PHYSICS 204A: MATH METHODS

IAN LIM
LAST UPDATED OCTOBER 30, 2019

These notes were taken for Physics 204A, *Math Methods*, 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

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

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.

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

$$\begin{aligned}0 &< \langle \Delta u|\Delta u\rangle && (2.20)\\ &= \langle u - \textstyle\sum' c_n^u \phi_n | u - \sum' c_n^u \phi_n\rangle && (2.21)\\ &= \langle u|u\rangle - \textstyle\sum' |c_n^u|^2. && (2.22)\end{aligned}$$

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.

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

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

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

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.

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

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

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

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

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

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

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

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

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

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

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

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

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

or equivalently

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

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

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

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:

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

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.

[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.

Lecture 5.

Wednesday, October 9, 2019

"If you can ask a general-sounding question that will help you with the homework, by all means do it. I encourage you to be sneaky." –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?

$$|\phi_i'\rangle = U|\phi_i\rangle \tag{5.3}$$

$$= \sum_j |\phi_j\rangle\langle\phi_j|U|\phi_i\rangle \tag{5.4}$$

$$= \sum_j U_{ji}|\phi_j\rangle. \tag{5.5}$$

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

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

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

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

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:

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

[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.

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

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

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

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

[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.

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

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

[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.

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

$$\ddot{x}_1 = -\frac{k}{m}(x_1 - x_2) \tag{6.46}$$

$$\ddot{x}_2 = -\frac{k}{M}(x_2 - x_1) - \frac{k}{M}(x_2 - x_3) \tag{6.47}$$

$$\ddot{x}_3 = -\frac{k}{m}(x_3 - x_2) \tag{6.48}$$

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

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

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 $d^2 = 0$.)

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

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

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

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

$$D + \lambda = e^{-\lambda x}De^{\lambda x}$$
$$D - \lambda = e^{\lambda x}De^{-\lambda x}$$
$$(D-\lambda)^n = e^{\lambda x}D^ne^{-\lambda x}.$$

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.

[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.

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

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 $\lambda_\pm = \mu + i\omega$, so that the equation factors as

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

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 we want some z such that $D^2 z = 0$, and hence

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

which gives

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

This is equivalent to the statement 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.

Example 7.31. Consider

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

so

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

so now $(D - 1)z(x)e^x$ has solution $z(x) = Ae^x + e^x$ and so

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

Then

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

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}. 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.

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

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

Hence the equation we were given has a regular singular point at $x = \infty$.

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

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,[14] 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

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

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-Lindelof, which placed constraints on how far we could extend our solution around the initial condition. 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,

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

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

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:

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

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

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

Hence the exponential factor $e^{\alpha x}$ will cancel the dependence $e^{-\alpha x}$ from the derivatives and we have $e^{\alpha x} D^2 g = 0$ where $g = e^{-\alpha x} f$. We can solve this very easily; $D^2 g = 0 \implies g$ 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 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 $D^2 y = 0$ has solutions $y_1 = 1$ and $y_2 = x$, so the general solution is $c_1 y_1 + c_2 y_2 = c_1 + c_2 x$.

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.

Notice that for linear dependence, $\sum c_p y_p = 0$ 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.[15] 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.[16]

[15] 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.

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

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.

Lecture 9.

Wednesday, October 23, 2019

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

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

Then its Wronskian is

$$W = \begin{vmatrix} y_1 & y_2 & y \\ y_1' & y_2' & y' \\ y_1'' & y_2'' & y'' \end{vmatrix} = \begin{vmatrix} y_1 & y_2 & y \\ y_1' & y_2' & y' \\ -Py_1' - Qy_1 & -Py_2' - Qy_2 & -Py' - Qy \end{vmatrix}. \tag{9.10}$$

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, $W = 0$ 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.[17]

Moreover, suppose we have one solution in hand, 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

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

[17] 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.

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 $y_2 = uy_1$, we have

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

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.[18]

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.

Let us now try to solve the harmonic oscillator potential

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

by a series method,

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

Hence

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

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

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

such that

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

[18] 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.

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:

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

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

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.

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 we do not take s to a priori 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*.

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.[19]

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.[20] 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 $a_1, a_3, \ldots$.

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.

[19] This is an exercise. Notice that $\log x = p \implies dx/x = dp$. Find p in terms of α and β.

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

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 $a_1 = 0$.[21] 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

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

[21] 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.[22]

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

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

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

[22] 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

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

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

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