PHYSICS 204A: METHODS OF MATHEMATICAL PHYSICS I

IAN LIM
LAST UPDATED JANUARY 5, 2020

These notes were taken for Physics 204A, *Methods of Mathematical Physics I*, as taught by Nemanja Kaloper at the University of California, Davis in fall quarter 2019. I live-TEXed them using Overleaf, and as such there may be typos; please send questions, comments, complaints, and corrections to itlim@ucdavis.edu.
Many thanks to Arun Debray for the LATEX template for these lecture notes: as of the time of writing, you can find him at https://web.ma.utexas.edu/users/a.debray/.

Contents

Lecture 1.

Wednesday, September 25, 2019

Reading assignment: read Ch. 2 and 3 of the course text (Arfken/Weber). This is basic linear algebra and vector analysis.

The purpose of this course is to learn formal aspects of quantum mechanics. We'll focus on doing analysis in Hilbert space. It's a remarkable fact about the natural world that most of our physical world is well-approximated by linear systems.

In the simplest form, we may think of vectors as arrays of numbers,

$$(v_1, v_2, \dots). \tag{1.1}$$

But we can also think of some real function $f(x) : \mathbb{R} \to \mathbb{R}$ as a collection of numbers too, just by taking its values at arbitrarily close points.

Definition 1.2. A *linear vector space* over a field F, denoted $L(F)$, is a set $\{|v\rangle\}$ with an addition operation $+$ such that

- for $|v\rangle, |u\rangle \in L, |u\rangle + |v\rangle = |w\rangle \in L$ (closure)
- for $c \in F$, $c|v\rangle \in L$ (scalar multiplication).

These axioms directly imply that any linear combination of vectors in L is also in the vector space:

$$c_v|v\rangle + c_u|u\rangle \in L. \tag{1.3}$$

This leads us naturally to the notion of *linear (in)dependence*. Suppose we take some vectors $|v_k\rangle \in L$ and make a linear combo,

$$\sum_k c_k|v_k\rangle. \tag{1.4}$$

Definition 1.5. A set of vectors $\{|v_k\rangle\}$ is *linearly dependent* if there exists some $\{c_k\}$ not all zero such that

$$\sum_k c_k|v_k\rangle = |0\rangle, \tag{1.6}$$

and such that $|0\rangle \notin \{|v_k\rangle\}$.

We need this last condition because otherwise we could simply take the coefficient of the $|0\rangle$ vector to be 1 and then arrive at a trivial solution.

Definition 1.7. If a set of vectors is not linearly dependent, it is *linearly independent*.

The next question we might ask is as follows: what is the size of the biggest set of linearly independent vectors we can construct for a given vector space?

Definition 1.8. The maximum number of linearly independent vectors associated to a given vector space is called the *dimension*.

Example 1.9. We may consider an infinitely differentiable (C^∞) function. It has a Taylor expansion

$$f(x) = \sum \frac{f^{(n)}(0}{n!}x^n, \tag{1.10}$$

which we may think of as an expansion in the basis $(1, x, x^2, \ldots)$.

So this is a vector space with countably infinite dimension. But we can have uncountably infinite-dimensional spaces too, e.g. the space of Fourier-transformable functions in a basis $e^{ikx}, k \in \mathbb{R}$. These factors are not just linearly independent; introducing an appropriate inner product, they are orthogonal.

It follows that for a vector space L_D of dimension D, any set with more than D vectors must be linearly dependent, i.e. $\exists \bar{c}_v, \bar{c}_k$ not all zero such that

$$\bar{c}_v|v\rangle + \sum_k \bar{c}_k|v_k\rangle = 0. \tag{1.11}$$

Moreover $\bar{c}_v \neq 0$ or else the original set $|v_k\rangle$ would be linearly dependent. Hence we can divide through, define $\hat{c}_k = \bar{c}_k/\bar{c}_v$, and write

$$|v\rangle = \sum_k \hat{c}_k|v_k\rangle. \tag{1.12}$$

That is, we have *decomposed* a general vector $|v\rangle$ in terms of its components $\hat{c}_k$ with respect to a basis $|v_k\rangle$.

We might now be interested in adding more structure to our vector space. Consider L_D with $|v\rangle, |u\rangle$.

Definition 1.13. We define an *inner product* by

$$\langle v|u\rangle : (|v\rangle, |u\rangle) \to F \tag{1.14}$$

as a map from the input vectors to the field over which the vector space is defined, with the following properties:

- $\langle u|(\lambda|v_1\rangle + \mu|v_2\rangle) = \lambda\langle u|v_1\rangle + \mu\langle u|v_2\rangle$ (linearity)
- $\langle v|u\rangle = \langle u|v\rangle^*$
- $\langle v|v\rangle \geq 0$, with equality only for the zero vector (positive semi-definite).

If we choose a basis $\{|v_k\rangle\}$, then if our vectors $|u\rangle, |v\rangle$ have some expansion in this basis then by linearity that

$$\langle u|v\rangle = \sum_k \hat{c}_k^v \langle u|v_k\rangle, \tag{1.15}$$

and we can expand each of these inner products as

$$\langle u|v_k\rangle = (\langle v_k|u\rangle)^* = \sum_n c_n^{u*}\langle v_n|v_k\rangle. \tag{1.16}$$

It follows that we can write a general inner product as

$$\langle u|v\rangle = \sum_{n,k} = c_k^v c_n^{u*}\langle v_n|v_k\rangle. \tag{1.17}$$

Moreover if we could choose a nice basis which had a special property of *orthogonality* or better yet *orthonormality*, we could reduce this to a single sum

$$\langle u|v\rangle = \sum_k c_k^v c_k^* \tag{1.18}$$

in terms of the components alone.

Lecture 2.

Monday, September 30, 2019

Last time, we wrote a general form for the dot (inner) product,

$$\langle u|v\rangle = \sum_{n,m} c_n^{*u} c_m^v \langle \phi_n|\phi_m\rangle, \tag{2.1}$$

where the $|\phi\rangle$s are basis vectors and u, v have expansions

$$|u\rangle = \sum_n c_n^u |\phi_n\rangle. \tag{2.2}$$

This is some quadratic form (it depends only quadratically on the components). And indeed it would be very nice if we could define $\langle \phi_N|\phi_m\rangle = \delta_{nm}$, so that our double-sum collapses to a single sum.

Orthonormality Let us suppose we start with a basis $\{|\phi_n\rangle\}$ for a vector space L_D. We shall show that we can construct a new basis $\{|\chi_n\rangle\}^D$ such that $\langle \chi_m|\chi_n\rangle = \delta_{mn}$ has the desired property.

WLOG let us number the basis vectors $|\phi_1\rangle, |\phi_2\rangle, \ldots$ and consider some inner products. The inner product

$$\langle \phi_1|\phi_1\rangle = N_1 \tag{2.3}$$

is some value N_1. If $N_1 = 1$ then we are done; otherwise, define

$$|\chi_1\rangle \equiv \frac{1}{\sqrt{N_1}}|\phi_1\rangle \tag{2.4}$$

so that

$$\langle \chi_1|\chi_1\rangle = \frac{\langle \phi_1|\phi_1\rangle}{N_1} = 1. \tag{2.5}$$

Hence $|\chi_1\rangle$ is a unit vector.

Consider the next vector $|\phi_2\rangle$. If

$$\langle \chi_1|\phi_2\rangle = 0, \tag{2.6}$$

then we can normalize and get

$$|\chi_2\rangle = \frac{|\phi_2\rangle}{\sqrt{N_2}}, \tag{2.7}$$

where $N_2 = \langle \phi_2|\phi_2\rangle$. Otherwise, we first subtract off the projection of the first normalized vector,

$$|\hat{\chi}_2\rangle = |\phi_2\rangle - \langle \chi_1|\phi_2\rangle|\chi_1\rangle, \tag{2.8}$$

so that

$$\langle \chi_1|\hat{\chi}_2\rangle = \langle \chi_1|\phi_2\rangle - \langle \chi_1|\phi_2\rangle \underbrace{\langle \chi_1|\chi_1\rangle}_{=1} = 0. \tag{2.9}$$

Hence by our definition, $|\chi_1\rangle$ and $|\hat{\chi}_2\rangle$ are orthogonal and we can just normalize. Defining

$$\hat{N}_2 = \langle \hat{\chi}_2 | \hat{\chi}_2 \rangle, \tag{2.10}$$

we have

$$|\chi_2\rangle = \frac{|\hat{\chi}_2\rangle}{\sqrt{\hat{N}_2}}, \tag{2.11}$$

which is a unit vector and normal to $|\chi_1\rangle$.

We continue by induction, subtracting off projections and normalizing. This is the *Gram-Schmidt procedure*.[1] Notice also that because $\langle \phi_n | \phi_m \rangle = \langle \phi_m | \phi_n \rangle^*$, we can consider values of m, n to be entries in a Hermitian matrix. Recalling that $\langle u | u \rangle \geq 0$, the norm is perfectly well-defined and indeed we can see that orthonormalization is equivalent to diagonalizing a Hermitian matrix.

More inner products On a function space, we can define an inner product

$$\langle f | g \rangle = \int_{-\infty}^{\infty} dx\, f^*(x) g(x). \tag{2.12}$$

These inner products come with strings attached; our functions usually have to satisfy some integrability properties in order for the inner products to be well-defined. Often the functions we're interested in come from differential equations. Most of the ones we encounter in physics are second-order so these functions ought to be twice-differentiable.[2] And we should also require that our functions are square-integrable so that the integral is well-defined.

We can then define

$$\begin{aligned} \langle u | u \rangle &= \sum_{n,m} c_n^* c_m \langle \phi_n | \phi_m \rangle \\ &= \sum_n |c_n|^2. \end{aligned} \tag{2.13}$$

And this is none other than the generalization of Pythagoras's theorem.

Schwarz inequality From the axioms, we can prove the following inequality.

$$|\langle f | g \rangle|^2 \leq \langle f | f \rangle \langle g | g \rangle. \tag{2.14}$$

We may define the linear combination

$$|f - \lambda g\rangle \tag{2.15}$$

and consider its norm (squared)

$$\langle f - \lambda g | f - \lambda g \rangle = \langle f | f \rangle - \lambda^* \langle g | f \rangle - \lambda \langle f | g \rangle + \lambda \lambda^* \langle g | g \rangle. \tag{2.16}$$

This is obviously non-negative, given the axioms. We can extremize this by taking derivatives with respect to λ, λ^* (which we may treat as linearly independent, since they are complex) and find that

$$\lambda^* = \frac{\langle g | f \rangle}{\langle g | g \rangle}, \quad \lambda = \frac{\langle f | g \rangle}{\langle g | g \rangle}. \tag{2.17}$$

A bit of manipulation yields the Schwarz inequality.

Let us also note that we can in general translate between the ket notation and vector notation. For a ket vector $|u\rangle$ we can associate the bra (row) vector $\langle u| = (c_1^*, c_2^* \ldots, c_D^*)$. Then the vector inner product is the same as old-fashioned row-column multiplication.

[1] Note that the procedure is a little more subtle in the infinite-dimensional case.

[2] This is a little too strong, actually. They can have finitely many discontinuities and this is still okay.

Bessel inequality Suppose we rewrite the vector $|u\rangle$ in a weird way, as

$$|u\rangle = \sum_n{}' c_n^u |\phi_n\rangle + |\Delta u\rangle, \tag{2.18}$$

where we take some terms and separate them out (so the sum $\sum'$ omits some indices). We know that

$$|\Delta u\rangle = u - \sum_n{}' c_n^8 |\phi_n\rangle \neq 0, \tag{2.19}$$

so that

$$0 < \langle \Delta u | \Delta u \rangle \tag{2.20}$$

$$= \langle u - \sum{}' c_n^u \phi_n | u - \sum{}' c_n^u \phi_n \rangle \tag{2.21}$$

$$= \langle u | u \rangle - \sum{}' |c_n^u|^2. \tag{2.22}$$

Check the cross-terms with the definition of u to get this final term. Rearranging, we get the Bessel inequality, which says that

$$\langle u | u \rangle > \sum{}' |c_n^u|^2, \tag{2.23}$$

i.e. the norm of a vector is greater than the partial sums of the squares of the components.

Linear operators We are primarily interested in linear operators, i.e. linear maps from the vector space to itself obeying

$$A(\lambda|\phi\rangle + \mu|\chi\rangle) = \lambda A|\phi\rangle + \mu A|\chi\rangle. \tag{2.24}$$

We can define two operators to be equal if they have the same action on all vectors, i.e.

$$A|\phi\rangle = B|\phi\rangle \tag{2.25}$$

for all $\phi \in L$.

In particular there's a nice way that we can rewrite the identity operator, as

$$\mathbb{I} = \sum_n |\phi_n\rangle\langle\phi_n|. \tag{2.26}$$

Let's prove this: by definition, $\mathbb{I}|u\rangle = |u\rangle$. On the other side, we see that

$$\sum_n |\phi_n\rangle\langle u| =\rangle \sum_n c_n^u |\phi_n\rangle \equiv |u\rangle. \tag{2.27}$$

Provided that $\{|\phi_n\rangle\}$ is a complete basis, this operator is indeed the identity.

The delta function The Dirac delta function is defined in such a way that

$$\int_a^b f(t)\delta(x-t)dt = f(x), \tag{2.28}$$

provided that x is in the interval (a, b). We could think of this as an inner product, however. We have a function and its shadow on the delta function picks out a value. The delta function isn't properly square-integrable, but we may consider it as having a good inner product with functions in our function space (as the limit of some sequence of square-integrable functions, if you like).

Now we'll do something strange. Let us express the delta function in a function basis,

$$\delta(x-t) = \sum_n c_n(t)\phi_n(x). \tag{2.29}$$

The t dependence must be in the coefficients since the functions themselves are just given. How do we find the coefficients? Just take the integral

$$\int dx\, \phi_m^*(x)\delta(x-t) = \int dx \sum_n c_n(t)\phi_n(x)\phi_n^*. \tag{2.30}$$

This is super easy to evaluate. On the RHS we have a Kronecker delta δ_{nm} by the orthonormality of the basis, and on the left side we have the evaluation of the basis vector ϕ_m^* at t, i.e.

$$c_m(t) = \phi_m^*(t). \tag{2.31}$$

Hence

$$\delta(x-t) = \sum_n \phi_n^*(t)\phi_n(x). \tag{2.32}$$

We can se that this had to be the case by substituting our expression for the delta function into an integral:

$$f(x) = \int dt\, \delta(x-t) f(t) = \sum_n \phi_n(x) \int dt\, \phi_n^*(t) f(t) \tag{2.33}$$

$$= \sum \langle \phi_n | f \rangle \phi_n(x), \tag{2.34}$$

which is none other than the components of f in the basis ϕ_n.

To make our discussion more concrete, let us consider analytic functions which have Taylor expansion

$$f(x) = \sum \frac{f^{(n)}(0)}{n!} x^n, \tag{2.35}$$

defined over the interval $[-1, 1]$. Hence $\{1, x, x^2, x^3, \ldots\}$ form a complete basis set for arbitrarily differentiable functions. They are certainly not orthogonal in general, e.g. $\int_{-1}^{1} dx\, 1 \cdot x^2 \neq 0$. But we can make them orthonormal with Gram-Schmidt.

Under this inner product, we have

$$\int_{-1}^{1} dx\, 1 \cdot 1 = 2, \tag{2.36}$$

so our first normalized vector is $1/\sqrt{2}$. We can check x:

$$\int_{-1}^{1} dx\, x \cdot x = 2/3, \tag{2.37}$$

so the next normalized vector is $\sqrt{3/2}x$. Continuing this way, we see that x^2 and x are already orthogonal but

$$\int_{-1}^{1} dx\, \frac{1}{\sqrt{2}} x^2 = \frac{2}{3\sqrt{2}}, \tag{2.38}$$

so our first unit vector is not orthogonal to x^2. We can instead define

$$\hat{x}^2 = x^2 - \sqrt{\frac{2}{3}}. \tag{2.39}$$

which is now orthogonal to the first unit vector $1/\sqrt{2}$ and to the second unit vector $\sqrt{3/2}x$. We can normalize $\hat{x}^2$ and determine the third unit vector in this set, which is $\frac{3x^2}{2} - 1$ (we think).

Let us remark that the most general operators that can be diagonalized are *normal* operators, i.e. those satisfying

$$[A, A^\dagger] = 0. \tag{2.40}$$

Clearly, one set of operators that are not normal are the raising and lowering operators, whose commutator is $[a, a^\dagger] = 1$.

Lecture 3.

Wednesday, October 2, 2019

Today we'll continue discussing operators. We've discussed the identity operator,

$$\mathbb{I} = \sum |\phi_i\rangle\langle\phi_i|, \tag{3.1}$$

which maps any vector into itself. More generally, we can define an operator as follows:

Definition 3.2. An *operator* is a map $A : |\psi\rangle \to |\bar{\psi}\rangle$ where $|\psi\rangle, |\bar{\psi}\rangle \in L_D$ are in the same vector space L_D.[3] A *linear operator* is an operator obeying the linearity property

$$A(\mu|\psi\rangle + \nu|\chi\rangle) = \mu A|\psi\rangle + \nu A|\chi\rangle. \tag{3.3}$$

[3] The math-inclined among us may talk about the space of inner automorphisms on the vector space.

If we have a set of linear operators, we may define an addition operation on operators as

$$(A+B)|\psi\rangle = A|\psi\rangle + B|\psi\rangle \tag{3.4}$$

and scalar multiplication as

$$(kA)|\psi\rangle = k(A|\psi\rangle). \tag{3.5}$$

Hence we can take linear combinations of linear operators as

$$kA + lB, \tag{3.6}$$

and under this definition we see that linear operators form a vector space.

But there's another way to combine operators, namely by *composition*. That is, given operators A and B we can define a new operator AB defined by the composition

$$AB|\psi\rangle = A(B|\psi\rangle). \tag{3.7}$$

Composition must satisfy certain properties with respect to the other operations we've defined, namely distributivity with respect to addition:

$$A(B+C) = AB + AC. \tag{3.8}$$

Let us note that the product of operators (composition) is generally not commutative,

$$AB \neq BA \tag{3.9}$$

in general. We know this from matrix multiplication.

Alternately, we could define a composition rule using a commutator (bracket),

$$[A,B] = AB - BA. \tag{3.10}$$

It's not to hard to check that this rule also satisfies distributivity over addition. If we wished, we could also prove (by crunching through the commutators) the Jacobi identity,

$$[[A,B],C] + [[B,C],A] + [[C,A],B] = 0. \tag{3.11}$$

This is related to the Bianchi identity in differential geometry.

Under our standard composition rule, we can define *inverses*. That is, if an operator A acts as

$$|\psi'\rangle = A|\psi\rangle, \tag{3.12}$$

then the inverse A^{-1} (if it exists) is the operator such that

$$A^{-1}|\psi'\rangle = |0\rangle. \tag{3.13}$$

That is, $A^{-1}A|\psi\rangle = |\psi\rangle$.

As we know, not every operator is invertible. Consider the operator which just sends some vector to zero (i.e. it has a null eigenvector). If

$$A|\psi\rangle = 0 \tag{3.14}$$

then the inverse is not well-defined: $A^{-1}|0\rangle =?$.

Definition 3.15. The adjoint $A^\dagger$ of an operator A is defined by

$$\langle\chi|A^\dagger|\psi\rangle = \langle\psi|A|\chi\rangle^*. \tag{3.16}$$

Definition 3.17. If an operator is self-adjoint, $H^\dagger = H$, then we call it Hermitian.

Notice that

$$\langle\psi|H|\psi\rangle = \langle\psi|H^\dagger|\psi\rangle = \langle\psi|H|\psi\rangle^* \tag{3.18}$$

by the definition of the adjoint and hermiticity. Therefore the diagonal matrix elements of H are *real numbers*, i.e. their eigenvalues are real.

We could have also taken an operator which was *anti-Hermitian*, $H^\dagger = -H$, which implies that the diagonal elements are instead purely imaginary by the same argument.

Definition 3.19. A *unitary* operator is an operator obeying the property

$$U^{-1} = U^\dagger. \tag{3.20}$$

It's clear that we can restrict to the real case, in which case Hermitian matrices become symmetric matrices and anti-Hermitian matrices become antisymmetric (sometimes called skew-symmetric). Our unitary matrices reduce to orthogonal matrices.

Example 3.21. Consider the Hilbert space of smooth square-integrable functions over the real line, $f \in L$. Our inner product is the integral

$$\int dx\, f^*(x) g(x). \tag{3.22}$$

Define the operator $D = -\frac{d}{dx}$. What is the adjoint $D^\dagger$? We have

$$\int f^*(-\frac{d}{dx})g = f^* g|_a^b + \int (\frac{d}{dx} f^*) g. \tag{3.23}$$

The boundary term vanishes based on the boundary conditions, i.e. given that f, g vanish at infinity. Taking the complex conjugate to get the adjoint, what is left is

$$\left[\int (\frac{d}{dx} f^*) g\right]^* = -\int g^* D f. \tag{3.24}$$

So D is not Hermitian but iD is (it adds a minus sign to fix the sign in the integration by parts).

Note also that when the integration region is finite, the boundary conditions become nontrivial. However, if f and g vanish at the boundary (e.g. $[-1, 1]$) then we restore hermiticity.

Recall we said that we could assign matrix elements to an operator with respect to some set of vectors,

$$\langle \chi | A | \psi \rangle. \tag{3.25}$$

In fact, it's a Sisyphean task to do this for all sets of vectors, but fortunately (thanks to linearity) it suffices to compute the matrix elements in some (complete) basis. With a basis $|\psi_i\rangle$ we can define

$$\langle \psi_n | A | \psi_k \rangle = A_{nk}. \tag{3.26}$$

For recall that the identity can be written as $\mathbb{I} = \sum_k |\psi_k\rangle\langle\psi_k|$, and suppose $|\psi\rangle$ has some decomposition in the basis

$$|\psi\rangle = \sum_k \langle \psi_k | \psi \rangle |\psi_k\rangle \tag{3.27}$$

Then

$$\begin{aligned} A|\psi\rangle &= \mathbb{I}A\mathbb{I}|\psi\rangle \\ &= \sum_{k,n} |\psi_k\rangle\langle\psi_k|A|\psi_n\rangle\langle\psi_n|\psi\rangle. \end{aligned}$$

Thus we can recognize the components of $|\psi\rangle$ in our basis, which are given by $\langle\psi_n|\psi\rangle = C_n$. Hence

$$\begin{aligned} A\left(\sum_k \langle\psi_k|\psi\rangle|\psi_k\rangle\right) &= \sum_k |\psi_k\rangle \left(\sum_n \langle\psi_k|A|\psi_n\rangle\langle\psi_n|\psi\rangle\right) \\ &= \sum_k |\psi_k\rangle \left(\sum_n A_{kn} C_n\right) \\ &= \sum |\psi_k\rangle C_k^{\bar{\psi}} \end{aligned}$$

in terms of the components of some new vector $|\bar{\psi}\rangle$. That is, if we know the matrix elements of A in some basis and we know the components of the vector in that basis, we can uniquely determine the components of its image under A in the same basis.

In the end, this is just abstract matrix multiplication. That is,

$$C_n^{\bar{\psi}} = \sum_k A_{nk} C_k^{\psi}. \tag{3.28}$$

We can also write the operator A in terms of its matrix elements:

$$\begin{aligned} A &= \mathbb{I}A\mathbb{I} \\ &= \sum_{n,k} |\psi_k\rangle\langle\psi_k|A|\psi_n\rangle\langle\psi_n| \\ &= \sum_{n,k} A_{kn}|\psi_k\rangle\langle\psi_n|. \end{aligned}$$

This also tells us immediately that the matrix elements of the identity in any orthonormal basis are as we could have guessed– $\mathbb{I}_{nk} = \delta_{nk}$, the Kronecker delta.

It also follows that the matrix elements of the adjoint of an operator obey

$$(A^\dagger)_{nk} = A^*_{kn}. \tag{3.29}$$

This gives us another statement of hermiticity– equivalently, a hermitian operator is one whose matrix elements obey

$$A^*_{kn} = A_{nk}. \tag{3.30}$$

And thus

$$A^\dagger = \sum_{k,n} A^*_{nk}|\psi_k\rangle\langle\psi_n|. \tag{3.31}$$

Let's check that for hermitian operators, the expectation value is non-negative,

$$\langle\psi|A|\psi\rangle \geq 0. \tag{3.32}$$

Writing A in terms of its matrix elements, we have

$$\begin{aligned} \langle\psi|A|\psi\rangle &= \sum_{n,k} \langle\psi|\psi_k\rangle A_{kn}\langle\psi_n|\psi\rangle \\ &= C^{*\psi}_k A_{kn} C^\psi_n. \end{aligned}$$

Suppose we have two orthonormal bases for the same space, $\{|\psi_k\rangle\}, \{|\psi'_k\rangle\}$. It follows that the new basis has some decomposition in the old basis. That is, the set $\{|\psi_k\rangle, |\psi'_1\rangle\}$ is linearly dependent and so

$$|\psi'_n\rangle = \sum c^{n\prime}_k |\psi_k\rangle, \tag{3.33}$$

in terms of some coefficients $c^{n\prime}_k$. It's also true that we can go back,

$$|\psi_n\rangle = \sum c^n_k |\psi'_k\rangle. \tag{3.34}$$

Certainly we can write this decomposition as the action of an operator U:

$$U|\psi_n\rangle = \sum c^{n\prime}_k |\psi_k\rangle, \tag{3.35}$$

and moreover U must be invertible.

Lecture 4.

Monday, October 7, 2019

When we choose coordinates, the name of the game is to exploit the *symmetries* of the problem. That is, to find coordinates which respect the dynamical symmetries of the Hamiltonian. If we choose our coordinates well enough, the equations of motion become trivial to solve.

Suppose we have two bases $\{|\psi_i\rangle\}, \{|\psi'_i\rangle\}$. Then we can certainly write

$$U|\psi_i\rangle = |\psi'_i\rangle = \sum_j C^{\psi'}_{ij} |\psi_j\rangle \tag{4.1}$$

or equivalently

$$\bar{U}|\psi'_j\rangle = |\psi_j\rangle = \sum_i C^\psi_{jk} |\psi'_i\rangle. \tag{4.2}$$

That is, we can express an element of one basis in another basis. It is evident that the inverse exists, since it doesn't matter what we call the first and the second basis. So take the first equation and act on it with U^{-1}. Then

$$|\psi_i\rangle = U^{-1}U|\psi_i\rangle = U^{-1}|\psi_i'\rangle = \sum_j C_{ij}^{\psi'} U^{-1}|\psi_j\rangle. \tag{4.3}$$

What are the matrix elements $\langle\psi_k|U|\psi_i\rangle = \langle\psi_k|\psi_i'\rangle$?

$$\langle\psi_k|\psi_i'\rangle = \sum_j C_{ij}^{\psi'}\langle\psi_k|\psi_j'\rangle, \tag{4.4}$$

or equivalently

$$|\psi_i'\rangle = \sum_j |\psi_j\rangle\langle\psi_j|\psi_i'\rangle. \tag{4.5}$$

Similarly

$$|\psi_j\rangle = \sum_i |\psi_i'\rangle\langle\psi_i'|\psi_j\rangle. \tag{4.6}$$

Notice also that these coefficients are therefore clearly related by

$$\langle\psi_i'|\psi_j\rangle = \langle\psi_j|\psi_i'\rangle^*. \tag{4.7}$$

Consider now the inner product

$$\begin{aligned}
\delta_{kj} &= \langle\psi_k|\psi_j\rangle \\
&= \sum_{i,l}\langle\psi_l'|\psi_k\rangle^* \underbrace{\langle\psi_l'|\psi_i'\rangle}_{\delta_{li}}\langle\psi_i'|\psi_j\rangle \\
&= \sum_i \langle\psi_i'|\psi_k\rangle^*\langle\psi_i'|\psi_j\rangle \\
&= \sum_i \langle\psi_k|\psi_i'\rangle\langle\psi_i'|\psi_j\rangle.
\end{aligned}$$

Define

$$U_{ij} = \langle\psi_i'|\psi_j\rangle. \tag{4.8}$$

And similarly

$$\bar{U}_{ki} = \langle\psi_k|\psi_i'\rangle = U_{ik}^*. \tag{4.9}$$

Hence

$$\delta_{kj} = \sum_i U_{ki}' U_{ij}, \tag{4.10}$$

so since $(U^{-1})_{kj} = U_{jk}^*$, we see that transformations between two orthonormal bases are unitary, i.e. they satisfy

$$UU^\dagger = \mathbb{I}. \tag{4.11}$$

We could have seen this in a basis-free way by requiring that inner products do not depend on the choice of basis. Hence

$$\langle\chi'|\psi'\rangle = \langle\chi|U^\dagger U|\psi\rangle = \langle\chi|\psi\rangle = \langle\chi|\mathbb{I}|\psi\rangle, \tag{4.12}$$

and hence it must be that $U^\dagger U = \mathbb{I}$ since they agree on *any* vectors $|\chi\rangle, |\psi\rangle$.

We've concluded that changes of basis can be written as unitary transformations and moreover that inner products must be invariant under such transformations. What about operators? Consider

$$A = \sum_{i,j} a_{ij}|\psi_i\rangle\langle\psi_j|. \tag{4.13}$$

Suppose we have a new vector $|\phi\rangle$ given by

$$|\phi\rangle = \sum_i |\psi_i\rangle\langle\psi_i|\phi\rangle, \tag{4.14}$$

and we act on it with a unitary U to get some new $|\phi'\rangle$. Hence

$$U|\phi\rangle = \sum_i U|\psi_i\rangle\langle\psi_i|\phi\rangle. \tag{4.15}$$

Let U take us between bases, such that $|\psi_i'\rangle = U|\psi_i\rangle$ or equivalently

$$U^\dagger|\psi_i'\rangle = |\psi_i\rangle. \tag{4.16}$$

The easy way to see what happens to operators is to recognize that

$$U|\phi'\rangle = UA|\phi\rangle = UAU^\dagger(U|\phi\rangle), \tag{4.17}$$

so that

$$A' = UAU^\dagger \tag{4.18}$$

is the corresponding operator to A in the new basis. Looking at the spectral decomposition,

$$A' = UAU^\dagger = \sum U|\psi_i\rangle A_{ij}\langle\psi_j|U^\dagger \tag{4.19}$$

$$= \sum |\psi_i'\rangle A_{ij}\langle\psi_j'|, \tag{4.20}$$

so the matrix elements are left unchanged in the new basis. In fact, it must have been so, since

$$\langle\psi_i|A|\psi_j\rangle = \left(\langle\psi_i|U^\dagger\right)\left(UAU^\dagger\right)(U|\psi_j\rangle). \tag{4.21}$$

Let us now make a connection to quantum mechanics. If the Hamiltonian is time-independenct, recall that we can write the solution of

$$i\partial_t|\psi\rangle = H|\psi\rangle \tag{4.22}$$

as

$$|\psi(t)\rangle = e^{-itH}|\psi(0)\rangle. \tag{4.23}$$

That is, we may think of time evolution as a complex rotation of the initial condition.[4]

Consider now a set of vectors which are not orthonormal but are linearly independent, $\{|\chi_i\rangle\}$. By Gram-Schmidt we can therefore define $\{|\psi_\mu\rangle\}$ as the result of acting on $|\psi_i\rangle$ with an operator T such that

$$|\psi_\mu\rangle = \sum_{i=1}^{\mu} T_{\mu i}|\chi_i\rangle. \tag{4.24}$$

In particular we may take $T_{\mu i}$ to be upper-triangular because we limit the sum to run only up to μ.

Now observe that

$$\begin{aligned}\delta_{\nu\mu} &= \langle\psi_\nu|\psi_\mu\rangle \\ &= \sum_{ij} T^*_{\nu j}T_{\mu i}\langle\chi_j|\chi_i\rangle \\ &= \sum_{ij} T^*_{\nu j}\langle\chi_j|\chi_i\rangle T_{\mu i} \\ &= (T^\dagger S T)_{\nu\mu}.\end{aligned}$$

These final inner products are in general not 1; they are like a metric, in that they relate the original basis vectors. We see that

$$\mathbb{I} = T^\dagger S T. \tag{4.25}$$

But notice that T is invertible, and therefore so is $T^\dagger$. We find that

$$(T^\dagger)^{-1}T^{-1} = S, \tag{4.26}$$

so the matrix S relating the original vectors can be written in terms of the matrix T which performs the orthonormalization process. In general, this allows us to perform a *similarity transforamation* so that we can write some operator $A' = TAT^{-1}$ and

$$A = (T^\dagger)^{-1}\bar{A}T^{-1}, \tag{4.27}$$

which says that an operator may be written as the product of an upper triangular matrix, a diagonal matrix, and a lower triangular matrix. This is a special case of the Jordan decomposition of an operator.

Notice also that unitaries leave the trace and the determinant unchanged:

$$\mathrm{Tr}(UAU^\dagger) = \mathrm{Tr}(AU^\dagger U) = \mathrm{Tr}(A), \tag{4.28}$$

[4]There is a book describing the structure of path integrals in quantum mechanics, by Feynman and Hibbs, related to the exponentiation formula for the Hamiltonian. There is also a text by Messiah which contains lots of good quantum mechanics.

and determinants are unchanged since

$$UU^\dagger = \mathbb{I} \implies 1 = \det \mathbb{I} = \det U \det U^\dagger = (\det U)(\det U)^*, \tag{4.29}$$

so the determinant of U is a phase $e^{i\delta}$ and

$$\det(UAU^\dagger) = \det(A)\det(U^\dagger U) = \det A. \tag{4.30}$$

This helps our intuition that unitary matrices are really complex generalizations of rotations (orthogonal matrices).

It will often be our interest in quantum mechanics of solving the eigenvector and eigenvalue problem, i.e.

$$A|\phi\rangle = \lambda|\phi\rangle. \tag{4.31}$$

Thanks to the fundamental theorem of algebra, we are guaranteed at least one eigenvalue, some λ satisfying the characteristic equation.

Lecture 5.

Wednesday, October 9, 2019

> *"$\hbar$ is 1, the only reason we didn't realize it earlier is because we had to start with bananas and rocks and stuff."*
>
> –Nemanja Kaloper

We said last time that operators in a new basis $|\phi_i'\rangle = U|\phi_i\rangle$ are given by

$$A' = UAU^\dagger = \sum a_{ik}|\phi_i'\rangle\langle\phi_k'|. \tag{5.1}$$

Hence the *components* of the corresponding operator A' are left unchanged in the new basis. However, it is also true that the original operator has some components[5]

$$A = \sum a_{ik}|\phi_i\rangle\langle\phi_k| = \sum a'_{ik}|\phi_i'\rangle\langle\phi_k'|. \tag{5.2}$$

That is, we can either define a change of basis and also rotate the operator along with it ($A' = UAU^\dagger$) or we can just change the basis and leave the operator fixed. Hence its projections onto the new basis will change and therefore its components will change. This is the difference between *active and passive transformations*. Incidentally, this is related to the Schrödinger and Heisenberg pictures. There is also a third hybrid[6] picture known as the interaction picture, and this ends up being the most practical one to do perturbative calculations in.

We will work mostly in the picture where operators are fixed and the bases change. How do we find the components in the new basis?

$$\begin{aligned} |\phi_i'\rangle &= U|\phi_i\rangle && (5.3)\\ &= \sum_j |\phi_j\rangle\langle\phi_j|U|\phi_i\rangle && (5.4)\\ &= \sum_j U_{ji}|\phi_j\rangle. && (5.5) \end{aligned}$$

Hence

$$a'_{ik} = \sum_{jl}\langle\phi_i'|\phi_j\rangle a_{jl}\langle\phi_l|\phi_k'\rangle = \sum_{jl} U^*_{ji}a_{jl}U_{lk}, \tag{5.6}$$

which tells us that the components transform in the opposite way from the operator. That is, in this picture we have

$$A^{\phi'}_{ik} = a'_{ik} = (U^\dagger A U)_{ik}. \tag{5.7}$$

Recall from last time that for some general operator A, we are guaranteed at least one eigenvector,

$$A|\psi\rangle = \lambda|\psi\rangle. \tag{5.8}$$

[5] I tend to write this as components times $|\phi_i\rangle\langle\phi_k|$ to emphasize that this is really a linear combination of operators.

[6] "Bastard," as per Nemanja.

Suppose we have a coupled harmonic oscillator potential

$$V = ax^2 + bxy + cy^2, \tag{5.9}$$

which results in the forces

$$F_x = -2ax - by \tag{5.10}$$

$$F_y = -2cy - bx. \tag{5.11}$$

This looks like an ellipsoidal well, and particles will form closed orbits (Lissajous figures) in this well.

In general orbits will be 2-dimensional, changing in x and y. But! Sometimes if we set it up right, we will find that our orbits become 1-dimensional. And this tells us that we've found the *characteristic directions* of the oscillator.

We can now write this as an eigenvalue problem:

$$\begin{pmatrix} a_x \\ a_y \end{pmatrix} = \frac{1}{m}\begin{pmatrix} F_x \\ F_y \end{pmatrix} = \begin{pmatrix} -2ax - by \\ -2cy - bx \end{pmatrix} = \begin{pmatrix} -2a & -b \\ -b & -2c \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}. \tag{5.12}$$

We could imagine an orbit where the acceleration becomes parallel to the displacement, in which case the motion just becomes one-dimensional.

If we project down the equipotentials, we can read off the principal axes, which tells us the right directions in which the motion separates.

We can abstract the problem: consider

$$\begin{pmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \lambda\begin{pmatrix} x \\ y \end{pmatrix}. \tag{5.13}$$

We solve the eigenvalue problem by finding the characteristic equation. That is,

$$(h_{11} - \lambda)(h_{22} - \lambda) - h_{12}^2 = 0. \tag{5.14}$$

These could be real eigenvalues, and potentially degenerate.[7] In this example, the matrix is real and symmetric and therefore Hermitian, which implies its eigenvalues are real. In the general case the eigenvalues could be complex if the matrix entries were not all real.[8] Note that we get the constraint equations,

$$(h_{11} - \lambda_1)x + h_{12} = 0 \tag{5.15}$$

$$h_{12}x + (h_{22} - \lambda_1)y = 0. \tag{5.16}$$

Notice that the determinant of $H - \lambda\mathbb{I}$ vanished, which tells us that the rows are linearly *dependent*. So we will not get any extra information out of the second equation, i.e. our solution is not uniquely determined.[9] We can solve

$$y = -\frac{h_{11} - \lambda_1}{h_{12}}x, \tag{5.17}$$

but we cannot a priori fix the value of x. For notice that to any eigenvector, i.e. a solution to

$$\begin{pmatrix} h_{11} - \lambda_1 & h_{12} \\ h_{12} & h_{22} - \lambda_1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 0, \tag{5.18}$$

another (nontrivial) solution to this is clearly $\begin{pmatrix} cx \\ cy \end{pmatrix}$ for some $c \neq 0$. Hence eigenvectors are only determined up to an overall normalization factor.

For instance, let us compute the eigenvectors of

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{5.19}$$

The characteristic equation is

$$\lambda^2 - 1 \implies \lambda = \pm 1. \tag{5.20}$$

[7] In this case, if the eigenvalues are degenerate, then our elliptical bowl becomes a circle. There are not just two characteristic directions but a continuum of characteristic directions.

[8] Incidentally they would be complex conjugates, since complex roots to real equations must come in pairs.

[9] In fancier language, we get a one-parameter family of eigenvalues, i.e. a 1D subspace of eigenvectors.

Hence we have

$$\begin{pmatrix} y \\ x \end{pmatrix} = \pm \begin{pmatrix} x \\ y \end{pmatrix}, \tag{5.21}$$

giving $y = x$ or $y = -x$. The normalized eigenvectors are

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{5.22}$$

Moreover, notice that we can build a matrix built out of the components of the eigenvectors, namely

$$E = \begin{pmatrix} x_+ & x_- \\ y_+ & y_- \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{5.23}$$

Observe that

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} E = E \begin{pmatrix} +1 & 0 \\ 0 & 1 \end{pmatrix} = E \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix}. \tag{5.24}$$

Moreover, $E^T E = \mathbb{I}$ since the eigenvectors are orthonormal. So if we now multiply on the left by E^T we find that

$$E^T A E = A' \tag{5.25}$$

where A' is now the diagonal matrix of the eigenvalues of A.

Lecture 6.

Monday, October 14, 2019

Last time, we showed that certain kinds of matrices can be diagonalized, i.e. we can write them as a similarity transform of a diagonal matrix. The goal is basically to project onto the (normalized) eigenvectors (E), multiply by eigenvalues in the eigenbasis (D), and then perform the inverse transformation to take us back to the original basis (E^T).

Recall we are interested in the eigenvalue problem for Hermitian matrices,

$$H|\psi\rangle = \lambda_\psi |\psi\rangle \tag{6.1}$$

where $H = H^\dagger$. In the finite dimensional case, we are guaranteed at least one (simply write down the characteristic equation, and it must have at least one root over the complex numbers). Suppose moreover we have another eigenvalue

$$H|\phi\rangle = \lambda_\phi |\phi\rangle. \tag{6.2}$$

Let us now conjugate the second equation to get

$$\langle\phi|H = \lambda_\phi^* \langle\phi|. \tag{6.3}$$

We can multiply Eqn. 6.1 by $\langle\phi|$ and Eqn. 6.3 by $|\psi\rangle$. Hence

$$\langle\phi|H|\psi\rangle = \lambda_\psi \langle\phi|\psi\rangle \tag{6.4}$$

$$\langle\phi|H|\psi\rangle = \lambda_\phi^* \langle\phi|\psi\rangle. \tag{6.5}$$

We can now subtract these two from each other to find that

$$0 = (\lambda_\psi - \lambda_\phi^*)\langle\phi|\psi\rangle. \tag{6.6}$$

If $|\phi\rangle = |\psi\rangle$ with $|\psi\rangle \neq 0$, then we immediately find that

$$\lambda_\phi = \lambda_\phi^*, \tag{6.7}$$

which tells us the eigenvalues are real.

Conversely if the eigenvalues are distinct ($\lambda_\psi \neq \lambda_\phi$) then

$$0 = (\lambda_\psi - \lambda_\phi)\langle\phi|\psi\rangle \implies \langle\phi|\psi\rangle = 0. \tag{6.8}$$

Hence eigenvectors associated to distinct eigenvalues *must be orthogonal.*

Finally, suppose we have a degenerate case where multiple eigenvectors are associated to the same eigenvalue, $\lambda_\phi = \lambda_\psi$. Then this equation cannot help us because it automatically vanishes. We get a degenerate subspace associated to the eigenvalue λ_ϕ, and we can then perform Gram-Schmidt within the subspace.

This completes the process of diagonalization. Distinct eigenvalues correspond to orthogonal eigenvectors, and degenerate eigenvalues give us a subspace whose basis vectors can be made orthogonal by Gram-Schmidt.

Moreover, there is a completeness relation for the eigenvectors:

$$\sum_i |\psi_i\rangle\langle\psi_i| = \mathbb{I}. \tag{6.9}$$

This is fairly trivial in the finite-dimensional case (in n dimensions we can construct n orthogonal vectors from the eigenvectors, so we get a basis). However, there is a convergence requirement in the infinite-dimensional case. It is not enough to just count and show orthogonality; we must show that all the same information from the original function is contained in the Fourier coefficients, for example.

We will keep in the back of our mind that this completeness relation is really something we have to prove, but for the most part we will perform calculations as though it is true. That is, recall we can write a decomposition in some basis as

$$|\phi\rangle = \sum_i \langle\psi_i|\phi\rangle|\psi_i\rangle. \tag{6.10}$$

If

$$\lim_{N\to\infty} |||\phi\rangle - \sum_i^N \langle\psi_i|\phi\rangle|\psi_i\rangle||^2 = 0 \tag{6.11}$$

then we may say that our basis is complete in the sense that a decomposition into components $\langle\psi_i|\phi\rangle$ converges to the "real" vector $|\phi\rangle$.

Now

$$H = H\sum_i |\psi_i\rangle\langle\psi_i| \tag{6.12}$$

$$= \sum_i H|\psi_i\rangle\langle\psi_i| \tag{6.13}$$

$$= \sum_i \lambda_i|\psi_i\rangle\langle\psi_i|. \tag{6.14}$$

That is, H is diagonal in its eigenbasis and its elements are the eigenvalues. Let us note that

$$H\begin{pmatrix} \psi_1^1 & \psi_1^2 \dots \\ \vdots & \vdots \\ \psi_N^1 & \psi_N^2 \dots \end{pmatrix} = \begin{pmatrix} \psi_1^1 & \psi_2^1 \dots \\ \vdots & \vdots \\ \psi_1^N & \psi_2^N \dots \end{pmatrix}\begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \end{pmatrix}, \tag{6.15}$$

where lower indices indicate which eigenvector we're looking at and upper indices indicate the component of that eigenvector. That is, this is just a matrix form of

$$H|\psi_i\rangle = \lambda_i|\psi_i\rangle. \tag{6.16}$$

More generally we have constructed

$$HU = UH_d \implies H_d = U^\dagger HU, \tag{6.17}$$

which tells us that U gives the complex rotation to take us to the eigenbasis of the hermitian operator H.

Now suppose that two operators commute, $[A, B] = 0$. Then a mutual eigenbasis can be found:

$$[A, B] \iff A|\phi\rangle = a|\phi\rangle, B|\phi\rangle = b|\phi\rangle, \tag{6.18}$$

for all eigenvectors $|\phi\rangle$. The $\impliedby$ direction is easy:

$$AB|\phi\rangle = bA|\phi\rangle = ab|\phi\rangle, \tag{6.19}$$

and similarly

$$BA|\phi\rangle = ba|\phi\rangle = ab|\phi\rangle \tag{6.20}$$

for scalars. By completeness,

$$(AB - BA)|\psi\rangle = 0 \tag{6.21}$$

for a general vector $|\psi\rangle$.

Let us now prove the $\implies$ direction. If A and B commute, then

$$AB|\phi\rangle = BA|\phi\rangle \tag{6.22}$$

for all $|\phi\rangle$. Let us suppose $|\phi\rangle$ is an eigenvector of B,

$$B|\phi\rangle = b|\phi\rangle. \tag{6.23}$$

Then

$$bA|\phi\rangle = BA|\phi\rangle. \tag{6.24}$$

But this tells us that $A|\phi\rangle$ is an eigenvector of B with eigenvalue b.

There are a few possibilities here. It could be that $A|\phi\rangle = 0$, which tells us that $|\phi\rangle$ is an eigenvector of A with eigenvalue zero,

$$A|\phi\rangle = 0|\phi\rangle. \tag{6.25}$$

The next case is where b is a unique eigenvalue of B. Then

$$A|\phi\rangle \propto |\phi\rangle, \tag{6.26}$$

i.e. it lives in the same 1D subspace as $|\phi\rangle$. It is simply proportional to the original eigenvector $|\phi\rangle$, and hence

$$A|\phi\rangle = a|\phi\rangle. \tag{6.27}$$

The only subtle case is when b is a degenerate eigenvalue. What then? In this case, $A|\phi\rangle$ simply lives in the degenerate subspace corresponding to eigenvalue b.

Normal operators Suppose we have an operator A obeying

$$[A, A^\dagger] = 0. \tag{6.28}$$

Suppose A has an eigenvalue,

$$A|\phi_1\rangle = \lambda_1|\phi_1\rangle \implies (A - \lambda_1)|\phi_1\rangle = 0. \tag{6.29}$$

Taking the dagger,

$$\langle\phi_1|(A^\dagger - \lambda^*) = 0. \tag{6.30}$$

Now since these are individually zero, we can multiply them to get

$$\langle\phi_1|(A^\dagger - \lambda_1^*)(A - \lambda_1)|\phi_1\rangle = 0. \tag{6.31}$$

Expanding out the multiplication we have

$$\begin{aligned} 0 &= \langle\phi_1|(A^\dagger A - \lambda^* A - \lambda A^\dagger + \lambda_1^*\lambda_1|\phi_1\rangle \\ &= \langle\phi_1|(AA^\dagger - \lambda^* A - \lambda A^\dagger + \lambda_1^*\lambda_1|\phi_1\rangle \\ &= \langle\phi_1|(A - \lambda_1)(A^\dagger - \lambda_1^*)|\phi_1\rangle \end{aligned}$$

where in the second line we have used the fact that A and $A^\dagger$ commute. This is just the modulus $||(A^\dagger - \lambda_1^*)|\phi_1\rangle||^2$, so since it is zero,

$$(A^\dagger - \lambda_1^*)|\phi_1\rangle = 0 \implies A^\dagger|\phi_1\rangle = \lambda_1^*|\phi_1\rangle. \tag{6.32}$$

We conclude that $|\phi_1\rangle$ is an eignevector of $A^\dagger$ with eigenvalue λ_1^*.

Consider now some vector $|\psi\rangle$ in the orthogonal complement of $|\phi_1\rangle$,

$$|\psi\rangle \in L_D \setminus \{|\phi\rangle | |\phi\rangle = a|\phi_1\rangle, a \in \mathbb{C}\}. \tag{6.33}$$

Then

$$0 = \langle\phi_1|\psi\rangle \tag{6.34}$$

by definition, and

$$\langle\phi_a|A|\psi\rangle = \langle\psi|A^\dagger|\phi_1\rangle^* = [\lambda_1^*\langle\psi|\phi_1\rangle]^* = \lambda_1\langle\phi_1|\psi\rangle = 0, \tag{6.35}$$

since $|\phi_1\rangle$ is an eigenvector of $A^\dagger$. Hence it is not just $|\psi\rangle$ that is in the orthogonal complement of $|\phi_1\rangle$– in fact, its image $A|\psi\rangle$ is also in the orthogonal complement. Hence this orthogonal complement is well-defined. Its image $A|\psi\rangle$ does not mix with any of the vectors we just took away.

Now we have reduced the dimension of the space by one, which tells us that we can repeat this process with the resulting subspace and therefore find a complete set of eigenvectors and eigenvalues. This completes the proof that normal operators can be diagonalized.[10]

As a practical aside, these sorts of techniques appear in the case of perturbation theory. Sometimes we will have a degenerate eigenvalue of the zeroth order Hamiltonian due to certain symmetries of the problem, and if we are lucky, the perturbation will lift (break) this symmetry, splitting the degenerate energy levels. This is the case with the Stark effect (electric field aligning the electric dipole of an electron) and Zeeman effect (magnetic equivalent).

Now suppose we have an eigenbasis in hand and can write

$$H = \sum |\phi_i\rangle \lambda_i \langle \phi_i|. \tag{6.36}$$

Hence

$$H^n |\phi_i\rangle = \lambda_i^n |\phi_i\rangle, \tag{6.37}$$

so in general

$$H^n = \sum |\phi_i\rangle \lambda_i^n \langle \phi_i| \tag{6.38}$$

and we can define a function of an operator by its Taylor expansion,

$$f(H) = \sum_n \frac{f^{(n)}(0)}{n!} H^n \tag{6.39}$$

$$= \sum_i |\phi_i\rangle \sum_n \frac{f^n(0)}{n!} \lambda_i \langle \phi_i| \tag{6.40}$$

$$= \sum_i |\phi_i\rangle f(\lambda_i) \langle \phi_i|, \tag{6.41}$$

provided that the set is complete (so we can change the order of summation over i, n) and the λ_is are in the radius of convergence of f (so we can perform the sum over n).

We can then define the inverse of a Hermitian operator acting on an eigenvector of H:

$$H^{-1} |\phi_i\rangle = \frac{1}{\lambda_i} |\phi_i\rangle, \tag{6.42}$$

which is valid so long as *none of the eigenvalues are zero*. Then its action on any vector is well-defined since the $|\phi_i\rangle$ are complete.

Inserting the identity, we can also find that

$$\langle \phi | H | \phi \rangle = \sum \langle \phi | \phi_i \rangle \lambda_i \langle \phi_i | \phi \rangle = \sum_i \lambda_i |\langle \phi_i | \phi \rangle|^2. \tag{6.43}$$

When our states are normalized, this gives us a probabilistic interpretation of the expectation value $\langle \phi | H | \phi \rangle$ in terms of the probability $|\langle \phi_i | \phi \rangle|^2$ of measuring λ_i.

Example 6.44. Suppose we have three masses attached on springs, with two small masses m at the sides and a larger mass M in the center. The spring constants are k. Label their positions by x_1, x_2, x_3. Then the potential is

$$V = \frac{k}{2}\Big[(x_1 - x_2)^2 + (x_2 - x_3)^2\Big], \tag{6.45}$$

and the restoring forces are given by

$$\ddot{x_1} = -\frac{k}{m}(x_1 - x_2) \tag{6.46}$$

$$\ddot{x_2} = -\frac{k}{M}(x_2 - x_1) - \frac{k}{M}(x_2 - x_3) \tag{6.47}$$

$$\ddot{x_3} = -\frac{k}{m}(x_3 - x_2) \tag{6.48}$$

[10] In the continuous case we must also be careful about questions of convergence in the inner products so they are well-defined, but this otherwise totally defines the procedure.

We take the ansatz

$$x_i \sim e^{i\omega t}|\phi_i\rangle, \tag{6.49}$$

so that all the derivatives become $-\omega^2$s, and then we can easily write this as a matrix equation. The solution for the normal modes reduces to finding the eigenvalues of the matrix

$$\begin{pmatrix} \frac{k}{m} & -\frac{k}{m} & 0 \\ -\frac{k}{M} & \frac{2k}{M} & -\frac{k}{M} \\ 0 & -\frac{k}{m} & \frac{k}{m} \end{pmatrix}. \tag{6.50}$$

Note this matrix is normal (it commutes with its transpose) but not hermitian. That's the standard way to solve this. But we'll try something different. We have translational symmetry– the forces depend only on relative separations. Hence there is an $\omega^2 = 0$ mode corresponding to the eigenvector $(1, 1, 1)$.

There is also another momentum-conserving mode $(1, 0, -1)$ corresponding to the outer two masses moving, and we can read off the eigenvalue immediately: $\omega^2 = k/m \implies \omega = \sqrt{k/m}$.

Finally, there is a third mode. It is orthogonal to the other two, and it takes more work to find: $(1, -2m/M, 1)$, with eigenvalue $\frac{k}{m} + \frac{2k}{M}$. Hence we could take a cross product and normalize, or we could use Gram-Schmidt to calculate the final eigenvector.

Lecture 7.

Wednesday, October 16, 2019

Today we'll look at roughly the material from Arfken 7.1-7.4. The topic of this section will be ordinary differential equations. We basically know that the way to solve differential equations is to either perform some integral transformation to make it easier or alternately to know a special function which solves the equation.

A linear ordinary differential equation takes the form

$$\mathcal{L}y(x) = f(x), \tag{7.1}$$

where $\mathcal{L}$ is a linear operator,

$$\mathcal{L} = p_0(x) + p_1(x)D + \ldots + p_n(x)D^n \tag{7.2}$$

where $D = \frac{d}{dx}$ and f is the source function. A solution $y(x)$ can be written as

$$y(x) = \sum_{i=1}^{n} c_i y_n^{(i)}(x) + y_p(x), \tag{7.3}$$

where the y_n are solutions to the homogeneous equation (i.e. they are in the kernel of the operator $\mathcal{L}$, $\mathcal{L}y_n(x) = 0$) and y_p is the "particular solution" fitting the source, $\mathcal{L}y_p(x) = f(x)$. It's a general fact that $\dim(\ker\mathcal{L}) = n$, i.e. there are n linearly independent solutions to the homogeneous equation.

First-order equations A first order equation looks like

$$y'(x) = f(x, y) = -\frac{P(x, y)}{Q(x, y)}, \tag{7.4}$$

where P and Q are some functions. The simplest case is

$$y' = -\frac{P(x)}{Q(y)}, \tag{7.5}$$

which is *separable* and has implicit solution

$$\int dy\, Q(y) = -\int dx\, P(x). \tag{7.6}$$

But we can also rearrange this equation without integrating it[11] to get

$$P(x, y)dx + Q(x, y)dy = 0. \tag{7.7}$$

[11] Strictly, this manipulation is justified by differential forms.

Suppose there exists $\phi(x,y)$ s.t.

$$d\phi(x,y) = P(x,y)dx + Q(x,y)dy = 0. \tag{7.8}$$

Then it must be that $\phi(x,y) =$ constant, which provides an implicit relation between x and y. Notice that since

$$P = \frac{\partial\phi}{\partial x}, Q = \frac{\partial\phi}{\partial y}, \tag{7.9}$$

it must be that since mixed partials commute,

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} \tag{7.10}$$

is a necessary and sufficient condition for the equation to be exact. In the language of differential forms, we may say that the equation is closed,

$$d\omega = Pdx + Qdy = 0, \tag{7.11}$$

but that closed does not imply exact, $\omega = d\mu$ for some μ. (Exact implies closed since $d^2 = 0$.)

However, if the equation is not exact, we could think of multiplying by some unknown function $\alpha(x,y)$ such that it becomes exact:

$$\alpha(x,y)P(x,y)dx + \alpha(x,y)Q(x,y)dy = 0. \tag{7.12}$$

In general there's no formula for finding such an α, but there is one case where we can do this, namely the linear case. Let

$$y' + p(x)y = Q(x). \tag{7.13}$$

We then write

$$D(\alpha y) = \alpha' y + \alpha y', \tag{7.14}$$

If we multiply our linear equation by

$$\alpha(x) = \exp\left(+\int^x dx'\, p(x')\right) \tag{7.15}$$

then we claim the equation becomes exact (a total derivative), since

$$D(\alpha y) = p(x)\alpha y + \alpha y'. \tag{7.16}$$

Hence

$$D[\alpha y] = \alpha(x)Q(x), \tag{7.17}$$

so we can just integrate:

$$y(x) = \frac{1}{\alpha(x)}\left[\int^x dx'\, \alpha(x')Q(x') + C\right]. \tag{7.18}$$

We can think of this as a the sum of a homogeneous solution and a particular solution, i.e. when $Q(x) = 0$ we just get $y(x) = \frac{C}{\alpha(x)}$ and there is a particular solution from dealing with $Q(x) \neq 0$.

Second-order equations Let us consider the case of second-order equations with constant coefficients,

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f(x). \tag{7.19}$$

There are various ways of solving such equations. As before, we can solve the homogeneous equation and then add back in the particular solution turning on the source term. There are some methods in the literature such as the method of undetermined coefficients or variation of parameters.

We can denote our equation of order n by

$$\frac{d^ny}{dx^n} + a_{n-1}\frac{d^{n-1}y}{dx^{n-1}} + \ldots + a_0 y = f(x), \tag{7.20}$$

where we have divided through by the coefficient of D^n since it must be nonzero.[12] In the simplified notation

$$p_n(D)y = f(x), \quad p_n(D) = D^n + a_{n-1}D^{n-1} + \ldots \tag{7.21}$$

[12]Otherwise, the equation would be of some lower order. The definition of it being order n is that the lowest non-vanishing derivative term is D^n.

Then let us note that derivatives obey the following identities[13]

$$\begin{aligned} D + \lambda &= e^{-\lambda x} D e^{\lambda x} \\ D - \lambda &= e^{\lambda x} D e^{-\lambda x} \\ (D - \lambda)^n &= e^{\lambda x} D^n e^{-\lambda x}. \end{aligned}$$

This is equivalent to writing that

$$e^{-\lambda x} D(e^{\lambda x} y) = y' - \lambda y. \tag{7.22}$$

Now the homogeneous second order equation takes the form

$$(aD^2 + bD + c)y = 0. \tag{7.23}$$

We know how to factor expressions. If we can factor the differential operator, then the solutions are straightforward because we know what $D \pm \lambda$ does.

The nature of solutions depends on how nicely the operator factors. There are three cases depending on the discriminant

$$\Delta = b^2 - 4ac. \tag{7.24}$$

If $\Delta > 0$ then the equation has two real roots λ_1, λ_2, so that

$$(D - \lambda_1)(D - \lambda_2)y = 0, \tag{7.25}$$

with solution

$$y = Ae^{\lambda_1 x} + Be^{\lambda_2 x}. \tag{7.26}$$

If $\Delta < 0$ then there are two complex roots $\lambda_\pm = \mu \pm i\omega$, so that the equation can be rewritten as[14]

$$\left[(D - \mu)^2 + \omega^2\right] y = 0. \tag{7.27}$$

Recall that $(D^2 + \omega^2)y = 0$ has solutions $y = Ae^{i\omega x} + Be^{-i\omega x}$, so then

$$\begin{aligned} p(D)y &= (e^{\mu x} D^2 e^{-\mu x} + \omega^2)y \\ &= e^{\mu x}(D^2 + \omega^2)e^{-\mu x} y = 0 \\ &\implies e^{-\mu x} y = Ae^{i\omega x} + Be^{-i\omega x}. \end{aligned}$$

Hence the solutions are both oscillating and growing or decaying exponentially with time,

$$y = e^{\mu}(Ae^{i\omega x} + Be^{-i\omega x}) \tag{7.28}$$

Finally, we could have $\Delta = 0$, which gives a double root at λ. Then

$$\begin{aligned} p(D)y &= (D - \lambda)^2 y = 0 \\ &= e^{\lambda x} D^2 e^{-\lambda x} y = 0. \end{aligned}$$

But this tells us $e^{-\lambda x} y$ is equal to some function z such that $D^2 z = 0$ (it is in the kernel of the operator D^2), and hence

$$e^{-\lambda x} y = A + Bx, \tag{7.29}$$

which gives

$$y = e^{\lambda x}(A + Bx). \tag{7.30}$$

Equivalently, we have found that $e^{\lambda x}, xe^{\lambda x}$ are in the kernel of $(D - \lambda)^2$. In each of these cases, we look for the kernel of some differential operator and set it equal to a known expression.

Now can we do this for an inhomogenous equation? Indeed we can.

[13] That is, we are using the generator of translations and exponentiating.

[14] A little more slowly: $(D - \lambda_+)(D - \lambda_-)y = (D - \mu - i\omega)(D - \mu + i\omega)y = [(D - \mu)^2 + \omega^2]y$, since this is a difference of squares.

Example 7.31. Consider

$$y'' - 3y' + 2y = e^x, \tag{7.32}$$

so

$$(D^2 - 3D + 2)y = (D-1)(D-2)y = e^x, \tag{7.33}$$

so now we can rewrite this as $(D-1)z(x) = e^x$, with $z = (D-2)y$. This equation has a solution[15] $z(x) = Ae^x + xe^x$ and so

$$(D-2)y = z(x) = Ae^x + xe^x. \tag{7.34}$$

Then[16]

$$y = -e^x - xe^x - Ae^x + Be^{2x}. \tag{7.35}$$

Example 7.36. Suppose we want to find a particular solution to

$$3y'' - 2y' + 6y = 5e^{3x}. \tag{7.37}$$

Then my ansatz might be that the solution should be in the form of some constant times e^{3x}.[17] But alternately, we could invert the differential operator. We know how to do this. With

$$(3D^2 - 2D + 6)y = 5e^{3x}, \tag{7.38}$$

we can write

$$y = \frac{1}{3D^2 - 2D + 6}(5e^{3x}) = \frac{1}{3(3)^2 - 2(3) + 6}5e^{3x} = \frac{5}{27}e^{3x}. \tag{7.39}$$

In general for

$$p_n(D)y = Ae^{\lambda x}, \tag{7.40}$$

if $p_n(\lambda) \neq 0$ then

$$y = \frac{1}{p_n(\lambda)}Ae^{\lambda x} \tag{7.41}$$

and if $p_n(\lambda) = 0$ then

$$y = \frac{1}{p_n'(\lambda)}xAe^{\lambda x}. \tag{7.42}$$

This is known as the exponential input theorem.

General linear 2nd order equation Let's consider the general case,

$$y'' + P(x)y' + Q(x)y = 0. \tag{7.43}$$

This is generally hard to solve. It's a little easier if $P(x), Q(x)$ are analytic except at a finite number of poles. If we wanted to solve the equation near a nonsingular (ordinary) point, we could expand P and Q as power series and solve locally.

Alternately, we might have a *regular singular point* where $P(x)$ has no worse than a single pole ($P(x) \sim 1/x$) and $Q(x)$ has no worse than a double pole ($Q(x) \sim 1/x^2$). Beyond this there are also *irregular singular points* or sometimes essential singularities.

---

[15] This follows because $(D-1)z = e^x$ is the same as $e^x De^{-x}z = e^x$. Solving, we have $De^{-x}z = 1 \implies z = Ae^x + xe^x$.

[16] Let's do this one more time. Our equation is now $e^{2x}De^{-2x}y = Ae^x + xe^x$, so $De^{-2x}y = Ae^{-x} + xe^{-x}$. Now we integrate. The first term is easy, that's just $-Ae^{-x}$. The second term can be done by parts: $\int dx\, xe^{-x} = -xe^{-x} - \int dx\,(-e^{-x}) = -xe^{-x} - e^{-x}$. Thus $e^{-2x}y = -Ae^{-x} - xe^{-x} - e^{-x} + B$, picking up a constant of integration, so $y = -e^x - xe^x - Ae^x + Be^{2x}$. We can also absorb the $-e^x$ term into A if we wish. The inhomogeneous solution is therefore just $-xe^x$, and we have two free parameters A and B since this is a second-order equation.

[17] In general such differential equations will have particular solutions of this form, i.e. if the inhomogeneous part is e^{ax} then the particular solution will probably be Ae^{ax}. If that doesn't work, the next thing to try is Axe^{ax}.

Example 7.44. Consider the following differential equation:

$$x(x-1)y'' + [(1+a+b)x + c]y' + aby = 0, \tag{7.45}$$

the hypergeometric equation. This equation has regular singular points at $x = 0, x = 1$. But we could also analyze the point at $x \to \infty$, where if we replace $x = \frac{1}{z}$ and write $w(z) = y(1/z)$, then with a bit of chain rule, our standard second order equation becomes[18]

$$w'' + \left[\frac{2z - P(1/z)}{z^2}\right]w' + \frac{Q(1/z)}{z^4}w = 0. \tag{7.46}$$

If we analyze the particular form of P and Q here, we find that the hypergeometric equation has a regular singular point at $x = \infty$.[19]

There's one more nice fact about general linear second-order equations, which is that we can generally get rid of the order y term. Set $y(x) = \alpha(x)z(x)$ for some α. In particular if we choose

$$\alpha(x) = \exp\left(-\frac{1}{2}\int^x dx'\, P(x')\right) \tag{7.47}$$

then the equation (exercise) becomes

$$z''(x) + I(x)z(x) = 0 \tag{7.48}$$

for some I given by

$$I(x) = Q(x) - \frac{1}{2}P'(x) - \frac{1}{4}P(x)^2. \tag{7.49}$$

This expression I is actually an invariant– if two different-looking differential equations have the same value of I, they are really the same equation.

Lecture 8.

Monday, October 21, 2019

"Ignorance is not a sin. Refusing to learn is. The problem isn't making a mistake. The problem is defending a mistake."
–Nemanja Kaloper

Today we will continue our discussion of differential equations. A brief note on the history of differential equations– it all was invented/discovered about the same time as calculus by (glorious) Sir Isaac Newton as a mathematical tool to do physics. Differential equations tell us how physical observables change in space and time. The focus was originally on finding solutions. Our equations depend on some variables that are in principle all observable, and when we measure them we can find solutions that describe their dependence on one another.

But one can start to see commonalities between equations. Sometimes the same technique works to solve multiple equations. Some equations are really the same. There are books like Gradshteyn and Ryzhik,[20] and Kamke's book, which have solutions of many, many integrals and such equations. There are also methods of numerical integration to solve equations which are not readily integrable.

Let us note that given a differential equation

$$f(x, y, y', \ldots, y^{(n)}) = 0, \tag{8.1}$$

if this function is sufficiently nice in terms of its highest derivative $y^{(n)}$ ($\delta/\delta y^{(n)} f$ exists), then

$$y^{(n)} = \hat{f}(x, y, y', \ldots, y^{(n-1)}) \tag{8.2}$$

[18] See Arfken 344 for the details.

[19] This is a little bit too quick. Here, $P(x) = \frac{(1+a+b)x+c}{x(x-1)}$. Hence $P(1/z) = \frac{(1+a+b)/z+c}{(1/z)(1/z-1)} = \frac{(1+a+b)z+cz^2}{1-z}$. Now $P(1/z)/z^2$ diverges only as $1/z$ as $z \to 0$. Likewise $Q(1/z) = \frac{ab}{(1/z)(1/z-1)} = \frac{abz^2}{1-z}$, so that $\frac{Q(1/z)}{z^4} \sim 1/z^2$, so these both diverge slowly enough for $x \to \infty$ to be a regular singular point.

[20] https://en.wikipedia.org/wiki/Gradshteyn_and_Ryzhik

gives the highest derivative in terms of the lower derivatives. If indeed our function is given to be sufficiently differentiable then we may write an initial condition given by the first $n-1$ derivatives:

$$x_0 : y_0, y_0', \ldots, y_0^{(n-1)}. \tag{8.3}$$

That is, if we have a Taylor expansion

$$y = \sum_k a_k (x - x_0)^k \tag{8.4}$$

then we can get the next derivative and therefore higher derivatives, provided that the functions are sufficiently differentiable. Hence the problem of finding a solution reduces to the process of differentiation itself.

This line of thinking produced the theorems of Peano and Picard-Lindelöf, which placed constraints on how far we could extend our solution around the initial condition.[21] And we can take our extended solution and extend a bit more. In general we get curves through phase space. If the equation is sufficiently nice, curves in phase space do not intersect. When they do, we call it a singularity. One example of a really bad singularity is $e^{1/z}$ as $z \to 0$. No derivative of this function exists as $z \to 0$.

In general, we know that an nth order differential equation depends on n integration parameters. That fixes what point we start at in phase space. Moreover, sometimes our equations have symmetries, conserved quantities which can reduce the dimension of our problem in phase space. It was Emmy Noether who figured out that continuous symmetries imply the existence of conserved charges. Noether's theorem tells us that the invariance under complex phases of quantum mechanics results in the conservation of probability in quantum mechanics.

Let us consider again the general linear differential equation of order N with constant coefficients,

$$\sum_{n=0}^{N} a_n \left(\frac{d}{dx}\right)^n f(x) = 0. \tag{8.5}$$

We may of course rewrite this as an operator equation in terms of the Heaviside D operators,

$$D = \frac{d}{dx} : f \mapsto f'. \tag{8.6}$$

As long as f is sufficiently differentiable, its image under D exists. The equation now looks like a polynomial equation, and this suggests we can factor it:

$$\prod_{k=1}^{N} (D - \alpha_k) f = 0, \tag{8.7}$$

in terms of N (possibly degenerate) roots α_k, such that

$$\sum_{n=0}^{N} a_n \alpha_k^n = 0. \tag{8.8}$$

Note the index on the product– a polynomial of order N should have N roots.

Now take one factor in the expansion,

$$(D - \alpha_k) f \equiv e^{\alpha_k x} \left(D e^{-\alpha_k x} f\right). \tag{8.9}$$

Hence we can rewrite our equation as

$$\prod_{k=1}^{N} \left(e^{\alpha_k x} D e^{-\alpha_k x}\right) f \tag{8.10}$$

Note that we can certainly evaluate these derivatives in whatever order we want, since derivatives commute:

$$[(D - \alpha_k), (D - \alpha_n)] = 0. \tag{8.11}$$

[21] There's a very readable write-up on the Picard-Lindelöf theorem here: https://ptolemy.berkeley.edu/projects/embedded/eecsx44/lectures/Spring2013/Picard.pdf with a nice little sojourn into the world of operator fixed-point theory. Would recommend if you have a basic understanding of mathematical analysis.

In order for the big product to vanish, we must have $(D-\alpha_k)f = 0$ for any such root (assuming they are nondegenerate). Hence

$$0 = (D-\alpha_k)f \equiv e^{\alpha_k x}\left(De^{-\alpha_k x}f\right) \implies f \propto e^{\alpha_k x}, \tag{8.12}$$

so we may write a general solution as

$$f = \sum_{k=1}^{N} c_k e^{\alpha_k x}, \tag{8.13}$$

in the case of *nondegenerate roots*.

What about when roots are degenerate? Then we write

$$c_1 e^{\alpha x} + c_2 x e^{\alpha x}, \tag{8.14}$$

the exponential times a linear polynomial in x. In general a root of degeneracy n will have an exponential times a polynomial in x of order $n-1$. For notice that

$$0 = (D-\alpha)(D-\alpha)f = \left[e^{\alpha x}De^{-\alpha x}e^{\alpha x}De^{-\alpha x}\right]f = e^{\alpha x}D^2e^{-\alpha x}f. \tag{8.15}$$

Hence the exponential factor $e^{\alpha x}$ will cancel the dependence $e^{-\alpha x}$ from the derivatives and we have $e^{\alpha x}D^2g = 0$ where $g = e^{-\alpha x}f$. We can solve this very easily; $D^2g = 0 \implies g$ is a linear polynomial in x. This sort of behavior shows up in resonance in classical mechanics, as well as in Pais & Uhlenbeck effects in quantum mechanics.

We see that the homogeneous linear ODE with constant coefficients is therefore completely solved. Note that we will not discuss nonlinear ODEs for the time being, i.e. equations involving differential operators F such that

$$F(f+g) \neq F(f) + F(g). \tag{8.16}$$

For instance, consider the equation

$$y' + ay + by^2 + c = 0. \tag{8.17}$$

If a, b, c are all constants then this equation is separable. If $c = 0$ we can change variables and perhaps find a solution by dividing by y^2 and rewriting in terms of $w = -1/y$ as

$$\frac{1}{y^2}\frac{dy}{dx} + \frac{a}{y} + b = 0, \tag{8.18}$$

where this is now just linear. But when a, b, c depend on x, this is much harder to solve.

If we can find a *particular* solution to the inhomogeneous equation, then we can write $y = \frac{1}{w} + y_p$ where y_p is the particular solution, such that

$$w' = (a + 2by_p)w + b. \tag{8.19}$$

That is, solving for the particular solution allows us to turn the remaining equation into a linear equation.

But now we add a cubic term,

$$y' + ay + by^2 + dy^3 + c = 0. \tag{8.20}$$

This is called *Abel's equation*. There is no general solution known.

Second-order equations, revisited Let us write our general linear equation where the coefficients may now depend on x:

$$\sum_n a_n(x)D^n y = 0. \tag{8.21}$$

Let us suppose that there exists $y_1(x)$ a solution, i.e. it is in the kernel of this differential operator. Suppose moreover that $y_2(x)$ is also a solution. Because the operator is still linear, we can clearly add solutions, i.e.

$$c_1 y_1(x) + c_2 y_2(x) \tag{8.22}$$

is also a solution or more generally

$$\sum_{p=1}^{N} c_p y_p \tag{8.23}$$

for some set of solutions y_p. This is the generalization of the fact that $D^2y = 0$ has solutions $y_1 = 1$ and $y_2 = x$, so the general solution is $c_1y_1 + c_2y_2 = c_1 + c_2x$.

One can then ask how many solutions we need to entirely span the space of solutions. The way to do this is to (re)introduce the notion of linear independence. That is, if there exists c_p not all zero such that $\sum_{p=1}^{N} c_p y_p = 0$ at *every point in the interval*, then the y_ps are linearly *dependent*. If no such c_p exist, then the y_p are linearly independent.[22]

Notice that for a set of functions to be linearly dependent, we require that $\sum c_p y_p = 0$ at *every point* x in the interval, which means in particular that their derivatives in this linear combination must also vanish. That is,

$$\sum c_p y_p(x) = 0 \implies \sum c_p y_p'(x) = 0 \tag{8.24}$$

and in general

$$\sum c_p y_p^{(N-1)}(x) = 0. \tag{8.25}$$

But now this is just a system of N linear equations on the coefficients c_p:

$$0 = \begin{pmatrix} y_1 & y_2 & \cdots & y_n \\ y_1' & y_2' & & \\ \vdots & \vdots & \ddots & \\ y_1^{(N-1)} & y_2^{(N-1)} & \cdots & y_N^{(N-1)} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix}. \tag{8.26}$$

The condition for a nontrivial solution of c_ps to exist (linear dependence) is that this matrix must be singular, i.e.

$$\begin{vmatrix} y_1 & y_2 & \cdots & y_n \\ y_1' & y_2' & & \\ \vdots & \vdots & \ddots & \\ y_1^{(N-1)} & y_2^{(N-1)} & \cdots & y_N^{(N-1)} \end{vmatrix} = 0. \tag{8.27}$$

Otherwise, the matrix would be invertible and force all the coefficients to be zero, which is the condition for linear independence.[23] This determinant is called the *Wronskian*. One can then check that if we write the set of solutions $y_1, \ldots, y_N$ and compute the Wronskian, it will be non-vanishing. If we add a y_{N+1} solution, we must get linear dependence; the Wronskian will vanish.

Conversely, sometimes it is easy to find one solution to our differential equation and hard to find a second. In that case, we may use the Wronskian to find the second solution. We'll do this later.

Let us now specialize to the second-order case,

$$y'' + P(x)y' + Q(x)y = J(x), \tag{8.28}$$

where J is some source term. Our game plan will be as follows. We first solve the homogeneous equation for two solutions y_1, y_2, setting J to zero, and then add in a particular solution y_p.[24]

For the homogeneous case, suppose that $P(x), Q(x)$ are both infinitely differentiable. Then we may suppose that

$$y = \sum_{k=0}^{\infty} a_k x^k \tag{8.29}$$

has a Taylor expansion, and we can find a recursion relation between the coefficients a_k. In fact, it is not required that P and Q be infinitely differentiable; they are only required to not blow up too badly. So long as the quantities $(x - x_0)P(x)$ and $(x - x_0)^2 Q(x)$ remain finite at some singular points x_0, we can show that there exists at least one solution that satisfies a form similar to Taylor expansion, but with a compensating factor in order to make the equation integrable even at the singularities. That is, a solution exists of the form $y = \sum_{k=0}^{\infty} a_k x^{k+s}$ for $s \in \mathbb{R}$.

Such points x_0 are called regular singular points. Anything that diverges worse than this gives us irregular singular points. We may note that $(x - x_0)^2 Q(x)$ is somewhat like Gauss's law. We have a thing that's divergent but integrating it over some "area" in the limit makes it regular, like a point charge.

[22] In other words, linear dependence says that we can build a perfect match for (at least) one of the y_ps out of linear combinations of all the others, so there is redundancy in our set of solutions.

[23] Technically, a nonzero Wronskian implies linear independence. A zero Wronskian does not by itself imply linear dependence, see https://en.wikipedia.org/wiki/Wronskian#The_Wronskian_and_linear_independence.

[24] There are degenerate cases where y_p is proportional to one of the y_is, but we'll consider them later.

Lecture 9.

Wednesday, October 23, 2019

"I like to reward you if you want to be living dangerously and be smart like that."
–Nemanja Kaloper

Last time, we were discussing second-order equations of the form

$$y'' + P(x)y' + Q(x)y = 0. \tag{9.1}$$

Recall that two functions are linearly dependent if a nontrivial linear combination of those functions vanishes at every point on the interval:

$$c_1 y_1(x) + c_2 y_2(x) = 0 \quad \forall x. \tag{9.2}$$

Since this holds at the level of functions, if our solutions are differentiable then we may take derivatives of this expression to get

$$c_1 y_1'(x) + c_2 y_2'(x) = 0 \tag{9.3}$$

and then we can compute the Wronskian

$$W = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} \stackrel{?}{=} 0. \tag{9.4}$$

Of course in the case of two solutions, this can be written down explicitly as $y_1 y_2' - y_2 y_1'$. Let us restrict to solutions to Eqn. 9.2, and let us also take the derivative of the Wronskian,

$$W' = (y_1 y_2' - y_2 y_1')' = y_1 y_2'' + y_1' y_2' - y_2 y_1'' - y_1' y_2'. \tag{9.5}$$

But notice that the first derivative terms now cancel, and we can go back to Eqn. 9.2 to find expressions for y_i''. That is,

$$W' = -P y_1 y_2' - Q y_1 y_2 + P y_2 y_1' + Q y_1 y_2. \tag{9.6}$$

The Q terms drop out and what we see is that

$$W' = -P(y_1 y_2' - y_2 y_1') = -P(x)W. \tag{9.7}$$

Hence this is a separable differential equation for the Wronskian W. That is,

$$\frac{W'}{W} = -P \implies W(x) = W(x_0) e^{-\int_{x_0}^{x} dx\, P(x)}. \tag{9.8}$$

That is, the Wronskian will generically depend on x for some general $P(x)$; the only way (almost) this can vanish is if it initially vanishes at some x_0, i.e. $W(x_0) = 0$. This tells us something we already knew– two functions which are linearly dependent at one point in the interval must indeed be linearly dependent everywhere.

There is one caveat– what if $P(x)$ has a singular point? For instance, $P(x) \sim \frac{\alpha}{x - x_0}$. Thus the integral of $P(x)$ is a log, and taking the exponential of the log, we could get a Wronskian

$$W(x) = W(x_0)(x - x_0)^{\alpha}, \tag{9.9}$$

which simply tells us that regular singular points act like charges and will give us the equivalent of field lines beginning/ending on charges.

So this tells us that the vanishing of the Wronskian does imply linear dependence in the second-order case. But can we construct a third linearly independent solution? Take some third solution y, with

$$\begin{aligned} c y + c_1 y_1 + c_2 y_2 &= 0 \\ c y' + c_1 y_1' + c_2 y_2' &= 0 \\ c y'' + c_1 y_1'' + c_2 y_2'' &= 0. \end{aligned}$$

Then its Wronskian is

$$W = \begin{vmatrix} y_1 & y_2 & y \\ y_1' & y_2' & y' \\ y_1'' & y_2'' & y'' \end{vmatrix} = \begin{vmatrix} y_1 & y_2 & y \\ y_1' & y_2' & y' \\ -P y_1' - Q y_1 & -P y_2' - Q y_2 & -P y' - Q y \end{vmatrix}, \tag{9.10}$$

where in the third row we have used the fact that y_1, y_2, and y are all solutions to the differential equation. Now we recall an important fact about determinants. The determinant picks up sign changes under interchange of rows and columns, and more generally, if any row/column is a linear combination of the other rows/columns, then the whole determinant vanishes. That is, because the third row is a linear combination of the other two, $W = 0$ for three functions.

We conclude that $\exists c, c_1, c_2$ not all zero such that

$$cy + c_1 y_1 + c_2 y_2 = 0. \tag{9.11}$$

Suppose we had y_1, y_2 in hand linearly independent. In particular, $c \neq 0$ since if it were zero, this would reduce to the previous case and imply y_1, y_2 were linearly dependent. Hence

$$y = -c_1 y_1 - c_2 y_2, \tag{9.12}$$

so y is a linear combination of the other two solutions y_1, y_2. We see that there are two degrees of freedom for us to fix, and we can do this by using initial conditions.

Our general solution is a linear combination of the two linearly independent solutions y_1, y_2. How do we choose these solutions? In a way that makes our lives easiest.[25]

The Wronskian method for a second solution Suppose we have one solution in hand, some y_1 satisfying 9.2. We can compute the Wronskian,

$$W(x) = W(x_0) e^{-\int_{x_0}^{x} dx\, P(x)} = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}. \tag{9.13}$$

Let us instead divide through by $W(x_0)$, which we can take to be nonzero since we're looking for another linearly independent function. Hence we can write

$$W(x)/W(x_0) = e^{-\int_{x_0}^{x} dx\, P(x)} = \begin{vmatrix} y_1/W(x_0) & y_2 \\ y_1'/W(x_0) & y_2' \end{vmatrix} = \begin{vmatrix} \tilde{y}_1 & y_2 \\ \tilde{y}_1' & y_2' \end{vmatrix}, \tag{9.14}$$

where we've absorbed the constant $1/W(x_0)$ into $\tilde{y}_1$. What's left is again the Wronskian. We have an equation

$$y_1 y_2' - y_2 y_1' = \exp\left(- \int_{x_0}^{x} dx\, P(x) \right). \tag{9.15}$$

And now we see that our problem reduces to solving a first-order equation, which we may rewrite as

$$y_2' - y_2 y_1'/y_1 = \exp\left(- \int_{x_0}^{x} dx\, P(x) \right)/y_1. \tag{9.16}$$

Let us moreover write $y_2 = u y_1$ in terms of some unknown function u, such that $y_2' = u' y_1 + u y_1'$ and then

$$u' y_1 + u y_1' - u y_1' = u' y_1 = \frac{e^{-\int_{x_0}^{x} dx\, P(x)}}{y_1}, \tag{9.17}$$

and therefore we see that

$$u' = \frac{e^{-\int_{x_0}^{x} dx\, P(x)}}{y_1^2(x)}, \tag{9.18}$$

which is separable with general solution

$$U(x) = U_0 + \int_{x_0}^{x} dx' \, \frac{e^{-\int_{x_0}^{x'} dx''\, P(x'')}}{y_1^2(x')}. \tag{9.19}$$

Plugging back into our expression $y_2 = u y_1$, we have

$$y_2 = y_1(x) \int_{x_0}^{x} dx' \, \frac{e^{-\int_{x_0}^{x'} dx''\, P(x'')}}{y_1^2(x')}, \tag{9.20}$$

[25] For instance, we could write a solution in terms of sines and cosines or complex exponentials– totally equivalent. For circuits, the complex exponential might be nicer; for some real waves, sines and cosines might be better.

where we have WLOG dropped the U_0 term since that term is simply a multiple of our old solution y_1. If we like, we're just Gram-Schmidting away the U_0 term. In general we might like to have some nontrivial U_0 in order to make these solutions y_1, y_2 orthonormal.

Let us note also that if we sit at a regular point (e.g. $x = 0$) then our second-order equation gives us a recursion relation on the expansion coefficients of

$$y = \sum a_n x^n. \tag{9.21}$$

Just take derivatives and we get equations relating $y^{(n)}, y^{(n-1)}, y^{(n-2)}$ and so on. It turns out that many special functions solving (physically) interesting differential equations are simply special cases of 9.2 where P and Q are polynomials of no higher than second order. These can be rewritten as examples of the hypergeometric equation, which has known and catalogued solutions.[26]

We should also note that sometimes we must consider the point at infinity, i.e. as $x \to \infty$, define $z = 1/x$ and rewrite the equation using the chain rule so that

$$y' = \frac{dy}{dx} = \frac{dy}{dz}\frac{dz}{dx} = \frac{dy}{dz}\frac{1}{\frac{dx}{dz}} = -z^2\frac{dy}{dz}, \tag{9.22}$$

and something similar holds for y'',

$$y'' = \frac{d}{dz}\left(-z^2\frac{dy}{dz}\right). \tag{9.23}$$

Hence there might be singular points at ∞ in the $z \to 0$ limit. We get a new equation in terms of z and some $\bar{P}, \bar{Q}$ which are made of the original functions:

$$\bar{P} = \frac{2z - P(1/z)}{z^2}, \quad \bar{Q} = \frac{Q(1/z)}{z^4}, \tag{9.24}$$

which tells you that P cannot diverge worse than linearly and Q cannot diverge worse than quadratically in order to maintain regular singular points.

Introducing the series method Let us now try to solve the harmonic oscillator potential

$$y'' + \omega^2 y = 0 \tag{9.25}$$

by a series method, i.e. we take the solution to have a Taylor expansion

$$y = \sum_{n=0}^{\infty} a_n x^n. \tag{9.26}$$

We have already solved it by our previous methods of factorization since the equation has constant coefficients, but this is the simplest interesting case to practice a new method on. We can take derivatives of this Taylor expansion to find

$$y'' = \sum_{n=0}^{\infty}(n)(n-1)a_n x^{n-2} = \sum_{n=2}^{\infty}(n)(n-1)a_n x^{n-2}, \tag{9.27}$$

since the first two terms are really zero. We redefine the dummy index

$$n = m + 2, \tag{9.28}$$

such that

$$y'' = \sum_{m=0}^{\infty}(m+2)(m+1)a_{m+2}x^m. \tag{9.29}$$

But m is just a dummy variable, so we can relabel it to n and plug back into our harmonic oscillator equation, Eqn. 9.25. Since we have two sums, each of which are convergent and running over the same domain, we can now combine them and compare terms:

$$0 = \sum_{n=0}^{\infty}\Big((n+2)(n+1)a_{n+2} + \omega^2 a_n\Big)x^n. \tag{9.30}$$

[26] A nice reference is L. Elsgolts on differential equations. "It's very good, very clear, very methodical." –Nemanja. As for physical applications, Born & Wolfe wrote a book on optics (really wave mechanics) and this is also on the internet, probably.

And since the x^n are linearly independent, we can get rid of the sum and look at the *recursion relation* between coefficients:

$$a_{n+2} = -\frac{\omega^2}{(n+2)(n+1)} a_n. \tag{9.31}$$

Notice that the first two coefficients are set by

$$y(0) = a_0, \quad y'(0) = a_1. \tag{9.32}$$

All other coefficients are then given by these two.

Notice that the original equation 9.25 is in fact invariant under parity, $x \to -x$. Hence our solutions separate into even and odd solutions. In particular, if we write down the recursion relations for a_{2n+2} and $a_{(2n+1)+2}$, we get precisely the expansion coefficients for sines and cosines. Hence

$$a_{2n} = \frac{(-1)^n \omega^{2n}}{(2n)!} a_0, \tag{9.33}$$

so that

$$y = a_0 \sum_{n=0}^{\infty} \frac{(-1)^n \omega^{2n}}{(2n)!} x^{2n} = a_0 \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} (\omega x)^{2n} = a_0 \cos(\omega x), \tag{9.34}$$

and similarly the other solution is a sine,

$$y = a_0 \cos(\omega x) + a_1 \sin(\omega x). \tag{9.35}$$

Lecture 10.

Monday, October 28, 2019

"Dumb things are good because they're repeatable. Just don't do something dumb too often if you can do something clever on the side."
–Nemanja Kaloper

Last time, we began discussing series solutions to differential equations. We considered solutions to the classical harmonic oscillator,

$$y'' + \omega^2 y = 0, \tag{10.1}$$

and made the ansatz that such equations have a series expansion

$$y = \sum_n a_n x^n. \tag{10.2}$$

We got the recursion relation

$$a_{n+2} = -\frac{\omega^2}{(n+2)(n+1)} a_n, \tag{10.3}$$

which gave us an expression for the coefficients,

$$a_{2n+1} = \frac{(-1)^n \omega^{2n+1}}{(2n+1)!}, \tag{10.4}$$

and we got solutions which were sines and cosines,

$$y_1 = a_0 \cos(\omega x), \quad y_2 = a_1 \sin(\omega x). \tag{10.5}$$

Let us notice that if $y(x)$ and $y(-x)$ both solve a differential equation, then we can create the even and odd (symmetric and antisymmetric) combinations

$$y_\pm = y(x) \pm y(-x). \tag{10.6}$$

This sort of trick works for potentials that are themselves even.

The Frobenius method Let us now try to extend this power series trick. Maybe our differential equation includes regular singular points and forces us to start our series expansion at the power x^s so that we get terms $x^s, x^{s+1}, \ldots$. Here $s \in \mathbb{R}$ but a priori, we do not require s to be an integer. Its values will in general be fixed by an *indicial equation*, as we'll see shortly.

That is, let us write a series expansion

$$y = \sum_{n=0}^{\infty} a_n x^{n+s}. \tag{10.7}$$

This is called the *Frobenius method*.

Using this method, we could solve equations of the form

$$y'' + \frac{\alpha}{x} y' + \frac{\beta}{x^2} y = 0 \tag{10.8}$$

in this way, where primes indicate derivatives with respect to x. This is called the *Euler equation*. Note that if we replace the independent variable with $x = e^t$ then solutions take the form e^{pt} or in terms of the original variable x^p for some p.[27]

Let us try to apply this to the harmonic oscillator. Thus

$$y'' = \sum_{n=0}^{\infty} (n+s)(n+s-1) a_n x^{n+s-2}. \tag{10.9}$$

Hence the differential equation becomes

$$0 = \sum_{n=0}^{\infty} (n+s)(n+s-1) a_n x^{n+s-2} + \omega^2 \sum_{n=0}^{\infty} a_n x^{n+s} \tag{10.10}$$

$$= \sum_{n=-2}^{\infty} (n+s+2)(n+s+1) a_{n+2} x^{n+s} + \omega^2 \sum_{n=0}^{\infty} a_n x^{n+s}. \tag{10.11}$$

Notice that the first sum here now comes with two extra terms,

$$s(s-1) a_0 x^s + (s+1)(s) a_1 x^{s+1}, \tag{10.12}$$

Hence we have

$$0 = s(s-1) a_0 x^s + (s+1)(s) a_1 x^{s+1} + \sum_{n=0}^{\infty} \Big[(n+s+2)(n+s+1) a_{n+2} + \omega^2 a_n \Big] x^{n+s}. \tag{10.13}$$

The series term gives us almost the ordinary recursion relation on the coefficients a_n, up to a shift by s. And we know that because the individual x^ns are orthogonal, the coefficients in terms of s and a_0, a_1 must themselves vanish:

$$s(s-1) a_0 = 0 \tag{10.14}$$

$$(s+1) s a_1 = 0. \tag{10.15}$$

These are called the *indicial equations*. If we take $a_0 = 0, a_1 = 0$ then all higher coefficients are zero.[28] So if $a_0 \neq 0$ then we could either have $s = 0$ or $s - 1 = 0$.

Notice that there's a shared root $s = 0$ between the two indicial equations above. A priori we could have tried setting $s = -1$ in the second equation, which says that the lowest order term of the expansion goes as $1/x$. This is certainly not analytic, which simply forces us to set $a_0 = 0$ and then we get all the same behavior in terms of $a_1, a_3, \ldots$.

---

[27] This is an exercise. Notice that $\log x = t \implies dx/x = dt$. Find t in terms of α and β. If you'd like to check your work, see https://en.wikipedia.org/wiki/Cauchy%E2%80%93Euler_equation. There are actually three (two and a half) ways you could solve this. One is to guess the solution as some x^m and then figure out m from the equation. The next is to make the change of variables suggested and then solve the equation with constant coefficients. The last is to assume a Frobenius series solution. You won't get an indicial equation, and you actually won't get a recursion relation. But you will get a constraint involving n, s, α, and β. Since it must hold for all n, you can then take $n = 0$ to find a value for s. This reduces to the first way, so it only counts as half.

[28] "We have just committed suicide by zero." –Nemanja

Hence there are really only two interesting roots, $s = 0$ and $s = 1$. The recursion relation becomes the same as in the ordinary series expansion of the harmonic oscillator and the sines and cosines fall out as before.

Bessel's equation The equation

$$x^2y'' + xy' + (x^2 - \nu^2)y = 0 \tag{10.16}$$

is called *Bessel's equation*. Normally ν is an integer, but we may as well take it to be a real number. We shall see that as ν becomes an integer, something goes "very bad in a very fun way."

Let's write our series expansion in the Frobenius way. Thus

$$y = \sum_{n=0} a_n x^{n+s} \tag{10.17}$$

$$y' = \sum_{n=0} (n+s) a_n x^{n+s-1} \tag{10.18}$$

$$y'' = \sum_{n=0} (n+s)(n+s-1) a_n x^{n+s-2}. \tag{10.19}$$

Notice that if the x^2y term were zero, we would just get back the Euler equation from earlier. Note also that this equation enjoys a scale symmetry of sorts– as x scales, y' scales inversely and so the first two terms of this equation are scale invariant under $x \to \alpha x$.

This equation has a regular singular point at $x = 0$, as we can see if we put the equation in the standard form. Substituting our Frobenius expansion, we have

$$0 = \sum_{n=0}^{\infty} \{x^{n+s} a_n \Big[(n+s)(n+s-1) + (n+s) - \nu^2\Big] + x^{n+s+2} a_n\} \tag{10.20}$$

$$= \sum_{n=0}^{\infty} \{x^{n+s} a_n \Big[(n+s)^2 - \nu^2\Big] + x^{n+s+2} a_n\}. \tag{10.21}$$

We can see that since one of these stars at order $s+2$ and the other starts at order s, we will get again our "orphans" and therefore indicial equations. That is, let us rewrite as

$$0 = x^s a_0 (s^2 - \nu^2) + x^{s+1} a_1 \Big[(s-1)^2 - \nu^2\Big] + \sum_{n=0}^{\infty} \{a_{n+2} \Big[(n+s+2)^2 - \nu^2\Big] + a_n\} x^{n+s+2}. \tag{10.22}$$

These first terms have no counterpart, so their coefficients must identically vanish. Hence

$$s_{\pm} = \pm\nu \tag{10.23}$$

are the solutions to the indicial equation. This generally forces $a_1 = 0$.[29] Take $s = +\nu$ and label the coefficients as a_n^+. Then

$$a_{n+2}^+ = -\frac{a_n^+}{(n+\nu+2)^2 - \nu^2} \tag{10.24}$$

is the recursion relation for the coefficients. We can certainly factorize this denominator as

$$a_{n+2}^+ = -\frac{a_n^+}{(n+2)(n+2+2\nu)}. \tag{10.25}$$

In the first term, a ν cancelled out. The same thing would have happened if we took the negative root instead. We took a_1 to be zero, so the first few even coefficients are

$$a_2 = \frac{-1}{2(2+2\nu)} a_0 \tag{10.26}$$

and

$$a_4 = \frac{-1}{(2+2)(4+2\nu)} a_2 = \frac{(-1)^2}{(2\cdot 2\cdot 2\cdot 1)(2+2\nu)(4+2\nu)} a_0. \tag{10.27}$$

If we keep going, we get in general

$$a_{2n} = \frac{(-1)^n}{2^{2n} n! (1+\nu)(2+\nu)\dots(n+\nu)} a_0. \tag{10.28}$$

[29] I'm not sure what happens if $\nu = 1/2$.

Suppose ν was an integer N. Then the sequence $(1+\nu)(2+\nu)\dots(n+\nu)$ is almost a factorial; it is

$$\frac{1}{(1+\nu)(2+\nu)\dots(n+\nu)} = \frac{N!}{(n+N)!}. \tag{10.29}$$

Hence the series expansion is

$$J_N(x) = \sum_n \frac{(-1)^n N!}{n!(n+N)!}\left(\frac{x}{2}\right)^{N+2n}. \tag{10.30}$$

More generally these factorials are replaced by gamma functions,

$$J_\nu = \sum_n \frac{(-1)^n \Gamma(\nu+1)}{n!\Gamma(n+\nu+1)}\left(\frac{x}{2}\right)^{\nu+2n}, \tag{10.31}$$

where the Γ function is the generalization of the factorial (its argument is shifted by 1 by convention).

Notice that the factor $N!$ (in the integer case) or $\Gamma(\nu+1)$ (in the real case) can be pulled out, so we can normalized. Moreover, when ν is not an integer, we get the $-\nu$ solutions for free. That is, we have a second set of solutions!

However, this is *not* the case when $\nu = N$ is an integer. For at some point, $n+N$ will be zero, and so $1/(n+N)!$ blows up. If you like, this is actually a proof that the Γ function blows up at negative integers. This is the statement that the Γ function has regular singular points.

If we did this carefully, we would in fact find that the second solution for $-N$ would give something proportional to the $+N$ solution. Hence we need to go to the Wronskian method to construct the second solution.

What broke down? The Frobenius method gave of both non-analytic solutions when ν was real but non-integer, since the powers of x appearing in this expansion turn out to be non-integer; it only gave us one analytic solution when ν was an integer. Why? It gave us the non-singular solutions. In general so long as the difference between the roots of the indicial equation is not an integer, the Frobenius method will give us both solutions. This is the content of Fuch's theorem.

When Frobenius fails, we will need to use

$$y_2(x) = y_1(x)\int_a^x dx' \frac{e^{-\int_a^{x'} P(x'')dx''}}{y_1^2(x')} \tag{10.32}$$

to construct the second solution. This might look asymmetric in P and Q but in fact the information from Q is hidden inside y_1; we needed Q to solve for the first solution.

Let us take an equation of the form

$$y'' + P(x)y' + Q(x)y = 0, \tag{10.33}$$

which has a regular singular point at $x = 0$. Hence

$$P(x) = \frac{p_{-1}}{x} + p(x) \tag{10.34}$$

$$Q(x) = \frac{q_{-2}}{x^2} + \frac{q_{-1}}{x} + q(x), \tag{10.35}$$

where $p(x), q(x)$ are analytic.[30]

Let us take a solution

$$y_1 = x^s \sum_n a_n x^n, \tag{10.36}$$

where s is the larger of the roots of the indicial equation. n.b. we can always take a_0 to be the lowest nonvanishing coefficient. Its derivatives are

$$y' = \sum_n a_n s x^{s+n-1}, \tag{10.37}$$

$$y'' = \sum_n a_n s(s+n-1)x^{s+n-2}. \tag{10.38}$$

[30] We should not include a log-divergent term like $q_0 \log(x)$. This would be really bad because there is no convergent expansion of the log around $x = 0$.

Hence the indicial equation is given by the coefficient of the lowest power "orphan." In our case, we have second derivatives, so the lowest order terms are the order x^{s-2} terms– taking the $n = 0$ terms from each, we get

$$Q(x)y \sim q_{-2}x^{s-2}a_0, \quad P(x)y' \sim p_{-1}x^{s-2}sa_0, \quad y'' \sim x^{s-2}s(s-1)a_0. \tag{10.39}$$

Dividing through by a_0, the indicial equation will come out to

$$s(s-1) + sp_{-1} + q_{-2} = 0. \tag{10.40}$$

which is quadratic in s. Hence we can solve this for s explicitly!

Let us now suppose that the indicial equation is one of the bad ones. That is, its roots are

$$s = \alpha, \alpha - n \text{ for } n \in \mathbb{Z}. \tag{10.41}$$

Hence Frobenius will fail to give us both solutions. What we will do is "zoom in" on the bad case and consider what sorts of equations can give these bad roots. There must be some relation between p_{-1} and q_{-2} that causes Frobenius to break down. If we write an equation with the desired roots

$$0 = (s-\alpha)(s-(\alpha-n)) = s^2 - (2\alpha - n)s + (\alpha^2 - n\alpha), \tag{10.42}$$

then by comparing this to our indicial equation Eqn. 10.40, we have

$$p_{-1} - 1 = n - 2\alpha. \tag{10.43}$$

Lecture 11.

Wednesday, October 30, 2019

"If you're woken up in the middle of night in cold sweat on deserted island, this is the one thing you must know. Oranges and apples do not compare well... but their weights might."
–Nemanja Kaloper, on dimensional analysis

Last time, we studied the Frobenius method for the Bessel equation,

$$x^2y'' + xy' + (x^2 - \nu^2)y = 0, \tag{11.1}$$

and we found that it breaks down for integer ν. We are interested in studying second-order equations of the form

$$y'' + P(x)y' + Q(x)y = 0, \tag{11.2}$$

about a regular singular point (WLOG take it to be at $x = 0$) such that

$$P = \frac{p_{-1}}{x} + p(x) \tag{11.3a}$$

$$Q = \frac{q_{-2}}{x^2} + \frac{q_{-1}}{x} + q(x). \tag{11.3b}$$

These are early examples of Laurent series, a generalization of Taylor series for expansions which include the equivalent of charge terms (i.e. terms which are finite when multiplied by an "area" like x^2 before taking the $x \to 0$ limit). Let us also suppose that we have one solution y_1 in hand of the (Frobenius) form

$$y_1 = \sum_{n=0}^{\infty} a_n x^{n+s}. \tag{11.4}$$

That is, we solved the equation with the Frobenius method and got some value s for the lowest power of x in the expansion. By plugging in such an expansion into the differential equation 11.2 and using the forms of P and Q given by 11.3a and 11.3b, we can derive a general indicial equation

$$s(s-1) + sp_{-1} + q_{-2} = 0, \tag{11.5}$$

and by comparison to our "bad" indicial equation where the roots differ by an integer,

$$(s-\alpha)(s-\alpha+n) = 0, \tag{11.6}$$

we found a condition on p_{-1}:

$$p_{-1} - 1 = n - 2\alpha. \tag{11.7}$$

When the roots for s differ by an integer, the recursion relations "bomb" because the coefficients blow up at some finite power. Let us note that in the Bessel equation, we could take

$$\nu = \pm\frac{2k+1}{2} \tag{11.8}$$

where $k \in \mathbb{Z}$, and then the roots will again be separated by an integer. However, such solutions will still be okay in the Frobenius method. So apparently the separation of the roots by an integer is a necessary but not sufficient condition for the breakdown of the series method.

We use the Wronskian method to construct the second solution:

$$y_2(x) = y_1(x)\int_a^x dx' \frac{e^{-\int_a^{x'} P(x'')dx''}}{y_1^2(x')}. \tag{11.9}$$

If our first series solution y_1 only contains terms of order s or higher, then we can write it as

$$y_1 = \sum_{n=0}^{\infty} a_n x^{n+s} \equiv x^s A(x) \tag{11.10}$$

for some function $A(x)$.

Let us look at the integral in the definition of the second solution 11.9, and consider the larger root for s, i.e. $s = \alpha$. Then substituting in $y_1 = x^\alpha A(x)$, we have

$$\int_a^x dx' \frac{\exp\Big(-\int_a^{x'} dx''\big(\frac{p_{-1}}{x''} - p(x'')\big)\Big)}{x'^{2\alpha}A^2(x')} \tag{11.11}$$

The integral of the sum is the sum of the integral,[31] so let us perform the integral of $p(x'')$ and call $\exp(-\int_a^{x'} dx''\, p(x'')) = B(x')$, some arbitrary (known) function of x'. Hence

$$\begin{aligned}
\int_a^x dx' \frac{e^{-\int_a^{x'} P(x'')dx''}}{y_1^2(x')} &= \int_a^x dx' \frac{\exp\Big(-\int_a^{x'} dx''\big(\frac{p_{-1}}{x''} - p(x'')\big)\Big)}{x'^{2\alpha}A^2(x')} \\
&= \int_a^x dx' \frac{\exp\Big(-\int_a^{x'} dx''\big(\frac{p_{-1}}{x''}\big)B(x')\Big)}{x'^{2\alpha}A^2(x')} \\
&= \int_a^x dx' \frac{\exp(-p_{-1}\log(x'/a))B(x')}{x'^{2\alpha}A^2(x')} \\
&= \int_a^x dx' \frac{(a/x')^{p_{-1}}B(x')}{x'^{2\alpha}A^2(x')}.
\end{aligned}$$

where we have used the fact that $p(x'')$ is analytic to replace its integral with $B(x')$, and we can absorb the constant $a^{p_{-1}}$ into the overall normalization. Hence we can combine the powers of x' and the happy functions $B(x'), A^2(x')$ to get

$$y_2 = y_1(x)\int_a^{x'} dx' \frac{C(x')}{x'^{p_{-1}+2\alpha}} \tag{11.12}$$

$$= y_1(x)\int_a^{x'} dx' \frac{C(x')}{x'^{n+1}}, \tag{11.13}$$

where we said the roots were separated by an integer. Let us also expand $C(x')$ in powers of x' and write the equation as

$$y_2(x) = y_1(x)\int_a^x dx' \frac{(c_0 + c_1x' + c_2x'^2 + \ldots + c_nx'^n + x^{n+1}\hat{C}(x')}{x'^{n+1}}, \tag{11.14}$$

[31] At least in the absolutely convergent case!

where $\hat{C}(x')$ is now a happy function (analytic). Hence we get the integral of an analytic function $\hat{\hat{C}} \equiv \int dx' \hat{C}(x')$ plus n singular terms which form a sum (schematically)[32]

$$y_2 \sim y_1(x)\left[\left(\frac{1}{x^n} + \frac{1}{x^{n-1}} + \ldots + \frac{1}{x}\right) + c_n \log(x/a) + \hat{\hat{C}}(x)\right]. \tag{11.15}$$

We can of course absorb the $\log(1/a)$ into our $\hat{\hat{C}}(x)$ function. Hence our solution takes the overall structure

$$y_2 = c_n y_1(x)\log(x) + y_1(x)\left[\sum_{m=1}^{n} \frac{1}{x^m} + \hat{\hat{C}}(x)\right]. \tag{11.16}$$

Something nice happens when $c_n = 0$, though. This could certainly happen due to the symmetry of the problem. When this happens, the remaining terms are exactly of the form of the Frobenius solution for the other root with $s = \alpha - n$, since $y_1 \sim x^\alpha$. This tells us that when Frobenius fails, we should try adding a log-divergent term $y_1(x)\ln(x)$, and that the Frobenius expansion is actually a special case of this general method when the solution does not include a log term.

For the Bessel equation, if we crunch through the values of $B(x'), A(x'), C(x')$ we find that when ν is an odd half-integer, our solution has the form

$$C(x) \sim \frac{1}{(x^{1/2})^2\left[\sum a_k x^{2k}\right]^2}, \tag{11.17}$$

and we see that the Frobenius method gives us the right solution without a log divergence because the function we're integrating starts at power $1/x^2$.[33]

Recall that we previously saw we can redefine a second order equation to cancel the order y' term. That is, take

$$\begin{aligned} y &= x^\alpha W \\ y' &= x^\alpha W' + \alpha x^{\alpha-1} W \\ y'' &= x^\alpha W'' + 2\alpha x^{\alpha-1} W' + \alpha(\alpha-1)x^{\alpha-2} W. \end{aligned}$$

We might call this an integrating factor. In the Bessel equation

$$y'' + \frac{y'}{x} + (1 - \nu^2/x^2)y = 0, \tag{11.18}$$

we must pick $\alpha = 1/2$ and this makes our equation in terms of W

$$W'' + \left(1 - \frac{\nu^2}{x^2} - \frac{5}{4x^2}\right)W = 0. \tag{11.19}$$

If we write

$$W'' + \left(1 - \frac{4\nu^2 + 5}{4x^2}\right)W = 0, \tag{11.20}$$

we see that this is a 1D Schrodinger equation in a potential like $1 - 1/x^2$, at least for real ν.

This is a quadratic hyperbolic well. Conversely, nothing prevents us from taking ν to be imaginary, in which case we could have a barrier instead of a well. We can see now that in the case of imaginary ν, we know what to expect. We send in a wavepacket towards the barrier which gets increasingly redshifted as it approaches the barrier. As it hits its classical turning point, the sines and cosines (complex exponentials) turn into sinhs and coshs (real exponentials). Hence we know what the asymptotic behavior of Bessel functions should be based on our intuition for the Schrödinger equation.

[32] This is schematic because there are some constants attached to the divergent pieces. If we care about dimensions, those constants should carry the right dimensions to cancel out x^n, x^{n-1}, etc.

[33] There's a nice book on special functions from Dover by the mathematician Lebedev. "This should be your *vade mecum*, your manual." –Nemanja

Inhomogeneous second-order linear equations Finally, the last thing to do with these sorts of equations is to introduce a source term $J(x)$, such that

$$y'' + P(x)y' + Q(x)y = J(x). \tag{11.21}$$

This will lead us to the wonderful phenomenon of *parametric resonance*, which occurs in fly fishing and playground swings. We have an oscillator and if we kick it just right, we can drive it to higher amplitudes.

Let us consider an equation of the form

$$y' + f(x)y = J(x). \tag{11.22}$$

Maybe we don't know how to solve this. But let us write $y \equiv uv$ in terms of some unknown functions u, v. Computing y', we get

$$v'u + (u' + fu)v = J(x). \tag{11.23}$$

We see now that we can always cancel the order v term by making the choice

$$u(x) = e^{-\int^x f(x')dx'}. \tag{11.24}$$

Just calculate u' and apply the fundamental theorem of calculus to see this is true. (Or separate variables and integrate, if you prefer.) If we make this choice, what remains is the equation[34]

$$v'u = J(x) \implies v'(x) = J(x)e^{\int^x f(x')dx'}, \tag{11.25}$$

with solution

$$v = v_0 + \int^x dx\, J(x')e^{\int^{x'} f(x'')dx''}, \tag{11.26}$$

and so

$$y = uv = v_0 e^{-\int^x dx' f} + e^{-\int^x dx' f} \int^x dx'\, J(x')e^{\int^{x'} f}. \tag{11.27}$$

Why did this work? If we just solve the homogeneous equation $y' + fy = 0$, we get

$$y_b = v_0 e^{-\int f}. \tag{11.28}$$

We saw this term in the solution above. But we still need a particular solution to fit the source term J and write

$$y = y_p + y_b. \tag{11.29}$$

If we promote v_0 to a *function* of x, we can employ the method of variation of parameters. That is, we might guess that

$$y_p = v(x)e^{-\int f} \tag{11.30}$$

so that derivatives of the homogeneous solution will cancel out, and derivatives of $V(x)$ will let us cancel J. Hence all our work with the homogeneous equation was *not wasted*; variation of parameters will let us cancel the source term.

Let us suppose we have two solutions y_1, y_2 to the homogeneous second-order equation. Then promote their coefficients in the homogeneous solution to functions,

$$y_p = c_1(x)y_1(x) + c_2(x)y_2(x), \tag{11.31}$$

so that

$$y_p' = c_1 y_1' + c_2 y_2' + c_1' y_1 + c_2' y_2, \tag{11.32}$$

$$y_p'' = c_1 y_1'' + c_2 y_2'' + c_1' y_1' + c_2' y_2'. \tag{11.33}$$

But wait, you say. Where are the terms involving c_1'' and c_2'' in the second derivative? It has to do with the fact we only really needed one extra function to solve the particular solution. Two functions gives us some extra freedom, so let us impose the constraint that $c_1' y_1 + c_2' y_2 = 0$. Then the second derivative we've

[34] I'm writing the integrals at least with their variables of integration and limits of integration. This was a little quick in class, so I've rewritten it here to make clear(er) what we're integrating over and what the value of the integral depends on (e.g., is x an argument of a function or a limit of integration?). Also, it is $Je^{\int f}$, not Jeff. I know it looks like Jeff.

written is true. What happens if we also make the extra terms in the second derivative zero, $c_1'y_1' + c_2'y_2' = 0$? Then we get

$$\begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix} \begin{pmatrix} c_1' \\ c_2' \end{pmatrix} = 0, \tag{11.34}$$

and since this matrix is just the matrix whose determinant is the Wronskian of two linearly independent solutions y_1, y_2, it must be invertible. Hence $c_1' = c_2' = 0 \implies c_1, c_2$ are constants, which gives us back our old homogeneous solution. Instead, we should write

$$\begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix} \begin{pmatrix} c_1' \\ c_2' \end{pmatrix} = \begin{pmatrix} 0 \\ J \end{pmatrix}. \tag{11.35}$$

Now the fact that $\left(\begin{smallmatrix} y_1 & y_2 \\ y_1' & y_2' \end{smallmatrix}\right)$ is invertible actually helps us; we can multiply both sides by its inverse, which lets us solve for c_1', c_2' individually in terms of J, y_1, y_2, y_1', and y_2' and just integrate.[35]

Lecture 12.

Monday, November 4, 2019

"There is physics everywhere, you just need to open your eyes to it."
–Nemanja Kaloper

We were studying the differential equation of the form

$$y'' + py' + qy = J. \tag{12.1}$$

Second-order equations are ubiquitous in physics. Let us note that we can trivially rewrite as $w = y'$ where

$$w = y' \tag{12.2}$$

$$w' = -Pw - Qy + J, \tag{12.3}$$

which looks much like Hamilton's equations, i.e. a set of two coupled first-order equations rather than a single second-order equation. We could also discuss this from the perspective of the Hamilton-Jacobi equations, if we like. But in general higher-derivative equations could be put into the form of Hamilton's equations with a suitable Legendre transform.

We hope to define physical systems leading to Hamiltonians which are bounded from below so that our systems are in some sense stable; they have well-defined ground states.

How to solve second-order equations Here's our recipe for solving second order equations, in order of least to most painful.[36]

(a) Write it down by inspection.[37]
(b) Use Frobenius (series expansion). Check the indicial equation and make sure the roots are not separated by an integer.
(c) If they are separated by an integer, guess the second solution is $y_2 \sim y_1 \log(x) + \sum a_n x^{n+s_-}$, where s_- was the smaller root from Frobenius.
(d) Use the Wronskian method to construct the second solution by taking some integrals.

The story does not end here, for these methods give solutions to the homogeneous equation. To figure out the solution to the equation *with a source*, we can construct particular solutions by variation of parameters, i.e. by promoting the coefficients c_1, c_2 to functions:

$$y_p = c_1(x)y_1 + c_2(x)y_2. \tag{12.4}$$

In general we expect this approach to work since we showed that by a simple change of variables, we could make our second order equation look like two first-order equations. We can generalize this to a system of n first-order equations and n unknown functions, and then start imposing boundary conditions. That is,

$$y_p' = c_1y_1' + c_2y_2' + [c_1'y_1 + c_2'y_2], \tag{12.5}$$

[35] No one said these integrals were going to be easy, but at least we can write them down explicitly.
[36] A good resource for performing complicated integrals by contour integration is in Whittaker and Watson.
[37] This includes our trick of factoring equations with constant coefficients, making clever substitutions, etc.

where we can add the constraint that the bracket term vanishes, $c_1'y_1 + c_2'y_2 \equiv 0$. We only needed one particular solution, so we gave ourselves some extra freedom in introducing two functions $c_1(x), c_2(x)$. Hence

$$y'' = c_1y_1'' + c_2y_2'' + c_1'y_1' + c_2'y_2', \tag{12.6}$$

and when we plug this back into the original equation, all the terms without derivatives of c satisfy the homogeneous equation, while the other terms give us

$$c_1'y_1 + c_2'y_2 = 0 \text{ by construction} \tag{12.7}$$

$$c_1'y_1' + c_2'y_2' = J(x) \text{ because } y_1, y_2 \text{ solve the homogeneous equation.} \tag{12.8}$$

The matrix of coefficients here is simply the Wronskian. y_1, y_2 are linearly independent, so this matrix is invertible by construction. Hence we can solve the first equation for c_2' as

$$c_2' = -c_1'\frac{y_1}{y_2}, \tag{12.9}$$

and then by substituting into the second,

$$\frac{c_1'}{y_2}(y_1'y_2 - y_1y_2') = J. \tag{12.10}$$

This is a first-order eqaution for c_1 in terms of y_1, y_2 and their derivatives, and also J. Hence

$$c_1 = -\int^x \frac{J(x')}{W(x')}y_2(x') + c_1^0, \tag{12.11}$$

$$c_2 = \int^x \frac{J(x')}{W(x')}y_1(x') + c_2^0. \tag{12.12}$$

Let's reassemble our solution, which we said was of the form $y(x) = c_1y_1 + c_2y_2$. Plugging in, it is

$$y(x) = c_1^0y_1(x) + c_2^0y_2(x) + \int^x dx' \frac{J(x')}{W(x')}\left[y_1(x')y_2(x) - y_2(x')y_1(x)\right]. \tag{12.13}$$

Our solution was smart enough to give us back not only the homogeneous solution but also a particular solution which gave us exactly the source term. This construction is so important that we give it a name– a *Green's function*. That is, a Green's function is a function $G(x, x')$ which lives under an integral,

$$\int^x dx' G(x, x')J(x'), \tag{12.14}$$

which allows us to precisely fit the source term.[38]

A small diversion: Bernoulli's equation is

$$y'/y = p(x) + q(x)y^n. \tag{12.15}$$

We claim this is really a fake nonlinear equation, for this equation can be put in the form

$$-\frac{1}{n}\left(\frac{1}{y^n}\right)' = p(x)\frac{1}{y^n} + q(x), \tag{12.16}$$

and then under the substitution $w = 1/y^n$ this is just first-order in w. This is about as far as we will discuss nonlinear equations.

Note. A logistic note– the midterm content ends here, with Chapters 5, 6, and 7. The midterm is in class on Wednesday, November 20, and open book, open notes, open solutions. The expectation will be that the problems are of comparable difficulty to the homework.

[38] I tend to think of Green's functions as the inverse of differential operators. That is, when you hit a Green's function with a differential operator, you get back a delta function. Hence when you integrate a Green's function against a source and then hit it with the differential operator, you get the source term back.

Green's functions Let's go back to our inhomogeneous equation

$$y'' + Py' + Qy = J. \tag{12.17}$$

Let's do a trivial thing to this equation, and change how we write this equation. In particular, let us define a function p by

$$\frac{p'}{p} = P. \tag{12.18}$$

Hence $p = e^{\int P}$, and we can write

$$y'' + \frac{p'}{p}y' + Qy = J. \tag{12.19}$$

Note that this is a safe substitution unless we try to expand about a singular point, in which case we simply construct p piecewise. Now we can write

$$py'' + p'y' + pQy = Jp, \tag{12.20}$$

and define

$$q = pQ. \tag{12.21}$$

The first two terms in our rewritten equation are now a total derivative,

$$(py')' + qy = j, \tag{12.22}$$

where we have similarly defined

$$j \equiv pJ. \tag{12.23}$$

Let us now rewrite this in terms of a derivative operator D:

$$Ly \equiv [D(pD) + q]y = j, \tag{12.24}$$

where we have defined a linear operator L which now takes y in the Hilbert space and maps it to some other vector j. It would be very nice if we could invert the linear operator, i.e. solving the equation is equivalent to

$$y = L^{-1}j. \tag{12.25}$$

We must be careful– our source cannot contain any component of the zero eigenfunction, or our solutions will be terribly degenerate. That is, $Lu = 0 \implies L^{-1}(j + \alpha u)$ will be ill-defined. For

$$L^{-1} = \sum \frac{1}{\lambda_\alpha} |u_\alpha\rangle\langle u_\alpha| . \tag{12.26}$$

One way to do this is to explicitly construct the eigenbasis of L and then just invert it. That is, we diagonalize and then take the inverse of a diagonal matrix. A Green's function is therefore an object G such that

$$LG = \mathbb{I}, \tag{12.27}$$

i.e. when we hit it with the linear operator L, we get back the identity. Moreover, we can expand it in a (coordinate) basis as

$$\langle x'|LG|x\rangle = \langle x'|x\rangle. \tag{12.28}$$

This means that

$$\int dx'' \, \langle x'|L|x''\rangle\langle x''|G|x\rangle = \delta(x - x'). \tag{12.29}$$

Let us note that the matrix elements of L will in general depend on both x' and x''. However, we can restrict to local operators L (i.e. those with only diagonal matrix elements $\langle x'|L|x'\rangle\delta(x' - x'')$) such that

$$\int dx'' \, \langle x'|L|x'\rangle\delta(x' - x'')\langle x''|G|x\rangle = \delta(x - x'). \tag{12.30}$$

If we now perform the x'' integral (thanks to locality), we get back

$$\langle x|L|x\rangle\langle x|G|x'\rangle = \delta(x - x'), \tag{12.31}$$

where we have relabeled variables $x \leftrightarrow x'$. This tells us that we must solve a very special differential equation for a point source, and from our point source solution we can build solutions for arbitrary sources.

This 1D problem is actually analogous to solving a capacitor problem, perhaps with some charged plate inserted between. The charge sources a field. The Green's function tells you what field is sourced by the charge. This is precisely the theory of electric charges and boundary conditions in one dimension.

The scaling of flux density in 3 dimensions tells us how the electric field scales (given that the total flux is constant)– this is the basis of Gauss's law. This tells us that in 3 dimensions, the electric field from a point source drops off as $1/r^2$ (inverse to the area of a sphere) and the electric field from a long charged wire drops off as $1/r$ (inverse to the area of a cylinder). Hence the potentials go as $1/r$ and $\log r$ respectively. In 1D, the potential is actually a positive power of r, which is thought to be responsible for the phenomenon of *confinement* in quantum chromodynamics (the theory of the strong force).

Once we have a Green's function in hand, we can build arbitrary sources. That's the name of the game. That is,

$$V[\rho] = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d^3\mathbf{r}' = \int G(\mathbf{r},\mathbf{r}')\rho(\mathbf{r}')d^3\mathbf{r}'. \tag{12.32}$$

We could rewrite this in terms of a generalized *source* function $j(\mathbf{r}')$. That is,

$$\int \frac{j(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d^3\mathbf{r}' = \int G(\mathbf{r},\mathbf{r}')j(\mathbf{r}')d^3\mathbf{r}'. \tag{12.33}$$

Let us now show that the Green's function is actually useful. For notice that

$$(p(x)G')' + q(x)G(x,x') = \delta(x-x'), \tag{12.34}$$

where derivatives are taken with respect to the x coordinate, and

$$(py')' + qy = 0. \tag{12.35}$$

Hence

$$y\big((pG')' + qG\big) = \delta(x-x')y(x) \tag{12.36}$$

and

$$G(x,x')\big[(py')' + qy\big] = 0, \tag{12.37}$$

where we've multiplied the second equation by the Green's function. Subtracting now gives

$$y(pG')' - G(py')' = (ypG' - Gpy')' \tag{12.38}$$

on the LHS. This is an example of Green's theorem– we shall see what happens when we now integrate over x. That is, the LHS gives us

$$\big[ypG' - Gpy'\big]_a^b, \tag{12.39}$$

which must vanish if we set appropriate boundary conditions. There are Dirichlet boundary conditions (the function vanishes at the boundaries), Neumann boundary conditions (the function or its derivative are fixed at the boundaries), or mixed boundary conditions (all others).

If we impose the same boundary conditions on the Green's function, we can see that the bracket term vanishes when evaluated at the boundaries. For the RHS (i.e. $\int dx\,\delta(x-x')y(x) = y(x') = \int G(x,x')j(x)dx$) we see that we are in fact forced to impose the same boundary conditions on G as on the solution. That is, the Green's function must satisfy the same boundary conditions as the general solution (for an arbitrary source) because it is itself a solution for a point source.

There is one small caveat– in general, we ought to have taken the complex conjugate of G and then used the property that the Green's function is hermitian. The result is the same, but in this case we simply restricted ourselves to real functions.

Lecture 13.

Wednesday, November 6, 2019

"I mean, someone introduces an inch. How narcissistic, right? At least use an electron."
–Nemanja Kaloper

Let's continue our discussion of Green's functions and boundary conditions. We can revisit our discussion of Wronskians– as mentioned previously in these notes, the vanishing of the Wronskian is sometimes not enough to imply linear dependence. For if we take the functions x^2 and $|x|x$, these functions have vanishing Wronskian everywhere but they are not in general linearly dependent– the problem is the smoothness of the second function at zero. Most of the time, we never need to worry about such pathologies because our differential equations have differentiable coefficients. But sometimes the boundary conditions introduce complications.

For instance, we have different charges and boundary conditions for electromagnetism. In particular, we can have positive and negative charges. Why don't we have negative masses? The reason is simply that a theory with negative masses is unstable to small perturbations. Instead of a harmonic oscillator potential tending to return the mass to equilibrium, we will get an imaginary frequency corresponding to not sines and cosines but exponential growth. This instability usually suggests that we've simply expanded around the wrong vacuum.[39]

In particular, recall that $a = F/m$, so if we take two objects, one of positive and one of negative mass, these masses would simply chase each other around the universe. One exerts a pull and the other exerts a push, and moreover Heisenberg's uncertainty principle suggests that such particle pairs could pop out of the vacuum and accelerate to arbitrarily high energies. This example also covariantizes[40] properly, and it is called the Bondi dipole.

More on Green's functions We have a Green's function, which is Hermitian in its arguments:

$$G(x,t) = G^*(t,x). \tag{13.1}$$

That is, we start with a differential equation

$$(py')' + qy = j \tag{13.2}$$

for some source j, and we define the differential operator

$$Ly = [D(pD) + q]y, \tag{13.3}$$

such that the Green's function inverts this operator:

$$LG = \mathbb{I}. \tag{13.4}$$

Suppose the differential operator has an eigenspectrum:

$$L\phi_m = \lambda_m \phi_m, \tag{13.5}$$

and moreover let us take the operator to be self-adjoint, such that $\lambda_m = \lambda_m^*$, i.e. there are real eigenvalues. In a basis, we may compute

$$\langle x'| \left(\sum_m |\phi_m\rangle\langle\phi_m| \right) |x\rangle \langle \phi_m | \phi_n \rangle = \delta_{nm}. \tag{13.6}$$

Hence we see that

$$\sum \phi_m^*(x')\phi_m(x) = \delta(x - x'), \tag{13.7}$$

i.e. the eigenfunctions are delta-function normalized.

Now

$$L_x G(x,x') = \delta(x-x') = \sum_m \phi_m^*(x')\phi_m(x), \tag{13.8}$$

which tells us that

$$G(x,x') = \sum g_m(x')\phi_m(x), \tag{13.9}$$

where these $g_m(x')$s can be thought of as components of the Green's function in the basis ϕ_m. That is, because the ϕ_ms were eigenfunctions of L_x, it follows that G can be decomposed in the basis ϕ_m. That is, we now compute

$$L_x G(x,x') = L_x \sum g_m(x')\phi_m(x) \tag{13.10}$$

$$= \sum \lambda_m g_m(x')\phi_m(x), \tag{13.11}$$

[39] Compare the Higgs potential or the Ising model.

[40] Becomes compatible with general relativity

so by comparison to our general expression for $L_x G(x,x')$ we see that

$$g_m(x') = \frac{\phi_m^*(x')}{\lambda_m}. \tag{13.12}$$

That is, we can recover exactly these coefficients $g_m(x')$ by comparing these expressions component-wise. We seem to have a problem when $\lambda_m = 0$, but in fact if we're a little careful we can apply the same analysis to the operator $L - z$ where z is complex, and then take the limit as $z \to 0$.

Substituting back in these coefficients, we now see that

$$G(x,t) = \sum_m \frac{\phi_m^*(t)\phi_m(x)}{\lambda_m}, \tag{13.13}$$

where we have relabeled x' to t. In particular if the differential equation and its boundary conditions are both translationally invariant, then the Green's function can only depend on the difference of the two position in the arguments:

$$G(x,x') = G(x-x'), \tag{13.14}$$

e.g. the $1/|\mathbf{r}-\mathbf{r}'|$ Coulombic potential.

Let us now study the differential equation for the Green's function:

$$(pG')' + qG = \delta(x-x'). \tag{13.15}$$

This looks scary. Delta functions are a little tricky to work with. However, away from where the delta function lives ($x \neq x'$), we can just solve the homogeneous equation in each region. Suppose we fix some x'. Then we can solve the equation in $x < x'$ and $x > x'$, and patch together by continuity.[41]

That is, we have

$$G = \begin{cases} A_L y_1 + B_L y_2 & x < x' \\ A_R y_1 + B_R y_2 & x > x'. \end{cases} \tag{13.16}$$

We must have boundary conditions– usually one at each endpoint of the interval. These place two constraints on our a priori 4 coefficients A_L, B_L, A_R, B_R. However, the last two constraints will come from setting boundary conditions *at* the delta function, $x = x'$.

Suppose we take Dirichlet boundary conditions on the interval $[a,b]$, i.e. the function vanishes at a and at b. In particular, suppose we choose $y_1(a) = 0$ and $y_2(b) = 0$ so that our Green's function takes the form

$$G = \begin{cases} A_L y_1(x) & x < x' \\ B_R y_2(x) & x > x'. \end{cases} \tag{13.17}$$

We now impose continuity,

$$G_L(x',x') = G_R(x',x'). \tag{13.18}$$

That is, the Green's function must be continuous. It must, a fortiori, since it is differentiable. Its derivative may not be continuous– in general it will not, since we need to pick up a step function jump. For suppose we integrate the differential equation around x', i.e. over the interval $[x'-\epsilon, x'+\epsilon]$. Then

$$\int_{x'-\epsilon}^{x'+\epsilon} \big[(pG')' + qG\big]dx = \int_{x'-\epsilon}^{x'+\epsilon} \delta(x-x')dx = 1. \tag{13.19}$$

The second term is finite by the mean value theorem, and it is of order ϵ. The other term, however, is a total derivative. We can evaluate it by the fundamental theorem of calculus. It is therefore just

$$(pG'(x,x'))\big|_{x'-\epsilon}^{x'+\epsilon} = p(x)\big[G_R'(x',x') - G_L'(x',x')\big] = 1 \tag{13.20}$$

taking the $\epsilon \to 0$ limit. This is our final boundary condition, a condition on the derivative of G at $x = x'$. Hence

$$y_1(x')A_L - y_2(x')B_R = 0 \tag{13.21}$$

$$y_1'(x')A_L - y_2'(x')B_R = -\frac{1}{p(x')}. \tag{13.22}$$

[41] This should feel a lot like the delta-function scattering in quantum mechanics.

In matrix form, we see that

$$\begin{pmatrix} y_1 & -y_2 \\ y_1' & -y_2' \end{pmatrix} \begin{pmatrix} A_L \\ B_R \end{pmatrix} = - \begin{pmatrix} 0 \\ \frac{1}{p(x')} \end{pmatrix}. \tag{13.23}$$

Notice the matrix here is in fact the Wronskian, up to a sign flip on the second column. We could always abosrb this into B_R if we wish, so it's not too important until it comes time to write down the solution. But this matrix is invertible by construction– we took y_1, y_2 to be linearly independent, so that the Wronskian is nonzero.

If we work out the matrix inverse, we find that

$$G = \begin{cases} Ay_1(x)y_2(x') & x < x' \\ Ay_1(x')y_2(x) & x > x', \end{cases} \tag{13.24}$$

where this constant is given to be

$$A = \frac{1}{p(x')(y_1(x')y_2'(x') - y_2(x')y_1'(x')}. \tag{13.25}$$

How do we know it's really a constant? Well, if A is constant then we see G is properly symmetric under interchange of x and x', hence real and symmetric $\implies$ hermitian. If A were not constant, this would break the hermiticity of G.

Let's show that A really is constant. Multiply the differential equation for one solution by the other solution, i.e.

$$y_{2,1}\left[(py_{1,2}')' + qy_{1,2}\right] = 0. \tag{13.26}$$

If we expand out and subtract, we find that

$$p(x)\big(y_2(x)y_1'(x) - y_1y_2'\big) = \text{constant}. \tag{13.27}$$

This defines a generalized Wronskian, and this is none other than the inverse of our normalization constant A.

Example 13.28. Consider the equation

$$y'' = -f(x), \tag{13.29}$$

on an interval $[0,1]$. We impose Dirichlet boundary conditions, $y(0) = y(1) = 0$. The solutions to the homogeneous equation are

$$y = A + Bx. \tag{13.30}$$

Now the left solution (vanishing at $x = 0$) is

$$y_L = Bx, \tag{13.31}$$

while the right solution (vanishing at $x = 1$ is

$$y_R = A(x-1). \tag{13.32}$$

These are clearly linearly independent; they differ by a constant. Hence the Green's function takes the form

$$G = \begin{cases} Bx & x < x' \\ A(x-1) & x > x'. \end{cases} \tag{13.33}$$

These must be equal at $x = x'$, so

$$Ax' - A = Bx' \implies B = \frac{x'-1}{x'}. \tag{13.34}$$

Doing the Gaussian pillbox in a little epsilon interval about x' gives the discontinuity in the derivative. That is, $(Bx)'|_{x=x'} = B$ and similarly $(A(x-1))'|_{x=x'} = A$, so the condition on the derivative (note that $p = 1$ in this case) is

$$A - B = 1. \tag{13.35}$$

We have two equations in two variables. If we substitute and solve, we get

$$A\left(1 - \frac{x'-1}{x'}\right) = 1, \tag{13.36}$$

we see that

$$A = x', B = (x' - 1). \tag{13.37}$$

Hence the Green's function is really

$$G(x, x') = \begin{cases} x(x' - 1) & x < x' \\ x'(x - 1) & x > x'. \end{cases} \tag{13.38}$$

We can now solve the general problem

$$y'' = -f(x) \tag{13.39}$$

using our Green's function. That is,

$$y(x) = \int_0^1 dx'\, G(x, x')\big(-f(x')\big) \tag{13.40}$$

Really, we have to keep in mind that the Green's function is defined piecewise, so we must also perform the integral piecewise, from $[0, x]$ and then from $[x, 1]$. That is,

$$y = -\int_0^x dx'(x - 1)x'f(x') - \int_x^1 dx'x(x' - 1)f(x'). \tag{13.41}$$

We can check by taking derivatives with respect to x that this is a solution. What's a bit more fun is to notice that the boundary conditions are correctly satisfied as $x \to 0$ and as $x \to 1$. As $x \to 0$, the first vanishes by the fact that the integration limits coincide and the second vanishes by the fact that $x \to 0$ makes the entire second integral vanish, thanks to our construction $y_L(0) = 0$.

We can also define Green's functions for other boundary conditions like $y(a) = y_i, y(b) = y_f$). The procedure is similar, and the physical interpretation is simply that we are solving the equation (e.g. the Laplace equation) in the presence of a background field/potential.

Lecture 14.

Wednesday, November 13, 2019

Sturm-Liouville problems Today we shall introduce Sturm-Liouville problems, corresponding roughly to Chapter 8 of Arfken. Let us define what we mean by a Sturm-Liouville problem, and motivate why we might be interested in one.

Our focus is differential equations of the form

$$[\partial_x(p_0(x)\partial_x) + p_2(x)]y(x) = \lambda w(x)y. \tag{14.1}$$

This is a second-order nonlinear equation where $p_0(x), p_2(x), w(x)$ are given functions of x and λ is a parameter fixed by the boundary conditions. When we discussed Green's functions, we wanted to solve equations of the form

$$\mathcal{L}y(x) = f(x) \tag{14.2}$$

for $y(x)$ given some source $f(x)$ and a differential operator $\mathcal{L}$. However, we needed a boundary condition to find the Green's function; there are in general infinitely many solutions to such an equation, so we need boundary conditions to fix a particular solution.

Not every set of boundary conditions is equally good. Some will be contradictory and indeed yield no solution. Others might not constrain the problem enough. Let us therefore write our problem as

$$\mathcal{L}y(x) = \lambda y(x) \tag{14.3}$$

where

$$\mathcal{L} = p_0\partial_x^2 + p_1\partial_x + p_2 \tag{14.4}$$

is the most general second-order differential operator.

Maybe you're not convinced by this yet.[42] We can present another argument. Where do (families of) orthogonal functions come from? For instance, we have

$$f_n(x) = \sin(nx), n = 1, 2, \ldots \tag{14.5}$$

[42] I should point out that the time-independent Schrödinger equation has this form. So that's an infinity of good reasons to study such problems.

such that

$$\int_0^\pi dx\, f_n^*(x) f_m(x) = \delta_{nm}. \tag{14.6}$$

We claim that orthogonal functions naturally arise as solutions to Sturm-Liouville problems. The goal of today will be to understand how to solve such problems in general.

We learned previously[43] that certain (differential) operators have nice spectral properties. We called these operators hermitian, and such operators had real eigenvalues and orthogonal eigenvectors.[44] Let's also recall that operators must be specified on domains, i.e. the space of vectors they act on.

Definition 14.7. An operator $\mathcal{L}$ with domain $D(\mathcal{L})$ is *Hermitian* if

$$\langle \mathcal{L}u|v\rangle = \langle u|\mathcal{L}v\rangle \quad \forall u, v \in D(\mathcal{L}). \tag{14.8}$$

That is, our vector space is equipped with an inner product, and we can move our operator freely between acting on either argument of the inner product.

Definition 14.9. The adjoint of $\mathcal{L}$, denoted by $\mathcal{L}^\dagger$, is defined by the relation

$$\langle \mathcal{L}^\dagger u|v\rangle = \langle u|\mathcal{L}v\rangle \quad \forall v \in D(\mathcal{L}). \tag{14.10}$$

Notice we haven't said anything about the domain of $\mathcal{L}^\dagger$. In particular, the domain $D(\mathcal{L}^\dagger)$ may not be identical to original domain $D(\mathcal{L})$. We therefore make the following refinement.

Definition 14.11. An operator $\mathcal{L}$ is *self-adjoint* if

$$\mathcal{L}u = \mathcal{L}^\dagger u \quad \forall u \in D(\mathcal{L}), \tag{14.12}$$

and

$$D(\mathcal{L}) = D(\mathcal{L}^\dagger). \tag{14.13}$$

Self-adjoint operators have real eigenvalues and orthogonal and complete sets of eigenvectors.

For instance, the momentum operator in quantum mechanics is a good example of an operator which is Hermitian but not self-adjoint when defined on a bounded domain.

Example 14.14. Consider $p = -i\partial_x$ on the domain $[0,1]$ subject to $y(0) = y(1) = 0$. Hence

$$\begin{aligned}\langle u|pv\rangle &= \int_0^1 dx u^*(x)(-iv'(x)) \\ &= (-i)u^*(x)v(x)|_0^1 - \int_0^1 dx u^{*\prime}(-i)v(x) \\ &= \langle pu|v\rangle.\end{aligned}$$

Hence p is Hermitian. But p is *not* self-adjoint for this set of boundary conditions. For suppose p had an eigenvector,

$$py = \lambda y, \tag{14.15}$$

with $\lambda \in \mathbb{R}$. Then

$$y = e^{i\lambda x}, \tag{14.16}$$

which certainly does not vanish at $x = 0$.

Let us make the following definition.

Definition 14.17. For a differential operator

$$\mathcal{L} = p_0(x)\partial_x^2 + p_1(x)\partial_x + p_2(x), \tag{14.18}$$

we say that $\mathcal{L}$ is "self-adjoint in the differential equation sense" if $p_0' = p_1$.

[43] Arfken Ch. 5.

[44] There are slight caveats here. Sometimes self-adjointness depends on the boundary conditions.

This definition makes no mention of boundary conditions, so it does not exactly agree with the definition we've seen already. Let us see how we can reconcile the two.

For such an operator, we can rewrite it as

$$\mathcal{L} = p_0\partial_x^2 + p_0'\partial_x + p_2 = \partial_x(p_0\partial_x) + p_2. \tag{14.19}$$

Suppose $\mathcal{L}$ lives on a domain $[a,b]$. If we act on vectors, we find that

$$\begin{aligned}\langle v|\mathcal{L}u\rangle &= \int_a^b dx\left[v^*\partial_x(p_0u'(x)) + v^*p_2u\right]\\ &= v^*p_0u'|_a^b + \int_a^b dx\left[-v^{*\prime}p_0u' + v^*p_2u\right]\\ &= (v^*p_0u' - v^{*\prime}p_0u)|_a^b + \underbrace{\int_a^b dx\left[\partial_x(p_0v^{*\prime})u + v^*p_2u\right]}_{\langle\mathcal{L}v|u\rangle}.\end{aligned} \tag{14.20}$$

Integrating by parts, we pick up some boundary terms evaluated at a,b and we see that this operator is Hermitian if the boundary term vanishes, i.e.

$$(v^*p_0u' - v^{*\prime}p_0u)|_a^b = 0. \tag{14.21}$$

This leads us naturally to the Dirichlet boundary conditions (where the functions themselves vanish at the boundaries) and the Neumann boundary conditions (where the derivatives vanish). We can also mix them if we're a bit clever.

The most general boundary conditions which lead to self-adjoint operators are

$$\alpha_1y(a) + \beta_1y'(a) = 0 \tag{14.22a}$$

$$\alpha_2y(b) + \beta_2y'(b) = 0. \tag{14.22b}$$

Example 14.23. The Legendre equation is one common differential equation that's already in the Sturm-Liouville form. That is,

$$\mathcal{L}y = \lambda y \tag{14.24}$$

where

$$\mathcal{L} = \underbrace{-(1-x^2)}_{p_0}\partial_x + \underbrace{2x}_{p_0'}\partial_x + l(l+1). \tag{14.25}$$

More generally, if we have

$$\mathcal{L}y = \lambda y, \quad \mathcal{L} = p_0\partial_x^2 + p_1\partial_x + p_2, \tag{14.26}$$

where $p_1 \neq p_0'$, we claim that we can put it into self-adjoint form. To see this, let us multiply both sides by a factor $w(x)$ to get

$$w(x)\mathcal{L}(x)y(x) = \lambda w(x)y(x). \tag{14.27}$$

Hence

$$w(x)\mathcal{L} = wp_0\partial_x^2 + wp_1\partial_x + wp_2 \tag{14.28}$$

$$= \bar{p}_0\partial_x^2 + \bar{p}_0'\partial_x + \bar{p}_2 \tag{14.29}$$

in terms of some new barred functions. That is, we would like to put this into self-adjoint form, which means that

$$\bar{p}_0 = wp_0, \quad \bar{p}_0' = wp_1. \tag{14.30}$$

Hence

$$wp_1 = w'p_0 + wp_0', \tag{14.31}$$

which we can rearrange (since it is separable) into the form

$$\frac{w'}{w} = \frac{p_1}{p_0} - \frac{p_0'}{p_0}. \tag{14.32}$$

Integrating, we find that

$$\log w = \int^x dx'\frac{p_1}{p_0} - \log p_0, \tag{14.33}$$

or equivalently

$$w = \frac{1}{p_0} \exp\left(\int^x dx' \frac{p_1}{p_0} \right). \tag{14.34}$$

This function $w(x)$ is called the *weight function*. It follows that

$$\bar{p}_0 = \exp\left(\int^x dx' \frac{p_1}{p_0} \right), \quad \bar{p}_0' = \frac{p_1}{p_0} \exp\left(\int^x dx' \frac{p_1}{p_0} \right), \tag{14.35}$$

exactly as desired. This is why we can write our problem as

$$\mathcal{L}y = \lambda w(x) y, \tag{14.36}$$

where w is fixed but $\mathcal{L}$ is taken to be self-adjoint.

We can interpret w as defining a new, weighted inner product,

$$\langle v | u \rangle = \int_a^b dx\, w(x) v^*(x) u(x). \tag{14.37}$$

Properties of Sturm-Liouville problems Some important properties of these Sturm-Liouville (SL) problems, i.e. equations of the form

$$\mathcal{L}y = \lambda w(x) y, \tag{14.38}$$

are the following.

- The eigenvalues λ are real.
- The eigenfunctions are orthogonal with respect to the weighted inner product,

$$\int_a^b dx\, w(x) \psi_n^*(x) \psi_m(x) = \delta_{mn}. \tag{14.39}$$

- The eigenvalues are simple, i.e. to each eigenvalue there is a unique eigenfunction up to normalization.
- The eigenvalues of SL problems are ordered,

$$\lambda_1 < \lambda_2 < \ldots < \lambda_n < \ldots, \tag{14.40}$$

  such that

$$\lambda_n \to \infty \text{ as } n \to \infty. \tag{14.41}$$

Example 14.42. Suppose we want to solve

$$y'' + \lambda y = 0 \tag{14.43}$$

on the interval $[0, \pi]$ subject to the constraint that $y(0) = 0, y'(\pi) = 0$.

This equation is already in SL form, with weight function $w = 1$. Hence we can just consider $\lambda \in \mathbb{R}$. We now have some cases to consider.

Case 1: $\lambda < 0$. If λ is negative, then let $\lambda = -\omega^2$. The general solution is

$$y = Ae^{\omega x} + Be^{-\omega x}. \tag{14.44}$$

We now impose the boundary conditions. The first tells us that

$$A + B = 0, \tag{14.45}$$

and applying the second, we have

$$A(e^{\omega p} + e^{-\omega \pi}) = 0 \implies A = 0 \implies B = 0. \tag{14.46}$$

So the only solution is $y = 0$, the trivial solution.

Case 2: $\lambda = 0$. We again get $y = 0$ (the solution is $Ax + B$ and $A = B = 0$).

Case 3: $\lambda > 0$. Let $\lambda = \omega^2$ so our general solution is

$$y = A \sin \omega x + B \cos \omega x. \tag{14.47}$$

The boundary conditions tell us now that $B = 0$ and

$$A\omega \cos \omega \pi = 0 \implies \omega = n + \frac{1}{2}, n \in \mathbb{Z}. \tag{14.48}$$

We conclude that only certain values of λ have nontrivial solutions.

Example 14.49. Let's modify this example. We want to solve

$$y'' + \lambda y = 0 \tag{14.50}$$

on the interval $[0, 1]$ where now the boundary conditions are

$$y(0) + y'(0) = 0 \tag{14.51a}$$

$$y(1) - y'(1) = 0. \tag{14.51b}$$

In the $\lambda = 0$ case we have a general solution $y = Ax + B$, where the boundary conditions fix $A = B = 0$.

For $\lambda < 0$, something more interesting happens. We get solutions

$$y = Ae^{\omega x} + Be^{-\omega x}, \tag{14.52}$$

and imposing BCs gives us

$$A + B + \omega(A - B) = 0, Ae^{\omega} + Be^{-\omega} - \omega(Ae^{\omega} - Be^{-\omega}) = 0. \tag{14.53}$$

We can write this as a matrix equation,

$$\begin{pmatrix} 1+\omega & 1-\omega \\ (1-\omega)e^{\omega} & (1+\omega)e^{-\omega} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = 0. \tag{14.54}$$

Such an equation has nontrivial solutions if the determinant of the coefficient matrix is zero, i.e.

$$(1+\omega)^2 e^{-\omega} - (1-\omega)^2 e^{\omega} = 0. \tag{14.55}$$

Equivalently

$$e^{\omega} = \left| \frac{1+\omega}{1-\omega} \right|, \tag{14.56}$$

which is now a transcendental equation. We can solve this graphically; it turns out to have two solutions, one of which is trivial ($\omega = 0$) and the other of which we can numerically determine to be about $\omega \approx 1.54$.

For the $\lambda > 0$ case we have

$$y = A \sin \omega x + B \cos \omega x. \tag{14.57}$$

The boundary conditions are now

$$B + \omega A = 0 \tag{14.58}$$

$$B \cos \omega + A \sin \omega - (A\omega \cos \omega - B\omega \sin \omega) = 0. \tag{14.59}$$

Taking the determinant and performing a bit of algebra gives another transcendental equation

$$\tan \omega = \frac{2\omega}{1 - \omega^2}. \tag{14.60}$$

Instead of one solution, we now have infinitely many.

Lecture 15.

Monday, November 18, 2019

Last time we started discussing Sturm-Liouville problems. Let's take a step back and revisit a few examples of Green's functions.

Green's functions on infinite intervals Previously, we had self-adjoint operators which defined differential equations, such that their Green's functions G satisfy the (schematically written) equation

$$(pG')' + qG = \delta. \tag{15.1}$$

Before, we had boundary conditions on a finite interval $[a, b]$, like $y'(a) = y'(b) = 0$. What if we instead have an infinite interval? Consider

$$y'' - k^2 y = 0 \tag{15.2}$$

on the whole real line. The general solution is

$$y = Ae^{kx} + Be^{-kx}. \tag{15.3}$$

We can fit boundary conditions on finite intervals or indeed on semi-infinite intervals like $[0, +\infty)$. For instance on $[0, +\infty)$ we could take $A = 0$. We certainly don't want the e^{+kx} solution for physical problems on such an interval because it's not normalizable– it blows up as $x \to +\infty$.

If we now take the full $(-\infty, \infty)$ interval we cannot construct an individual smooth solution on this interval– we can write down instead

$$y = e^{-k|x|}, \tag{15.4}$$

but our boundary conditions now exclude $e^{k|x|}$.

Let us construct the Green's function y_g satisfying

$$y_g'' + k^2 y_g = \delta(x - x'). \tag{15.5}$$

That is, our solution takes the form

$$y_{L,R}(x) = A_{L,R} e^{ikx} + B_{L,R} e^{-ikx}. \tag{15.6}$$

On either side of x', it is a linear combination of the solutions to the homogeneous equation.

The convention for which wave solutions are right-movers and which are left-movers comes from Kirchhoff. To understand this, we need to recognize that the equation we've written down is really the Fourier-transformed version of a time-dependent PDE. By convention, positive-frequency modes have a time dependence

$$e^{-ikt}. \tag{15.7}$$

Hence our modes with time dependence are really

$$e^{ik(x-t)}, \quad e^{-ik(x+t)}. \tag{15.8}$$

We see that the first solution has constant phase on lines

$$k(x - t) = \delta \implies x = t + \frac{\delta}{k}, \tag{15.9}$$

so these are the right-movers. Conversely points of constant phase move to more negative x as t increases, so the solutions e^{-ikx} are the left-movers.

Let us now impose the boundary conditions that the wave is purely right-going on the right of the δ-function $\delta(x - x')$ and left-going on the left. That is,

$$y = \begin{cases} y_L = Be^{-ikx} & x > x' \\ y_R = Ae^{ikx} & x < x'. \end{cases} \tag{15.10}$$

Now we must fix these constants based on the boundary at $x = x'$. Continuity gives us

$$y_L(x') = y_R(x') \implies B = Ae^{2ikx'} \tag{15.11}$$

and integrating the equation 15.5 over a $(x' - \epsilon, x' + \epsilon)$ interval gives the discontinuity in the derivative. That is,

$$y_R'(x') - y_L'(x') = 1 \implies ikAe^{ikx'} + ikBe^{-ikx'} = 1. \tag{15.12}$$

Plugging in and solving we find that

$$2ikAe^{ikx'} = 1, \tag{15.13}$$

so

$$A = \frac{1}{2ik} e^{-ikx'}, \quad B = \frac{1}{2ik} e^{ikx'} \tag{15.14}$$

and therefore

$$y = \frac{1}{2ik} e^{ik|x-x'|}. \tag{15.15}$$

We can in fact do this calculation in a different, faster way via the method of contour integrals. But we see the analysis extends in a straightforward way to the infinite interval.

Green's functions with one-sided boundary conditions So far our examples have had two-sided boundary conditions, but we can also examine one-sided boundary problems. Consider

$$y'' + y = f(t). \tag{15.16}$$

This is just a driven oscillator. Let us suppose the force begins driving the oscillator (a pendulum, or something) at some time $t = 0$, and that the oscillator was at rest before,

$$y(0) = y'(0) = 0. \tag{15.17}$$

These are one-sided boundary conditions. We can analyze this with the Green's function method, taking our driving force to be a delta function impulse spike at time t,

$$y'' + y = \delta(x - t). \tag{15.18}$$

Hence we can write

$$y_L = A_L \sin x + B_L \cos x, \tag{15.19}$$

$$y_R = A_R \sin x + B_R \cos x. \tag{15.20}$$

Our initial condition that $y(0) = y'(0) = 0$ implies that $A_L = B_L = 0$. The particle stays at rest until hit by the impulse at time t, so

$$y_L = 0. \tag{15.21}$$

Both the integration constants are therefore in y_R. By continuity,

$$y_L(t) = y_R(t) \implies A_R \sin t + B_R \cos t = 0. \tag{15.22}$$

The other boundary condition gives us

$$y_R'(t) - y_L'(t) = 1 \implies A_R \cos t - B_R \sin t = 1. \tag{15.23}$$

Solving we find that

$$G = \begin{cases} 0 & x < t \\ \sin(x - t) & x > t. \end{cases} \tag{15.24}$$

Hence the solution for a general driving force is

$$y(x) = \int_0^x dt f(t) \sin(x - t). \tag{15.25}$$

That is, we say that the final motion is nothing more than the sum of the impulses from the driving force, added up only until the time we are looking at. This is the *retarded Green's function*.

We could have changed the boundary conditions, though. Instead of having an oscillator at rest that gets a kick from the delta function, we could have had an oscillator moving with one unit of momentum, and then we hit it exactly right to cancel its momentum and leave it stopped at all future times.

Formal integral solutions There are three main mathematical formalisms we use to do quantum mechanics. We can use infinite-dimensional matrix equations, differential equations, or integral formulations (path integrals). Let us discuss the integral idea a bit more.

Let us consider the equation

$$\mathcal{L}y = \lambda y. \tag{15.26}$$

We are used to thinking of this as an eigenvalue problem. But suppose instead we considered it as an inhomogeneous equation, i.e. it is certainly of the form

$$\mathcal{L}y = f, \quad f \equiv \lambda y. \tag{15.27}$$

Hence we can write down the solution in terms of the Green's function as

$$y(x) = \int_a^b dt\, G(x,t) f(t) = \lambda \int_a^b dt\, G(x,t) y(t). \tag{15.28}$$

Such equations are known as Fredholm integral equations. A closely related variant is the Volterra integral equation, which allows x to be one of the integration limits as well.[45] Hence we have rewritten this as an inverse problem,

$$Ay = \lambda y \implies y = \lambda \frac{1}{A} y. \tag{15.29}$$

For finite matrices we don't really gain anything from doing this but in the infinite-dimensional case sometimes integrals produce quantities that are nicer to work with.

Notice that if we hit both sides with the operator $\mathcal{L}$, we get

$$\mathcal{L}_x y(x) = \lambda \int_a^b dt\, \mathcal{L}_x G(x,t) y(t). \tag{15.30}$$

Provided that the operator is self-adjoint and G obeys the boundary conditions, we see that this formal expression precisely solves the original differential equation, i.e.

$$\mathcal{L}_x y(x) = \lambda \int_a^b dt\, \mathcal{L}_x G(x,t) y(t) = \lambda \int_a^b dx\, \delta(x-t) y(t) = \lambda y(t). \tag{15.31}$$

The story of the path integral is well-documented in Feynman-Hibbs, which establishes how Feynman turned Dirac's idea about infinitesimal propagation of the wavefunction into the language of Green's functions and from there to an integral equation compactly expressing quantum mechanics in the language of a sum over amplitudes.

Consider the free wave equation

$$y'' + k^2 y = 0. \tag{15.32}$$

We might imagine a guitar player plucking a string, giving it some initial displacement and maybe a bit of initial velocity too. An interesting question might then be to decompose this initial disturbance into some set of complete modes– namely, the harmonics of the string. The ends of the string fix some boundary conditions, that the wave must be zero at $x = 0$ and $x = L$. Hence we get the standing waves of the string, which take the form

$$\sin\Big(\frac{n\pi x}{L}\Big), \tag{15.33}$$

and we find that the permitted wavelengths take discrete values. $\lambda = L/n$. Treating this now as a Schrödinger problem, the energy is proportional to $(y'(x))^2$, so we see that the modes with large n (more nodes) have larger energy, and this holds more generally. Certain modes are excluded by the boundary conditions because they interfere destructively at the boundaries, and modes with more nodes are higher-energy. This analysis works so long as our operators are Hermitian. If we permit strange boundary conditions or worse, regular singular points, then we could have examples of momenta leaking into the nucleus and other nonphysical behavior.

Suppose we have an operator of the form

$$\mathcal{L} y = (py')' + qy. \tag{15.34}$$

Under what conditions is it self-adjoint? We see that

$$\int_a^b z^* \mathcal{L} y = \int_a^b qyz^* + \int \left([z^*(py')]' - (z^{*\prime} py)' + y(z^{*\prime} p)' \right) \tag{15.35}$$

$$= \int_a^b y \mathcal{L} z^* + \big[p(z^* y' - y z^{*\prime}) \big]_a^b. \tag{15.36}$$

That means that if this boundary term vanishes, then the operator is Hermitian and we get all the nice properties about orthogonality of eigenfunctions and so on.[46]

[45] We saw such an equation in the retarded Green's function for the driven oscillator, above.

[46] One more comment on references– Ince is unsurpassed in his discussion of Sturm-Liouville problems.

Lecture 16.

Monday, November 25, 2019

"I like to abuse you some, but in a way that allows you to get stronger and ultimately survive."
–Nemanja Kaloper

Note. We're figuring out the logistics for the final now. It seems that we'll try to reschedule for Thursday of finals week but approval has to be unanimous. Also, midterm solutions will be posted a bit later.

Let's return to the Sturm-Liouville problem. We were looking at the difference

$$y^* \mathcal{L} z - (\mathcal{L} y^*) z. \tag{16.1}$$

If we write our operator in self-adjoint form

$$\mathcal{L} = \frac{d}{dx}\left(p \frac{d}{dx}\right) + q(x) \tag{16.2}$$

then what remains in this difference is a total deriviatve,

$$\int (y^* \mathcal{L} z - (\mathcal{L} y^*) z) = \left(y^* p z' - y^{*\prime} p z\right)|_a^b, \tag{16.3}$$

where we have recognized a total derivative.

Let us notice moreover that if y is an eigenfunction with eigenvalue λ_y, then

$$\int (y^* \mathcal{L} z - (\mathcal{L} y^*) z) = (\lambda_z - \lambda_y^*) \int y^* z. \tag{16.4}$$

Hence if the boundary conditions are satisfied (this integral vanishes) then we are guaranteed that the eigenvalues are real ($\lambda_z - \lambda_z^* = 0$), and the eigenfunctions corresponding to distinct eigenvalues are orthogonal ($\int z y^* = 0$ for $y \neq z$).

If the boundary conditions are *not* satisfied then our normalizable functions couple to *non-normalizable* functions, which means at the level of the Schrödinger equation that if you prepare an electron at some time, it has a non-zero probability of just vanishing and never coming back. This is a very bad violation of unitarity.

Example 16.5. Consider Laguerre's equation

$$x D^2 y + (1 - x) y' = \lambda y, \tag{16.6}$$

where solutions obey the boundary conditions

$$y(0) < \infty, \quad y \to 0 \text{ as } x \to \infty. \tag{16.7}$$

This is the radial equation of the hydrogen atom. The first boundary condition tells us that our wavefunction is sufficiently smeared away from $r = 0$ in order to avoid feeling the infinite $1/r$ potential at the origin. In fact, one may argue that the function solves Neumann boundary conditions at the origin. This stops the electron from simply disappearing into the nucleus; this is forbidden by the uncertainty principle, since this would be an extremely localized electron state.

The second boundary condition is in fact our ballistics problem from last time. If we overshoot or undershoot, our function will not "hit" zero at ∞. For large x, this equation has asymptotic behavior

$$x D^2 y - x D y \simeq 0. \tag{16.8}$$

We're not interested in arbitrarily large λ; we can always take x larger. Hence if we define $P = y'$ then

$$P' = P \implies P = e^x, y = C + C_1 e^x. \tag{16.9}$$

This tells us that a generic solution will have a finite piece and an exponential piece as $x \to \infty$, so we must pick λ to turn off the exponential part. We can equivalently do the Frobenius expansion at $x \to \infty$ or $x' = 1/x \to 0$. If we do this expansion, we would see that some terms generically blow up *unless* λ takes special values. This is precisely the quantization condition.

That is, we get some recursion relation

$$a_n = \frac{w(n) - \lambda}{\hat{w}(n)} a_{n-2} \tag{16.10}$$

and if $w(N) = \lambda$ for some N then all higher-order coefficients a_N and above vanish identically.

Example 16.11. Let us now look at Legendre's equation,

$$-(1-x^2)y'' + 2xy' = \lambda y \tag{16.12}$$

or equivalently

$$(x^2-1)y'' + (x^2-1)'y' = \lambda y. \tag{16.13}$$

This is the angular equation of hydrogen if we substitute $x = \cos\theta$, and its solutions are spherical harmonics, etc. Notice this equation is bound to have issues when x is near ± 1. These are poles of the y'' coefficient, corresponding to the north and south poles of the sphere.

We can study this bad behavior by making the change of variables $z = x - 1$ and studying $z \to 0$. Then the factors become

$$(1-x)(1+x) = 2z + O(z^2). \tag{16.14}$$

The derivative doesn't change so our equation becomes

$$2zy'' + 2y' = \lambda y. \tag{16.15}$$

We now multiply by z to get

$$2z^2y'' + 2zy' = z\lambda y \tag{16.16}$$

so that this looks like a hypergeometric equation and its solutions will again be singular unless we tune the parameter λ.

The function should be good at the north *and* south poles, though. This imposes a stricter condition on our solutions. We can certainly write down solutions that are nice at the north pole and not at the south. But there's nothing special about either pole so indeed we must be careful to pick our solutions.

If we Frobenius this guy as

$$\sum a_n (z-1)^{n+s} \tag{16.17}$$

and write the indicial equation we will find

$$s(s-1) = 0. \tag{16.18}$$

Taking the $s = 0$ solution we find

$$a_{j+2} = \frac{j(j+1) - \lambda}{(j+1)(j+2)} a_j. \tag{16.19}$$

For a random value of λ, we'll get an infinite series solution rather than a polynomial. But notice that for large j, $j(j+1) - \lambda \approx j(j+1)$ and so the coefficients go to

$$a_{j+2} \approx \frac{j(j+1)}{(j+1)(j+2)} a_j \approx a_j. \tag{16.20}$$

As $x \to 1$ we see that we get an infinite sum of 1s in the Taylor expansion. Unless! Unless we pick

$$\lambda = l(l+1) \tag{16.21}$$

for some l, and then a_{l+2} and a_{l+4} and so on all vanish. Our infinite sum has now terminated and we have a polynomial which is perfectly regular at $x = 1$.

This is nothing more than the quantization of the angular wavefunctions (the azimuthal part).

Example 16.22. Suppose we now consider a spherically symmetric potential which is constant $V_0 < 0$ inside some radius R and $V = 0$ outside. We set up the Schrödinger equation:

$$-\frac{1}{2m}\nabla^2\psi + V\psi = E\psi. \tag{16.23}$$

This has the appearance of a Sturm-Liouville problem. We have one boundary condition that $\psi(0) < \infty$, and another that $\psi \to 0$ as $r \to \infty$. But we also know that the more nodes there are, the higher energy a state will have. This tells us that we should suppose our ground state has spherical symmetry (the angular wavefunction is trivial). If we change variables to write $\psi(r) = u/r$, then

$$\nabla^2\psi = \frac{1}{r}u'' \tag{16.24}$$

and

$$u'' + \underbrace{2m(E - V_0)}_{k_1^2} u = 0, r < R \tag{16.25}$$

and

$$u'' + \underbrace{2mE}_{-k_2^2} u = 0, r > R. \tag{16.26}$$

Hence these are just oscillators which we need to sew together at $r = R$. In particular, note that

$$V_0 < E < 0 \tag{16.27}$$

so that our solution actually fits in the potential well. These look like a pair of oscillators

$$u'' + k_1^2 u = 0 \tag{16.28}$$

$$u'' - k_2^2 u = 0. \tag{16.29}$$

That is, inside the well, we have

$$A \sin k_1 r + B \cos k_1 r, r < R, \tag{16.30}$$

and outside

$$Ce^{-k_2 r} + De^{+k_2 r}, r > R. \tag{16.31}$$

We must set $D = 0$ so our solutions don't blow up at infinity. In fact, we also need $u/r = \psi$ to be finite at $r \to 0$, which tells us that the cosine (being finite at $r \to 0$) is no good, so we must match

$$u = \begin{cases} A \sin k_1 r & r < R \\ Ce^{-k_2 r} & r > R. \end{cases} \tag{16.32}$$

We can patch these solutions together by continuity at R and continuity of the derivative,

$$u_L = u_R \tag{16.33}$$

$$u'_L = u'_R. \tag{16.34}$$

Note that we are free to fix overall normalization so WLOG we could take $A = 1$. Then we have two boundary conditions, so one fixes C and the other provides a relation between

$$k_1 = \sqrt{2m(E - V_0)}, \quad k_2 = \sqrt{2m|E|}. \tag{16.35}$$

This will be our quantization condition.

Hence

$$\sin(k_1 R) = Ce^{-k_2 R} \tag{16.36}$$

$$k_1 \cos(k_1 R) = -k_2 Ce^{-k_2 R}. \tag{16.37}$$

Eliminate C by dividing these equations and we find that

$$-k_2 = k_1 \cot(k_1 R), \tag{16.38}$$

or equivalently

$$\tan(k_1 R) = -\frac{k_1}{k_2} \tag{16.39}$$

We can then solve this graphically since $k_2 = \sqrt{2m|E|} \sim -\sqrt{2m(k_1)^2 + |V_0|}$. Hence the intersection of the graphs

$$k_1 \cot(k_1 R) = -\sqrt{2m(k_1)^2 + |V_0|} \tag{16.40}$$

give us the legitimate quantized solutions which obey the boundary conditions.

Lecture 17.

Wednesday, November 27, 2019

Today we proceed to the topic of partial differential equations (PDEs). These are the creatures we use to describe physical laws.

PDEs require us to know how to take derivatives on spaces of dimension greater than 1. If our space M where the function lives is sufficiently nice (it admits local coordinates) then we can define a directional derivative by defining a parametrized curve on M and looking at how a function f changes as we compare nearby points. That is, we can define

$$\frac{\delta f}{\delta \mathbf{l}} = \mathbf{l} \cdot \boldsymbol{\nabla} f \tag{17.1}$$

where

$$\boldsymbol{\nabla} f = (\partial_x f, \partial_y f, \partial_z f, \dots) \tag{17.2}$$

and the partial derivatives are given in the usual way by

$$\partial_x f(x,y,z) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, y, z) - f(x,y,z)}{\Delta x}. \tag{17.3}$$

Definition 17.4. A partial differential equation is an equation of the form

$$U(x_i, f, \partial_{x_i} f, \partial_{x_i}\partial_{x_j} f, \dots) = 0. \tag{17.5}$$

That is, our equation relates a function f, the coordinates x_i, and derivatives of f with respect to the coordinates.

Definition 17.6. We shall call a PDE *linear* if upon replacing $f \to \lambda f$, our equation becomes

$$\lambda U + F(x), \tag{17.7}$$

i.e. our equation just rescales up to an additive piece that does not depend on λ.

Let us recall that we can always turn a single nth-order differential equation into a system of n first-order equations by defining new variables (sort of like conjugate momenta) and rewriting. That is,

$$\mathcal{L}(D)X = 0 \leftrightarrow y_i' = f_i(x, y_i). \tag{17.8}$$

We'll also restrict our interest to homogeneous linear first-order PDEs.[47] That is, our equation takes the form

$$\sum f_k(x) \partial_k U = 0, \tag{17.9}$$

where $\{f_k(x)\}$ is some set of C^∞ smooth functions. A solution of this equation is a function $U = U(x) = C$ with some number of integration constants. That is, $U(x)$ maps points in the original space to a scalar value C. These are nothing more than equipotential surfaces in the original space. Sufficient data would then be a single point in the original space. Moreover, we can think about small variations in the "potential" value C. That is, $C \to C + \delta C$. Clearly, the function changes the most normal to the surface.

If we consider the tangent plane to the surface, we can also define the normal to the surface

$$\mathbf{n} = \boldsymbol{\nabla} U. \tag{17.10}$$

The tangent vectors on the surface are orthogonal to the normal,

$$\mathbf{f} \cdot \mathbf{n} = 0, \tag{17.11}$$

so that

$$0 = \mathbf{f} \cdot \boldsymbol{\nabla} U = \sum_k f_k \partial_k U. \tag{17.12}$$

[47] Lit reference: Myskys on PDEs.

This tells you that the $f_k(x)$ functions define a set of tangent vectors to a surface.[48] Most of the integration constants will fix the point we live at with respect to our surface. One constant will fix which surface we're on.[49]

Example 17.13. Consider the PDE

$$a\partial_x\phi + b\partial_y\phi = 0. \tag{17.14}$$

It is clear that we can translate between

$$\phi(x,y) = C \tag{17.15}$$

and the parametric representation

$$x = x(C), y = y(C). \tag{17.16}$$

If a, b are constants then we are simply looking for the flows of a vector field which is constant everywhere. Its equipotentials are constant lines where $a\partial_x + b\partial_y$ is normal to the surface.

The natural thing to do in the constant coefficient case is to change coordinates

$$s = ax + by \tag{17.17}$$

$$t = bx - ay. \tag{17.18}$$

We can rewrite our equation in terms of the new variables s, t by the chain rule, i.e. because ϕ implicitly depends on s and t via

$$\phi(x,y) = \phi(x(s,t), y(s,t)), \tag{17.19}$$

we can therefore write

$$0 = a\partial_x\phi + b\partial_y\phi \tag{17.20}$$

$$= a\left(\frac{\partial\phi}{\partial s}a + \frac{\partial\phi}{\partial t}b\right) + b\left(\frac{\partial\phi}{\partial s}b - a\frac{\partial\phi}{\partial t}\right) \tag{17.21}$$

$$= (a^2+b^2)\frac{\partial\phi}{\partial s}. \tag{17.22}$$

Hence any function that depends only on t is a solution,

$$\phi = f(t). \tag{17.23}$$

Note that this worked since the new equation for s was positive-definite. It can only become degenerate if $a^2 + b^2 = 0$, in which case there's no equation.

In other words, along curves of constant t, the function itself is constant since the solution is independent of s. Lines of constant t are called *characteristics*. If you specify initial data on a line of constant s, then this sets initial data for the entire space. Conversely we are not free to specify arbitrary data on curves of constant t. In general, we must give initial data on a surface that crosses every characteristic at least once.

Another way to look at it is like this: suppose we have a parametrized curve $x = x_0(\lambda), y = y_0(\lambda)$, and we moreover are given initial data $\phi(x_0(\lambda), y_0(\lambda))$ on this curve. Then we can write the full solution as

$$\phi(x,y) = \phi(x_0,y_0) + \partial_x\phi|_{x_0,y_0}\Delta x + \partial_y\phi|_{x_0,y_0}\Delta y. \tag{17.24}$$

That is,

$$\frac{d\phi}{dl} = \partial_x\phi\frac{dx}{dl} + \partial_y\phi\frac{dy}{dl}, \tag{17.25}$$

by the chain rule, and we also know that

$$a\partial_x\phi + b\partial_y\phi = 0 \tag{17.26}$$

is our differential equation, so this provides us a system of equations for $\partial_x\phi, \partial_y\phi$. For this equation to have a solution, we must have

$$\begin{vmatrix} \frac{dx}{dl} & \frac{dy}{dl} \\ a & b \end{vmatrix} \neq 0. \tag{17.27}$$

[48] Equivalently we're specifying a n-dimensional vector field on an n-dimensional space. We specify one constraint, the value of the function, so the surface is of codimension 1.

[49] Curiously, there's a link between the Frobenius theorem from differential geometry to the cosmological constant problem in effective field theory. See Weinberg '89, Reviews of Modern Physics

This tells us that

$$b\frac{dx}{dl} - a\frac{dy}{dl} \neq 0, \tag{17.28}$$

or equivalently our initial data must not be specified on a curve parallel to the characteristics. Otherwise, we are bound to run into pathologies.

Let us now consider something more exotic:

$$a\partial_x\phi + b\partial_y\phi + q(x,y)\phi = F(x,y). \tag{17.29}$$

If we replace x and y by s,t as before, we find that

$$(a^2+b^2)\partial_s\phi + \hat{q}(s,t)\phi = \hat{F}(s,t) \tag{17.30}$$

where

$$\hat{q}(s,t) = q(x(s,t),y(s,t)) \tag{17.31}$$

and F is similar. Notice that this new equation has only one derivative, so now our equation has reduced to an ODE problem where t just parametrizes which curve we are on.

Another natural generalization is to three variables,

$$a\partial_x\phi + b\partial_y\phi + c\partial_z\phi = 0. \tag{17.32}$$

If a,b,c are constant we can certainly just define s,t,u with say

$$s = ax + by + cz \tag{17.33}$$

and t,u in the plane orthogonal to s. Then our equation simplifies as before, i.e.

$$(a^2+b^2+c^2)\partial_s\phi = 0 \tag{17.34}$$

and $\phi = f(t,u) =$ constant gives the solutions. Characteristics run along s and take us off our initial data surface. We just follow the flow to extend our solution to the entire space.

In the 3D case, we now need initial data along two curves l, l' in the surface, and then we can solve for $\partial_x\phi, \partial_y\phi, \partial_z\phi$. The following determinant must be nonvanishing:

$$\begin{vmatrix} \frac{dx}{dl} & \frac{dy}{dl} & \frac{dz}{dl} \\ \frac{dx}{dl'} & \frac{dy}{dl'} & \frac{dz}{dl'} \\ a & b & c \end{vmatrix} \tag{17.35}$$

Finally, we will preview our first second-order equation,

$$(\partial_x^2 - \partial_y^2)\phi = 0. \tag{17.36}$$

This is a wave equation. Clearly, this factorizes as

$$(\partial_x - \partial_y)(\partial_x + \partial_y)\phi = 0. \tag{17.37}$$

If we now introduce new variables

$$u = x + y \tag{17.38}$$

$$v = x - y, \tag{17.39}$$

then in terms of the new coordinates, we can rewrite this as

$$\partial_u\partial_v\phi = 0. \tag{17.40}$$

These are none other than the null coordinates in Minkowski space. This tells us that functions of u or v alone are solutions to the equation, and so

$$\phi(u,v) = f(u) + g(v) \tag{17.41}$$

solve the wave equation, such that f, g are at least twice differentiable.

Lecture 18.

Monday, December 2, 2019

"There may be some mumble about non-linear partial differential equations but God forbid you ever have to do that."

–Nemanja Kaloper

Last time, we discussed the fact that once we rewrite the wave equation as

$$\partial_u \partial_v \phi = 0 \tag{18.1}$$

in terms of null coordinates u, v, then our general solutions take the famous d'Alembert form

$$\phi = g(x-t) + f(x+t) = g(u) + f(v). \tag{18.2}$$

That is, a solution is specified by two arbitrary functions, or equivalently five integration constants. We must specify initial data on a good surface (formally a codimension one spacelike surface), and then we can solve the equation within some domain of dependence.

Consider now

$$(\partial_x^2 + \partial_y^2)\phi = 0, \tag{18.3}$$

a 2D Laplace equation. It is clear we can now factor this in terms of complex combinations of the original variables,

$$(\partial_x + i\partial_y)(\partial_x - i\partial_y)\phi = \partial_z \partial_{z^*}\phi = 0, \tag{18.4}$$

which tells us that ϕ is a sum of *holomorphic* and *anti-holomorphic parts*

$$\phi = f(z) + g(z^*), \tag{18.5}$$

where we have defined

$$z = x + iy. \tag{18.6}$$

Alternately we may write this as a sum of a real and an imaginary part. We'll discuss this more in our complex analysis unit next quarter.[50]

Consider some equation

$$a\partial_x^2\phi + 2b\partial_x\partial_y + c\partial_y^2 = 0 \tag{18.7}$$

In principle we could have other first- and zeroth-order terms $d\partial_x\phi + e\partial_y\phi + f\phi$, but we can always turn off the first-order terms by a judicious change of variables. In general we can't get rid of the inhomogeneous part. But with our second-order term we can define new coordinates by completing the square as

$$a(\partial_x^2 + 2\frac{b}{a}\partial_x\partial_y) = a\left((\partial_x + \frac{b}{a}\partial_y)^2 - \left(\frac{b}{a}\right)^2\partial_y^2\right) \tag{18.8}$$

and then this second term can be absorbed into the existing ∂_y^2 term. That is, by the quadratic equation,

$$0 = \left(\frac{b+\sqrt{\Delta}}{\sqrt{c}}\partial_x + \sqrt{c}\partial_y\right)\left(\frac{b-\sqrt{\Delta}}{\sqrt{c}}\partial_x + \sqrt{c}\partial_y\right)\phi, \tag{18.9}$$

where

$$\Delta = b^2 - ac \tag{18.10}$$

is the discriminant.

- If $\Delta > 0$ then this behaves like the wave (hyperbolic) equation.
- On the other hand, if $\Delta < 0$ then this is a Laplace (elliptic) equation.

But there is a third, degenerate case where $\Delta = 0$. In that case, we have an equation which can be written as

$$\partial_{x'}^2\phi = 0, \tag{18.11}$$

which is therefore just an ordinary differential equation. If we turn back on the ∂_y dependence then we have the heat (parabolic) equation, which is

$$\partial_{x'}^2\phi = \partial_y\phi. \tag{18.12}$$

[50] A good reference on this is Ruell Churchill, which "drops you in media res" and just goes from there.

The initial data for the wave equation is usually called the Cauchy problem. That is, we take something with the interpretation of a time slice and we evolve it to the future. On the other hand, initial data is much harder to specify for elliptic and parabolic equations.

In the case of the elliptic equation, we must specify either the function ϕ on the boundary or its normal derivative $\hat{n} \cdot \nabla\phi$ on the boundary. This is none other than electrostatics. Either we specify the potential itself (as sourced by some battery) or the electric field (as sourced by some charges) on the boundary. Sometimes the boundary is infinite, so we modify the boundary conditions to restrict us to having no charges at infinity.

In the case of the parabolic equation, we instead provide conditions at the endpoints of a (possibly half-infinite) interval. We fix the "temperature" at the endpoints and let heat flow over time.

Separation of variables

Example 18.13. Suppose we have a (rectangular) box, some region where we want to solve the equation

$$(\partial_x^2 + \partial_y^2 + \partial_z^2 + k^2)\phi = 0. \tag{18.14}$$

This could be a Fourier transform of a wave equation, for instance. Some values are given on the boundary. How do we solve this? We'll take advantange of linearity and completeness by expanding our solutions in a complete set that respects the symmetries of the boundary conditions. That is, let us look for *separable* solutions of the form

$$\phi(x, y, z) = \sum_{i,j,k} c_{ijk} X_i(x) Y_j(y) Z_k(z) \tag{18.15}$$

where c_{ijk} are some constants. Hence when we plug this ansatz into the equation, we get

$$X''(x)YZ + XY''(y)Z + XYZ''(z) + k^2 XYZ = 0. \tag{18.16}$$

Suppose we now divide the whole expression by XYZ. Then

$$\frac{X''}{X} + \frac{Y''}{Y} + \frac{Z''}{Z} + k^2 = 0. \tag{18.17}$$

We can trivially rewrite this as

$$-\frac{X''(x)}{X} = \frac{Y''(y)}{Y} + \frac{Z''(z)}{Z} + k^2. \tag{18.18}$$

But the LHS can depend only on x, while the RHS can depend only on y and z. Hence the expression on the LHS is in fact a constant, i.e.

$$-\frac{X''}{X} = l^2. \tag{18.19}$$

For now, l can be complex. It's just a placeholder until we fix the boundary conditions. But of course this equation is simply

$$X'' + l^2 X = 0, \tag{18.20}$$

an ordinary differential equation which we can readily solve. The RHS must also be a constant, so

$$\frac{Y''}{Y} + \frac{Z''}{Z} + k^2 - l^2 = 0. \tag{18.21}$$

This is just another equation of the same form, so we can write

$$m^2 = -\frac{Y''}{Y} = \frac{Z''}{Z} + k^2 - l^2 \tag{18.22}$$

and then solve for the modes of Y and Z. That is, Z obeys

$$0 = \frac{Z''}{Z} + k^2 - m^2 - l^2 = \frac{Z''}{Z} + n^2 \tag{18.23}$$

Hence we have turned our one PDE into three ODEs:

$$X'' + l^2 X = 0 \tag{18.24}$$

$$Y'' + m^2 Y = 0 \tag{18.25}$$

$$Z'' + n^2 Z = 0. \tag{18.26}$$

If we fix Dirichlet boundary conditions (ϕ vanishes on the sides of the box) then we get sines as our solutions,[51] where

$$l, m, n = (L/a, M/b, N/c)\pi \tag{18.27}$$

where the box has dimensions $a \times b \times c$, and the solution is

$$\sin(L\pi x/a)\sin(M\pi y/b)\sin(N\pi z/c) \tag{18.28}$$

where

$$k^2 = \left[(L/a)^2 + (M/b)^2 + (N/c)^2\right]\pi^2. \tag{18.29}$$

Let us now look at the same equation but in cylindrical symmetry,

$$\Delta\phi + k^2\phi = 0. \tag{18.30}$$

The Laplacian is generally pretty horrible to work out in curvilinear coordinates, but we can do it with a trick.[52] The volume of a little cube in cylindrical coordinates is

$$(dr)(dz)(rd\phi). \tag{18.31}$$

The trick is now to compare two vectors $\mathbf{r}$ and $\mathbf{r} + d\mathbf{r}$. In cylindrical coordinates this is

$$d\mathbf{r} = \mathbf{e}_r dr + \mathbf{e}_\phi r d\phi + \mathbf{e}_z dz. \tag{18.32}$$

If we forgot these relations, we could rederive them by taking the differential of r. Thus

$$\begin{aligned} d\mathbf{r}^2 &= dr^2 + r^2 d\phi^2 dz^2 \\ &= \sum h_i^2 dq_i^2. \end{aligned}$$

These functions h_i tell you how the coordinate "unit" vectors scale up or down as a function of the point we're looking at.

Now the Laplacian is like div grad, i.e. for a little box, it is the flux of ϕ divided by the volume of this box:

$$\boldsymbol{\nabla}^2\phi = \frac{\text{flux}[\phi]}{\Delta V}. \tag{18.33}$$

These fluxes will give us second derivatives. That is,

$$\boldsymbol{\nabla}^2\phi = \sum_i \frac{1}{h_1 h_2 h_3}\partial_i\left(\frac{h_1 h_2 h_3}{h_i^2}\partial_i\phi\right). \tag{18.34}$$

Using this trick, we get

$$0 = \frac{1}{r}\partial_r(r\partial_r\varphi) + \frac{1}{r^2}\partial_\phi^2\varphi + \partial_z^2\varphi + k^2\varphi \tag{18.35}$$

as our equation in cylindrical coordinates. Now we appeal to separation of variables again, as

$$\varphi = R\Phi Z \tag{18.36}$$

and we get

$$0 = \frac{1}{rR}(rR')' + \frac{1}{r^2\Phi}\Phi''(\phi) + \frac{Z''(z)}{Z} + k^2. \tag{18.37}$$

The z variable separates out nicely as

$$Z''(z) + n^2 Z = 0. \tag{18.38}$$

Hence we are left with

$$0 = \frac{1}{rR}(rR')' + \frac{1}{r^2\Phi}\Phi''(\phi) + k^2 - n^2. \tag{18.39}$$

If we define $k^2 - n^2 = l^2$ and multiply through by r^2 then we get

$$0 = \frac{r}{R}(rR')' + \frac{\Phi''(\phi)}{\Phi} + l^2 r^2. \tag{18.40}$$

Now Φ is a constant on its own, satisfying

$$\Phi''(\phi) + m^2\Phi = 0 \tag{18.41}$$

[51] We could have seen this just as well by passing to Fourier space first and gotten the separation constants right away.

[52] In GR notation, this is just $\boldsymbol{\nabla}^2\phi = \frac{1}{\sqrt{-g}}\partial^\mu(\sqrt{-g}\partial_\mu\phi)$.

and what remains is

$$r(rR')' + (l^2r^2 - m^2)R = 0. \tag{18.42}$$

Simplifying,

$$0 = r^2R'' + rR' + (l^2r^2 - m^2)R, \tag{18.43}$$

which is the Bessel equation. We see that z is still behaving like a Cartesian coordinate. The variable ϕ is our magnetic quantum number m. We may impose periodic boundary conditions if ϕ takes its full range $[0, 2\pi]$, which gives sines and cosines. On the other hand, if the problem is specified on some angular slice (not necessarily periodic), then we may not have these nice boundary conditions. For m an integer, we may write

$$0 = \rho^2R'' + \rho R' + (\rho^2 - m^2)R, \tag{18.44}$$

where there are now two linearly independent sets of solutions, $J_m(lm) = J_m(\rho)$ and the Neumann solutions $N_m(\rho r)$. We can throw away the Neumann solutions since they diverge as $r \to 0$ but in solutions for exterior regions (say, $r_a < r < r_b$) we must keep them and be more careful.[53]

In spherical coordinates, we can again use our trick to write down the Laplacian

$$\frac{1}{r^2}\partial_r(r^2\partial_r\varphi) + \frac{1}{r^2\sin\theta}\partial_\theta(\sin\theta\partial_\theta\varphi) + \frac{1}{r^2\sin^2\theta}\partial_\phi^2\varphi + k^2\varphi. \tag{18.45}$$

One can check that separation of variables gives us the spherical harmonics for the angular dependence, and that the radial dependence turns out to be (spherical) Bessel functions.

Lecture 19.

Wednesday, December 4, 2019

"People could not live inside the shell of the earth. They would just float. Unless of course they lived on the equator. Real estate would be extremely precious."

–Nemanja Kaloper

Let's finish this. We were talking about separation of variables in spherical equations. We wanted to solve

$$(\boldsymbol{\nabla}^2 + \mathbf{k}^2)\varphi = 0 \tag{19.1}$$

in terms of a solution which takes the form

$$\varphi = R\Theta\Phi. \tag{19.2}$$

Separation of variables gives us two angular equations:

$$-\frac{1}{\Phi}\Phi''(\phi) = m^2, \tag{19.3}$$

where we take m to be positive and quantized in units of the period (for the regular periodic boundary conditions), and

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right) - \frac{m^2 + \sin^2\theta}{\sin^2\theta}\Theta = 0, \tag{19.4}$$

which is just the Legendre equation in terms of a variable $x = \cos\theta$.

Finally, we have the radial equation

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + k^2R - \frac{\lambda R}{r^2} = 0. \tag{19.5}$$

Solving the Φ equation readily gives us a phase

$$\Phi = e^{im\phi} \tag{19.6}$$

The Θ equation gives us (associated) Legendre polynomials,

$$\Theta = P_l^m(x = \cos\theta) \tag{19.7}$$

[53] These Bessel functions are modified if we study the equation on a cone, i.e. if we make ϕ periodic on $[0, \theta_0]$ with $\theta_0 < 2\pi$. We might have to worry about the conical singularity as $r \to 0$. In that case, it's like modifying the charge– the field lines are denser than they would be with the regular 2π periodicity.

which are derivatives of the standard Legendre polynomials $D_x^m P_l(x)$. Once we impose the quantization condition, the Legendre polynomials (and therefore their derivatives) are perfectly regular at the poles.[54] Finally, rewriting the radial equation yields

$$r^2R'' + 2rR' + (k^2r^2 - \lambda)R = 0, \tag{19.8}$$

where

$$\lambda = l(l+1). \tag{19.9}$$

The solutions are Bessel functions J_ν for half-integer values of ν.

Suppose we have a pure Laplace equation rather than Poisson. That is, the radial equation is

$$r^2R'' + 2rR' - l(l+1)R = 0. \tag{19.10}$$

Then the solutions are just of the form r^α: we see that

$$\alpha(\alpha+1) = l(l+1), \tag{19.11}$$

so that the solutions are $r^l, r^{-(l+1)}$. Hence our full solution to the Laplace equation in spherical coordinates is

$$\varphi = \sum_{l,m}\left(A_{lm}r^l + \frac{B_{lm}}{r^{l+1}}\right)P_l^m(\cos\theta)e^{im\phi} \tag{19.12}$$

where $-l \le m \le l$ and $l = 0, 1, 2, \ldots$. Of course, this is the multipole expansion. For $l = 0$ the first term is constant (r^0) and the second term is $1/r$. When $l = 0$ we must have $m = 0$ and therefore the $l = m = 0$ term gives us a constant plus the standard $1/r$ Coulombic potential.[55]

Look now at the $l = 0, m = 1$ term. The polar angular dependence is now $P_0^1(\cos\theta) = \cos\theta$ and the $1/r^{l+1}$ radial dependence becomes $1/r^2$. This gives us a $\cos\theta/r^2$ dependence, which we should recognize as the potential of an electric dipole. Higher l will give us the quadrupole, octupole, and so on. The dominant behavior at large r is the monopole behavior, unless the monopole moment vanishes.

However, notice that the first term A_{lm} blows up at infinity, while the second term B_{lm} blows up as $r \to 0$. We say the second describes an "external solution" outside some distribution of charges, while the first is an internal one.

More generally we might imagine patching together solutions based on some boundary conditions. Suppose for instance we impose a boundary condition on a sphere of some radius a, such that

$$\mathbf{E} = \frac{\partial V}{\partial \mathbf{n}} = -V_0\cos\theta. \tag{19.13}$$

These specify the normal derivative on a closed surface, so they are Neumann boundary conditions. We may check that

$$\int_0^\pi V_0\cos\theta\sin\theta d\theta d\phi = \int_1^{-1} V_0\cos\theta d(\cos\theta)d\phi = 0. \tag{19.14}$$

That is, there is no net charge on the sphere. This allows for the Gauss law to be satisfied; more generally, there are complications arising from such a calculation when these sorts of integrals do not vanish.

We now see by comparison to our general multipole expansion that $m = 0$, since there must be no azimuthal (axial) angular dependence of our problem. We throw away the growing solution in the exterior region since it diverges as $r \to \infty$, and so our solutions are constrained to take the form

$$\frac{B_{l0}}{r^{l+1}}P_l(\cos\theta). \tag{19.15}$$

To match the other angular dependence with the boundary condition, we see that $l = 1$, which means our solution is just

$$\frac{B_{10}}{a^2}\cos\theta, \tag{19.16}$$

on the boundary, and therefore

$$V = -V_0\left(\frac{a}{r}\right)^2\cos\theta. \tag{19.17}$$

[54] We can see this from the Rodriguez formula for the Legendre polynomials.

[55] There's a cute property of associated Legendre polynomials, that $P_l^{-m} = (-1)^m P_l^m$.

We recognize this as the dipole potential. Apparently the field of such a distribution is exactly the dipole potential outside!

What about the interior? We would keep the r^l terms, and if we again keep only the $m = 0, l = 1$ terms, we now get

$$V = -V_0 r \cos\theta = -V_0 z \tag{19.18}$$

inside. The key element here is the orthonormality of the angular dependence eigenfunctions. Because these eigenfunctions are orthonormal and complete, we can expand a general solution in their basis and just match coefficients to the boundary conditions.

We can look at the Laplace equation in two dimensions,

$$\Delta\Phi = 0. \tag{19.19}$$

Note that solutions Φ can have no extrema in the interior, for otherwise the Laplacian could not vanish– suppose we chose axes such that $(\partial_x^2 + \partial_y^2)\Phi = 0$. A maximum would have $\partial_x^2 < 0$ and $\partial_y^2 < 0$, which would imply that $\Delta\Phi > 0$.

We shall now prove uniqueness of solutions. Suppose we had two solutions Φ_1, Φ_2 satisfying the boundary conditions. Notice that

$$\Delta\phi = \Delta\Phi_1 - \Delta\Phi_2 = 0, \tag{19.20}$$

where

$$\phi|_b = 0 \tag{19.21}$$

vanishes on the boundary. Now notice that

$$0 = \int \phi \boldsymbol{\nabla}^2 \phi \tag{19.22}$$

$$= \int \boldsymbol{\nabla}(\phi \boldsymbol{\nabla}\phi) - (\boldsymbol{\nabla}\phi)^2 \tag{19.23}$$

$$= \int_b d\mathbf{A} \cdot (\phi \boldsymbol{\nabla}\phi) - \int (\boldsymbol{\nabla}\phi)^2. \tag{19.24}$$

Notice this quantity is like an energy, since $\boldsymbol{\nabla}^2\phi \sim \rho$. The first term vanishes since ϕ vanishes on the boundary. The second one is negative semi-definite, so this whole expression can only be zero if $\boldsymbol{\nabla}\phi = 0$ identically. We conclude that ϕ is at most a constant, so these solutions differ at most by a constant; this is just our gauge symmetry. This argument works with Neumann boundary equations also; it's just the $\boldsymbol{\nabla}\phi$ term which vanishes on the boundary instead, which forces $\boldsymbol{\nabla}\phi$ to vanish everywhere.

For the wave equation, initial conditions are much simpler. Recall that solutions are a sum of a left-mover and a right-mover, i.e.

$$\phi = f(x - t) + g(x + t), \tag{19.25}$$

such that

$$\phi_0(x, 0) = f(x) + g(x) \tag{19.26}$$

$$\dot{\phi}_0(x, 0) = -f'(x) + g'(x). \tag{19.27}$$

We can therefore integrate the second equation from $x - t$ to $x + t$ as

$$\int_{x-t}^{x+t} dx\, \dot{\phi}_0(x, 0) = -\int_{x-t}^{x+t} dx\, f'(x) + \int_{x-t}^{x+t} dx\, g'(x) \tag{19.28}$$

$$= -f(x + t) + f(x - t) + g(x + t) - g(x - t). \tag{19.29}$$

We can then evaluate the first equation at $x + t, x - t$ to get

$$\phi(x + t, 0) + \phi(x - t, 0) = f(x + t) + f(x - t) + g(x + t) + g(x - t). \tag{19.30}$$

Adding this with the lat equation gives us

$$\phi_0(x + t, 0) + \phi_0(x - t, 0) + \int_{x-t}^{x+t} dx\, \dot{\phi}_0(x, 0) = 2[f(x - t) + g(x + t)] = 2\phi(x, t). \tag{19.31}$$

Equivalently,

$$\phi(x, t) = \frac{1}{2}\phi_0(x - t) + \frac{1}{2}\phi_0(x + t) + \frac{1}{2}\int_{x-t}^{x+t} dx\, \dot{\phi}_0(x). \tag{19.32}$$

This solution is somewhat intuitive– setting $t = 0$ gives back $\phi_0(x)$, while taking the time derivative makes the first two terms cancel and the second one exactly match $\dot{\phi}_0$. The first two terms propagate the initial condition off to the left and right, while the integral adds up the influence of the initial "velocities" we specified at $t = 0$.

Finally, we conclude with the heat equation,

$$\partial_t \phi = \partial_x^2 \phi, \tag{19.33}$$

which we will just treat in one spatial dimension. Let us again try separation of variables, $\Phi = TX$. Then

$$\frac{T'(t)}{T} = \frac{X''(x)}{X} = \beta \tag{19.34}$$

and each of these must be constant, where we have called the separation constant β. Then

$$T = e^{\beta t}, \quad X = \bar{A} e^{\sqrt{\beta} x} + \bar{B} e^{-\sqrt{\beta} x}. \tag{19.35}$$

We can also define $\beta = -\alpha^2$. We ought to take $\beta \in \mathbb{R}$ so that temperature does not oscillate in time. In fact, we should (on physical grounds) take $\beta < 0$ in order to say our system is cooling down over time, and thus α is real and positive, which means that the X dependence is just sines and cosines. That is,

$$\Phi = [A \sin(\alpha x) + B \cos(\alpha x)] e^{-\alpha^2 t} \tag{19.36}$$

or more generally

$$\Phi = \int d\alpha C(\alpha) [A \sin(\alpha x) + B \cos(\alpha x)] e^{-\alpha^2 t}. \tag{19.37}$$

We could therefore set initial conditions by specifying the heat distribution at $t = 0$, and let it evolve. We might also have rewritten our solution in terms of a new dimensionless scale $u = \alpha x$, in which case we would see that the $e^{-\alpha^2 t}$ dependence becomes e^{-t/u^2}. If we suppose that Φ is therefore only a function of the coordinates in the combination

$$U(x/\sqrt{t}), \tag{19.38}$$

then

$$\partial_t \Phi = -\frac{1}{2} \Phi \frac{x}{t^{3/2}}. \tag{19.39}$$

We arrive at an equation of the form

$$2U'' + \zeta U' = 0, \tag{19.40}$$

which is first-order in U'. Hence the solution is

$$U(\zeta) = C_1 \int_0^\zeta d\zeta' e^{-\zeta'^2/4} + C_2. \tag{19.41}$$

The limit $\zeta \to \infty$ corresponds to $t \to 0$, and we can explicitly calculate

$$U(\zeta \to \infty) = C_1 \int_0^\infty d\zeta' e^{-\zeta'^2/4} + C_2 = C_1 \sqrt{\pi} + C_2. \tag{19.42}$$

Similarly we can evaluate $U(-\infty) = -C_1 \sqrt{\pi}$, which means that we can solve for C_1, C_2 in terms of initial conditions.

It follows that we can take derivatives with respect to x, t of our original solution and generate new solutions. The solution is therefore

$$\frac{1}{\sqrt{\pi}} \int_{-\infty}^\infty d\zeta e^{-\zeta^2} C(x - 2\zeta\sqrt{t}). \tag{19.43}$$