 |w_n\rangle \tag{339}$$

In what follows the $|w_n\rangle$ are taken to be orthonormal.

Define

$$|v_n\rangle := K|w_n\rangle/\lambda_n \tag{340}$$

for n with $\lambda_n > 0$. Then

$$\langle v_n|v_m\rangle = \frac{1}{\lambda_m \lambda_n}\langle w_n|K^\dagger K|w_m\rangle = \frac{\lambda_m^2}{\lambda_m\lambda_n}\delta_{mn} = \delta_{mn} \tag{341}$$

Note

$$A|\xi\rangle = A\sum_n |w_n\rangle\langle w_n|\xi\rangle =$$

$$\sum_n (A|w_n\rangle/\lambda_n)\lambda_n\langle w_n|\xi\rangle =$$

$$\left(\sum_n \langle v_n\rangle \lambda_n \langle w_n|\right)|\xi\rangle \tag{342}$$

Which shows

$$A = \sum_n \langle v_n|\lambda_n\langle w_n| \tag{343}$$

as desired.

The converse of this theorem is also true, because if

$$A = \sum_{n=1}^{\infty} |v_n\rangle \lambda_n \langle w_n| \tag{344}$$

and I define

$$A_N := \sum_{n=1}^{N} |v_n\rangle \lambda_n \langle w_n| \tag{345}$$

then

$$A - A_N = \sum_{n=N+1}^{\infty} |v_n\rangle \lambda_n \langle w_n| \tag{346}$$

and

$$|||A - A_N||| = max_{n>N} |\lambda_n| \tag{347}$$

which vanishes by the Reisz Schauder theorem as $N \to \infty$.

In applications it is important to know that an operator is compact before on attempts to make finite dimensional approximations.

There are two important cases

A linear operator K is Hilbert Schmidt if and only if

$$\mathrm{Tr}(K^\dagger K) < \infty \tag{348}$$

A linear operator K is Trace class if and only if

$$\mathrm{Tr}(K^\dagger K)^{1/2} < \infty \tag{349}$$

Here the trace of a linear operator on an infinite dimensional vector space is

$$\mathrm{Tr}(A) = \sum_n \langle n|A|n\rangle \tag{350}$$

where $\{|n\rangle\}$ is any orthonormal basis.

Note that the trace is independent of the choice of basis because

$$\mathrm{Tr}'(A) = \sum_n \langle n'|A|n'\rangle =$$

$$\sum_{lmn} \langle n'|m\rangle \langle m|A|l\rangle \langle l|n'\rangle =$$

$$\sum_{lm} (\sum n\langle l|n'\rangle\langle n'|m\rangle)\langle m|A|l\rangle =$$

$$\sum_{lm} \delta_{ml}\langle m|A|l\rangle =$$

$$\sum_{m} \langle m|A|m\rangle = \mathrm{Tr}(A) \tag{351}$$

To show that both trace class and Hilbert Schmidt operators are compact first note that

$$\mathrm{Tr}(A^\dagger A) \geq \mathrm{Tr}(A^\dagger P_N A) \tag{352}$$

where P_N is any orthogonal projector. If this is the orthogonal projector on the first N eigenstates of $K^\dagger K$, the condition of being trace class or Hilbert Schmidt ensures that the sum of any number of eigenvalues is bounded.

This requires that $\lambda_n \to 0$, which means that K is compact.

In the Hilbert Schmidt case

$$\sum_n \lambda_n^2 < \infty \tag{353}$$

For the trace class case

$$\sum_n \lambda_n < \infty \tag{354}$$

and in general if

$$\sum_n \lambda_n^p < \infty \tag{355}$$

for $p > 0$ the operator is still compact. In practice the condition of being Hilbert Schmidt is the easiest condition to show.

Example: Compact operators are often associated with integral equations. Consider the integral equation

$$f(x) = g(x) + \int_a^b K(x,y) f(y) dy \tag{356}$$

We can check to see if $K(x,y)$ is the kernel of a Hilbert Schmidt operator by calculating

$$\mathrm{Tr}(K^\dagger K) = \int_a^b dx \int_a^b dy K^*(x,y) K(y,x) \tag{357}$$

If this is finite then it follows that we can uniformly approximate K by

$$K(x,y) \approx \sum_{n=1}^{N} \phi_n(x)\lambda_n\chi_n^*(y) \tag{358}$$

Replacing the actual kernel by the approximate kernel gives

$$f(x) \approx g(x) + \int_a^b \sum_{n=1}^{N} \phi_n(x)\lambda_n\chi_n^*(y)f(y)dy \tag{359}$$

This can be reduced to a matrix equation if I multiply by $\chi_m^*(x)$ and integrate over x

$$c_n = g_n + \sum_{m=1}^{N} K_{nm}c_m \tag{360}$$

where

$$c_n = \int_a^b dx\chi_n^*(x)f(x) \tag{361}$$

$$g_n = \int_a^b dx\chi_n^*(x)g(x) \tag{362}$$

$$K_{nm} = \int_a^b dx\chi_n^*(x)\lambda_m\phi_m(x) \tag{363}$$

leading to the approximate solution

$$f(x) = g(x) + \int_a^b K(x,y)\chi_n(y)dyc_ndy \tag{364}$$

If a linear operator A is compact we know that it can be uniformly approximated by a finite matrix. In principle any finite dimensional approximation will eventually converge. If A is compact and Hermitian then

$$A = \sum_{n=1}^{\infty} |n\rangle\lambda_n\langle n| \tag{365}$$

with

$$|\lambda_n| \geq |\lambda_{n+1}| \tag{366}$$

In this basis the approximation

$$A \approx A_N := \sum_{n=1}^{N} |n\rangle \lambda_n \langle n| \tag{367}$$

satisfies

$$|||A - A_N||| < |\lambda_{N+1}| \to 0 \tag{368}$$

It is clear that this is the best choice of basis. For a large matrix it may be too much work to diagonalize A. Note that if we take a large power m of A

$$A^m = \sum_{n=1}^{\infty} |n\rangle \lambda_n^m \langle n| \tag{369}$$

then λ_n^m will have larger coefficients for the larger eigenvalues.

One way to take advantage of this is to generate a basis using different powers of K applied to a single vector. After normalization, as long as the starting vector has some overlap with the eigenvector with the largest eigenvalue, it will get emphasized. The problem is all of the vectors generated this way will be almost parallel. To fix this we use the Gram-Schmidt method, and orthogonalize the vectors at each step.

To illustrate this method consider the linear equation

$$|f\rangle = |g\rangle + K|f\rangle \tag{370}$$

where $|g\rangle$ is known and K is compact and hermetian.

We generate and orthonormal basis as follows

$$|\hat{1}\rangle = \frac{|g\rangle}{\||g\rangle\|} \tag{371}$$

$$|2\rangle = K|\hat{1}\rangle - |\hat{1}\rangle\langle\hat{1}|K|\hat{1}\rangle \tag{372}$$

$$|\hat{2}\rangle = \frac{|2\rangle}{\||2\rangle\|} \tag{373}$$

$$\vdots \tag{374}$$

$$|n\rangle = K|\widehat{n-1}\rangle - |\widehat{n-1}\rangle\langle\widehat{n-1}|K|\widehat{n-1}\rangle - |\widehat{n-2}\rangle\langle\widehat{n-2}|K|\widehat{n-1}\rangle \tag{375}$$

$$|\hat{n}\rangle = \frac{|n\rangle}{\||2\rangle\|} \tag{376}$$

Vectors generated this algorithm have the properties

$$\langle \hat{m} | \hat{n} \rangle = \delta_{mn} \tag{377}$$

$$\langle \hat{m} | K | \hat{n} \rangle = 0 \qquad |m - n| > 1 \tag{378}$$

In this basis the equation has the form

$$\langle \hat{n} | f \rangle = \| |g\rangle \| \delta_{n1} + \sum_{m=n-1, m \geq 0}^{n+1} \langle \hat{n} | K | \hat{m} \rangle \langle \hat{m} | f \rangle \tag{379}$$

This gives a tridiagonal matrix. More important, if K has a small number of large eigenvalues, this solution to the equation can often be accurately approximated using a small number of these basis functions, even when the original matrix is very large.

This same method can also be used to find approximate solutions to the eigenvalue equation. If K is not Hermetian one there are related methods that use two sets of basis vectors. One can also convert the equation to a compact kernel equation with a Hermetian kernel using

$$(1 - K)|f\rangle = |g\rangle \rightarrow$$

$$(1 - K^\dagger)(1 - K)|f\rangle = (1 - K^\dagger)|g\rangle \tag{380}$$

$$(1 - K^\dagger)(1 - K) = I - K' \qquad K' := K + K^\dagger - K^\dagger K \tag{381}$$

where $K'^\dagger = K'$ and

$$|f\rangle = (1 - K')^{-1}(1 - K^\dagger)|g\rangle \tag{382}$$

These are the methods of choice for solving very large linear systems based on approximating compact kernels.

0.18 Lecture 18

Definition: An n^{th} degree ordinary differential equation is an equation of the form

$$F(x, u, \frac{du}{dx}, \cdots, \frac{d^n u}{dx^n}) = 0 \tag{383}$$

The differential equation is linear if F is linear in u and all of its derivatives

$$F = q(x) + r_0(x)u(x) + r_1(x)\frac{du}{dx} + \cdots + r_n(x)\frac{d^n u}{dx^n} = 0 \tag{384}$$

The linear differential equation is called homogeneous if $q(x) = 0$, otherwise it called inhomogeneous.

The existence of a solution to any of these equations depends on the space of functions where one looks for solutions. The functions may have N derivatives, or they could be tempered distributions, they may satisfy certain boundary or asymptotic conditions.

It is difficult to find results that apply to general equations of the form (383). The partial derivative of F with respect the the highest derivative is not zero at point where all of the other arguments are fixed, then the implicit function theorem allows us to replace (383) by an equation of the form

$$\frac{d^n u}{dx^n} = G(x, u, \frac{d^1 u}{dx^1}, \cdots, \frac{d^{n-1}u}{dx^{n-1}}) \tag{385}$$

which is valid near the point defined by

$$x = x_0, u(x_0) = u_0, \cdots \frac{d^{n-1}u}{dx^{n-1}}(x_0) = u_{0,n-1} \tag{386}$$

There is no guarantee the this inversion is valid far form this point, but we will show that under mild conditions on G this equation alway has local solutions.

The mild conditions are the following. None that G depends on $n+1$ variable. The function $G(x_1, \cdots, x_{n+1})$ satisfies a Lipschitz condition x_i if there exists an $\eta > 0$ such that $x_i, x_i' \in [c_i - \eta, c_i + \eta]$ implies

$$|G(x_1, \cdots, x_i, \cdots, x_n) - G(x_1, \cdots, x_i', \cdots, x_n)| < k|x_i - x_i'| \tag{387}$$

If G satisfies a Lipschitz condition in all $n+1$ variables the differential equation, with initial conditions

$$u(x_0) = c_0 \cdots \frac{d^{n-1}u}{dx^{n-1}} = c_{n-1} \tag{388}$$

I outline the proof. The differential equation can be converted to a system of first order equations using

$$y_0 := u, y_1 = \frac{du}{dx} \cdots y_{n-1} := \frac{d^{n-1}u}{dx^{n-1}} \tag{389}$$

$$\begin{pmatrix} \frac{dy_0}{dx} = y_1 \\ \frac{dy_1}{dx} = y_2 \\ \vdots \\ \frac{dy_{n-2}}{dx} = y_{n-1} \\ \frac{dy_{n-1}}{dx} = G(x_0, y_0, \cdots, y_{n-1}) \end{pmatrix} \tag{390}$$

Next we convert this to a system of integral equations by integrating from x_0 to x:

$$\begin{pmatrix} y_0(x) \\ y_1(x) \\ \vdots \\ y_{n-2}(x) \\ y_{n-1}(x) \end{pmatrix} = \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-2} \\ c_{n-1} \end{pmatrix} + \begin{pmatrix} \int_{x_0}^{x} y_0(x')dx' \\ \int_{x_0}^{x} y_1(x')dx' \\ \vdots \\ \int_{x_0}^{x} y_{n-1}(x')dx' \\ \int_{x_0}^{x} G(x', y_0, \cdots, y_{n-1})dx' \end{pmatrix} \tag{391}$$

We can attempt to solve this system by successive approximations. To do this write the integral equation as

$$\mathbf{y}(x) = \mathbf{c} + \int_{x}^{x_0} \mathbf{G}(x', \mathbf{y})dx' \tag{392}$$

$$\mathbf{y}_{(1)}(x) = \mathbf{c} \tag{393}$$

$$\mathbf{y}_{(2)}(x) = \mathbf{c} + \int_{x}^{x_0} \mathbf{G}(x', \mathbf{y}_{(1)}dx' \tag{394}$$

$$\vdots \tag{395}$$

$$\mathbf{y}_{(n)}(x) = \mathbf{c} + \int_{x}^{x_0} \mathbf{G}(x', \mathbf{y}_{(n-1)})dx' \tag{396}$$

0.19 Lectue 19

I define a norm on this system by

$$|\mathbf{y}(x) - \mathbf{y}'(x)| = \sum_{m=0}^{n-1} |y_m(x) - y'_m(x)| \tag{397}$$

The Lipschitz condition means that on a sufficiently small $n+1$ dimensional volume there is a constant C such that

$$|\mathbf{y}_{(1)} - \mathbf{y}_{(0)}| \leq |x - x_0| C \tag{398}$$

$$|\mathbf{y}_{(2)} - \mathbf{y}_{(1)}| \leq \frac{1}{2} |(x - x_0)^2| C^2 \tag{399}$$

$$|\mathbf{y}_{(n)} - \mathbf{y}_{(n-1)}| \leq \frac{1}{n!} |(x - x_0)^n| C^n \tag{400}$$

$$|\mathbf{y}_{(n)}| = \tag{401}$$

$$|\sum_{k=1}^{n} (\mathbf{y}_{(k)} - \mathbf{y}_{(k-1)}) + \mathbf{f}_0| \leq \sum_{k=0}^{n} \frac{1}{n!} ||x - x_0|^n C^n + |\mathbf{f}_0| \tag{402}$$

$$\lim_{n \to \infty} |\mathbf{y}_{(n)}| \leq e^{C|x - x_0|} + |\mathbf{f}_0| \tag{403}$$

This shows that not only does the series for the function converge, but the series for all of the lower order derivatives converges. This leads to the local existence of the function and its first $n-1$ derivatives. The local existence does not mean that the solution can be extended.

This method of converting the differential equation to an integral equation is called Picard's method. Most local existence theorems for solutions to ordinary differential equations are based on this result. It is of limited value because typical functions $\mathbf{G}$ will satisfy a Lipschitz condition in all variables in only a small region about an initial point.

In what follows I will discuss the solution of linear differential equations of first and second order. These equations have the general form

$$\alpha(x) \frac{du}{dx} + \beta(x) u(x) = f(x) \qquad \alpha(x) \neq 0 \tag{404}$$

$$\alpha(x) \frac{d^2 u}{dx^2} + \beta(x) \frac{du}{dx} + \gamma(x) u(x) = f(x) \qquad \alpha(x) \neq 0 \tag{405}$$

First I show that any first order linear differential equation can be solved in terms of integrals of known functions. To do this define the function $p(x)$ as the solution to

$$\frac{1}{p(x)}\frac{dp}{dx} = \frac{\beta(x)}{\alpha(x)} \tag{406}$$

which can be solved to give

$$p(x) = p(x_0)e^{\int_{x_0}^{x} dx'\beta(x')/\alpha(x')} \tag{407}$$

The differential equation can then be written as

$$0 = \frac{du}{dx} + \frac{\beta(x)}{\alpha(x)}u(x) - \frac{f(x)}{\alpha(x)} =$$

$$p(x)\frac{du}{dx} + u(x)\frac{dp}{dx} - \frac{f(x)p(x)}{\alpha(x)} \rightarrow$$

$$\frac{d}{dx}(u(x)p(x)) = \frac{f(x)p(x)}{\alpha(x)}$$

which can be integrated to get the solution

$$u(x) = \frac{1}{p(x)}(u(x_0)p(x_0) + \int_{x_0}^{x} dx' \frac{f(x')p(x')}{\alpha(x')}) \tag{408}$$

Next I discuss some general properties of second order differential equations of the form (89).

First consider the case that $q(x) = 0$. In this case if $u_1(x)$ and $u_2(x)$ are both solutions of (89) then so is

$$u_3(x) = c_1u_1(x) + c_2u_2(x). \tag{409}$$

This is an elementary consequence of the linearity of the linearity of the differential operator.

I use this to show that two solutions of a second order homogeneous differential equation are independent if and only if

$$W(u_1, u_2) := \det[\begin{pmatrix} u_1(x) & u_2(x) \\ \frac{du_1}{dx}(x) & \frac{du_2}{dx} \end{pmatrix}] \neq 0 \tag{410}$$

This determinant is called the Wronskian.

Clearly dependence means that there are constants c_1 and c_2 satisfying

$$\begin{pmatrix} u_1(x) & u_2(x) \\ \frac{du_1}{dx}(x) & \frac{du_2}{dx} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{411}$$

which implies the matrix has 0 as an eigenvalue and the determinant vanishes. Conversely, if the determinant does not vanish the solutions cannot be dependent.

Clearly if $W = 0$ then

$$\frac{1}{u_1}\frac{du_1}{dx} = \frac{1}{u_2}\frac{du_2}{dx} \tag{412}$$

which can be integrated to get

$$\ln(u_1/u_2) = const \tag{413}$$

or

$$u_1 = u_2 \times constant \tag{414}$$

Note that

$$\frac{dW(x)}{dx} = u_1(x)\frac{d^2u_2}{dx^2} - u_2(x)\frac{d^2u_1}{dx^2} = -\frac{\beta(x)}{\alpha(x)}W(x). \tag{415}$$

Integrating this gives

$$W(x) = W(x_0)e^{-\int_{x_0}^{x} dx'\beta(x')/\alpha(x')} \tag{416}$$

which shows if $W(x)$ is zero at one point, (x_0), then it vanishes everywhere, otherwise it vanishes nowhere (recall $\alpha(x) > 0$).

It follows that if the Wronskian is zero at one point, it vanishes at all points and the solutions must be proportional. If it is non-zero at one point is cannot be zero at any point, so the solutions are independent.

Assume that a second order linear homogeneous differential equation has three independent solutions. This means that there no non-zero coefficients c_1, c_2, c_3 satisfying

$$c_1u_1(x) + c_2u_2(x) + c_3u_3(x) = 0 \tag{417}$$

Differentiating I also have

$$c_1\frac{du_1}{dx}(x) + c_2\frac{du_2}{dx}(x) + c_3\frac{du_3}{dx}(x) = 0 \tag{418}$$

and

$$c_1\frac{d^2u_1}{dx^2}(x)+c_2\frac{d^2u_2}{dx^2}(x)+c_3\frac{d^2u_3}{dx^2}(x)=0 \tag{419}$$

$$c_1u_1(x)+c_2u_2(x)+c_3u_3(x)=0 \tag{420}$$

Consider

$$\det\begin{pmatrix} u_1 & u_2 & u_2 \\ \frac{du_1}{dx} & \frac{du_2}{dx} & \frac{du_3}{dx} \\ -\frac{\beta}{\alpha}\frac{du_1}{dx}-\frac{\gamma}{\alpha}\frac{du_1}{dx} & -\frac{\beta}{\alpha}\frac{du_2}{dx}-\frac{\gamma}{\alpha}\frac{du_2}{dx} & -\frac{\beta}{\alpha}\frac{du_3}{dx}-\frac{\gamma}{\alpha}\frac{du_3}{dx} \end{pmatrix}=0 \tag{421}$$

This vanishes because the last row is a linear combination of the first two rows for every value of x. This means that it is possible to find non-zero coefficients c_1, c_2, c_3 that make this vanish. It follows that the solutions are necessarily linearly dependent.

It follows that second order linear homogeneous differential equations have at most two independent solutions.

Next I show that if a second order linear homogeneous differential equation has one solution $u_1(x)$ then it necessarily has a second independent solution, $u_2(x)$.

To construct $u_2(x)$ assume that

$$u_2(x)=u_1(x)h(x) \tag{422}$$

The differential equation implies

$$\alpha(x)\frac{d^2}{dx^2}(u_1(x)h(x))+\beta(x)\frac{d}{dx}(u_1(x)h(x))+\gamma(x)(u_1(x)h(x))=0 \tag{423}$$

The important observation is that all of the terms with no derivatives on h cancel by the differential equation when applied to u_1. What survives involves only first and second derivatives of h:

$$\alpha(x)(\frac{d^2h}{dx^2}u_1(x)+2\frac{dh}{dx}\frac{du_1}{dx})+\beta\frac{d}{dx}(u_1(x)\frac{dh}{dx})=0 \tag{424}$$

$$\frac{d^2h}{dx^2}+(\frac{2}{u_1(x)}\frac{du_1}{dx}+\frac{\beta(x)}{\alpha(x)}\frac{dh}{dx}=0 \tag{425}$$

This is a first order linear homogeneous equation for $\frac{dh}{dx}$. It has the form

$$\frac{d}{dx}(\ln(\frac{dh}{dx}))=-2\frac{d}{dx}\ln(u_1(x))-p(x) \tag{426}$$

where

$$p(x) = \int^{x} \frac{\beta(x')}{\alpha(x')} dx' \tag{427}$$

Integrating this gives

$$\frac{dh}{dx}(x) = \frac{c}{u_1(x)^2} e^{\int_{-x_0}^{x} p(x')dx'} \tag{428}$$

Integrating a second times gives

$$h(x) = \int_{x_0}^{x} \frac{c}{u_1(x')^2} e^{-\int_{x_0}^{x'} p(x'')dx''} dx' \tag{429}$$

$$u_2(x) = h(x)u_1(x) \tag{430}$$

The structure of the solution leads to a Wronskian of the form

$$W(u_1, u_2) = u_1 \frac{d}{dx}(u_1 h) - u_1 h \frac{d}{dx} u_1 = u_1^2(x) \frac{dh}{dx} = ce^{\int_{-x_0}^{x} p(x')dx'} \tag{431}$$

which is not zero. This shows that it is alway possible to construct a second independent solution to the homogeneous equation is one is already known.

Next assume that a solution to the homogeneous equation are known. I show that is possible to construct a solution to the inhomogeneous equation.

Assume a solution $v(x)$ or the form $v(x) = u(x)h(x)$. Using this in the differential equation, all of the terms involving no derivatives of h cancel because u_1 satisfies the homogeneous form of the differential equation. What remains is

$$\frac{d^2h}{dx^2} + (\frac{2}{u_1(x)} \frac{du_1}{dx} + \frac{\beta(x)}{\alpha(x)}) \frac{dh}{dx} = \frac{q(x)}{\alpha(x) u_1(x)} \tag{432}$$

To solve this let $R(x)$ be a solution to

$$\frac{1}{R} \frac{dR}{dx} = \frac{2}{u_1(x)} \frac{du_1}{dx} + \frac{\beta(x)}{\alpha(x)} \tag{433}$$

leading to the equation

$$R \frac{d^2h}{dx^2} + \frac{dR}{dx} \frac{dh}{dx} = \frac{R(x)q(x)}{\alpha(x)u_1(x)} \tag{434}$$

where

$$R(x) = cu_1(x)^2 e^{\int_{x_0}^{x} p(x')dx'} = \frac{cu^2(x)}{W(x)} \tag{435}$$

The differential equation can be written as

$$\frac{d}{dx}(R\frac{dh}{dx}) = \frac{R(x)q(x)}{\alpha(x)u_1(x)} \tag{436}$$

which has the solution

$$\frac{dh}{dx}(x) = \frac{1}{R(x)}\int_{x_0}^{x}\frac{R(x')q(x')}{\alpha(x')u_1(x')}dx' \tag{437}$$

This can be integrated to get

$$dh(x) = \int_{x_0}^{x}\frac{1}{R(x')}\int_{x_0}^{x'}\frac{R(x'')q(x'')}{\alpha(x'')u_1(x'')}dx''dx'' \tag{438}$$

$$v(x) = u_1(x)h(x) \tag{439}$$

All of the methods just discussed require that one solution is known. This is used to express the other solution as the solution of a first order equation that can be solved to get the desired solution.

In the general setting one is not typically fortunate enough to have a solution available.

0.20 Lecture 20

In this section I begin the treatment of linear differential operators.

Define the differential operator

$$L_x := \sum_{n=0}^{N} a_n(x) \frac{d^n}{dx^n}. \tag{440}$$

This is really a formal operator until we specify the vector space on which it acts.

Suppose that for a a given formal differential operator L_x there is another formal differential operator $L_x^\dagger$ with the property that for any sufficiently differentiable functions $u(x)$ and $v(x)$ and positive weight $w(x)$ on the interval $[a, b]$, that the quantity

$$w(x)[v^*(x)L_x u(x) - u(x)(L_x^\dagger v(x))^*] = \frac{d}{dx}Q[u(x), v^*(x), \frac{du}{dx}, \frac{dv^*}{dx}] \tag{441}$$

where $Q[x, u(x), v^*(x), \frac{du}{dx}, \frac{dv^*}{dx}]$ is a function of the form

$$Q = A(x)u(x)v^*(x) + B(x)u(x)\frac{dv^*}{dx}(x) +$$

$$C(x)\frac{du}{dx}(x)v^*(x) + D(x)\frac{du}{dx}(x)\frac{dv^*}{dx}(x) \tag{442}$$

Equation (**??**) is called the Lagrange identity and $L_x^\dagger$ is called the formal adjoint of L_x with respect to the weight $w(x)$.

Integrating the Lagrange identity from a to b gives

$$\int_a^b w(x)[v^*(x)L_x u(x)dx - \int_a^b u(x)(L_x^\dagger v(x))^* dx =$$

$$Q[b, u(b), v^*(b), \frac{du}{dx}(b), \frac{dv^*}{dx}(b)] - Q[b, u(a), v^*(a), \frac{du}{dx}(a), \frac{dv^*}{dx}(a)] \tag{443}$$

Equation (**??**) is called a generalized Green's identity. When $L_x = L_x^\dagger$ then L_x is called self-adjoint.

Next I discuss boundary conditions. Assume conditions of the general form

$$B_1(u) = \alpha_1 u(a) + \beta_1 \frac{du}{dx}(a) + \gamma_1 u(b) + \delta_1 \frac{du}{dx}(b) = 0 \tag{444}$$

$$B_2(u) = \alpha_2 u(a) + \beta_2 \frac{du}{dx}(a) + \gamma_2 u(b) + \delta_2 \frac{du}{dx}(b) = 0 \tag{445}$$

When these boundary conditions are combined with the requirement

$$Q[b, u(b), v^*(b), \frac{du}{dx}(b), \frac{dv^*}{dx}(b)] - Q[b, u(a), v^*(a), \frac{du}{dx}(a), \frac{dv^*}{dx}(a)] = 0 \tag{446}$$

one can eliminate two of the four quantities $u(a), \frac{du}{dx}(a), u(b), \frac{du}{dx}(b)$. The coefficients of the remaining two quantities with be linear combinations of

$$C_1(v^*) = \alpha_3 v^*(a) + \beta_3 \frac{dv^*}{dx}(a) + \gamma_3 v^*(b) + \delta_3 \frac{dv^*}{dx}(b) = 0 \tag{447}$$

$$C_2(v^*) = \alpha_4 v^*(a) + \beta_4 \frac{dv^*}{dx}(a) + \gamma_4 v^*(b) + \delta_4 \frac{dv^*}{dx}(b) = 0 \tag{448}$$

These are called adjoint boundary conditions.

When the functions u satisfy the homogeneous boundary conditions (**??**) and (**??**) and the function v^* satisfies the adjoint homogeneous boundary conditions then we obtain Green's identity

$$\int_a^b dx w(x) v^*(x) L_x u(x) = \int_a^b dx w(x) u(x) (L^\dagger v)^*(x) \tag{449}$$

The choice of homogeneous boundary conditions, along with sufficient differentiability defines the domain of the formal differential operator operator L_x. If $L_x = L_x^\dagger$ and the domains of both operators are identical then the operator is Hermitian.

In what follows I illustrate these concepts using second order differential operators. These have the form

$$L_x = a(x)\frac{d^2}{dx^2} + b(x)\frac{d}{dx} + c(x) \tag{450}$$

First I construct the formal adjoint operator. I consider the case $w(x) = 1$. Consider the following two quantities

$$a(x)v^*(x)\frac{d^2}{dx^2}u(x) - u(x)\frac{d^2}{dx^2}(a(x)v^*(x)) =$$

$$\frac{d}{dx}\left(a(x)v^*(x)\frac{d}{dx}u(x) - u(x)\frac{d}{dx}(a(x)v^*(x))\right) \tag{451}$$

and

$$v^*(x)b(x)\frac{d}{dx}u(x) + u(x)\frac{d}{dx}(b(x)v^*(x)) =$$
$$\frac{d}{dx}(v^*(x)b(x)u(x)) \tag{452}$$

Since these are both total derivatives, their sum is a total derivative, which gives after adding and subtracting $v^*(x)c(x)u(x)$:

$$v^*(x)[a(x)\frac{d^2}{dx^2}u(x)b(x)\frac{d}{dx}u(x) + c(x)u(x)] -$$
$$u(x)[\frac{d^2}{dx^2}(a^*(x)v(x)) - \frac{d}{dx}(b^*(x)v(x)) + v(x)c^*(x)]^* =$$
$$\frac{d}{dx}\left(a(x)v^*(x)\frac{d}{dx}u(x) - u(x)\frac{d}{dx}(a(x)v^*(x)) + v^*(x)b(x)u(x)\right) \tag{453}$$

In this case the formal adjoint to L_x is the operator

$$L_x^\dagger = a^*(x)\frac{d^2}{dx^2} + (2\frac{da^*}{dx} - b^*(x))\frac{d}{dx} + (\frac{d^2a^*}{dx^2} - \frac{db^*}{dx} + c^*(x)) \tag{454}$$

and the operator Q is

$$Q = a(x)v^*(x)\frac{d}{dx}u(x) - u(x)a(x)\frac{d}{dx}v^*(x) + v^*(x)(b(x) - \frac{da}{dx})u(x) \tag{455}$$

The conditions for L_x to be self-adjoint are

$$a(x) = a^*(x) \tag{456}$$

$$b(x) = 2\frac{da^*}{dx} - b^*(x) \tag{457}$$

$$c(x) = (\frac{d^2a^*}{dx^2} - \frac{db^*}{dx} + c^*(x)) \tag{458}$$

The first condition requires that a is real. The second condition requires that $b(x)$ is real and given by

$$b(x) = \frac{da}{dx} \tag{459}$$

The third condition requires that c is real. When L_x is self adjoint then

$$Q = a(x)(v^*(x)\frac{d}{dx}u(x) - u(x)\frac{d}{dx}v^*(x)) \tag{460}$$

In this case L_x can be expressed in the compact form

$$L_xu = \frac{d}{dx}\left(a(x)\frac{du}{dx}\right) + c(x)u(x) \tag{461}$$

0.21 Lecture 21

Next I show that given any second order differential equation with real coefficient functions that it is possible to choose a weight functions so it is self-adjoint.

Let

$$L = a(x)\frac{d^2}{dx^2} + b(x)\frac{d}{dx} + c(x) \tag{462}$$

and consider

$$w(x)g^*(x)a(x)\frac{d^2 f}{dx^2} - f(x)\frac{d^2}{dx^2}(w(x)g^*(x)a(x)) =$$

$$\frac{d}{dx}\{w(x)g^*(x)a(x)\frac{df}{dx} - f(x)\frac{d}{dx}(w(x)g^*(x)a(x))\} \tag{463}$$

and

$$w(x)g^*(x)b(x)\frac{df}{dx}+f(x)\frac{d}{dx}(w(x)g^*(x)b(x)) = \frac{d}{dx}(w(x)g^*(x)b(x)f(x)) \tag{464}$$

Subtracting these terms gives

$$w(x)g^*(x)\{a(x)\frac{d^2 f}{dx^2} + b(x)\frac{df}{dx} + c(x)f(x)\}-$$

$$f(x)\{\frac{d^2}{dx^2}(a(x)g^*(x)w(x)) + \frac{d}{dx}(b(x)g^*(x)w(x)) + c(x)g^*(x)w(x)\} =$$

$$\frac{d}{dx}\{w(x)a(x)g^*(x)\frac{df}{dx}-f(x)\frac{d}{dx}(w(x)g^*(x)a(x0)+w(x)b(x)g^*(x)f(x)\} \tag{465}$$

The left side of this equation can be put in the form

$$w(x)g^*(x)(L_x f)(x) - w(x)f(x)\{a(x)\frac{d^2}{dx^2} + (\frac{2}{w}\frac{d}{dx}(w(x)a(x)) - b(x))\frac{d}{dx}+$$

$$(c(x) - \frac{1}{w}\frac{d}{dx}(w(x)b(x)) + \frac{1}{w(x)}\frac{d^2}{dx^2}(a(x)w(x)))\}g^*(x) \tag{466}$$

From this expression it is possible to read off $L^\dagger$ and Q:

$$L^\dagger = a^*(x)\frac{d^2}{dx^2} + (\frac{2}{w}\frac{d}{dx}(w(x)a^*(x)) - b^*(x))\frac{d}{dx}+$$

$$(c^*(x) - \frac{1}{w}\frac{d}{dx}(w(x)b^*(x)) + \frac{1}{w(x)}\frac{d^2}{dx^2}(a^*(x)w(x)))\} \tag{467}$$

and

$$Q = w(x)a(x)g^*(x)\frac{df}{dx} - f(x)w(x)a(x)\frac{dg^*}{dx} -$$
$$f(x)\frac{d}{dx}(w(x)a(x))g^*(x) + w(x)b(x)g^*(x)f(x) \tag{468}$$

In arriving at this result the weight $w(x)$ was assumed to be known. It is possible to try to choose $w(x)$ so $L = L^\dagger$. I show that this can be done if all of the coefficient functions are real: $a(x) = a^*(x), \quad b(x) = b^*(x), \quad c(x) = c^*(x)$. The requirement that $L = L^\dagger$ is equivalent to the following relations

$$a(x) = a(x) \tag{469}$$

$$b(x) = \frac{2}{w}\frac{d}{dx}(w(x)a(x)) - b(x) \tag{470}$$

$$c(x) = c(x) - \frac{1}{w}\frac{d}{dx}(w(x)b(x)) + \frac{1}{w(x)}\frac{d^2}{dx^2}(a(x)w(x)) \tag{471}$$

The first equation is trivially satisfied, the second and third equations are satisfied provided

$$b(x) = \frac{1}{w}\frac{d}{dx}(w(x)a(x)) \tag{472}$$

This equation can be integrated to find the desired weight:

$$\frac{b(x)}{a(x)} - \frac{1}{a(x)}\frac{da}{dx} = \frac{1}{w}\frac{d}{dx}w(x) \tag{473}$$

$$\Downarrow$$

$$\ln(\frac{w(x)}{w(x_0)}) = -\ln(\frac{a(x)}{a(x_0)}) + \int_{x_0}^{x} \frac{b(x')}{a(x')}dx' \tag{474}$$

$$w(x) = \frac{w(x_0)a(x_0)}{a(x)} e^{\int_{x_0}^{x} \frac{b(x')}{a(x')}dx'} \tag{475}$$

which can be made positive by a suitable choice of $w(x_0)$.

With this choice the function Q becomes

$$Q = w(x)a(x)(g^*(x)\frac{df}{dx} - f(x)\frac{dg^*}{dx}) \tag{476}$$

We can also investigate the boundary conditions. The differential operator is Hermitian if the boundary conditions are identical to the adjoint boundary conditions. I consider some useful cases:

Dirichlet Boundary conditions: Dirichlet boundary conditions at $x = a$ and $x = b$ are $f(a) = f(b) = 0$. The adjoint boundary conditions can be read off from (476):

$$w(a)a(a)\frac{df}{dx}(a)g^*(a) = 0 \qquad w(b)a(b)\frac{df}{dx}(b)g^*(b) = 0 \tag{477}$$

which are satisfied for $g(a) = g(b) = 0$. Since these are identical to the boundary conditions on f, it follows that L with Dirichlet boundary conditions is a Hermitian operator.

Neumann Boundary conditions: Neumann boundary conditions at $x = a$ and $x = b$ are $\frac{df}{dx}(a) = \frac{df}{dx}(b) = 0$. The adjoint boundary conditions can be read off from (476):

$$w(a)a(a)f(a)\frac{dg}{dx}^*(a) = 0 \qquad w(b)a(b)f(b)\frac{dg}{dx}^*(b) = 0 \tag{478}$$

which are satisfied for $\frac{dg}{dx}(a) = \frac{dg}{dx}(b) = 0$. Since these are identical to the boundary conditions on f, it follows that L with Neumann boundary conditions is a Hermitian operator.

Mixed Boundary conditions: Mixed boundary conditions at $x = a$ and $x = b$ are

$$\alpha f(a) + \frac{df}{dx}(a) = 0 \beta f(b) + \frac{df}{dx}(b) = 0 \tag{479}$$

Now I compute the adjoint boundary conditions note that the mixed conditions lead to the relations

$$w(a)a(a)(g^*(a)(-\alpha f(a)) - f(a)\frac{dg^*}{dx}(a)) = 0 \tag{480}$$

$$w(b)a(b)(g^*(b)(-\beta f(b)) - f(b)\frac{dg^*}{dx}(b)) = 0 \tag{481}$$

Factoring out f gives

$$w(a)a(a)f(a)(-g(a)\alpha^* - \frac{dg}{dx}(a))^* = 0 \tag{482}$$

$$w(b)a(b)f(b)(-g(b)\beta^* - \frac{dg}{dx}(b))^* = 0 \tag{483}$$

For real α and β these are equivalent to

$$g(a)\alpha + \frac{dg}{dx}(a) = 0 \qquad g(b)\beta + \frac{dg}{dx}(b) = 0 \tag{484}$$

which have the exact same form as the original boundary conditions. This means the L with mixed boundary conditions and real coefficients (α, β) is a Hermitian operator. Dirichlet and Neumann boundary conditions are special cases of mixed boundary coefficients.

Periodic Boundary conditions: Periodic boundary conditions at $x = a$ and $x = b$ are $f(a) = f(b)$ and $\frac{df}{dx}(a) = \frac{df}{dx}(b)$ The adjoint boundary conditions can be read off from (476):

$$0 = w(b)a(b)(g^*(b)\frac{df}{dx}(b) - f(b)\frac{dg^*}{dx}(b)) - w(a)a(a)(g^*(a)\frac{df}{dx}(a) - f(a)\frac{dg^*}{dx}(a)) \tag{485}$$

$$0 = f(a)(w(a)a(a)\frac{dg^*}{dx}(a) - w(b)a(b)\frac{dg^*}{dx}(b)) +$$
$$\frac{df}{dx}(a)(w(b)a(b)g^*(b) - w(a)a(a)g^*(a)) \tag{486}$$

This will vanish provided $w(a)a(a) = w(b)a(b)$ and $g(a) = g(b)$ and $\frac{dg}{dx}(a) = \frac{dg}{dx}(b)$. Thus L is Hermitian with periodic boundary conditions provided $w(x)a(x)$ is periodic. Returning to the expression (475) for the weight

$$w(b)a(b) = w(a)a(a)e^{\int_a^b \frac{b(x')}{a(x')}dx'} \tag{487}$$

The periodic boundary conditions requires

$$1 = e^{\int_a^b \frac{b(x')}{a(x')}dx'} \tag{488}$$

0.22 Lecture 22:

In this section I discuss Green functions. Many of the concepts will be illustrated using Hermitian differential operators, however some of the methods deal with larger classes of Hermitian operators.

The discussion starts by considering the problems

$$L_x|u\rangle = |f\rangle \tag{489}$$

with homogeneous boundary conditions and

$$L_x^\dagger|v\rangle = |h\rangle \tag{490}$$

with the adjoint boundary conditions

Assume that one can find a linear operators G and g that satisfy

$$L_x G = I \tag{491}$$

where G has the same homogeneous boundary conditions as $u(x)$ and

$$L_x^\dagger g = I \tag{492}$$

where g has the same adjoint boundary conditions as $g(x)$ and

Then

$$L_x G|f\rangle = I|f\rangle = |f\rangle \tag{493}$$

which gives

$$|u\rangle = G|f\rangle \tag{494}$$

Similarly Then

$$L_x^\dagger g|h\rangle = I|h\rangle = |h\rangle \tag{495}$$

which gives

$$|v\rangle = g|h\rangle \tag{496}$$

Next I introduce more of the quantum mechanical notation that was used previously

$$\langle x|f\rangle = f(x) \tag{497}$$

$$\langle f|g\rangle = \int_a^b w(x) f^*(x) g(x) dx \tag{498}$$

The condition

$$\langle f|g\rangle = \langle f|I|g\rangle \tag{499}$$

leads to

$$I = \int dx|x\rangle w(x)\langle x| \tag{500}$$

and

$$\langle x|y\rangle = \frac{1}{w(x)}\delta(x-y) \tag{501}$$

$$\langle x|A|f\rangle = \int \langle x|A|y\rangle w(y)f(y)dy \tag{502}$$

$$\langle g|A|f\rangle = \int g(x)^* w(x)\langle x|A|y\rangle w(y)f(y)dy \tag{503}$$

In this notation the Green's functions G and g are denoted by

$$G(x,y) = \langle x|G|y\rangle \tag{504}$$

$$g(x,y) = \langle x|g|y\rangle \tag{505}$$

The equation $LG = I$ reads

$$L_x\langle x|G|y\rangle = \frac{1}{w(x)}\delta(x-y) \tag{506}$$

and $L^\dagger g = I$ reads

$$L_x^\dagger\langle x|g|y\rangle = \frac{1}{w(x)}\delta(x-y) \tag{507}$$

Note that if G and g satisfy the boundary conditions mentioned above then

$$\langle g, LG\rangle = \langle^\dagger g, G\rangle \tag{508}$$

where the meaning of this equation is

$$\int_a^b \langle x|g|z\rangle^* w(x)L_x\langle x|G|y\rangle dx = \int_a^b (L_x^\dagger\langle x|g|z\rangle)^* w(x)\langle x|G|y\rangle dx \tag{509}$$

Note that equations (506) and (507) give

$$\int_a^b \langle x|g|z\rangle^* w(x)\frac{1}{w(x)}\delta(x-y)dx = \int_a^b (\frac{1}{w(x)}\delta(x-z))^* w(x)\langle x|G|y\rangle dx \tag{510}$$

which gives

$$\langle y|g|z\rangle^* = \langle z|G|y\rangle \tag{511}$$

This means that second variable in G satisfies the adjoint boundary conditions.

In most cases we will consider G that Hermitian. Then $G = g$ and both variables satisfy the same boundary conditions.

If the two independent solutions of the homogeneous form of the differential equation are f_1 and f_2 then when $x \neq y$ the kernel $\langle x|G|y$ satisfies the homogeneous form of the differential equation. It necessarily has the form

$$\langle x|G|y\rangle = a_>(y)f_1(x) + b_>(y)f_2(x) \qquad x > y \tag{512}$$

$$\langle x|G|y\rangle = a_<(y)f_1(x) + b_<(y)f_2(x) \qquad x < y \tag{513}$$

The problem the then to find the four coefficient functions $a_>(y), a_<(y), b_>(y), b_<(y)$. To find this relation I use the form of the differential operator and the relation

$$L_x\langle x|G|y\rangle = \frac{1}{w(x)}\delta(x-y) \tag{514}$$

The desired relations and be obtained by integrating the general solutions for $x > y$ and $x <$ over the point $x = y$ in the above relation.

To do this write (514) as

$$\left(\frac{d^2}{dx^2} + \frac{b(x)}{a(x)}\frac{d}{dx} + \frac{c(x)}{a(x)}\right)\langle x|G|y\rangle = \frac{1}{a(x)w(x)}\delta(x-y) \tag{515}$$

Next let

$$\frac{1}{p(x)}\frac{dp}{dx}(x) = \frac{b(x)}{a(x)} \tag{516}$$

which can be integrated to get

$$p(x) = p(x_0)e^{\int_{x_0}^{x}\frac{b(x')}{a(x')}dx'} \tag{517}$$

Using (517) in (515) gives

$$\left(p(x)\frac{d^2}{dx^2} + \frac{dp}{dx}(x)\frac{d}{dx} + \frac{p(x)c(x)}{a(x)}\right)\langle x|G|y\rangle = \frac{p(x)}{a(x)w(x)}\delta(x-y) \tag{518}$$

$$\frac{d}{dx}(p(x)\frac{d}{dx}\langle x|G|y\rangle = -\frac{p(x)c(x)}{a(x)}\langle x|G|y\rangle + \frac{p(x)}{a(x)w(x)}\delta(x-y) \tag{519}$$

Next I use the relation

$$\delta(x-y) = \frac{d}{dx}\theta(x-y) \tag{520}$$

where

$$\theta(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases} \tag{521}$$

Next I write

$$\frac{p(x)}{a(x)w(x)}\delta(x-y) = \frac{p(y)}{a(y)w(y)}\delta(x-y) = \frac{d}{dx}\frac{p(y)}{a(y)w(y)}\theta(x-y) \tag{522}$$

Using (522) in (519) gives

$$\frac{d}{dx}\left(p(x)\frac{d}{dx}\langle x|G|y\rangle - \frac{p(y)}{a(y)w(y)}\theta(x-y)\right) = -\frac{p(x)c(x)}{a(x)}\langle x|G|y\rangle \tag{523}$$

Integrating (523) from x_0 to x gives

$$p(x)\frac{d}{dx}\langle x|G|y\rangle - \frac{p(y)}{a(y)w(y)}\theta(x-y)$$

$$-p(x_0)\frac{d}{dx}\langle x|G|y\rangle_{|x=x_0} + \frac{p(y)}{a(y)w(y)}\theta(x_0-y)$$

$$= -\int_{x_0}^{x} dx'\frac{p(x')c(x')}{a(x')}\langle x'|G|y\rangle \tag{524}$$

Next I evaluate this for $x = y + \epsilon$ and $x_0 = y - \epsilon$ assuming that $\langle x'|G|y\rangle$ is bounded and measurable. In the limit that $\epsilon \to 0$ the right hand side of the equation vanishes under this assumption. The left hand side becomes, canceling $p(x)$

$$\frac{d}{dx}\langle x|G|y\rangle_{|x=x+\epsilon} - \frac{d}{dx}\langle x|G|y\rangle_{|x=x-\epsilon} = \frac{1}{a(x)w(x)} \tag{525}$$

This gives the discontinuity in the derivative of $\langle x|G|y\rangle$ is x at $x = y$.

If I use (524) again write the derivative of $\langle x|G|y\rangle$ as

$$\frac{d}{dx}\langle x|G|y\rangle = \frac{p(y)}{p(x)a(y)w(y)}\theta(x-y) + \frac{p(x_0)}{p(x)}\frac{d}{dx}\langle x|G|y\rangle_{|x=x_0}]$$

$$-\frac{p(y)}{p(x)a(y)w(y)}\theta(x_0-y)-\frac{1}{p(x)}\int_{x_0}^{x}dx'\frac{p(x')c(x')}{a(x')}\langle x'|G|y\rangle \qquad (526)$$

This has the form

$$\frac{d}{dx}\langle x|G|y\rangle = f(x) \qquad (527)$$

where $f(x)$ is a bounded function, which gives

$$\langle x|G|y\rangle = \langle x'|G|y\rangle + \int_{x_0}^{x} f(x')dx' \qquad (528)$$

This implies that $\langle x|G|y\rangle$ is continuous at $x=y$. Equations (525) and (527) relate the coefficient functions $a_>(y), a_<(y), b_>(y), b_<(y)$. The relations are

$$\langle x|G|y\rangle = (a_>(y)-a_<(y))f_1(y)+(b_>(y)-b_<(y))f_2(y)=0 \qquad (529)$$

$$(a_>(y)-a_<(y))\frac{df_1}{dx}(y)+(b_>(y)-b_<(y))\frac{df_2}{dx}(y)=\frac{1}{a(y)w(y)} \qquad (530)$$

It is useful to write this pair of equations as a matrix equation

$$\begin{pmatrix} f_1(y) & f_2(y) \\ \frac{df_1}{dx}(y) & \frac{df_2}{dx}(y) \end{pmatrix}\begin{pmatrix} a_>(y)-a_<(y) \\ b_>(y)-b_<(y) \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{a(y)w(y)} \end{pmatrix} \qquad (531)$$

which has a solution

$$\begin{pmatrix} a_>(y)-a_<(y) \\ b_>(y)-b_<(y) \end{pmatrix} = \begin{pmatrix} f_1(y) & f_2(y) \\ \frac{df_1}{dx}(y) & \frac{df_2}{dx}(y) \end{pmatrix}^{-1}\begin{pmatrix} 0 \\ \frac{1}{a(y)w(y)} \end{pmatrix} \qquad (532)$$

because the determinant of the matrix is the Wronskian of the solutions of the homogeneous equation which is never zero.

The remaining relations are determined by imposing the homogeneous boundary conditions on the x variable in $\langle x|G|y\rangle$. This works if $L_x f(x) =$ has no non-trivial solutions satisfying the homogeneous boundary conditions.

To understand this result first note that we can always solve for the differences $a_> - a_<$ and $b_> - b_<$ because the Wronskian is not zero and $a(y)w(y)$ is not zero. We can write

$$\langle x|G|y\rangle = a_>(y)f_1(x)+b_>(y)f_2(x)+\langle x|R|y\rangle \qquad (533)$$

where here $\langle x|R|y\rangle$ is defined by this equation assuming $\langle x|G|y\rangle$ exists.

Recall that the homogeneous boundary conditions can be considered as linear operators B_1 and B_2 on the solution, giving formally

$$B_1(G) = a_> B_1(f_1) + b_> B_1(f_2) + B_1(R) = 0 \tag{534}$$

$$B_2(G) = a_> B_2(f_1) + b_> B_2(f_2) + B_2(R) = 0 \tag{535}$$

This is a linear system for the coefficients $a_>$ and $b_>$

$$\begin{pmatrix} B_1(f_1) & B_1(f_2) \\ B_2(f_1) & B_2(f_2) \end{pmatrix} \begin{pmatrix} a_>(y)(y) \\ b_>(y)(y) \end{pmatrix} = \begin{pmatrix} -B_1(R) \\ -B_2(R) \end{pmatrix} \tag{536}$$

This system will have a unique solution if the matrix has a non-zero determinant. This shows that the absence of non-trivial solutions satisfying the homogeneous boundary conditions is sufficient for the existence and uniqueness of $\langle x|G|y\rangle$ If the determinant is zero then there is a non-trivial linear combination $g(x) = c_1 f_1(c) + c_2 f_2(x)$ that is a solution to the homogeneous equation satisfying $B_1(g) = B_2(g) = 0$.

Next we show that if G exists then the homogeneous equation has no non-trivial solutions. Consider

$$\langle v|LG\rangle - \langle L^\dagger v|G\rangle = 0 \tag{537}$$

This is equivalent to

$$v^*(y) = \int w(x)\langle x|L^\dagger|v\rangle^* \langle x|G|y\rangle = \tag{538}$$

From this equation of $L^\dagger|v\rangle = 0$ has a non-trivial solution satisfying the adjoint homogeneous boundary conditions then by (538) this solution must vanish. Thus, $L^\dagger|v\rangle = 0$ has no non-trivial solutions. By our previous argument to g and $L^\dagger$ it follows that $\langle x|g|y\rangle$ exists and is unique. Next consider

$$\langle g|Lf\rangle = \langle L^\dagger g|f\rangle = f(y) \tag{539}$$

which shows that if $Lf = 0$ has a solution it must be zero, which means the homogeneous equation has no non-trivial solutions.

Next I consider an example. The differential equation is

$$L\langle x|u\rangle = \langle x|f\rangle \tag{540}$$

where

$$L = \frac{d^2}{dx^2} \tag{541}$$

on the interval $[0,1]$ satisfying Dirichlet boundary conditions

$$\langle 0|u\rangle = \langle 1|u\rangle = 0 \tag{542}$$

First note that there are two independent solutions to the homogeneous equation

$$L\langle x|f\rangle = 0 \tag{543}$$

given by

$$\langle x|f_1\rangle = 1 \qquad \langle x|f_1\rangle = x \tag{544}$$

If we set $a + b\cdot 0 = 0$ and $a + b \cdot 1 = 0$ we conclude that there are no non-trivial solutions to the homogeneous equation satisfying the boundary conditions.

We also note that

$$\langle x|v\rangle^* L\langle x|u\rangle - \langle x|u\rangle L\langle x|v\rangle^* = \frac{d}{dx}(\langle x|v\rangle^* \frac{d}{dx}\langle x|u\rangle - \langle x|u\rangle \frac{d}{dx}\langle x|v\rangle^* \tag{545}$$

which shows that L is self adjoint with weight 1. It is also straightforward to show that the adjoint boundary conditions are

$$\langle 0|v\rangle = \langle 1|v\rangle = 0 \tag{546}$$

which are identical to the boundary conditions on $\langle x|u\rangle$, which show that L with these boundary conditions is Hermitian.

The Green's function, $\langle x|G|y\rangle$ has the general form

$$\langle x|G|y\rangle = \begin{cases} a_>(y)\cdot 1 + b_>(y)\cdot x & x > y \\ a_<(y)\cdot 1 + b_<(y)\cdot x & x < y \end{cases} \tag{547}$$

The boundary conditions at $x = y$ for $a(x) = w(x) = 1$ so $1/(a(y)w(y) = 1$ are

$$(a_>(y) - a_<(y)) + (b_>(y) - b_<(y))y = 0 \tag{548}$$

$$(b_>(y) - b_<(y))1 = 1 \tag{549}$$

which gives

$$(b_>(y) - b_<(y)) = 1 \tag{550}$$

$$(a_>(y) - a_<(y)) = -y \tag{551}$$

Next we impose Dirichlet boundary conditions 0 and 1.

On the left end we can assume that $y > x = 0$ so

$$0 = a_<(y) \cdot 1 + b_<(y) \cdot 0 = a_<(y) \tag{552}$$

On the right end we can assume that $1 = x > y$

$$0 = a_>(y) \cdot 1 + b_>(y) \cdot 1 \tag{553}$$

Taking these four equations together we get

$$a_<(y) = 0 \qquad a_>(y) = -y \qquad b_>(y) = y \qquad b_<(y) = y - 1 \tag{554}$$

Putting these equations together gives

$$\langle x|G|y\rangle = \begin{cases} y(x-1) & x > y \\ (y-1)x & x < y \end{cases} \tag{555}$$

We can now use this to construct the general solution to the equation

$$L\langle x|u\rangle = \langle x|f\rangle \tag{556}$$

which is

$$\langle x|u\rangle = \int_0^1 dy \langle x|G|y\rangle\langle y|f\rangle \tag{557}$$

which in this case is

$$\langle x|u\rangle = (x-1)\int_0^x dyy\langle y|f\rangle + x\int_x^1 (y-1)\langle y|f\rangle \tag{558}$$

This clearly vanishes at $x = 0$ and $x = 1$. We can check the differential equation

$$\frac{d^2}{dx^2}\langle x|u\rangle =$$

$$\frac{d}{dx}[\int_0^x dyy\langle y|f\rangle + (x-1)x\langle x|f\rangle + \int_x^1 (y-1)\langle y|f\rangle - x(x-1)\langle x|f\rangle] =$$

$$\frac{d}{dx}[-\int_0^x dyy\langle y|f\rangle + \int_x^1 (y-1)\langle y|f\rangle] =$$

$$-x\langle y|f\rangle - (1-x)\langle y|f\rangle = \langle y|f\rangle \tag{559}$$

as required.

To be more specific choose $\langle x|f\rangle = x^2$. Then the general formula becomes

$$\langle x|u\rangle = (x-1)\int_0^x dyy^3 + x\int_x^1 (y-1)y^2 =$$

$$\frac{1}{4}(x-1)x^4 + x(\frac{1}{4} - \frac{1}{3} - \frac{x^4}{4} + \frac{x^3}{3}) = \frac{1}{12}x(x^3-1) \tag{560}$$

Which obviously satisfies the differential equation and boundary condition.

In this example the solution could have been written down by inspections; the advantage of the Green function methods is that solutions for any choice of $\langle x|f\rangle$ can be calculated this way.

0.23 Lecture 23

In this section I consider the situation where the homogeneous equation

$$L_x|f\rangle = 0 \tag{561}$$

has non-vanishing solutions satisfying homogeneous boundary conditions. In this case it is still possible to construct a generalized Green's function. I start by assuming that

$$L_x^\dagger|v_i\rangle = 0 \tag{562}$$

$$L_x|u_i\rangle = 0 \tag{563}$$

If there are more than one solution to these equations we can without loss of generality choose linear combinations of these solutions that are orthonormal

$$\langle v_i|v_j\rangle = \delta_{ij} \qquad \langle u_i|u_j\rangle = \delta_{ij} \tag{564}$$

The equations

$$L_x G = I - \sum_i |v_i\rangle\langle v_i| \qquad \langle u_i|G = 0 \tag{565}$$

$$L_x^\dagger g = I - \sum_i |u_i\rangle\langle u_i| \qquad \langle v_i|g = 0 \tag{566}$$

can be solved. The solution is called a generalized Green's function.

As an example of how this works consider

$$L = \frac{d^2}{dx^2} \tag{567}$$

with periodic boundary conditions at $x = -a$ and $x = a$. Clear the function

$$\langle x|u_1\rangle = \frac{1}{\sqrt{2a}} \tag{568}$$

is a unit normalized solution to the homogeneous equation satisfying the homogeneous boundary conditions. In this case $G = g$,

$$\langle x|v_1\rangle = \langle x|u_1\rangle = \frac{1}{\sqrt{2a}} \tag{569}$$

and equations (565) and (566) become

$$\frac{d^2}{dx^2}\langle x|G|y\rangle = \delta(x-y) - \langle x|v_1\rangle\langle v_1|y\rangle =$$

$$\delta(x-y) - \frac{1}{2a} \tag{570}$$

The above equations can be solved for $x \neq y$. Using

$$\frac{d^2}{dx^2}\left(\frac{x^2}{4a}\right) = \frac{1}{2a} \tag{571}$$

gives

$$\frac{d^2}{dx^2}(\langle x|G|y\rangle + \frac{x^2}{4a}) = 0 \tag{572}$$

for $x \neq y$. The solutions of this equation are arbitrary linear combinations of the solutions of the homogeneous form of the differential equations. It follows that

$$\langle x|G|y\rangle = -\frac{x^2}{4a} + a_>(y)1 + b_>(y)x \qquad x > y \tag{573}$$

and

$$\langle x|G|y\rangle = -\frac{x^2}{4a} + a_<(y)1 + b_<(y)x \qquad x < y \tag{574}$$

The boundary conditions on the generalized green function remain unchanged because the subtracted terms is smooth across $x = y$. This means that

$$a_>(y) + b_>(y)y - \frac{y^2}{4a} - a_<(y) - b_<(y)y + \frac{y^2}{4a} = 0 \tag{575}$$

and

$$b_>(y) - \frac{y}{2a} - b_<(y) + \frac{y}{2a} = 1 \tag{576}$$

which are identical to the matching conditions on the regular Green functions. These can be solved to give

$$b_<(y) = b_>(y) - 1 \tag{577}$$

$$a_<(y) = a_>(y) + y \tag{578}$$

Next we impose the periodic boundary conditions and the condition that

$$\langle u_1|G = 0 \tag{579}$$

$$\langle x|G|y\rangle = -\frac{x^2}{4a} + a_>(y)1 + b_>(y)x \qquad x > y \tag{580}$$

and

$$\langle x|G|y\rangle = -\frac{x^2}{4a} + (a_>(y) + y)1 + (b_>(y) - 1)x \qquad x < y \tag{581}$$

The periodic boundary conditions require

$$-\frac{a^2}{4a} + a_>(y)1 + b_>(y)a = -\frac{(-a)^2}{4a} + (a_>(y) + y)1 + (b_>(y) - 1)(-a) \tag{582}$$

and

$$-\frac{a}{2a} + b_>(y) = -\frac{(-a)}{2a} + (b_>(y) - 1) \tag{583}$$

The second equation is trivially satisfied for any value of $b_>(y)$. The first equation gives

$$b_>(y) = \frac{y}{2a} + \frac{1}{2} = 1 + b_<(y) \tag{584}$$

To find $a_>(y)$ the requirement $\langle u|G = 0$ means that

$$0 = \int_{-a}^{a} \frac{1}{\sqrt{2a}} \langle x|G|y\rangle dx. \tag{585}$$

Multiply through by $\sqrt{2a}$ to get

$$0 = \int_{-a}^{y} [-\frac{x^2}{4a} + (a_>(y) + y)1 + (b_>(y) - 1)x]dx +$$

$$\int_{y}^{a} [-\frac{x^2}{4a} + a_>(y)1 + b_>(y)xdx =$$

$$-\frac{y^3}{12a} - \frac{a^3}{12a} + (a_> + y)(y + a) + (b_> - 1)\frac{y^2}{2} - (b_> - 1)\frac{a^2}{2}$$

$$-\frac{a^3}{12a} + \frac{y^3}{12a} + a_>(a - y) + b_>\frac{a^2}{2} - b_>\frac{y^2}{2} \tag{586}$$

Inserting

$$b_> = \frac{y}{2a} + \frac{1}{2} \tag{587}$$

in (**??**) gives

$$-\frac{y^3}{12a}-\frac{a^3}{12a}+(a_>+y)(y+a)((\frac{y}{2a}+\frac{1}{2})-1)\frac{y^2}{2}-((\frac{y}{2a}+\frac{1}{2})-1)\frac{a^2}{2}$$

$$-\frac{a^3}{12a}+\frac{y^3}{12a}+a_>(a-y)+(\frac{y}{2a}+\frac{1}{2})\frac{a^2}{2}-(\frac{y}{2a}+\frac{1}{2})\frac{y^2}{2} \tag{588}$$

which is a linear equation for $a >$ which can be solved to get

$$a >= -\frac{1}{2a}(\frac{a^2}{3}+\frac{y^2}{2}+ya) \tag{589}$$

along with the full solutions for the coefficients

$$a_< = -\frac{1}{2a}(\frac{a^2}{3}+\frac{y^2}{2}-ya) \tag{590}$$

$$b_> = \frac{y}{2a}+\frac{1}{2} \tag{591}$$

$$b_< = \frac{y}{2a}-\frac{1}{2} \tag{592}$$

This example shows how to handle the special case of generalized Green's functions. In order to understand the nature of Generalized Green functions consider a compact operator of the form

$$L=\sum_{m=1}|v_n\rangle\lambda_n\langle u_n| \tag{593}$$

and

$$L^\dagger=\sum_{m=1}|u_n\rangle\lambda_n\langle v_n| \tag{594}$$

Note that

$$L^\dagger L=\sum_{m=1}|u_n\rangle\lambda_n^2\langle u_n| \tag{595}$$

$$LL^\dagger=\sum_{m=1}|v_n\rangle\lambda_n^2\langle v_n| \tag{596}$$

so $\lambda_n=0$ implies

$$L|u_n\rangle=L^\dagger|v_n\rangle=0 \tag{597}$$

If none of the $\lambda_n = 0$ then it is possible to construct the inverse to these operators

$$G = \sum_{m=1} |u_n\rangle \frac{1}{\lambda_n} \langle v_n| \tag{598}$$

$$g = \sum_{m=1} |v_n\rangle \frac{1}{\lambda_n} \langle u_n| \tag{599}$$

If some of the λ_n are zero restrict the sum $\sum \to \sum'$ to a sum over the non-zero λ_n and define

$$G' = \sum_{m=1}^{\prime} |u_n\rangle \frac{1}{\lambda_n} \langle v_n| \tag{600}$$

$$g' = \sum_{m=1}^{\prime} |v_n\rangle \frac{1}{\lambda_n} \langle u_n| \tag{601}$$

In this case

$$LG' = \sum_{m=1}^{\prime} |v_n\rangle\langle v_n| = I - \sum_{m=1}^{\prime\prime} |v_n\rangle\langle v_n| \tag{602}$$

where second sum in (**??**) is over all $|v_n\rangle$ that satisfy

$$L^{\dagger}|v_n\rangle = 0 \tag{603}$$

Similarly

$$L^{\dagger}g' = \sum_{m=1}^{\prime} |u_n\rangle\langle u_n| = I - \sum_{m=1}^{\prime\prime} |u_n\rangle\langle u_n| \tag{604}$$

where second sum in (604) is over all $|u_n\rangle$ that satisfy

$$L^{\dagger}|u_n\rangle = 0 \tag{605}$$

Note that

$$LG'L = L = \sum_{n=1}^{\prime} |v_n\rangle \lambda_n \langle v_n| \tag{606}$$

$$G'LG' = G' = \sum_{n=1}^{\prime} |u_n\rangle \frac{1}{\lambda_n} \langle u_n| \tag{607}$$

$$(LG') = (LG')^\dagger = \sum_{n=1}^{\prime} |v_n\rangle\langle v_n| \tag{608}$$

$$(G'L) = (G'L)^\dagger = \sum_{n=1}^{\prime} |u_n\rangle\langle u_n| \tag{609}$$

These equations are called the Penrose equations. They always have a unique solution called the Moore Penrose Generalized inverse.

So far we have used Green functions to treat the case where the solution satisfies homogeneous boundary conditions. We can also treat the case where the solution satisfies inhomogeneous boundary conditions of the form

$$b_{i1}f(b) + b_{i2}f'(b)b_{i3}f(a) + b_{i4}f'(a) = \sigma_i \tag{610}$$

for $i = 1, 2$.

To solve this problem construct the Green function for the differential equation satisfying the homogeneous form of the boundary conditions.

Consider the generalized Green's identity:

$$\int [(\langle x|g|y\rangle)^* w(x) L_x \langle x|f\rangle - (L_x^\dagger \langle x|g|y\rangle)^* w(x)\langle x|f\rangle] dx =$$

$$Q(g^*, f, g^{*\prime}, f')(b) - Q(g^*, f, g^{*\prime}, f')(a) \tag{611}$$

This holds for any f. Using $L^\dagger g = I$ gives

$$\int (\langle x|g|y\rangle)^* w(x) L_x \langle x|f\rangle dx =$$

$$\langle x|f\rangle + Q(g^*, f, g^{*\prime}, f')(b) - Q(g^*, f, g^{*\prime}, f')(a) \tag{612}$$

Next choose the weight functions w that makes $L = L^\dagger$. In this case Q has the form

$$Q(f, g) = w(x)a(x)[f(x)\frac{dg}{dx} - g(x)\frac{df}{dx}]$$

and $\langle x|g|y\rangle^* = \langle y|G|x\rangle$. It follows that

$$\langle x|f\rangle =$$

$$\int \langle y|G|x\rangle w(x) L_x \langle x|f\rangle dx$$

$$-w(b)a(b)[\langle y|G|b\rangle \frac{df}{dx}(b) - f(b)\frac{d}{dx}\langle y|G|x=b\rangle]$$
$$+w(a)a(a)[\langle y|G|a\rangle \frac{df}{dx}(a) - f(a)\frac{d}{dx}\langle y|G|x=a\rangle] \qquad (613)$$

This equation is true for any function $\langle x|f\rangle$

Now assume that $L\langle x|f\rangle = \langle x|h\rangle$. The case $\langle x|h\rangle = 0$ is a special case where this method also works. It gives solution to the homogeneous form of the differential equation with inhomogeneous boundary conditions.

Also note that G satisfies adjoint boundary conditions in the second variable. What survives involves boundary values of f.

To be specific assume G is the Green function satisfying homogeneous Dirichlet boundary condition. The adjoint boundary conditions are also Dirichlet, which means

$$0 = \langle y|G|b\rangle = \langle y|G|a\rangle \qquad (614)$$

This leaves

$$\langle x|f\rangle =$$
$$\int \langle y|G|x\rangle w(x)\langle x|h\rangle dx +$$
$$w(b)a(b)[f(b)\frac{d}{dx}\langle y|G|x=b\rangle]$$
$$-w(a)a(a)[f(a)\frac{d}{dx}\langle y|G|x=a\rangle] \qquad (615)$$

where the inhomogeneous dirichlet boundary conditions are

$$f(x=a) = f(a) \qquad f(x=b) = f(b) \qquad (616)$$

which is specified input on the right side of this equation.

This equation gives the solution to

$$L\langle x|f\rangle = \langle x|h\rangle \qquad (617)$$

with values $\langle b|f\rangle = f(b)$ and $\langle a|f\rangle = f(a)$

Obviously there is nothing special about Dirichlet boundary conditions. A similar result could have been achieved using Neumann boundary condition. Setting $\langle x|h\rangle = 0$ gives a solution of homogeneous differential equation satisfying inhomogeneous boundary conditions.

0.24 Strum Liouville Equations

Consider second order differential operator with weight chosen so $L = L^\dagger$. In this case

$$L\langle x|f\rangle = \frac{d}{dx}(a(x)\frac{d\langle x|f\rangle}{dx}) + c(x)\langle x|f\rangle \tag{618}$$

Consider the eigenvalue problem for this operator

$$L\langle x|f\rangle = \lambda\langle x|f\rangle \tag{619}$$

In what follows we assume that $\langle x|G|y\rangle$ exists and satisfies $LG = I$. We also assume that while $\langle x|G|y\rangle$ may have a discontinuous derivative at $x = y$, the functions

$$|\langle x|G|y\rangle| < C < \infty \tag{620}$$

for all $a \leq x, y \leq b$.

Consider

$$L\langle x|f\rangle = \lambda\langle x|f\rangle \tag{621}$$

Then

$$G|f\rangle = \frac{1}{\lambda}GL|f\rangle = \frac{1}{\lambda}|f\rangle \tag{622}$$

which means that $|f\rangle$ is also an eigenvector of G with eigenvalue $\frac{1}{\lambda}$. The eigenvalue equation has the from

$$\langle x|G|y\rangle w(y)dy\langle y|f\rangle = \frac{1}{\lambda}\langle x|f\rangle \tag{623}$$

The bound $|\langle x|G|y\rangle| < C < \infty$ ensures that for a bounded interval $[a, b]$ that

$$\int |\langle x|G|y\rangle|^2 w(x)w(y)dxdy \leq C^2(b-a)^2W^2 < \infty \tag{624}$$

where W is any upper bound for $w(x)$ on $[a, b]$. This shows that $\langle x|G|y\rangle$ is Hilbert-Schmidt and hence G is compact.

It follows that $\frac{1}{|\lambda_n|} \to 0$ as $n \to \infty$. This means that the eigenvalues of L satisfy $|\lambda_n| \to \infty$ as $n \to \infty$.

Since we know

$$G = \sum_{n=0}^{\infty} |n\rangle\eta_n\langle n| \tag{625}$$

it follows that

$$L = \sum_{n=0}^{\infty} |n\rangle \frac{1}{\eta_n} \langle n| \tag{626}$$

Example

Let $\gamma(\lambda)$ denote a curve connecting points a and b as λ varies between 0 and 1. It follows that

$$\gamma(0) = 1 \qquad \gamma(1) = b \tag{627}$$

In classical mechanics a particle travels along a path between two points that is an extremum of the action functional defined by

$$A[\gamma] = \int_0^1 d\lambda [\dot{\gamma}_i \dot{\gamma}_i - V(\vec{\gamma})] \tag{628}$$

where V represents the "potential energy". The equations define the desired part involve writing a general path as the sum of the correct path $\vec{\gamma}_0(\lambda)$ and a correction $\delta\vec{\gamma}(\lambda)$:

$$\gamma_i = \gamma_{i0} + \eta \delta\gamma_i(\lambda) \tag{629}$$

where the requirement that both paths have the same endpoint means that

$$\delta\vec{\gamma}(0) = \delta\vec{\gamma}(1) = 0. \tag{630}$$

The condition that $\vec{\gamma}_0$ is an extremum of this action functional can be written as

$$\frac{dA[\gamma, \eta]}{d\eta} = 0 \tag{631}$$

when $\vec{\gamma} = \vec{\gamma}_0$. This gives the usual Lagranges equations

$$\frac{d}{d\lambda}\left(\frac{\partial L}{\partial \dot{\gamma}_i}\right) - \frac{\partial L}{\partial \gamma_i} = 0 \tag{632}$$

We can ask the question if the solution of this equation is a minimum of the action functional or simply an extremum. This is done by looking at the "second derivative" and checking to see that it is positive. I will show that this will be the case if a certain Strum Liouville differential equation has only positive eigenvalues.

The check is to look at the second derivative of the action with respect to η evaluated at the extreme "point" $\vec{\gamma} = \vec{\gamma}_0$:

$$\frac{d^2}{d\eta^2} A[\vec{\gamma}_0 + \eta\delta\vec{\gamma}] =$$

$$\int_0^1 \left\{ \left(\frac{\partial^2 L}{\partial\gamma_i\partial\gamma_j}\right)_0 \delta\gamma_i\delta\gamma_j + 2\left(\frac{\partial^2 L}{\partial\gamma_i\partial\dot{\gamma}_j}\right)_0 \delta\gamma_i\delta\dot{\gamma}_j + \left(\frac{\partial^2 L}{\partial\dot{\gamma}_i\partial\dot{\gamma}_j}\right)_0 \delta\dot{\gamma}_i\delta\dot{\gamma}_j \right\} \quad (633)$$

The partial derivative terms are well defined functions of the λ because they depend on the original curve γ_0. The equation is quadratic in the variations, but is is homogenous. The means the we can always scale the variations to make this have any desired value. In order to remove the freedom to rescale the variations we normalize them so

$$\int_0^1 \sum_i \delta\gamma_i\delta\gamma_i d\lambda = 1 \quad (634)$$

We then define a new functional of the *variations*

$$\int_0^1 \left\{ 2\left(\frac{\partial^2 L}{\partial\gamma_i\partial\gamma_j}\right)_0 \delta\gamma_i\delta\gamma_j + 2\left(\frac{\partial^2 L}{\partial\gamma_i\partial\dot{\gamma}_j}\right)_0 \delta\gamma_i\delta\dot{\gamma}_j + \right.$$
$$\left. 2\left(\frac{\partial^2 L}{\partial\dot{\gamma}_i\partial\dot{\gamma}_j}\right)_0 \delta\dot{\gamma}_i\delta\dot{\gamma}_j - 2\sigma\sum_i \delta\gamma_i\delta\gamma_i \right\} d\lambda \quad (635)$$

We seek extreme values of this functional subject to the constraint (**??**). In this case

$$\delta\vec{\gamma} = \delta\vec{\gamma}_0 + \rho\delta\vec{\gamma}_1) \quad (636)$$

Clearly the minimum will be an extremum of this functional. Taking the derivative with respect to ρ setting $\rho = 0$ gives

$$\int_0^1 \left\{ 2\left(\frac{\partial^2 L}{\partial\gamma_i\partial\gamma_j}\right)_0 \delta\gamma_{i0}\delta\gamma_{j1} + 2\left(\frac{\partial^2 L}{\partial\gamma_i\partial\dot{\gamma}_j}\right)_0 (\delta\gamma_{i0}\delta\dot{\gamma}_{j1} + \delta\gamma_{i1}\delta\dot{\gamma}_{j0}) \right.$$
$$\left. +2\left(\frac{\partial^2 L}{\partial\dot{\gamma}_i\partial\dot{\gamma}_j}\right)_0 \delta\dot{\gamma}_{i0}\delta\dot{\gamma}_{j1} - 2\sigma\delta\gamma_{i0}\gamma_{i1} \right\} \quad (637)$$

Integrating the $\delta\dot{\gamma}_{i1}$ by parts gives the following

$$\int_0^1 \left\{ 2\left(\frac{\partial^2 L}{\partial\gamma_i\partial\gamma_j}\right)_0 \delta\gamma_{i0} - 2\frac{d}{d\lambda}[\left(\frac{\partial^2 L}{\partial\gamma_i\partial\dot{\gamma}_j}\right)_0 \delta\gamma_{i0}] + 2\left(\frac{\partial^2 L}{\partial\gamma_j\partial\dot{\gamma}_i}\right)_0 \delta\dot{\gamma}_{i0} \right.$$
$$\left. -\frac{d}{d\lambda}[2\left(\frac{\partial^2 L}{\partial\dot{\gamma}_i\partial\dot{\gamma}_j}\right)_0 \delta\dot{\gamma}_{i0}] - 2\sigma\delta\gamma_{j0}\delta \right\} \delta\gamma_{j1} d\lambda \quad (638)$$

If this is required to be extremal then is should vanish for all $\delta\gamma_{i1}$. This condition requires

The resulting equation is

$$-\frac{d}{d\lambda}[\left(\frac{\partial^2 L}{\partial\dot{\gamma}_i\partial\dot{\gamma}_j}\right)_0 \frac{d\delta\gamma_{i0}}{d\lambda}] + \left(\frac{\partial^2 L}{\partial\gamma_i\partial\gamma_j}\right)_0 \delta\gamma_{i0} - \frac{d}{d\lambda}[\left(\frac{\partial^2 L}{\partial\gamma_i\partial\dot{\gamma}_j}]\right)_0 \delta\gamma_{i0}] = -\sigma\delta\gamma_{j0} \tag{639}$$

This equation has the general form

$$\frac{d}{d\lambda}\left(A_{ij}(\lambda)\frac{df_i}{d\lambda}\right) + C_{ij}(\lambda)f_j = \sigma f_i \tag{640}$$

This is an ordinary Strum Liouville equation when γ is a one dimensional curve. The f_i represent variations of the curve in extremal directions while the eigenvalue sigma represents the value of second order variation when the displacements are normalized to 1.

The condition for a minimum is that all of the eigenvalues of the Strum Liouville problem must be positive. If the time interval $[0, 1]$ is changed the eigenvalues will change. Normally for short times the eigenvalues are positive, however eventually for longer times one eventually may pass through zero.

Next I consider the relation between Green functions, Resolvents, and solutions to the Strum- Liouville problem.

Let L be a second order Strum Liouville differential operator with a given set of boundary conditions and define

$$L' = z - L \tag{641}$$

where z is any complex number. We know that L has a complete set of eigenvectors $|f_n\rangle$ with eigenvalues λ_n which get large in magnitude as $n \to \infty$

Clearly

$$L'|f_n\rangle = (z - L)|f_n\rangle = (z - \lambda_n)|f_n\rangle \tag{642}$$

so L' also has the same complete set of eigenvectors with different eigenvalues.

The Green's function for L' satisfies

$$L'G'(z) = I \tag{643}$$

and it exists whenever z is not an eigenvalue of L. Expand

$$\langle x|G'(z)|y\rangle = \sum_{n=0}^{\infty}\langle x|f_n\rangle a_n(y) \tag{644}$$

To find the expansion coefficients $a_n(y)$

$$L'\langle x|G'(z)|y\rangle = \sum_{n=0}^{\infty}\langle x|f_n\rangle(z-\lambda_n)a_n(y) = \delta(x-y) \tag{645}$$

If we expand the delta function the same way

$$\delta(x-y) = \sum_{n=0}^{\infty}\langle x|f_n\rangle b_n(y) \tag{646}$$

and multiply by $\langle f_m|x\rangle$ and integrate we get

$$\langle f_m|y\rangle = \sum_{n=0}^{\infty}\delta_{mn}\rangle b_n(y) = b_m(y) \tag{647}$$

giving

$$\delta(x-y) = \sum_{n=0}^{\infty}\langle x|f_n\rangle\langle f_n|y\rangle \tag{648}$$

Using (648) in (645) gives

$$\sum_{n=0}^{\infty}\langle x|f_n\rangle[(z-\lambda_n)a_n(y)-\langle f_n|y\rangle] = 0 = 0 \tag{649}$$

The independence of the basis functions gives

$$(z-\lambda_n)a_n(y)-\langle f_n|y\rangle = 0 \tag{650}$$

which and be solve for the coefficients $a_n(y)$:

$$a_n(y) = \frac{\langle f_n|y\rangle}{z-\lambda_n} \tag{651}$$

Using (651) in (644) gives

$$\langle x|G'(z)|y\rangle = \sum_{n=0}^{\infty}\frac{\langle x|f_n\rangle\langle f_n|y\rangle}{z-\lambda_n} \tag{652}$$

This expressions relates the resolvent

$$G(z) = (z-L)^{-1} \tag{653}$$

of L to both the Green functions and eigenfunction expansion.

In some problems it is easy to solve the Strum Liouville eigenvalue problem, and then the above method can be used to write the Greens function in terms of eigenfunctions and eigenvalues of L.

It also happens that if one can find the resolvent by other methods, it can be used to extract solutions of the Strum Liouville problem. Assume the λ_m is an isolated eigenvalue of L. We showed previously in the neighborhood of any point z not an eigenvalue of of L that $G(z)$ is analytic in z. For isolated eigenvalues it as simple poles at the eigenvalues. Consider

$$\frac{1}{2\pi}\int G(z)dz \tag{654}$$

about a contour that encloses $z = \lambda_n$, but no other eigenvalues. Then the above integral becomes

$$\frac{1}{2\pi}\int G(z)dz = |f_n\rangle\langle f_n| \tag{655}$$

which is the orthogonal projector on the $n^t h$ eigenstate of L.

The advantage of the representation of the Green functions in terms of eigenfunctions expansions is that they can be applied to any Hermetian Hilbert space operators, whether they are differential operators, partial differential operators, integral operators, or integrodifferential operators.

The resolvent $G(z)$ has poles at the eigenvalues of L. There are precisely at the points where L' has homegeneous solutions satisfying homegeneous boundary conditions.

0.25 Lecture 25

So far our discussion of second order linear differential equations was based on the assumption that we already had at least one solution to the homogeneous equation. In this section we discuss one method for solving a large class of second order differential equation using power series.

Consider a second order linear differential operator of the form

$$L = \frac{d^2}{dz^2} + p(z)\frac{d}{dz} + q(z) \tag{656}$$

where $p(z)$ and $q(z)$ are analytic in an open region R except possibly on a set of isolated points. Recall that an open region (pathwise connected)

$z_0 \in R$ is called an ordinary point if both $p(z)$ and $q(z)$ are analytic at $z = z_0$.

$z_0 \in R$ is called a singular point if $p(z)$ or $q(z)$ has a singularity at $z = z_0$.

$z_0 \in R$ is called a regular singular point if $p(z)$ has a pole of order ≤ 1 and $q(z)$ has a pole of order ≤ 2 at $z = z_0$.

$z_0 \in R$ is called an irregular singular point if it is a singular point but not a regular singular point.

Next I show that in a neighborhood of an ordinary point of L the equation $Lu(z) = 0$ has two independent analytic solutions. Consider the problem

$$Lu(z) = 0 \qquad u(z_0) = a \qquad u'(z_0) = b \tag{657}$$

The first step is to change this into a differential equation for a new function $f(z)$ defined by

$$u(z) = f(z)e^{-\frac{1}{2}\int_{z_0}^{z} p(z')dx'} \tag{658}$$

where the integral does not depend on path in a sufficiently small region about z_0. Using this definition

$$u'(z) = \left(f'(z) - f(z)\frac{1}{2}p(z) \right) e^{-\int_{z_0}^{z} p(z')dx'} \tag{659}$$

and

$$u''(z) = \left(f''(z) - 2\frac{1}{2}f'(z)p(z) - f(z)\frac{1}{2}p'(z) + f(z)\frac{1}{4}p(z)^2 \right) e^{-\int_{z_0}^{z} p(z')dx'} \tag{660}$$

Using these formulas in the differential equation gives

$$Lu(z) =$$

$$\left(f''(z) - f'(z)p(z) + f(z)\frac{1}{4}p(z)^2 f'(z)p(z) - f(z)\frac{1}{2}p'(z) - f(z)\frac{1}{2}p(z)^2 + q(z)f(z) \right) \times$$

$$e^{-\frac{1}{2}\int_{z_0}^{z} p(z')dx'} = 0 \tag{661}$$

which is equivalent to

$$f''(z) + \left(q(z) - \frac{1}{4}p(z)^2 - \frac{1}{2}p'(z) \right) f(z) = 0 \tag{662}$$

which I write as

$$f''(z) = -k(z)f(z) \tag{663}$$

with

$$k(z) = q(z) - \frac{1}{4}p(z)^2 - \frac{1}{2}p'(z) \tag{664}$$

To solve this integrate twice to get

$$f'(z) = f'(z_0) - \int_{z_0}^{z} k(z')f(z')dz' \tag{665}$$

$$f(z) = f(z_0) + f'(z_0)(z - z_0) - \int_{z_0}^{z} \int_{z_0}^{z'} k(z'')f(z'')dz'' =$$

$$a + b(z - z_0) - \int_{z_0}^{z} dz' \int_{z_0}^{z'} k(z'')f(z'')dz'' = \tag{666}$$

I reduce the integral to a one dimensional integral by integrating by parts as follows

$$a + b(z - z_0) - \int_{z_0}^{z} (\frac{d}{dz'}z') \int_{z_0}^{z'} k(z'')f(z'')dz'' =$$

$$a + b(z - z_0) + \int_{z_0}^{z} z'k(z')f(z')dz' - z\int_{z_0}^{z} k(z')f(z')dz' =$$

$$a + b(z - z_0) + \int_{z_0}^{z} (z' - z)k(z')f(z')dz' \tag{667}$$

I solve this by successive approximations. I assume that

$$\langle x|f\rangle = \sum_{n=0}^{\infty} \langle z|f_n\rangle \tag{668}$$

where

$$\langle z|f_0\rangle = a + b(z - z_0) \tag{669}$$

$$\langle z|f_n\rangle = \int_{z_0}^{z} (z' - z)k(z')f_{n-1}(z') \tag{670}$$

By induction I assume that

$$|\langle z|f_n\rangle| \leq F_0 \frac{K_0^n (z - z_0)^{2n}}{n!} \tag{671}$$

where K_0 is the maximum value of $k(x)$ along the curve between z_0 and z Using this induction assumption in equation (670) for $n-1$ using the path

$$z' = z_0 + (z - z_0)t \tag{672}$$

I get

$$|\langle z|f_n\rangle| \leq$$

$$\int_{z_0}^{z} |(z' - z_0 - (z - z_0))| K_0 F_0 \frac{K_0^{n-1}|z' - z_0|^{2n-2}}{(n-1)!} \leq$$

$$|z - z_0|^{2n} \int_0^1 dt(1-t)K_0 F_0 \frac{K_0^{n-1} t^{2n-2}}{(n-1)!} \leq \tag{673}$$

Next note

$$\int_0^1 dt(1-t)t^{2n-2}dt = B(2, 2n-1) = \frac{\Gamma(2)\Gamma(2n-1)}{\Gamma(2n+1)} =$$

$$\frac{1(2n-2)!}{(2n)!} = \frac{1}{2n(2n-1)} \leq \frac{1}{n} \tag{674}$$

Using (674) in (673) gives

$$|\langle z|f_n\rangle| \leq |z - z_0|^{2n} K_0^n F_0 \frac{1}{(n)!} \tag{675}$$

which verifies that if the $(n-1)$-st term satisfies the induction assumption then the $n - th$ term satisfies the induction assumption.

Note that the starting function in the iteration is analytic and $k(z)$ is analytic (for z_0 regular). At each stage the second derivative of the $n - th$ terms is $-k(z)$ times the $n-1$-st term, which is analytic. Consider a disk centered at $z = z_0$ where both $f_0(z)$ and $k(z)$ are analytic. Then the integral in analytic in this region (see equation 16.1 of the text). Integrating this function again we obtain another function analytic in the same region.

It follows that the series (667) is a uniformly convergent sum of analytic functions in a region where $k(z)$ is analytic.

Because of the uniform convergence I can change the order of sum and derivative to obtain

$$\frac{d^2}{dx^2}\langle x|f\rangle = \sum_{n=0}^{\infty}\langle z|f_n\rangle = \sum_{n=1}^{\infty}\frac{d^2}{dx^2}\langle z|f_n\rangle =$$

$$-\sum_{n=1}^{\infty}k(z)\langle z|f_{n-1}\rangle = -k(z)\sum_{n=0}^{\infty}\langle z|f_n\rangle \tag{676}$$

which shows that this series satisfies the differential equation. The series solution also clearly satisfies the boundary conditions.

If there is another solution satisfying the same boundary condition the difference must be a solution of the equation that vanishes and has vanishing derivative at $z = z_0$. If follows from the differential equation that all derivatives vanish. By Taylor's theorem the difference function is necessarily zero.

This shows that the series solution is the unique analytic solution of the differential equation satisfying the two boundary conditions at z_0.

Since the solution of the differential equation is analytic in a neighborhood of an ordinary point z_0 it can be expressed as a convergent power series about $z - z_0$

$$u(z) = \sum_{n=0}^{\infty} c_n (z-z_0)^n \tag{677}$$

The equation can be solved by recursion:

$$k(z) = \sum k_n (z-z_0)^n \tag{678}$$

$$\frac{d^2u}{dz^2} = -k(z)u(z) \tag{679}$$

$$\sum_{n=0}^{\infty} c_n n(n-1)(z-z_0)^{n-2} = \sum_{l=0}^{\infty}\sum_{m=0}^{\infty} k_l c_m (z-z_0)^{m+l}. \tag{680}$$

The first two terms on the left do not contribute, giving

$$\sum_{n=0}^{\infty} c_n n(n-1)(z-z_0)^{n-2} = \sum_{n=2}^{\infty} c_n n(n-1)(z-z_0)^{n-2} =$$

$$\sum_{n=0}^{\infty} c_{n+2}(n+2)(n+1)(z-z_0)^n, \tag{681}$$

while the sum on the right is

$$\sum_{l=0}^{\infty} \sum_{m=0}^{\infty} k_l c_m (z-z_0)^{m+l} =$$

$$\sum_{k=0}^{\infty} (z-z_0)^k \sum_{m=0}^{k} k_{k-m} c_m \tag{682}$$

Equating the coefficients of identical powers of $(z - z_0)$ gives the recursion relations

$$c_{n+2} = \sum_{m=0}^{n} \frac{k_{n-m}}{(n+2)(n+1)} c_m \tag{683}$$

which requires the boundary conditions c_0 and c_1 to start the recursion.

As an example consider Hermite's equation

$$\frac{d^2}{dx^2}\langle z|u\rangle - 2z\frac{d}{dz}\langle z|u\rangle + 2\lambda\langle z|u\rangle = 0 \tag{684}$$

Here I simply write the solution as a series

$$\langle z|u\rangle = \sum_{n=0}^{\infty} c_n z^n \tag{685}$$

Inserting this series into the differential equation gives

$$\sum_{n=2}^{\infty} c_n n(n-1) z^{n-2} - \sum_{n=1}^{\infty} 2n c_n z^n + 2\lambda \sum_{n=0}^{\infty} c_n z^n = \tag{686}$$

$$\sum_{n=0}^{\infty} c_{n+2}(n+2)(n+1) z^n - \sum_{n=1}^{\infty} 2n c_n z^n + 2\lambda \sum_{n=0}^{\infty} c_n z^n \tag{687}$$

This gives the recursion

$$c_{n+2} = \frac{2(n-\lambda)}{(n+2)(n+1)} c_n \tag{688}$$

Note that the sum is finite when λ is an integer, but analytic solutions exist for non-integer lambdas.

Next I consider solutions for the case that the differential equation is solved at a regular singular point. In this case

$$A(z) = (z - z_0)p(z) \qquad B(z) = (z - z_0)^2 q(z) \tag{689}$$

are analytic in a region that includes the point z_0 These functions can be expanded about z_0

$$A(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \qquad B(z) = \sum_{n=0}^{\infty} b_n (z - z_0)^n \tag{690}$$

where

$$a_n = \frac{1}{2\pi i} \oint A(z)(z - z_0)^{-n+1} \qquad b_n = \frac{1}{2\pi i} \oint B(z)(z - z_0)^{-n+1} \tag{691}$$

where the integral is a circle in the region of analyticity about z_0.

I try a solution of the differential of the form

$$\langle z|u\rangle = (z - z_0)^r \sum_{n=0}^{\infty} c_n (z - z_0)^n \tag{692}$$

The constant r will be chosen later.

Inserting all three series in the differential equation, and multiplying the result by $(z - z_0)^{2-r}$ gives

$$\sum_{n=0}^{\infty} (n + r)(n + r - 1) c_n (z - z_0)^n =$$

$$\sum_{k=0}^{\infty} \sum_{r=0}^{\infty} a_k c_m (m + r)(z - z_0)^{k+1+m-1} + \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} b_k c_m (z - z_0)^{k+m}$$

$$\sum_{n=0}^{\infty} \sum_{m=0}^{n} (a_{n-m} c_m (m + r) + b_{n-m} c_m)(z - z_0)^n \tag{693}$$

Equating coefficients of z_n gives

$$(n + r)(n + r - 1) c_n + \sum_{m=0}^{n} (a_{n-m}(m + r) + b_{n-m}) c_m) = 0 \tag{694}$$

Next I separate all terms involving c_n to get

$$c_n[(n+r)(n+r-1)+a_0(n+r)+b_0] = -\sum_{m=0}^{n-1}(a_{n-m}(m+r)+b_{n-m})c_m) \quad (695)$$

This recursion make sense as long as $n \neq 0$ We can now choose r so the coefficient of c_0 vanishes. This gives

$$\lambda_0(r) := r(r-1) + a_0 r + b_0 = 0 \quad (696)$$

This is called the indicial equation. It has two roots that allow us to choose c_0 as an arbitrary starting value for the recursion.

This equation has two roots. If they are distinct and do not differ by integers they lead to independent solutions of the differential equation.

The indicial equation has two roots. I call them r_+ and r_-. I distinguish them by requiring $Re(r_+) \geq Re(r_-)$.

First I demonstrate that (692) with c_n determined by (695) with $r = r_+$.

$$(n+r)(n+r-1) + a_0(n+r) + b_0 = \lambda_0(n+r) \quad (697)$$

Because r^+ is the root of this polynomial with the largest real part, this is necessarily non-sero for any $n > 0$ (otherwise there will be a root with larger real part). This gives

$$c_n = -\sum_{m=0}^{n-1} \frac{(a_{n-m}(m+r) + b_{n-m})c_m)}{\lambda(n+r_+)} \quad (698)$$

Next I argue that the series generated this way is uniformly convergent in a sufficiently small disk about $z = z_0$.

Start by observing that both $A(z)$ and $B(z)$ are analytic in a neighborhood of z_0. In a disk D in the region of analyticity both functions are bounded because they are analytic on the closed disk:

$$\|A(z)| < \Lambda_1 \quad (699)$$

$$\|B(z)| < \Lambda_2 \qquad z \in D \quad (700)$$

The integral representation (691) of the coefficients a_n and b_n imply the bounds

$$a_n \leq \frac{\Lambda_1}{R^n} \qquad b_n \leq \frac{\Lambda_1}{R^n} \quad (701)$$

Using these bounds

$$|c_n| = |\sum_{m=0}^{n-1} \frac{(a_{n-m}(m+r) + b_{n-m})c_m)}{\lambda(n+r_+)}| \leq$$

$$|\sum_{m=0}^{n-1} \frac{(\Lambda_1(m+r) + \Lambda_2))}{R^{n-m}\lambda(n+r_+)}||c_m| \tag{702}$$

The denominator can be bounded by noting that

$$|\lambda(n+r_+)| = n^2 + n(2r_+ - 1 + a_0) \tag{703}$$

in addition, solving the indicial equation gives

$$2r_\pm + a_0 - 1 = \pm\sqrt{(1-a_0)^2 - 4b_0} \tag{704}$$

where one of the roots has positive real part which means

$$2r_+ + a_0 - 1 \geq 0 \tag{705}$$

It follows that

$$|\lambda(n+r_+)| = n^2 + n(2r_+ - 1 + a + 0) < n^2 \tag{706}$$

which gives

$$|c_n| = |\sum_{m=0}^{n-1} \frac{(a_{n-m}(m+r) + b_{n-m})c_m)}{\lambda(n+r_+)}| \leq$$

$$|\sum_{m=0}^{n-1} \frac{(\Lambda_1(m+r_+) + \Lambda_2))}{n^2 R^{n-m}}||c_m| \tag{707}$$

giving

$$|R^n c_n| \leq \frac{K}{n}|\sum_{m=0}^{n-1}||R^m c_m| \leq K|R^m c_m| \tag{708}$$

where

$$K \geq (\Lambda_1(1+r_+) + \Lambda_2)) \geq \frac{(\Lambda_1(m+r_+) + \Lambda_2))}{n} \tag{709}$$

By induction assume that

$$R^m|c_m| \leq K^m F_0 \tag{710}$$

then it follows that

$$|R^n c_n| \leq F_0 \frac{K}{n} | \sum_{m=0}^{n-1} ||K^m| = \tag{711}$$

$$|R^n c_n| \leq F_0 K^n \frac{1}{n} ||1 + \frac{1}{K} + \cdots + \frac{1}{K^{n-1}}| \tag{712}$$

We can always choose $K > 1$ which gives

$$|R^n c_n| \leq F_0 K^n \tag{713}$$

It follows that

$$\langle z|u\rangle = (z - z_0)^r f(z) \tag{714}$$

$$|f(z)| \leq \sum_{n=0}^{\infty} |c_n (z - z_0)^n| \tag{715}$$

$$\sum_{n=0}^{\infty} |F_0 \frac{K(z - z_0)}{R}|^n \tag{716}$$

which converges whenever

$$|\frac{K(z - z_0)}{R}| < 1 \tag{717}$$

By starting with the root with the largest real part we ensured that $\lambda(n + r_+)$ never vanished. This also holds for the root with the smaller real part provided the two roots *do not differ by an integer*. In this case the second solution is determined by the same procedure as the above solution, but starting with the second solution of the indicial equation. The convergence proof is unchanged.

In the case the that two roots of the indical equation differ by an integer the series method only works for the root with the larger real part. The other solution can be constructed by writing the second solution in the form

$$u_2(z) = h(z) u_1(z) \tag{718}$$

where previously we found

$$\frac{dh}{dz} = \frac{c}{u_1(z)^2} e^{\int_{z_0}^{z} p(z') dz'} \tag{719}$$

Using the series expansion for $p(z)$ and $u_+(z)$

$$p(z) = \frac{1}{z - z_0} \sum_{n=0}^{\infty} a_n (z - z_0)^n \tag{720}$$

$$u_+(z) = (z - z_0)^{r_+} \sum_{n=0}^{\infty} c_n (z - z_0)^n \tag{721}$$

$$\frac{dh}{dz} = \frac{c}{((z - z_0)^{r_+} \sum_{n=0}^{\infty} c_n (z - z_0)^n)^2} e^{-a_0 \ln(z - z_0) - \sum_{n=1}^{\infty} \frac{1}{n} a_n (z - z_0)^n} \tag{722}$$

where we have absorbed an infinite constant in c, This does not change the z dependence.

The function $\frac{dh}{dz}$ has the general form

$$\frac{dh}{dz} = \frac{c}{(z - z_0)^{2r_+ + a_0}} F(z) \tag{723}$$

where $F(z)$ is analytic near $z = z_0$ provided that $c_0 \neq 0$.

We note that since $r_- = r^+ - N$ where N is a positive integer we have

$$2r_+ + a_0 = 2r_- + a_0 + 2N \tag{724}$$

Recall the indicial equation has the form

$$r(r-1) + a_0 r + b_0 = 0 \tag{725}$$

Subtracting the equation with each root gives

$$(r_+)^2 - (r_-)^2 + (a_0 - 1)(r_+ - r_-) = 0 \tag{726}$$

$$r_+ + r_- + a_0 - 1 = 0 \tag{727}$$

$$r_+ + r_+ - N + a_0 - 1 = 2r_+ + a_0 - N - 1 \tag{728}$$

Using this in the representation (**??**) gives

$$\frac{dh}{dz} = \frac{c}{(z - z_0)^{N+1}} F(z) \tag{729}$$

It follows that if

$$F(z) = \sum_{n=1}^{\infty} f_n (z - z_0)^n \tag{730}$$

then

$$h(z) = cf_N \ln(z - z_0) + (z - z_0)^{-N} \sum_{n=1,\neq N}^{\infty} \frac{cf_n}{n - N}(z - z_0)^n \tag{731}$$

is a solution of the above equation. We finally get the form of the general solution

$$u_-(z) = u_+(z)[d_0 \ln(z - z_0) + (z - z_0)^{-N} G(z)] \tag{732}$$

where $G(z)$ is analytic. This can be solved by the series method by expanding $u_+(z)$ and $G(z)$

0.26 Lecture 26

In this section I consider a class of differential equations with regular singular points. These are called Fuchsian equations. We limit our considerations to a class of Fuchsian equations with at most three singular points in the entire complex plane - including a possible singular point at infinity. This type of equation was studied by Riemann and includes a large class of equations that are of great importance in physics.

It will be useful to write these equations in the following form

$$\frac{d^2u}{dz^2} + \left(\frac{1-\alpha-\alpha'}{z-z_1} + \frac{1-\beta-\beta'}{z-z_2} + \frac{1-\gamma-\gamma'}{z-z_3} + \right) \frac{du}{dz} +$$

$$\left[\frac{(z_1-z_2)(z_1-z_3)\alpha\alpha'}{z-z_1} + \frac{(z_2-z_1)(z_2-z_3)\beta\beta'}{z-z_2} + \frac{(z_3-z_2)(z_3-z_1)\gamma\gamma'}{z-z_3} \right] \times$$

$$\frac{u(z)}{(z-z_1)(z-z_2)z-z_3)} \tag{733}$$

where

$$\alpha + \alpha' + \beta + \beta' + \gamma + \gamma' = 1 \tag{734}$$

The equation is put in this form because the indicial equations at each singular point are

$$r(r-1) + (1-\alpha-\alpha')r + \alpha\alpha' = 0 \tag{735}$$

$$r(r-1) + (1-\beta-\beta')r + \beta\beta' = 0 \tag{736}$$

$$r(r-1) + (1-\gamma-\gamma')r + \gamma\gamma' = 0 \tag{737}$$

which have roots α, α', β, β', and γ, γ'.

Solutions of this equation are is represented by the symbol

$$P \left\{ \begin{array}{cccc} z_1 & z_2 & z_3 & \\ \alpha & \beta & \gamma & z \\ \alpha' & \beta' & \gamma' & \end{array} \right\} \tag{738}$$

The form of this function can be represented will have a different form when z is in a neighborhood of each of the three singular points. This symbol has nine parameters.

If the symbol is multiplied by

$$(z - z_1)^r (z - z_2)^s (z - z_3)^t \tag{739}$$

$$(z - z_1)^r (z - z_2)^s (z - z_3)^t P \left\{ \begin{array}{cccc} z_1 & z_2 & z_3 & \\ \alpha & \beta & \gamma & z \\ \alpha' & \beta' & \gamma' & \end{array} \right\} =$$

$$\left(\frac{z - z_1}{z - z_3} \right)^r \left(\frac{z - z_2}{z - z_3} \right)^s P \left\{ \begin{array}{cccc} z_1 & z_2 & z_3 & \\ \alpha & \beta & \gamma & z \\ \alpha' & \beta' & \gamma' & \end{array} \right\} =$$

$$\left(\frac{z - z_2}{z - z_1} \right)^s \left(\frac{z - z_3}{z - z_1} \right)^t P \left\{ \begin{array}{cccc} z_1 & z_2 & z_3 & \\ \alpha & \beta & \gamma & z \\ \alpha' & \beta' & \gamma' & \end{array} \right\} =$$

$$\left(\frac{z - z_1}{z - z_2} \right)^r \left(\frac{z - z_3}{z - z_2} \right)^t P \left\{ \begin{array}{cccc} z_1 & z_2 & z_3 & \\ \alpha & \beta & \gamma & z \\ \alpha' & \beta' & \gamma' & \end{array} \right\} =$$

$$P \left\{ \begin{array}{cccc} z_1 & z_2 & z_3 & \\ \alpha + r & \beta + s & \gamma + t & z \\ \alpha' + r & \beta' + s & \gamma' + t & \end{array} \right\} \tag{740}$$

The proof of this result will be a homework exercise. This shows that given one solution of the differential equation corresponding to one set of the nine parameters (eight independent) we can relate it to a special class of solutions having only six independent parameters. This is because we are free to choose the parameters r, s

If we use the homographic transformations

$$z' = \frac{Az + B}{Cz + D} \qquad z'_i = \frac{Az_i + B}{Cz_i + D} \tag{741}$$

it is possible to show

$$P\left\{\begin{array}{cccc} z_1 & z_2 & z_3 & \\ \alpha & \beta & \gamma & z \\ \alpha' & \beta' & \gamma' & \end{array}\right\} = P\left\{\begin{array}{cccc} z_1' & z_2' & z_3' & \\ \alpha & \beta & \gamma & z' \\ \alpha' & \beta' & \gamma' & \end{array}\right\} \tag{742}$$

This will also be verified in the homework. The result means that the six parameters can be reduced to three by utilizing the freedom to use homographic transformations to move the positions of the singularities.

Using this freedom we choose

$$z_1 = 0, \qquad z_2 = \infty, \quad z_3 = 1 \tag{743}$$

$$r = -\alpha, \qquad s = \alpha + \gamma, \quad t = -\gamma \tag{744}$$

which means that the general Riemann P-symbol can be expressed in terms of the special P-symbol

$$P\left\{\begin{array}{cccc} 0 & \infty & 1 & \\ 0 & \alpha+\beta+\gamma & 0 & z \\ \alpha'-\alpha & \beta'+\alpha+\gamma & \gamma'-\gamma & \end{array}\right\} \tag{745}$$

We define

$$1 - c := \alpha' - \alpha \qquad b := \beta' + \alpha + \gamma \qquad c - a - b := \gamma' - \gamma \tag{746}$$

$$P\left\{\begin{array}{cccc} z_1 & z_2 & z_3 & \\ \alpha & \beta & \gamma & z \\ \alpha' & \beta' & \gamma' & \end{array}\right\} = \tag{747}$$

$$(z-z_1)^{-\alpha}(z-z_2)^{\alpha+\gamma}(z-z_3)^{-\gamma} P\left\{\begin{array}{cccc} 0 & \infty & 1 & \\ 0 & a & 0 & \frac{Az+B}{Cz+D} \\ 1-c & b & c-a-b & \end{array}\right\} \tag{748}$$

where

$$0 = \frac{Az_1 + B}{Cz_1 + D} \tag{749}$$

$$\infty = \frac{Az_2 + B}{Cz_2 + D} \tag{750}$$

$$1 = \frac{Az_3 + B}{Cz_3 + D} \tag{751}$$

can be solved to give

$$\frac{B}{A} = -z_1 \qquad \frac{D}{C} = -z_2 \qquad \frac{A}{D} = \frac{(z_2 - z_3)}{z - 2(z_3 - z - 1)} \tag{752}$$

$$z' = \left(\frac{z - z_1}{z - z_2}\right)\left(\frac{z_3 - z_2}{z_3 - z_1}\right). \tag{753}$$

The function

$$P\left\{\begin{array}{cccc} 0 & \infty & 1 & \\ 0 & a & 0 & z \\ 1-c & b & c-a-b & \end{array}\right\} \tag{754}$$

is a solution to the hypergeometric equation.

$$\frac{d^2u}{dz^2} + \left(\frac{c}{z} + 0 + \frac{1-c+a+b)}{z-1}\right)\frac{du}{dz} +$$

$$\frac{ab}{z(z-1)}u(z) \tag{755}$$

It is useful to multiply this equation by $z(1-z)$ to get

$$z(1-z)\frac{d^2u}{dz^2} + (z(1+a+b) - c)\frac{du}{dz} + abu = 0. \tag{756}$$

The specific series solution about the origin, corresponding to the root $\alpha = 0$ of the indicial equation, with value 1 at the origin is called the Hypergeometric function, which is denoted by $F(a, b, c, z)$ Because $\alpha = 0$ it is analytic about the origin thus can be expressed in the form

$$F(a, b, c, z) = \sum_{n=0}^{\infty} c_n z^n \qquad c_0 = 1 \tag{757}$$

Inserting this expansion in the differential equation above give

$$\sum_{n=0}^{\infty} n(n-1)(z^n - z^{n-1})c_n + \sum_{n=0}^{\infty} n(z^n(1+a+b) - cz^{n-1})c_n +$$

$$\sum_{n=0}^{\infty} abz^n c_n = 0 \tag{758}$$

Shifting to get common powers of z gives

$$\sum_{n=0}^{\infty} (n(n-1)c_n - (n+1)nc_{n+1} + n(1+a+b)c_n - c(n+1)c_{n+1} +$$

$$abc_n)\, z^n = 0 \tag{759}$$

which leads to the following relation between c_{n+1} and c_n:

$$c_{n+1}\left((n+1)(n+c)\right) = \left(n^2 + n(a+b) + ab\right) c_n \tag{760}$$

which gives the simple relationship

$$c_n = \frac{(n+a)(n+b)}{(n+1)(n+c)} c_n \tag{761}$$

This can be iterated to get

$$c_n = \frac{a(a+1)\cdots(n+a-1)b(b+1)\cdots(n+b-1)}{n!c(c+1)\cdots(n+c-1)} c_0 \tag{762}$$

$$a(a+1)\cdots(n+a-1) = \frac{\Gamma(n+a)}{\Gamma(a)} \tag{763}$$

$$b(b+1)\cdots(n+b-1) = \frac{\Gamma(n+b)}{\Gamma(b)} \tag{764}$$

$$c(c+1)\cdots(n+c-1) = \frac{\Gamma(n+c)}{\Gamma(c)} \tag{765}$$

Which gives the following expression for the hypergeometric function

$$F(a,b,c,z) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{n=0}^{\infty} \frac{\Gamma(n+a)\Gamma(n+b)}{n!\Gamma(n+c)} z^n \tag{766}$$

This series converges for $|z| < 1$ since the next singular point is at 1.

This function has a number of useful properties. They are derived in the text.

$$z^{1-c}F(b-c+1, a-c+1, 2-c, z)$$

satisfies the same differential equation as $F(a, b, c; z)$ in a neighborhood of $z = 0$. It gives the second independent solution to the Hypergeometric equation near $z = 0$.

Note that we can write equation (748), replacing the P-symbol on the right by the Hypergeometric function of the appropriate parameters to get

$$P \left\{ \begin{array}{cccc} z_1 & z_2 & z_3 & \\ \alpha & \beta & \gamma & z \\ \alpha' & \beta' & \gamma' & \end{array} \right\} = \tag{767}$$

$$\left(\frac{z - z_1}{z - z_2}\right)^{\alpha} \left(\frac{z - z_3}{z - z_2}\right)^{\gamma} \times$$

$$F(\alpha + \beta + \gamma, \alpha + \beta' + \gamma, 1 + \alpha - \alpha'; \frac{(z - z - 1)(z_3 - z_2)}{(z - z_2)(z_3 - z_1)} \tag{768}$$

Note that the differential equation is invariant with respect to the 3! permutations of the columns of the Riemann P symbol. On the other hand the right side of the above equation is not symmetric with respect to permutations of the columns.

The equation is also symmetric with respect to $\alpha \leftrightarrow \alpha'$, $\beta \leftrightarrow \beta'$ $\gamma \leftrightarrow \gamma'$, while the right side of (ref) is only symmetric with respect to $\beta \leftrightarrow \beta'$.

Using these properties it is possible to relate both solutions in a neighborhood of each singular point to the hypergeometric functions. The following solutions can be derived using the above symmetries, of simply checked by substituting them into the differential equation, using properties of the Hypergeometric functions.

The two solutions near the singular point $z = 1$ have the form

$$F(a, b, a + b + 1 - c, 1 - z) \tag{769}$$

$$(1 - z)^{c-a-b} F(c - b, c - a, 1 + c - a - b, 1 - z) \tag{770}$$

while the two solutions near the singular point at infinity have the form

$$z^{-a} F(a, a - c + 1, a - b + 1, \frac{1}{z}) \tag{771}$$

$$z^{-b} F(b, b - c + 1, b - a + 1, \frac{1}{z}) \tag{772}$$

This shows that all of the solutions of the equation of Riemann are related to the Hypergeometric function.

2. The series solution to the hypergeometric equation has a limited radius of convergence. The domain of analyticity can be extended using integral representations. Some of these representations are

$$F(a,b,c,z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_1^\infty (t-z)^{-a}(t^{a-c}(t-1)^{c-b-1}dt \quad (773)$$

$$F(a,b,c,z) = \frac{\Gamma(c)}{\Gamma(c-b)\Gamma(b)} \int_0^1 (1-tz)^{-a}t^{b-1}(1-t)^{c-b-1}dt \quad (774)$$

These can be checked by direct substitution into the differential equation. There are many other representations that are analytic in different regions.

Some special cases of Hypergeometric functions are

$$F(-a,b,b;-z) = (1+z)^a \quad (775)$$

$$F(1,1,2;-z) = \frac{1}{z}\ln(1+z) \quad (776)$$

$$F(-\lambda,\lambda+\alpha+\beta+1,\alpha+1,\frac{1-z}{2}) = \frac{\Gamma(\lambda+1)\Gamma(\alpha+1)}{\Gamma(\lambda+\alpha+1)}P_\lambda^{(\alpha,\beta)}(z) \quad (777)$$

$$F(\lambda+1,\lambda+\alpha+1,2\lambda+\alpha+\beta+2,\frac{2}{1-z}) =$$

$$\frac{\Gamma(2\lambda+\alpha+\beta+2)(z-1)^{\lambda+\alpha+1}(z+1)^\beta}{\Gamma(\lambda+\alpha+1)\Gamma(\lambda+\beta+1)2^{\lambda+\alpha+\beta}}Q_\lambda^{(\alpha,\beta)}(z) \quad (778)$$

which are the Jacobi functions of the first and second kind.

There are a number of special cases

$$C_\lambda^{\alpha+1/2}(z) = \frac{\Gamma(2\alpha+\lambda+1)\Gamma(\alpha+1)}{\Gamma(2\alpha+1)\Gamma(\lambda+\alpha+1)}P_\lambda^{(\alpha,\alpha)}(z) \quad (779)$$

$$Q_\lambda(z) = Q_\lambda^{00}(z) = \frac{2^\lambda\Gamma(1+\lambda)^2}{\Gamma(2\lambda+2)(z-1)^{1+\lambda}}F(\lambda+1,\lambda+1,2\lambda+2,\frac{2}{1-z} \quad (780)$$

The Riemann equation

$$\frac{d^2f}{dz^2}+\left(\frac{c}{z}+\frac{1-a-b}{z-z_2}+\frac{1-c-a+b}{z-z_3}\right)\frac{df}{dz}+\frac{abz_2(z_2-z_3)}{z(z-z_2)^2(z-z_3)}f=0 \quad (781)$$

has regular singular points are $z = 0$ and $z = z_2, z_3$. Set

$$z_2 = b = 2\alpha \qquad z_3 = \alpha \tag{782}$$

and let $\alpha \to \infty$ keeping a and c fixed. The resulting equation is

$$z\frac{d^2 f}{dz^2} + (c - z)\frac{df}{dz} - af = 0 \tag{783}$$

In this limit the distinct singular points at z_2 and z_3 are both transformed into singular points at infinity. In this limit the point $z = \infty$ is no longer a regular singular point - it is an irregular singular point.

This equation has a solution that is analytic near zero and 1 at zero. This can be constructed using the series method

$$\Phi(a, c; z) = \sum_{n=0}^{\infty} f_n z^n \tag{784}$$

Using this in the differential equation gives

$$\sum_{n=0}^{\infty} n(n-1) f_n z^{n-1} + \sum_{n=0}^{\infty} cn f_n z^{n-1} - \sum_{n=0}^{\infty} n f_n z^n - \sum_{n=0}^{\infty} a f_n z^n = 0 \tag{785}$$

Finding common powers of z gives

$$\sum_{n=0}^{\infty} \left(n + 1(n) f_{n+1} + c(n+1) f_{n+1} - n f_n - a f_n\right) z^n \tag{786}$$

which leads to the recursion

$$f_{n+1} = \frac{a + n}{(n+1)(n+c)} f_n \tag{787}$$

If we normalize this to 1 at $z = 0$ the series gives

$$\Phi(a, c; z) = \frac{\Gamma(c)}{\Gamma(a)} \sum_{n=0}^{\infty} \frac{\Gamma(a+n)}{\Gamma(c+n)\Gamma(n+1)} z^n \tag{788}$$

This is called the Confluent Hypergeometric function. This series is valid when c is not zero or a negative integer. When the series makes sense it is analytic in the complex plane.

Comparing the series for the Hypergeometric function shows that

$$\Phi(a, c; z) = \lim_{bto\infty} F(a, b, c; z/b) \tag{789}$$

When c is not an integer a second linearly independent solution of equation (**??**) is

$$z^{1-c}\Phi(a - c + 1, 2 - c, z) \tag{790}$$

It is conventional to define

$$\Psi(a, c; z) := \frac{\Gamma(1-c)}{\Gamma(a-c-1)}\Phi(a, c, z) + \frac{\Gamma(c-1)}{\Gamma(a)}z^{1-c}\Phi(a-c+1, 2-c, z) \tag{791}$$

There are a number of special functions that are special cases of the confluent Hypergeometric function.

Some of the more familiar ones are

$$H_n(\frac{z}{\sqrt{2}}) = 2^n\Phi(-\frac{n}{2}, \frac{1}{2}, \frac{z^2}{2}) \tag{792}$$

$$L_n^{\mu}(z) = \frac{\Gamma(n+\mu+1)}{\Gamma(n+1)\Gamma(\mu+1)}\Phi(-n, \mu+1, z) \tag{793}$$

$$Erf(z) = z\Phi(\frac{1}{2}, \frac{3}{2}, -z^2) \tag{794}$$

$$J_\nu(z) = \frac{1}{\Gamma(\nu+1)}\left(\frac{z}{2}\right)^{\nu} e^{-iz}\Phi(\nu + \frac{1}{2}, 2\nu + 1, 2iz) \tag{795}$$

Other related functions are

- Kelvin functions
- Coulomb Wave functions
- Incomplete Gamma Function
- Weber functions
- Error integral
- Airy functions
- Si and Ci functions
- Bateman functions
- Cunningham functions

0.27 Lecture 27

In this section I give a brief introduction to group theory, representation theory, and tensors.

Definition A group is a set G with a multiplication operator $\cdot$ with the following properties

1. $a, b \in G \Rightarrow a \cdot b \in G$
2. $\exists\, e \in G$ such that $\forall a \in G \quad e \cdot a = a \cdot e = a$. The unique element e is called the identity element of the group.
3. $\forall\, a \in G \quad \exists b$ satisfying $a \cdot b = b \cdot a = e$. The unique element b is called the inverse of a.
4. $\forall\, a, b, c \in G \quad (a \cdot b) \cdot c = a \cdot (b \cdot c)$.

There are many examples of groups.

Example 1: The set of permutations of 4 objects with the group product being composition of permutations.

$$p_1 := \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix} \qquad p_2 := \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix}$$

$$p_1 \cdot p_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$$

Example 2: The real numbers with the group product being addition.

Example 3: Invertible $N \times N$ matrices with the group product being matrix multiplication.

Definition A subgroup H of a group G consists of a subset of the elements of G that

1. Is closed under group multiplication
2. $a \in H \rightarrow a^{-1} \in H$.

Example 4: The group of integers under addition is a subgroup of the real numbers under addition.

Example 5: The group of $N \times N$ matrices with determinant 1 is a subgroup of the invertible matrices.

Definition A group G is Abelian $a \cdot b = b \cdot a$ for every $a, b \in G$. A group that is not-Abelian is called non-Abelian.
Definition Two groups G_1 and G_2 are isomorphic there is a 1 to 1 correspondence between the elements of the two groups that preserves the group product:

$$g_1 \leftrightarrow g_2, \quad h_1 \leftrightarrow h_2, \quad \Rightarrow g_1 \cdot_1 h_1 \leftrightarrow g_2 \cdot_2 h_2$$

Definition A group G is cyclic if all elements of G have the form g^m for some $g \in G$ (by definition $g^0 = e$).
Definition Let H be a subgroup of G. The elements of G of the form hg where g fixed in G has h is any element of H is called a right coset of H in G. The elements of G of the form gh where g is fixed in G and h is any element of H is called a left coset of H in G.
Theorem: The right (left) cosets divide the elements of G into disjoint equivalence classes.
Definition A subgroup H of G is normal if the left and right cosets are identical.
Theorem: The cosets of a normal subgroup H of G form a group. This group is called the quotient group, written G/H

Recall that linear operators on vector spaces can be represented by matrices.
Definition A linear representation of a group G on a vector space V is a mapping from the group G to the linear transformations on V satisfying

$$g \to D(g)$$

with the property

$$D(g_1 \cdot g_2) = D(g_1)D(g_2)$$

for all $g_i \in G$. In this expression $D(g_1)D(g_2)$ is matrix multiplication.
Definition Two representations $D_1(g)$ and $D_2(g)$ of a group G are equivalent if they are related by a similarity transformation S

$$D_1(g) = SD_2(g)S^{-1} \tag{796}$$

Definition A representation $D(g)$ is reducible if it is equivalent to a block diagonal representation

$$SD(g)S^{-1} = \begin{pmatrix} D_1(g) & 0 & \cdots \\ 0 & D_2(g) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \tag{797}$$

Definition A representation $D(g)$ is irreducible if it is not reducible.
Theorem: (Schur's Lemma) Any matrix that commutes with all elements of an irreducible representation is necessarily a multiple of the identity.

Proof: Let $Ax = \lambda x$. The $x' = D(g)x$ is also an eigenvector of A for any $g \in G$. Thus the subspace spanned by the eigenvectors of A with eigenvalue λ is an invariant subspace of $D(G)$. If $D(G)$ irreducible the only invariant subspace is whole space, which means that $A = \lambda I$

A Lie group is group where the group elements are analytic functions of some parameters and where the group multiplication law depends analytically on the parameters. Lie groups are some of the most important groups in physics.

Examples of Lie groups are
Example 1: $U(1)$ group of unitary matrices in 1 dimension.
Example 2: $U(N)$ group of unitary matrices in N dimension.
Example 3: $SU(N)$ group of unitary matrices in N dimension with determinant 1.
Example 4: $O(N)$ group of real orthogonal matrices in N dimension with determinant 1.
Example 5: $SO(M, N)$ group of real matrices in $N + M$ dimension with determinant 1 that preserve a diagonal metric of the form

$$\eta = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & 0 \\ 0 & 0 & \cdots & 0 & -1 \end{pmatrix} \tag{798}$$

with M factors of 1 and N factors of -1.

Tensors are associated with group representations. Consider a set of linear transformations on an N dimensional vector space that preserves the following quadratic from

$$Q(\mathbf{v}, \mathbf{w}) := \sum_{i=1,j}^{N} v_i^* M_{ij} w_j \tag{799}$$

where $w_j = \langle j|w\rangle$, where $v_i^* = \langle v|i\rangle$, and

$$M_{ij} := \langle i|M|j\rangle \tag{800}$$

where M_{ij} is understood to be a fixed invertible matrix. In this representation linear transformations G_{ij} that satisfy

$$M_{ij} = G^{\dagger}_{im} M_{mn} G_{nj} \tag{801}$$

leave $Q(\mathbf{v}, \mathbf{w})$ invariant.

It is a simple exercise to show that this set of transformations is a group under matrix multiplication. This is like a generalization of the unitarity condition where the identity is replaced by an arbitrary but fixed matrix.

In terms of matrix multiplication

$$Q(\mathbf{v}, \mathbf{w}) = \mathbf{v}^{\dagger} M \mathbf{w}$$

If I define

$$\mathbf{u}^t := \mathbf{v}^{\dagger} M$$

then the invariant quadratic form becomes

$$Q(\mathbf{v}, \mathbf{w}) = \mathbf{u}^t \mathbf{w}.$$

This is invariant under the combined transformations

$$\mathbf{w} \to \mathbf{w}' = G_{\mathbf{w}} \tag{802}$$

$$\mathbf{u}' = (M^t G^* (M^t)^{-1} \mathbf{u} = G_D \mathbf{u} \tag{803}$$

There are quantities in nature the transform like product of vectors under the action of a group of transformations. A Rank (m, n) Tensor is a multi-index quantity that transforms like

$$T_{a_1 \cdots a_m ; b_1 \cdots b_n} \to T'_{a_1 \cdots a_m ; b_1 \cdots b_n} G_{a_1 a'_1} \cdots G_{a_m a'_m} G_{D b_1 b'_1} \cdots G_{D b_n b'_n} T_{a'_1 \cdots a'_m ; b'_1 \cdots b'_n} \tag{804}$$

We see that there are two kinds of indices - sometimes called covariant and contravariant indices.

The transformation matrices are symmetric with respect the interchanges of the contravariant and covariant indices are interchanged among themselves. This means that each set of indices (covariant) (contravariant) can be chosen to have a specific symmetry with respect to permutations, which is preserved under the group transformation law.

If a covariant and contravariant component are summed the resulting quantity is a tensor of rank $(m-1, n-1)$. A tensor of rank 0 is a scalar with respect to the group transformation law.

A useful property of tensors is that if T_1 and T_2 are tensors of the same rank, and $T_1 = T_2$ for one value of $g \in G$, then it holds for all values.

For example, in special relativity, if we demand the Newton's second law holds in the rest frame of a particle, and if the right and left side of the equation are assumed to be rank (1,0) tensors with respect to the Lorentz group, then the transformed equation gives the generlaization of these equations in any frame.

Examples of tensors are:

29:172 Assignment 1 - Due Wed. Jan. 24

1.) Consider the space of continuous complex-valued functions on the interval $[a, b]$ with inner product

$$\int_a^b f^*(x)g(x)dx$$

a. Show that the sum of two Cauchy sequences is a Cauchy sequence.

b. Show that if two different Cauchy sequences converge to the same function that the difference of these sequences converges to zero.

2.) Consider the functions $\{f_n(x)\}$ defined by

$$f(x) = \begin{cases} 0 & : \quad x < -1 - 1/n \\ n + 1 + nx & : \quad -1 - 1/n < x < -1 \\ 1 & : \quad -1 \leq x \leq 1 \\ n + 1 - nx & : \quad 1 < x < 1 + 1/n \\ 0 & : \quad x \geq 1 + 1/n \end{cases}$$

Show that this sequence is a Cauchy sequence. Show that it converges to the discontinuous block function

$$b(x) = \begin{cases} 0 & : \quad |x| > 1 \\ 1 & : \quad |x| \leq 1 \end{cases}$$

in the $L_2(\mathbb{R})$ norm.

3.) Estimate the Lebesgue integral of

$$\int_0^1 x^2 dx$$

by dividing the range of this function into 10 equally spaced intervals between 0 and 1. Find upper and lower bounds for the integral by using the largest or smallest value of the function on each interval. Compare this to the exact value.

4.) Show that the set of rational numbers (numbers that can be expresses as ratios of integers m/n) between 0 and 1 are a set of Lebesgue measure zero.

5.) Consider the set of rational numbers between zero and one. Show that it is possible to place each rational in the interior of an open interval (i.e. $a < (m/n) < b$) in such a way that if we discard all rationals *and* all of the open intervals containing these rationals that what remains in the interval $[0, 1]$ has a measure as close to 1 as desired. [This exercise was a critical element in Kolomogorov, Arnold, and Moser's solution to the classical three-body problem which asks whether the solar system is stable]

6.) Let S_1 and S_2 be subsets of a larger set S. For a subset S' of S let S'^c the the complement of S' in S. This means the set of points in S that are not in S'.

a. Show
$$S_1 \cap S_2 = (S_1^c U S_2^c)^c$$

b. Show how does this show that the intersection of two Lebesgue measurable sets are is measurable?

c. Show that every closed interval, $[a, b]$ of the real line is Lebesgue measurable. Find the Lebesgue measure of this set?

29:172 Assignment 2 - Due Wednesday, Jan. 31

1.) Consider the function $f(x) = (1-x^2)$ on the interval $[-1,1]$. Calculate a 4-th degree Weierstrass polynomial approximation to this function. Compare the exact and approximate functions.

2.) Use the Rodrigues formula to calculate the first four Hermite polynomials.

3.) Use the Gram Schmid method to calculate the first three orthogonal polynomials on $[-1,1]$ with weight $w(x) = 1$. Compare these to first three Legendre polynomials generated

4. Assume that $|f(x) - p_n(x)| < \epsilon$ for all $x \in [a,b]$, where $p_n(x)$ is a polynomial. Let M be a 2×2 Hermitian matrix with real eigenvalues λ satisfying $a < \lambda < b$. Show that

$$|||(f(M) - p_n(M))||| < \epsilon$$

where $|||O|||$ is the matrix or operator norm of O.

29:172 Assignment 3 - Due Wednesday, Feb. 7

1.) Use the Rodrigues formula for the Laguerre polynomials to derive the constants k_n, k'_n on page 212 of the text.

2.) Let

$$g(x,t) = \frac{1}{\sqrt{1-2xt+t^2}}. \tag{805}$$

Show that

$$g(x,t) = \sum_{n=0} \infty t^n P_n(x) \tag{806}$$

where $P_n(x)$ is the $n-th$ degree Legendre polynomial. The function $g(x,t)$ is called a generating function.

3.) The Schrödinger equation for a harmonic oscillator has the form

$$(-\frac{d^2}{dx^2} + x^2)F_n(x) = (2n+1)F_n(x). \tag{807}$$

Assume that

$$F_n(x) = C_n(x)e^{-\frac{1}{2}x^2}. \tag{808}$$

Show that C_n satisfies the same differential equation as one of the classical orthogonal polynomials.

4.) Find the Gauss Legendre points and weights for a two point Gauss Legendre quadrature on $[-1,1]$. Show that

$$\int_{-1}^{1} x^n dx = \sum_{i=1}^{2} x_i^n w_i \tag{809}$$

is exact or $n = 0,1,2,3$. How good is the integral for x^6?

5. Use the recursion relations to generate the first four Laguerre Polynomials with $\nu = 0$.

6. Show that the Legendre Polynomials satisfy the following recursion relations involving first derivatives

$$(n+1)P_{n+1}(x) = (n+1)xP_n - (1-x^2)\frac{dP_n}{x}. \tag{810}$$

29:172 Assignment 4 - Due Wednesday, Feb. 14

1.) Fourier Series: In class I developed the Fourier series for periodic functions on the interval $[-\pi, \pi]$. I constructed the orthonormal basis functions

$$\langle\theta|n\rangle = \frac{1}{\sqrt{2\pi}} e^{i\theta n}$$

a. Find the corresponding orthonormal basis functions for periodic functions on the interval $[0, L]$?

b. If $f(\theta)$ is a real valued periodic function, how can you recognize that it is a real functions by looking at the expansion coefficients:

$$f(\theta) = \sum_n f_n \langle\theta|n\rangle$$

2.) Calculate $\int_0^\infty \frac{\sin(ax)}{x} dx$ for $a > 0$.

3.) Calculate the Fourier Transform of e^{-ax^2}

4.) Let $f(x)$ and $g(x)$ have Fourier Transforms.

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int dy e^{iky} f(y) \qquad \tilde{g}(k) = \frac{1}{\sqrt{2\pi}} \int dy e^{iky} g(y)$$

Find an expression for the Fourier transform of the product $f(x)g(x)$ in terms of their individual Fourier transforms, $\tilde{f}(k)$ and $\tilde{g}(k)$,

5.) Show

$$\lim_{\lambda \to \infty} \int_{-\infty}^{\infty} e^{i\lambda x} f(x) dx = 0$$

if f is absolutely integrable and differentiable for every x.

6.) Let $f(\theta)$ be 1 for $0 < \theta < \pi$ and -1 for $-\pi < \theta \leq 0$. Find the coefficients f_n in the Fourier series

$$f(\theta) = \sum_n f_n \langle\theta|n\rangle$$

29:172 Assignment 5 - Due Wednesday, Feb. 21

1.) Delta functions. Show

$$\delta(x-x_0) = \lim_{\epsilon\to 0+} \frac{1}{\pi}\frac{\epsilon}{(x-x_0)^2+\epsilon^2}$$

$$\delta(x-x_0) = \lim_{\lambda\to\infty} \frac{1}{\pi}\frac{1-\cos(\lambda(x-x_0}{\lambda(x-x_0)^2}$$

$$\delta(x-x_0) = \lim_{\lambda\to 0} \frac{\chi_\lambda(x-x_0)}{2\lambda}$$

where $\chi_\lambda(x)$ is 1 when $-\lambda < x < \lambda$ and zero otherwise.

2.) Prove the following

$$\delta(x) = \delta(-x)$$

$$x\delta(x) = 0$$

$$\delta(ax) = \frac{1}{|a|}\delta(x)$$

$$x\frac{d\delta(x)}{dx} = -\delta(x)$$

$$\delta(f(x)) = \sum_{x_i | f(x_i)=0} \frac{1}{|\frac{df}{dx}(x_i)|}\delta(x-x_i)$$

3.) Show that if $f(x)$ is a continuous function that is identically zero for $|x| > L$, then its Fourier transform is an entire function.

4.) Show that

$$\int \theta(x)\delta(x)f(x)dx$$

does not make sense on the space of Schwartz functions.

5. Calculate the Fourier transforms of

$$f(x) = \frac{1}{(x^2+a^2)(x^2+b^2)}$$

$$f(x) = \begin{cases} 1 & |x| < 1 \\ 0 & |x| \geq 1 \end{cases}$$

$$\theta(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases}$$

6. Show that the linear operator $\frac{d}{dx}$ is *not* bounded on the space of square integrable functions on the real line.

29:172 Assignment 6 - Due Wed, Feb. 28

1. Let A be a bounded operator with the property that A^2 is compact. Under what conditions does

$$(I - A)^{-1}$$

exist. Hint use the Fredholm alternative.

2. Show that if A is compact then $A^\dagger$ is also compact.

3. Let $\{|\chi_n\rangle\}_{n=1}^\infty$ be any orthonormal basis and let K be compact. Consider the finite rank approximations

$$K_N = \sum_{m,n=1}^{N} |\chi_n\rangle\langle\chi_n|K|\chi_m\rangle\langle\chi_m|$$

Show that

$$\||K - K_N\||$$

can be made as small as desired by choosing a large enough N.

4. Show that the integral operator K defined by

$$\langle x|K|f\rangle = \int_{-\infty}^{\infty} \langle x|K|y\rangle dy \langle y|f\rangle$$

with

$$\langle x|K|y\rangle = e^{-x^2-y^2}$$

is compact.

5. The harmonic oscillator Hamiltonian is defined by

$$H := -\frac{d^2}{dx^2} + x^2$$

   a. Show that H is a linear operator.
   b. Show that H is an unbounded operator.

c. H is known to be Hermitian and has a complete set of orthogonal eigenvectors with eigenvalues $2n+1$, $n = 0, 1, 2, \cdots$. The resolvent of H is the operator

$$R(z) = (z - H)^{-1}$$

where z is a complex number. Show that if $z \neq 2n + 1$ that $R(z)$ is compact.

6. It is possible for a unitary operator to be compact? Prove your result.

29:172 Assignment 7 - Due Wednesday, March. 21

1.) Implicit function theorem. Consider the function $w = e^{-(x^2+y^2)}$. We would like to solve this for x near $x = 2$ by constructing

$$x = g(w, y).$$

While this can be done analytically, write

$$w = e^{-(2^2+y^2)} + \frac{\partial e^{-(x^2+y^2)}}{\partial x}|_{x=2}(x-2) + R(x, y)$$

Write this as

$$x = 2 + \frac{1}{\frac{\partial e^{-(x^2+y^2)}}{\partial x}|_{x=2}}[w - e^{-(2^2+y^2)} - R(x, y)]$$

Try to approximate this by iteration

$$x_0 := 2$$

$$x_n := 2 + \frac{1}{\frac{\partial e^{-(x^2+y^2)}}{\partial x}|_{x=2}}[w - e^{-(2^2+y^2)} - R(x_{n-1}, y)]$$

for x near 2. Compare the result of this for a few iterations to the exact result for a selected value of y. The purpose of this exercise is to illustrate how the implicit function theorem works.

2. Consider Newton's equation for a linear harmonic oscillator

$$m\frac{d^2x}{dt^2} = -kx$$

a. Convert this to a system of two coupled first order differential equations.

b. Convert the system of first order linear differential equations to a pair of coupled integral equations.

c. Using the initial conditions, $x(0) = a$ and $\frac{dx}{dt}(0) = 0$, solve the coupled integral equations by successive approximations, using the initial conditions as the first approximation (method used in class to prove the existence of solutions of differential equations). While there are an infinite number of iterations needed to obtain the full solution, this problem is simple enough that you should be able to obtain the exact solution.

3. Consider Legendre's differential equation. If we do not specify boundary conditions then we know that it has two independent solutions. One of the two independent solutions is the Legendre polynomial. Find the second independent solution for the case $n = 1$.

4. Find the solution of the first order differential equation

$$\cosh(x)\frac{df}{dx} + \sinh(x)f(x) = 0$$

5. Find the solution of the first order differential equation

$$\cosh(x)\frac{df}{dx} + \sinh(x)f(x) = \sinh(2x)$$

6. Consider the Wronskian of the differential equation

$$a(x)\frac{d^2f}{dx^2} + b(x)\frac{df}{dx} + c(x)f(x) = 0$$

where $a(x) > 0$. Show that the Wronskian of this equation is non-zero for all x.

29:172 Assignment 8 - Due Wednesday, March. 28

1.) Let $L_x = \frac{d^2}{dx^2}$ on the interval $[a, b]$.

Find the weight $w(x)$ that makes this a hermitian operator.

Find the Green function for this operator corresponding to Dirichlet boundary conditions.

Find the Green function for this operator corresponding to Neumann boundary conditions.

2. Consider the linear differential operator

$$L_x = e^x \frac{d^2}{dx^2} - x^2 \frac{d}{dx} + 1$$

Find the adjoint operator on the interval $[0, 1]$

Find the boundary conditions that are adjoint to Dirichlet boundary conditions.

3. Assume that K is the compact operator

$$K = \sum_{n=0}^{\infty} |p_n\rangle \frac{1}{n+1} \langle p_n|$$

that acts on square integrable functions on the interval $[-1, 1]$. In this expression

$$\langle x|p_n\rangle = P_n(x)\sqrt{n+1/2}$$

where $P_n(x)$ is the $n-th$ Legendre polynomial. Find the Green function for K. Express your answer in terms of Legendre polynomials. Hint, use the abstract identity $KG = I$.

4. Consider the differential equation

$$L_f(x) = \frac{d^2 f(x)}{dx^2} - \omega^2 f(x) = g(x)$$

where ω is real.

Find independent solutions to the homogeneous equation.

Calculate the Wronskian of this system. Verify that it never vanishes.

Find the most general solution to the inhomogeneous equation for a general $g(x)$,

5.) Consider Legendre's differential equation for $n = 1$. Last week you calculated the a second independent solution to the homogeneous equation. Use that solution to find the differences $a_< - a_>$ and $b_< - b_>$

29:172 Assignment 9 - Due Wed, April 4

1.) Find the Green's function for the differential operator

$$L = \frac{d}{dx}[x\frac{d}{dx}] - \frac{1}{x}$$

satisfying Dirichlet boundary condition at $x = 0$ and $x = 1$ (hint - see text)

2.) Let A be a linear operator and let M be the Moore-Penrose generalized inverse of A.

a. Prove that M is unique.

b. Let A be compact. Let M be the Moore Penrose generalized inverse of A. Discuss conditions on A for M to be compact.

3.) Consider the differential operator $L = \frac{d^2}{dx^2}$ on the interval $[0, \pi]$ with periodic boundary conditions.

a. Are there any solutions to the homogeneous equation, $L|f\rangle = 0$ satisfying periodic boundary conditions?

b. Find the generalized Green function for L that satisfies periodic boundary conditions.

c. If we want to solve $L|f\rangle = |g\rangle$, are there any restrictions no $|g\rangle$ in order for a solution to exist.

d. Solve $L|f\rangle = |g\rangle$ for $\langle x|g\rangle = \cos(2x)$.

4.) Consider the differential operator $L = \frac{d^2}{dx^2}$ on the interval $[a, b]$.

Find a solution to the equation $L|f\rangle = 0$ satisfying $\langle x|f\rangle(a) = 1$, $\langle x|f\rangle(b) = -1$.

29:172 Assignment 10 - Due Monday, April. 11

1.) Let M denote a finite dimensional matrix satisfying $|||M||| < c < \infty$. Let $M' = \frac{1}{2c}M$.

a. Show that the Moore-Penrose Generalized inverse of M is $X = \frac{1}{2c}X'$ where X' is the Moore Penrose Generalized inverse of M'

b. Consider the series

$$X' = \sum_{n=0}^{\infty} (I - M^{\dagger\prime} M')^n M^{\dagger\prime}$$

Show that this series converges to the Moore-Penrose generalized inverse of M'

2. Consider the differential operator

$$L = \frac{1}{x}\frac{d^2}{dx^2}x + 1$$

a. Find independent solutions to the homogeneous equation $L|f\rangle = 0$ on the interval $[0, a]$.

b. Find the Green function satisfying Dirichlet boundary on $[0, a]$.

3.) Let L be a Hermitian differential operator with some specified homogeneous boundary conditions. Consider the resolvent operator

$$R(z) = (z - L)^{-1}$$

which is defined when

$$L|f\rangle = z|f\rangle$$

has no solutions satisfying the homogeneous boundary conditions. Let $\gamma(t)$, $t \in [0, 1]$ be a parameterized circle (counter clockwise) in the complex plane with radius r and center at $z = 0$ with the property that $R(z)$ is defined for all z on the curve.

a. Show that

$$P = \frac{1}{2\pi i}\int_0^t R(\gamma(t))\frac{d\gamma}{dt}dt$$

is an orthogonal projection operator.

b. If $L|f\rangle = \eta|f\rangle$, what can you say about $P|f\rangle$?

4.) The Lagrangian for a free particle in a one dimensional box is

$$L = \frac{1}{2}\left(\frac{dx}{dt}\right)^2$$

The solution that is a stationary point of the action functional

$$A[x] = \int_0^T L(\dot{x}(t))dt$$

for $x(0) = a$ and $x(T) = b$ is known to be

$$x_0(t) = a + \frac{b-a}{T}t$$

a. Use the method used in the example in class to determine whether this solution is a minimum (among all curves satisfying $x(0) = a$ and $x(T) = b$) of the action functional or not.

6.) Given a Strum Liouville operator of the form

$$L = \frac{d}{dx}g(x)\frac{d}{dx} - h(x)$$

with weight $w(x) = 1$, $x \in [a, b]$, and $g(x)$ and $h(x)$ real. Show explicitly that the eigenvalues of $L|f_n\rangle = \lambda_n|f_n\rangle$ are real and that eigenvectors corresponding different eigenvalues are orthogonal on $[a, b]$

29:172 Assignment 11 - Due Wed, April. 25

1.) Use the series method to solve the second order differential equation with constant coefficients,

$$L\langle x|f\rangle = 0$$

$$L = \frac{d^2}{dx^2} + a\frac{d}{dx} + b$$

with boundary conditions

$$\langle 0|f\rangle = 1$$

$$\frac{d}{dx}\langle x|f\rangle_{|x=0} = 1$$

b. What is the domain of analyticity of your solution?

c. Put this equation in the form (14.6) (see K&D) and verify equation (14.12) for this example.

2.) Consider the differential equation

$$\frac{d^2}{dz^2}f(z) - f(z) = 0$$

Find recursion relations that define the coefficients of the two-independent power series solutions of this equation about $z = 0$.

3.) Consider the differential equation

$$\frac{d^2}{dz^2}f(z) + \sin(z)\frac{d}{dz}f(z) - \cos(z)f(z) = 0$$

Find recursion relations that define the coefficients of the two-independent power series solutions of this equation about $z = 0$.

4.) Consider the differential equation

$$\frac{d^2}{dz^2}f(z) + \frac{2}{z}\frac{d}{dz}f(z) - zf(z) = 0$$

Find the indicial equation for a solution about $z = 0$. Find the roots.

5.) Find the recursion relation that defines the coefficients of the series solution to the differential equation in problem 4) associated with the root of the indicial equation with largest real part.

29:172 Assignment 12 - Due Wed., May 2

1.) Let $r+s+t=0$ and $\alpha+\alpha'+\beta+\beta'+\gamma+\gamma'=1$. Show that

$$(z-z_1)^r(z-z_2)^s(z-z_3)^t P\left\{\begin{matrix} z_1 & z_2 & z_3 & \\ \alpha & \beta & \gamma & z \\ \alpha' & \beta' & \gamma' & \end{matrix}\right\}$$

satisfies the Riemann equation satisfied by

$$P\left\{\begin{matrix} z_1 & z_2 & z_3 & \\ \alpha+r & \beta+s & \gamma+t & z \\ \alpha'+r & \beta'+s & \gamma'+t & \end{matrix}\right\}$$

2. Show that for

$$z' = \frac{Az+B}{Cz+D} \qquad z_i' = \frac{Az_i+B}{Cz_i+D}$$

that

$$P\left\{\begin{matrix} z_1 & z_2 & z_3 & \\ \alpha & \beta & \gamma & z \\ \alpha' & \beta' & \gamma' & \end{matrix}\right\} = P\left\{\begin{matrix} z_1' & z_2' & z_3' & \\ \alpha & \beta & \gamma & z' \\ \alpha' & \beta' & \gamma' & \end{matrix}\right\}$$

3. Verify that

$$z^{1-c}F(b-c+1, a-c+1, 2-c, z)$$

satisfies the same differential equation as $F(a,b,c,z)$

4. Verify that

$$F(a,b,c,z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_1^\infty (t-z)^{-a}(t^{a-c}(t-1)^{c-b-1}dt$$

$$F(a,b,c,z) = \frac{\Gamma(c)}{\Gamma(c-b)\Gamma(b)}\int_0^1 (1-tz)^{-a}t^{b-1}(1-t)^{c-b-1}dt$$

are solutions of the Hypergeometric equation.

5. Verify that

$$z^{1-c}\Phi(a-c+1, 2-c, z)$$

satisfies the same differential equation as the confluent Hypergeometric function (c not integer) $\Phi(a,c,z)$.