

Physics 212A Lecture Notes, Fall 2015

Lecturer: McGreevy

Last updated: 2022/03/21, 13:52:12

These notes are dedicated to the
memory of Prof. E. Commins,
to whom they owe a great deal.

Contents

1	Mathematical machinery for quantum mechanics	4
1.1	Dirac notation	6
1.2	Bases	9
1.3	Eigenspaces	12
1.4	Functions of operators	16
2	Rules of quantum mechanics for isolated systems	19
2.1	Illustrations with isolated two-state quantum systems	23
2.1.1	Where to get your qbits.	25
2.1.2	Amplitudes add, not probabilities	32
2.1.3	Quantum and classical uncertainty	33
2.1.4	Density operator	34
2.2	Composite quantum systems	37
3	∞-dimensional Hilbert spaces, wave mechanics	40
3.0.1	The position operator	42
3.0.2	The derivative operator	43
3.1	Uncertainty principle	48
3.2	Wave mechanics	49
3.2.1	Probability current	51
3.2.2	Free particle	52
3.2.3	Particle in potential well	55
3.2.4	Relative and center-of-mass coordinates	58
3.2.5	Scattering and bound states: piecewise-constant potentials	59
3.2.6	Index theorem	64
3.2.7	Harmonic oscillator	65
3.2.8	Shape-invariance	71
3.2.9	Particle in a uniform force field	73
3.3	WKB approximation	76

3.4	Wave mechanics of charged particles	81
3.4.1	Aharonov-Bohm effect	83
3.4.2	Magnetic monopoles and quantization of charge and angular momentum	84
3.5	Path integrals, briefly	87
4	Theory of Angular Momentum	94
4.1	Continuous symmetries and the quantum Noether theorem	94
4.2	Rotations in ordinary 3-space	97
4.3	Spectra of the angular momentum operators	100
4.4	Matrix elements of the angular momentum operators	103
4.5	Orbital angular momentum	106
4.6	Addition of angular momentum	110
4.7	Wigner-Eckhart Theorem	113
5	Entanglement	115
5.0.1	GHZM states	116

Initial Tentative Plan:

1. Mathematical machinery for quantum mechanics.

2. Axioms for isolated quantum systems.

Examples of (approximately) isolated two-state systems.

3. Infinite-dimensional Hilbert spaces.

This includes some discussion of potential problems, harmonic oscillator algebras, coherent states.

4. Composite quantum systems and entanglement.

5. Some theory of quantum information.

No cloning, “decoding theorem”, distinguishability, quantum teleportation.

6. Quantum mechanics versus local hidden variables (Bell’s inequalities, GHZ).

7. Measurement, decoherence, time evolution of open quantum systems.

8. Continuous symmetries.

Addition of angular momentum, spherical harmonics.

9. Perturbation theory.

As the title indicates, this list was my initial guess for what we’ll study. What actually happened, based on what I thought the class needed most urgently to learn, was 1,2,3,4,8.

Sources for these notes: E. Commins, *Quantum Mechanics, An Experimentalist’s Perspective*; G. Baym, *Lectures on Quantum Mechanics*; B. Schumacher and M. Westmoreland, *Quantum processes, systems and information*; M. Le Bellac, *Quantum Physics*; A. Manohar, notes (taken by Sheena Patel) from Physics 212A, Fall 2014.

1 Mathematical machinery for quantum mechanics

First we must take our mathematical equipment out of the closet.

For reasons that still elude the human scientists, the physical description of a collection of degrees of freedom passes through a Hilbert space, which is a vector space with some bonus material.

A vector space is a collection of things which can be added (with coefficients) to produce another element¹

$$\mathbf{v}, \mathbf{w} \in V \implies a\mathbf{v} + b\mathbf{w} \in V.$$

In quantum mechanics, the allowed coefficients (‘scalars’) a, b are complex numbers. This is the superposition principle.

More formally: the vectors themselves form an abelian group under addition. $\mathbf{v}, \mathbf{w} \in V \implies \mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v} \in V$. Notice that there is a thing called $\mathbf{0}$ which is the zero vector which is the identity element of this group; it is different from the number 0.

The vector space also enjoys an operation of scalar multiplication: $\mathbf{v} \in V \implies a\mathbf{v} \in V$, where a is an element of some field, which we’ll take to be the complex numbers $\mathbb{C}$.

example 1: polynomials: $\mathbf{v}_a \equiv a_0 + a_1x + a_2x^2, \mathbf{v}_b \equiv b_0 + b_1x + b_2x^2, a_i, b_i \in \mathbb{C}$.

example 2: complex-valued differentiable functions on the interval with $f(0) = f(L) = 0$. If you add two of these, you get another. The zero vector is the function which is identically zero: $x \mapsto 0$.

A collection of vectors $\{\mathbf{v}\}$ forms a *basis* for V if every vector in V is a linear combination of elements of the collection, in which case we write $V = \text{span}\{\mathbf{v}\}$.

A collection of vectors $\{\mathbf{v}_i\}_i$ is *linearly independent* if the relation

$$\sum_i a_i \mathbf{v}_i = \mathbf{0}$$

implies $a_i = 0$. The *dimension* of the vector space, $\dim(V)$, is the minimum number of linearly independent vectors needed to make a basis. For example, if we fix the degree of the polynomials in the example above to be $\leq n$, then the dimension is $n + 1$.

Now we add some more structure to our vector spaces. An *inner product space* is a vector space equipped with an inner product, which we temporarily denote $(\cdot, \cdot)$. The

¹The boldface vectors are a temporary notation, until §1.1.

·s here represent mouths via which this machine eats vectors and spits out scalars. So this is a map

$$\begin{aligned}(\cdot, \cdot) : V \otimes V &\rightarrow \mathbb{C} \\ \mathbf{v} \otimes \mathbf{w} &\mapsto (\mathbf{v}, \mathbf{w})\end{aligned}\tag{1.1}$$

(the notation $\otimes$ is just a placeholder here). We will make three assumptions about our inner product:

1. It is *linear* in the second argument:

$$(\mathbf{v}, a\mathbf{w} + b\mathbf{x}) = a(\mathbf{v}, \mathbf{w}) + b(\mathbf{v}, \mathbf{x}).$$

2. It is *skew-symmetric*:

$$(\mathbf{v}, \mathbf{w})^* = (\mathbf{w}, \mathbf{v})$$

where $\star$ is complex conjugation.²

3. Finally, for physics purposes, it will be important that the inner product is *positive*, in the sense that

$$(\mathbf{v}, \mathbf{v}) \geq 0, \quad \text{and} \quad (\mathbf{v}, \mathbf{v}) = 0 \Leftrightarrow \mathbf{v} = \mathbf{0}.$$

(We'll need this because these quantities represent probabilities.)

A positive inner product can be used to make a *norm*:

$$\|\mathbf{v}\| \equiv \sqrt{(\mathbf{v}, \mathbf{v})}.$$

Two important facts about such a norm:

- Cauchy-Schwarz inequality

$$(\mathbf{v}, \mathbf{w})^2 \leq \|\mathbf{v}\|^2 \|\mathbf{w}\|^2$$

This is a complex-vector-space version of the fact that $|\cos \theta| \leq 1$. It is saturated when the two vectors are parallel, $\mathbf{v} \propto \mathbf{w}$.

- Triangle inequality

$$\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|.$$

²Notice that combining assumptions (1) and (2) above we see that the inner product is 'antilinear' in its left argument:

$$(a\mathbf{w} + b\mathbf{x}, \mathbf{v}) = a^*(\mathbf{w}, \mathbf{v}) + b^*(\mathbf{x}, \mathbf{v}).$$

A *Hilbert space* is an inner product space with positive norm which is also *complete*: every Cauchy sequence converges³. For finite dimensional Hilbert spaces, on which I would like to focus, the completeness condition is automatic.

Linear maps. We are going to have a lot of use for linear maps between Hilbert spaces,

$$\mathbf{S} : \mathcal{H} \rightarrow \mathcal{H}', \quad \mathbf{S}(a\mathbf{v} + b\mathbf{w}) = a\mathbf{S}\mathbf{v} + b\mathbf{S}\mathbf{w}, \quad \forall a, b, \mathbf{v}, \mathbf{w}.$$

(Especially in the case when $\mathcal{H}' = \mathcal{H}$, I will sometimes also call them ‘operators,’ or sometimes, perhaps redundantly, ‘linear operators’.) We are often going to have to compose such things, like

$$V_1 \xrightarrow{\mathbf{S}_1} V_2 \xrightarrow{\mathbf{S}_2} V_3 \xrightarrow{\mathbf{S}_3} V_4$$

and we will write the resulting map as

$$\begin{aligned} \mathbf{S}_3\mathbf{S}_2\mathbf{S}_1 : V_1 &\rightarrow V_4 \\ v &\mapsto \mathbf{S}_3\mathbf{S}_2\mathbf{S}_1 v \end{aligned}$$

– it is an unfortunate fact of our lexicography that the operators pile up to the left, the one all the way on the right acts first. The product of operators is associative $(\mathbf{S}_3\mathbf{S}_2)\mathbf{S}_1 = \mathbf{S}_3(\mathbf{S}_2\mathbf{S}_1)$ but it is not necessarily even well-defined (*e.g.* if $V_i \neq V_j$) to act with them in a different order.

In a basis (for $\mathcal{H} = \text{span}\{\mathbf{v}_i\}$ and $\mathcal{H}' = \text{span}\{\mathbf{w}_n\}$), the operator $\mathbf{S}$ is represented by a matrix

$$S_{ni} \equiv (\mathbf{w}_n, \mathbf{S}\mathbf{v}_i).$$

The *adjoint* of $\mathbf{S} : \mathcal{H} \rightarrow \mathcal{H}'$ is a linear map which goes the other way $\mathbf{S}^\dagger : \mathcal{H}' \rightarrow \mathcal{H}$ defined by the condition

$$(\mathbf{w}, \mathbf{S}\mathbf{v}) = (\mathbf{S}^\dagger \mathbf{w}, \mathbf{v}).$$

In a basis, the matrix representation of the adjoint is the conjugate transpose of the matrix for $\mathbf{S}$:

$$S_{in}^\dagger = (\mathbf{S}^\dagger)_{in} \stackrel{\text{def of matrix el't}}{=} (\mathbf{w}_n, \mathbf{S}^\dagger \mathbf{v}_i) \stackrel{\text{def of } \dagger}{=} (\mathbf{S}\mathbf{v}_i, \mathbf{w}_n) \stackrel{\text{skew-sym}}{=} (\mathbf{w}_n, \mathbf{S}\mathbf{v}_i)^* = S_{ni}^*$$

1.1 Dirac notation

Pay careful attention here; it is so simple that it is difficult. We are going to give a precise definition of kets and bras, and will thereafter use them all the time.

³A Cauchy sequence $\{v_i\}$ is one where the successive elements get close together – this is where the norm comes in: $\forall \epsilon, \exists n$ such that $\|v_i - v_j\| < \epsilon$ when $i, j > n$.

We can canonically identify any vector $\mathbf{v} \in \mathcal{H}$ with a linear map

$$\begin{aligned} |v\rangle : \mathbb{C} &\rightarrow \mathcal{H} \\ a &\mapsto a\mathbf{v} \end{aligned}$$

This is a *ket*. Notice that $a|v\rangle = |v\rangle a$.

What is the adjoint $|v\rangle^\dagger$ of this map? It takes a vector and makes a number:

$$\begin{aligned} |v\rangle^\dagger \equiv \langle v| : \mathcal{H} &\rightarrow \mathbb{C} \\ \mathbf{w} &\mapsto (\mathbf{v}, \mathbf{w}) \end{aligned}$$

This is the definition of an element of the dual space $\mathcal{H}^*$, the space of linear functionals on $\mathcal{H}$.

What happens if we compose a ket map and then a bra map? Following the definitions, this is a map

$$\begin{array}{ccccc} \langle v| \langle w| : & \mathbb{C} & \xrightarrow{|w\rangle} & \mathcal{H} & \xrightarrow{\langle v|} & \mathbb{C} \\ & a & \mapsto & a\mathbf{v} & \mapsto & a(\mathbf{v}, \mathbf{w}) \end{array}$$

– it's just multiplication by the number

$$(\mathbf{v}, \mathbf{w}) \equiv \langle v|w\rangle.$$

Together the bra and ket form a bracket.

What if we compose the bra and the ket?

$$\begin{array}{ccccc} |w\rangle \langle v| : & \mathcal{H} & \xrightarrow{\langle v|} & \mathbb{C} & \xrightarrow{|w\rangle} & \mathcal{H} \\ & \mathbf{x} & \mapsto & (\mathbf{v}, \mathbf{x}) & \mapsto & (\mathbf{v}, \mathbf{x})\mathbf{w} \end{array}$$

– this produces a vector parallel to $\mathbf{w}$. Passing now behind the horizon of Dirac notation, consider its action on the ket representation of some vector $|x\rangle$:

$$\underbrace{(|w\rangle \langle v|)}_{\text{operator}} \underbrace{|x\rangle}_{\text{vector}} = |w\rangle \langle v|x\rangle = \underbrace{|w\rangle}_{\text{vector}} \underbrace{\langle v|x\rangle}_{\text{number}}.$$

The three expressions here are just the same thing parsed differently; the first great virtue of the notation is that it makes operations like this automatic. We will see many other virtues. I will use it always from now on. A second virtue that we can see is that we can make it clearer that the action of an operator on a vector produces another vector by writing

$$\mathbf{S}|v\rangle \equiv |\mathbf{S}v\rangle.$$

(In fact, Dirac is inviting us to be creative about what we put in the ket.) So the definition of the adjoint is

$$\langle w | \mathbf{S} | v \rangle = \langle w | \mathbf{S} v \rangle = \langle \mathbf{S}^\dagger w | v \rangle = (| \mathbf{S}^\dagger w \rangle, | v \rangle) = (\mathbf{S}^\dagger | w \rangle, | v \rangle).$$

A slippery point: notice that

$$\langle \mathbf{S}^\dagger w | = (| \mathbf{S}^\dagger w \rangle)^\dagger = (\mathbf{S}^\dagger | w \rangle)^\dagger = \langle w | \mathbf{S}$$

while

$$\langle \mathbf{S} w | = \langle w | \mathbf{S}^\dagger.$$

The skew property of the inner product means

$$\langle w | \mathbf{S} | v \rangle^* = (| w \rangle, | \mathbf{S} v \rangle)^* = (| \mathbf{S} v \rangle, | w \rangle) = \langle v | \mathbf{S}^\dagger | w \rangle.$$

Also:

$$(\mathbf{S} \mathbf{T})^\dagger = \mathbf{T}^\dagger \mathbf{S}^\dagger.$$

On the relation between $|a\rangle$ and $\langle a|$. This is a slippery point so I will belabor it a bit: our notation very strongly suggests that each ket vector $|a\rangle$ in our Hilbert space $\mathcal{H}$ has a corresponding bra vector $\langle a|$ in the dual space $\mathcal{H}^*$. Who is it?

Given a vector space V over $\mathbb{C}$, recall that its *dual space* V^* is the space of *linear functionals* on V ; this means

$$V^* \equiv \{\text{linear maps} : V \rightarrow \mathbb{C}\}.$$

This is also a vector space over $\mathbb{C}$, and it has the same dimension as V itself. (One way to see this is to choose a basis $\{e_i\}_{i=1}^N$ of V ; a basis of V^* is then $\{f_i\}_{i=1}^N$ defined by $f_i(e_j) = \delta_{ij}$.)

So far we haven't used the inner product. With an inner product, there is a *natural* (basis independent) way to identify V and V^* which I will temporarily call $\mathcal{I}$:

$$\begin{aligned} \mathcal{I} : V &\rightarrow V^* \\ v &\mapsto \mathcal{I}(v) \\ |a\rangle &\mapsto \langle a| \end{aligned}$$

The definition of $\mathcal{I}$ is so simple that it is hard to understand:

$$\underbrace{\mathcal{I}(v)}_{\in V^*}(\underbrace{w}_{\in V}) \equiv \langle v | w \rangle.$$

$\underbrace{\hspace{1.5cm}}_{\in \mathbb{C}}$

1.2 Bases

On an inner product space, there is no reason to suffer with a basis which is not orthonormal (ON). A basis $\{|j\rangle\}$ is orthonormal means

$$\langle j|k\rangle = \delta_{jk} \equiv \begin{cases} 1, & \text{if } j = k \\ 0, & \text{if } j \neq k \end{cases}$$

Given a less awesome basis $\{|1_b\rangle, |2_b\rangle, \dots\}$ (where b is for ‘bad’) we may make it ON by the

Gram-Schmidt procedure: Pick an order. First normalize the first vector:

$$|1\rangle \equiv \frac{1}{\sqrt{\langle 1_b|1_b\rangle}} |1_b\rangle.$$

Then make a second vector orthogonal to the first by subtracting the component of $|2_b\rangle$ along $|1\rangle$:

$$|2_a\rangle \equiv |2_b\rangle - |1\rangle \langle 1|2_b\rangle$$

(a is for ‘almost’). Then normalize it:

$$|2\rangle \equiv \frac{1}{\sqrt{\langle 2_a|2_a\rangle}} |2_a\rangle$$

Make the third vector orthogonal to the first two,

$$|3_a\rangle \equiv |3_b\rangle - \sum_{j=1,2} |j\rangle \langle j|3_b\rangle$$

and then normalize that one. Keep going until you run out of dimensions. When I say ‘basis’ I will usually mean ON basis.

[End of Lecture 1]

Notice that $|1\rangle \langle 1|$ is a projector onto the 1 direction. (It satisfies $(|1\rangle \langle 1|)^2 = |1\rangle \langle 1|$ and it is hermitian.) Also a projector is the sum of two orthogonal projectors: $\sum_{j=1,2} |j\rangle \langle j|$, the projector onto the subspace spanned by $|1\rangle$ and $|2\rangle$. The ultimate logical extension of this is if we sum over all the elements of the basis

$$\sum_{j=1}^{\dim \mathcal{H}} |j\rangle \langle j| = \mathbb{1} \tag{1.2}$$

is the identity operator. This equation is called a completeness relation.

As a crucial illustration of this, expand an arbitrary vector in an ON basis:

$$|v\rangle = \sum_j v_j |j\rangle .$$

Who are the coefficients v_j ? Act with an arbitrary bra operator from the basis:

$$\langle k|v\rangle = \sum_j v_j \underbrace{\langle k|j\rangle}_{=\delta_{kj}} = v_k$$

Reassembling the vector, we have

$$|v\rangle = \sum_k \langle k|v\rangle |k\rangle = \sum_k |k\rangle \langle k|v\rangle = \underbrace{\left(\sum_k |k\rangle \langle k| \right)}_{=\mathbb{1}} |v\rangle .$$

The lesson of this exercise is: to express the vector $|v\rangle$ in the basis $\{|k\rangle\}$, act with the identity, resolved in that basis.

What we've just done can be regarded as a 'derivation' of Dirac notation [Feynman III-8-1]: Consider the inner product of two vectors, which we denote by

$$\langle \chi | \phi \rangle .$$

For real vectors in 3-space, we are used to computing this using Pythagoras, that is, by projecting the two vectors onto a preferred orthonormal basis, and adding the components, which we can do by the linearity property. That is, we can expand each vector in the given basis: the ket is

$$|\phi\rangle = \underbrace{\phi_1}_{=\langle 1|\phi\rangle} |1\rangle + \phi_2 |2\rangle + \dots$$

and the bra is (recall that the inner product is *antilinear* in the bra vector)

$$\langle \chi | = \langle 1 | \underbrace{\chi_1^*}_{=\langle \chi | 2 \rangle} + \langle 2 | \chi_2^* + \dots$$

$$\text{Therefore, the inner product is } \langle \chi | \phi \rangle = \sum_i \langle \chi | i \rangle \langle i | \phi \rangle .$$

BUT: this is true for any χ , so we may as well erase the χ :

$$|\phi\rangle = \sum_i |i\rangle \langle i | \phi \rangle$$

and we are forced directly to Dirac's notation.

But there is one further logical step we can take here. This relation is also true for all ϕ , so we can erase the ϕ , too (!):

$$| = \sum_i |i\rangle \langle i| .$$

The little line $|$ stands for the identity operator. I will usually write it a little fancier, as $\mathbb{1}$, to make it clear that it is an operator.

Change of basis.

Notice that the identity operator is defined to do nothing to anyone :

$$\mathbb{1} |v\rangle = |v\rangle , \quad \forall |v\rangle .$$

This is a basis independent statement. But a completeness relation (1.2) is an expression for $\mathbb{1}$ in a particular basis. This is a powerful fact.

Consider the overlap between two arbitrary vectors:

$$\langle \beta | \alpha \rangle = \langle \beta | \mathbb{1} | \alpha \rangle = \sum_i \langle \beta | i \rangle \langle i | \alpha \rangle = \sum_i \beta_i^* \alpha_i .$$

In the second step, we inserted the identity in the i basis.

A linear operator in a basis is a matrix: $T_{kj} \equiv \langle k | \mathbf{T} | j \rangle$. The matrix elements of a linear operator change between bases as follows. Suppose we are given two bases of the same Hilbert space, $\{|j\rangle\}, \{|j'\rangle\}$. Then we may expand the elements of one basis in the other:

$$\begin{aligned} |j'\rangle &= \sum_k |k\rangle \underbrace{\langle k | j' \rangle}_{\equiv U_{kj}} \\ T'_{jk} &= \langle j' | \mathbf{T} | k' \rangle = \sum_l \langle j' | l \rangle \sum_m \underbrace{\langle l | \mathbf{T} | m \rangle}_{=T_{lm}} \langle m | k' \rangle = \sum_{m,l} U_{jl}^* T_{lm} U_{mk} \end{aligned} \quad (1.3)$$

This is matrix multiplication: conjugation by the basis-change matrix

$$T' = U^\dagger T U .$$

The basis change operator U is unitary. A *unitary* operator is one which preserves the norms of all states:

$$\| \hat{U} |\psi\rangle \|^2 = \| |\psi\rangle \|^2 \quad \forall |\psi\rangle .$$

This means that

$$\hat{U} \hat{U}^\dagger = 1, \quad \text{and} \quad \hat{U}^\dagger = \hat{U}^{-1} .$$

In general, **unitary operators implement changes of basis**. To see this more explicitly, suppose we are given two ON bases for our $\mathcal{H}$: $\{|n\rangle, n = 1..N\}$ and $\{|a_n\rangle, n = 1..N\}$, and a 1-to-1 correspondence between the two $|n\rangle \leftrightarrow |a_n\rangle$. Define $\mathbf{U}$ to be the linear operation which takes $\mathbf{U}|n\rangle = |a_n\rangle$. Taking the adjoint gives $\langle n|\mathbf{U}^\dagger = \langle a_n|$. Then

$$\mathbf{U} = \mathbf{U}\mathbb{1} = \mathbf{U} \sum_n |n\rangle \langle n| = \sum_n (\mathbf{U}|n\rangle) \langle n| = \sum_n |a_n\rangle \langle n|.$$

Similarly,

$$\begin{aligned} \mathbf{U}^\dagger &= \mathbb{1}\mathbf{U}^\dagger = \sum_n |n\rangle \langle n| \mathbf{U}^\dagger = \sum_n |n\rangle \langle a_n|, \\ \mathbf{U}\mathbf{U}^\dagger &= \sum_{nm} |a_n\rangle \langle n|m\rangle \langle a_m| = \sum_n |a_n\rangle \langle a_n| = \mathbb{1}. \end{aligned}$$

(A similar story holds for $\mathbf{U}^\dagger\mathbf{U}$.) It is a unitary operator whose matrix elements are the matrix we defined above in (1.3).

Compare this discussion with your experience of changing bases for $\mathbb{R}^2$.

1.3 Eigenspaces

In general, if we act with a linear operator $\mathbf{A}$ on some vector $|v\rangle$, we end up with $\mathbf{A}|v\rangle = |Av\rangle$ a vector pointing in a different direction. For a given $\mathbf{A}$, we can try to find a special set of characteristic vectors or *eigenvectors* with the property that the action of $\mathbf{A}$ preserves their direction (but not necessarily their magnitude): $\mathbf{A}|\alpha\rangle = \alpha|\alpha\rangle$.

Q: Where do we get a good basis for our Hilbert space? A: From the Hermitian operators (aka observables). An operator $\mathbf{A}$ is hermitian if it is self-adjoint $\mathbf{A} = \mathbf{A}^\dagger$ and *bounded*; the latter condition means it takes normalizable states to normalizable states. For finite-dimensional Hilbert spaces, this just means none of its matrix elements is ∞ .

Fact 1: Hermitian operators have real eigenvalues and orthonormal (ON) eigenbases.

Proof: Let $|\alpha\rangle, |\beta\rangle$ be normalized eigenvectors of $\mathbf{A}$. So:

$$\mathbf{A}|\alpha\rangle = \alpha|\alpha\rangle \tag{1.4}$$

$$\text{and } \mathbf{A}|\beta\rangle = \beta|\beta\rangle \implies \langle\beta|\mathbf{A}^\dagger = \langle\beta|\beta^*.$$

Notice that we are taking advantage of Dirac's offer to put whatever we want in the box, and labeling the states by the eigenvalue of $\mathbf{A}$. This is a really good idea. Notice also that the eigenvalue equation (1.4) is a linear condition on $|\alpha\rangle$, so it does not determine the normalization of $|\alpha\rangle$; we take advantage of this to assume that $|\alpha\rangle$ is normalized.

Hit (1.4) on the left with $\langle\beta|$:

$$\langle\beta|\mathbf{A}|\alpha\rangle = \alpha\langle\beta|\alpha\rangle = \beta^*\langle\beta|\alpha\rangle.$$

This implies

$$0 = (\beta^* - \alpha)\langle\beta|\alpha\rangle.$$

So: First consider the case where $|\beta\rangle = |\alpha\rangle$, and $\langle\alpha|\alpha\rangle = 1$ so we learn that $\alpha = \alpha^*$: the eigenvalues are real.

If $\alpha \neq \beta$, *i.e.* the eigenvalues are distinct, we learn that $\langle\beta|\alpha\rangle = 0$, the eigenvectors are orthogonal. (They are not automatically normalized, since $a|\alpha\rangle$ is also an eigenvector with the same eigenvalue. It is up to us to normalize them (for example when Mathematica spits them out.))

If the eigenvalues are degenerate (different vectors give the same eigenvalue), it is up to us to choose an orthonormal basis for the degenerate eigenspace (for example by finding a complete set of commuting operators, see below).

I emphasize again that the eigenvalue equation (1.4) is a linear condition. For each value of α , it defines a subspace of $\mathcal{H}$ (which may be empty), which we can call the eigenspace V_α . This space may have dimension zero, it may have dimension one, it may have dimension bigger than one, in which case we say the eigenvalue α is *degenerate*.

For which linear operators $\mathbf{A}$ is the eigenvalue equation (1.4) so nicely solvable? A similar situation for the eigenvectors obtains if $\mathbf{A}$ is unitary, but the eigenvalues are phases rather than real numbers. On the homework, you'll show that a general condition which encompasses both these examples is the condition that $\mathbf{A}$ is *normal*, which means $[\mathbf{A}, \mathbf{A}^\dagger] = 0$.

Fact 2: Hermitian operators which commute $\mathbf{AB} = \mathbf{BA}$ can be simultaneously diagonalized: *i.e.* we can find a basis in which they are both diagonal. (We will denote the *commutator* of two operators by $[\mathbf{A}, \mathbf{B}] \equiv \mathbf{AB} - \mathbf{BA}$.)

Idea of proof: Consider an eigenvector of $\mathbf{A}$, $\mathbf{A}|a\rangle = a|a\rangle$. If $[\mathbf{A}, \mathbf{B}] = 0$ then we have

$$\mathbf{A}(\mathbf{B}|a\rangle) = \mathbf{B}(\mathbf{A}|a\rangle) = \mathbf{B}(a|a\rangle) = a(\mathbf{B}|a\rangle)$$

(the parentheses are just to direct your attention). This equation says that $\mathbf{B}|a\rangle$ is ALSO an eigenvector of $\mathbf{A}$ with eigenvalue a . SO: by the theorem above, if the eigenvalue a is a non-degenerate eigenvalue of $\mathbf{A}$, then we learn that this vector must also point along $|a\rangle$: $\mathbf{B}|a\rangle \propto |a\rangle$ that is

$$\mathbf{B}|a\rangle = b|a\rangle$$

for some complex number b , which we see is an eigenvalue of $\mathbf{B}$.

If on the other hand, the eigenvalue a is a *degenerate* eigenvalue of $\mathbf{A}$, (that is: there is more than one linearly independent eigenvector of $\mathbf{A}$ with eigenvalue a and hence a degenerate eigenspace of $\mathbf{A}$ with eigenvalue a (with dimension > 1)) then the action of $\mathbf{B}$ can generate another element of the subspace V_a of eigenvectors of $\mathbf{A}$ of eigenvalue a . That is: $\mathbf{B}|a\rangle$ is not necessarily *parallel* to $|a\rangle$. But: we can then diagonalize $\mathbf{B}$ within this subspace and label a nice orthonormal basis for the subspace as $|a, b\rangle$, by the eigenvalue of $\mathbf{A}$ and those of $\mathbf{B}$. If there is still a degeneracy, you need to find another operator.

This leads us to the useful notion of a *complete set of commuting operators*. A complete set of commuting operators allow us to specify an orthonormal basis of $\mathcal{H}$ by their eigenvalues, using their simultaneous eigenvectors. If we have in our hands an operator with a completely non-degenerate spectrum (no two eigenvalues are equal), then it is a complete set by itself. For example: spectrum of the position operator $\mathbf{x}$ for a particle on a line provides a nice ON basis for that Hilbert space (as does the momentum operator). This suggests a way to resolve the problem of degeneracies encountered above at (1.6): we can write the projector onto the degenerate eigenspace of $\mathbf{A}$ as

$$\mathbf{P}_\alpha = \sum_{\beta} |\alpha, \beta\rangle \langle \alpha, \beta| \quad (1.5)$$

– we can find a basis of states for the α subspace by diagonalizing some operator $\mathbf{B}$ which commutes with $\mathbf{A}$: $\mathbf{B}|\alpha, \beta\rangle = \beta|\alpha, \beta\rangle$; the analysis above shows that if $\beta \neq \beta'$ then $\langle \alpha, \beta' | \alpha, \beta \rangle = 0$.

Spectral decomposition: in terms of such an ON basis of eigenvectors of a Hermitian operator, we can make a super-useful resolution of the identity: If

$$\mathbf{A}|\alpha\rangle = \alpha|\alpha\rangle$$

then the identity operator, the one which does nothing to everybody, is:

$$\mathbb{1} = \sum_{\text{all } \alpha} |\alpha\rangle \langle \alpha| = \sum_{\alpha} \mathbf{P}_\alpha.$$

The object $\mathbf{P}_\alpha \equiv |\alpha\rangle \langle \alpha|$ is a *projector* onto the eigenstate $|\alpha\rangle$: $\mathbf{P}_\alpha^2 = \mathbf{P}_\alpha$. Notice that in this basis

$$\mathbf{A} = \sum_{\text{all } \alpha} \alpha |\alpha\rangle \langle \alpha|.$$

This is what a diagonal matrix looks like in Dirac notation. Note that this really does depend on Hermiticity of $\mathbf{A}$; the eigenvectors of a general matrix are not orthogonal and the sum of their projectors will not give $\mathbb{1}$ (try it in Mathematica!).

Actually, the last paragraph was strictly true as long as the eigenvalue α is non-degenerate; if $\mathbf{A}$ has degenerate eigenvalues, this presents a problem for our scheme of labeling the states by eigenvalues of $\mathbf{A}$. We can still write

$$\mathbf{A} = \sum_{\text{all } \alpha} \alpha \mathbf{P}_\alpha \quad (1.6)$$

where now $\mathbf{P}_\alpha$ to be understood as the projector onto the (usually one-dimensional) space spanned by the eigenvectors with eigenvalue α . These operators satisfy

$$\mathbf{P}_n \mathbf{P}_m = \delta_{nm} \mathbf{P}_m, \quad \mathbf{P}_m^\dagger = \mathbf{P}_m \quad .$$

This is again the statement that eigenvectors of a Hermitian operator associated with different eigenvalues are orthogonal.

Finally, I would like to explain the statement “We can diagonalize a Hermitian operator by a unitary transformation.” According to the previous discussion, this is the same as saying that we can find a basis where a Hermitian operator is diagonal.

Suppose given a Hermitian operator $\mathbf{A} = \mathbf{A}^\dagger$. And suppose we are suffering along in some random basis $\{|n\rangle, n = 1..N\}$ in which $\mathbf{A}$ looks like

$$\mathbf{A} = \sum_{nm} |n\rangle \langle m| A_{nm}$$

where $\exists n \neq m$ such that $A_{nm} \neq 0$, *i.e.* the *matrix* A is not diagonal (notice that this is a basis-dependent statement, not an essential property of the *operator* $\mathbf{A}$!). Now consider the eigenvectors of $\mathbf{A}$, $\mathbf{A} |a_i\rangle = a_i |a_i\rangle, i = 1..N$; we can choose $\{|a_i\rangle\}_{i=1}^N$ so that $\langle a_i | a_j \rangle = \delta_{ij}$ they form an orthonormal basis of $\mathcal{H}$.⁴ What are the matrix elements of $\mathbf{A}$ in the $\{|a_i\rangle\}$ basis?

$$\langle a_i | \mathbf{A} | a_j \rangle = \delta_{ij} a_i$$

– this is a diagonal matrix. And how are these matrix elements related to the ones in the other basis? Using the resolution $\mathbb{1} = \sum_i |a_i\rangle \langle a_i|$,

$$\mathbf{A} = \sum_{nm} A_{nm} |n\rangle \langle m| \mathbb{1} = \sum_{ij} \sum_{nm} A_{nm} |a_i\rangle \underbrace{\langle a_i | n \rangle}_{(\mathbf{U}^\dagger)_{in}} \underbrace{\langle m | a_j \rangle}_{\mathbf{U}_{mj}} \langle a_j| = \sum_{ij} (U^\dagger \mathbf{A} U)_{ij} |a_i\rangle \langle a_j|$$

where I’ve left the matrix multiplication implicit in the last expression. We’ve shown that

$$(U^\dagger \mathbf{A} U)_{ij} = \delta_{ij} a_i$$

⁴I say “can choose” because: (1) the normalization is not determined by the eigenvalue equation; (2) if there is a *degeneracy* in the spectrum ($\exists i \neq j$ such that $a_i = a_j$), then it is up to us to make an orthonormal basis for this subspace (e.g. by the Gram-Schmidt process)).

is diagonal. (Multiplying both sides by U from the left and $U^\dagger$ from the right, this implies $A_{nm} = (UaU^\dagger)_{nm}$, where a is the diagonal matrix of eigenvalues.)

Notice that a statement of the form ‘ D is diagonal’ is basis dependent. This leads us to a discussion of what information is not basis dependent:

Basis Invariants.

The *rank* of a matrix or linear operator is the dimension of the space of states that it doesn’t kill. By ‘kill’ I mean give zero when acting upon: $\mathbf{A}$ kills $|v\rangle$ if $\mathbf{A}|v\rangle = 0$. The subspace of vectors killed by $\mathbf{A}$ is called the *kernel* of $\mathbf{A}$. For an operator $\mathbf{A}$ acting on an N -dimensional Hilbert space (representable by an $N \times N$ matrix), $\text{rank}(A) = N -$ the dimension of the kernel of $\mathbf{A}$.

An invertible matrix has no kernel (you can’t undo an operation that gives zero), and hence rank N . The *determinant* of $\mathbf{A}$ is the product of its eigenvalues: $\det \mathbf{A} = \prod_n a_n$, so vanishes if any of them vanish. A matrix with rank less than N has zero determinant.

A matrix like $|n\rangle\langle n|$ has rank 1: it only acts nontrivially on the one-dimensional space of states spanned by $|n\rangle$. $\sum_{n=1}^k |n\rangle\langle n|$ has rank k if $|n\rangle$ are not linearly dependent.

Finally, the *trace* is the sum of the diagonal matrix elements in any basis. This is independent of the choice of basis because

$$\text{tr} A' = \text{tr} UAU^{-1} = \text{tr} U^{-1}UA = \text{tr} A.$$

In the second step we used the cyclic property of the trace: $\text{tr} ABC = \text{tr} BCA = \text{tr} CAB$ which can be shown by inserting resolutions of the identity everywhere. For a diagonalizable operator, the trace may be taken in its eigenbasis, and is therefore the sum of the eigenvalues.

The set of eigenvalues of a diagonalizable operator is also an invariant.

1.4 Functions of operators

[Commins §2.18]

Here are three ways to define functions of operators which map $\mathcal{H}$ to itself. They might be called ‘square operators’, in the sense that their associated matrices are square matrices.

In the eigenbasis. The spectral decomposition

$$\mathbf{S} = UDU^\dagger = \sum_s s |s\rangle\langle s|$$

allows us to define functions of diagonalizable operators:

$$f(\mathbf{S}) \equiv \sum_s f(s) |s\rangle \langle s|$$

By series. Sometimes it is useful to use an infinite-series definition: if the function f as a power series representation $f(x) = \sum_n a_n x^n$, then we can consider the sum $f(\mathbf{S}) \equiv \sum_n a_n \mathbf{S}^n$. For example, $e^{\mathbf{A}} = \sum_{n=0}^{\infty} \frac{1}{n!} \mathbf{A}^n$. Ordering trouble can arise when multiple operators which don't commute get involved. For example, it is in general true that $e^{\mathbf{A}+\mathbf{B}} \stackrel{?}{=} e^{\mathbf{A}} e^{\mathbf{B}}$. [\[End of Lecture 2\]](#)

A useful identity in this context is the Baker-Campbell-Hausdorff (BCH) relation

$$e^{\mathbf{A}} \mathbf{B} e^{-\mathbf{A}} = \mathbf{B} + [\mathbf{A}, \mathbf{B}] + \frac{1}{2!} [\mathbf{A}, [\mathbf{A}, \mathbf{B}]] + \frac{1}{3!} [\mathbf{A}, [\mathbf{A}, [\mathbf{A}, \mathbf{B}]]] + \dots \quad (1.7)$$

Proof of BCH. The proof is instructive. Consider the function

$$f(s) = e^{s\mathbf{A}} \mathbf{B} e^{-s\mathbf{A}}$$

where s is a parameter we will vary from 0 to 1. Differentiating using the product rule gives

$$\partial_s f(s) = (e^{s\mathbf{A}} \mathbf{A}) \mathbf{B} e^{-s\mathbf{A}} - e^{s\mathbf{A}} \mathbf{B} (\mathbf{A} e^{-s\mathbf{A}}) = e^{s\mathbf{A}} [\mathbf{A}, \mathbf{B}] e^{-s\mathbf{A}}.$$

If we differentiate again, we take another commutator with $\mathbf{A}$:

$$\partial_s^2 f(s) = e^{s\mathbf{A}} [\mathbf{A}, [\mathbf{A}, \mathbf{B}]] e^{-s\mathbf{A}}.$$

Now use the series expansion:

$$f(s) = f(0) + s \partial_s f(0) + \frac{1}{2!} \partial_s^2 f(0) s^2 + \frac{1}{3!} \partial_s^3 f(0) s^3 + \dots$$

and set $s = 1$ to get (1.7).

Some useful special cases where the series on the right hand side (RHS) is summable are when $[\mathbf{A}, \mathbf{B}] = 0$, when $[\mathbf{A}, \mathbf{B}] = c\mathbb{1}$, when $[\mathbf{A}, \mathbf{B}] = \mathbf{A}$.

Another example is the geometric series $\frac{1}{1-x} = 1 + x + x^2 + \dots$. Consider

$$\mathbf{B} \equiv \mathbb{1} + \mathbf{A} + \mathbf{A}^2 + \dots$$

$$\mathbf{A}\mathbf{B} = \mathbf{A} + \mathbf{A}^2 + \mathbf{A}^3 = \mathbf{B} - \mathbb{1}.$$

$$\mathbb{1} = \mathbf{B} - \mathbf{A}\mathbf{B} = (\mathbb{1} - \mathbf{A})\mathbf{B} = \mathbf{B}^{-1}\mathbf{B}$$

So $\mathbf{B} = (\mathbb{1} - \mathbf{A})^{-1}$ if the inverse is well-defined.

By product. Finally, recall that another way to define the exponential function is by infinite product:

$$e^y = \lim_{n \rightarrow \infty} \left(1 + \frac{y}{n}\right)^n.$$

Consider a unitary operator that's close to the identity:

$$U(\epsilon) = \mathbb{1} + \mathbf{i}\epsilon\mathbf{K} + \mathcal{O}(\epsilon^2)$$

where $\epsilon \ll 1$ is a small parameter. The condition that U is unitary says

$$\mathbb{1} = U^\dagger U = \mathbb{1} + \mathbf{i}\epsilon(\mathbf{K} - \mathbf{K}^\dagger) + \mathcal{O}(\epsilon^2)$$

so that $\mathbf{K}$ is Hermitian. Then we can construct a unitary for finite $x \equiv n\epsilon$ by repeated action of the unitary labelled by ϵ :

$$U(x) \equiv \lim_{n \rightarrow \infty, \epsilon \rightarrow 0, x = n\epsilon \text{ fixed}} U(\epsilon)^n = e^{-\mathbf{i}x\mathbf{K}}.$$

2 Rules of quantum mechanics for isolated systems

Our next goal is to agree upon the structure of quantum mechanics. We must answer the following questions:

1. **States:** How do we represent our knowledge of a physical system?
2. **Observables:** What can we, in principle, measure about it?
3. **Time Evolution:** If we know the state now, what happens to it later?
4. **Measurement:** What happens when we actually make a measurement?
5. **Composition:** If we put two quantum systems together how do we describe the combined system?

The answers to these questions are quite different for an isolated quantum system (such as the entire universe, or an electron perfectly sequestered in a vacuum chamber) than they are in the more realistic situation where we describe only the part of the world. We'll discuss the idealized case first.

Axiom 1: States

By 'state' here, I mean a complete description of what's going on in a physical system at some instant – a specification of as much data about it as is physically possible to specify. For example, in classical mechanics of a bunch of particles, this is a specification of the coordinates and momenta of every particle. In quantum mechanics, the state of a system is a *ray* in the Hilbert space of the system. A ray is an equivalence class of vectors

$$|v\rangle \simeq a |v\rangle \text{ for } a \in \mathbb{C} . \quad (2.1)$$

So the first axiom really has two parts: first, associated to a physical system, there is a Hilbert space $\mathcal{H}$; second, a state of the the system (when we speak about open systems, we will call this a 'pure state') is an equivalence class of vectors.

By normalizing our vectors, we fix the ambiguity associated with the magnitude $|a|$ of a in (2.1). Even for normalized vectors, the overall phase is not physical. Let me take this opportunity to remind you that the fact that the state is associated with a vector space builds in the superposition principle: if $|v\rangle$ and $|w\rangle$ are possible states, then so is $a |v\rangle + b |w\rangle$. Although the overall phase $a \rightarrow e^{i\varphi}a, b \rightarrow e^{i\varphi}b$ is not physical, the *relative phase* $\arg(b/a)$ (not changed by the above operation) is very important physical information.

Axiom 2: Observables

By observables, we mean properties of a physical system that can in principle be measured. In quantum mechanics, an observable is a self-adjoint (more precisely, Hermitian) operator.

Important Conclusion. An important conclusion, combining the second axiom with the previous discussion of eigenbases, is the following. In quantum mechanics, the choices of basis come from observables. The labels that are supposed to go inside the little kets are possible values of observables – eigenvalues of Hermitian operators. This is why this feature of Dirac’s notation – that we can put whatever we want in the box – is important: there are many possible observables we may consider diagonalizing in order to use their eigenbasis as labels on our kets.

Axiom 3: Dynamics

By ‘dynamics’ I mean dependence on time. Time evolution of a quantum state is implemented by the action of the *Hamiltonian* of the system, via

$$\frac{d}{dt} |\psi(t)\rangle = -\mathbf{i} \frac{\mathbf{H}}{\hbar} |\psi(t)\rangle \quad (2.2)$$

with $\mathbf{i} \equiv \sqrt{-1}$. (This is sometimes called the Schrödinger equation.) To first order in a tiny time-interval, dt , this says:

$$|\psi(t + dt)\rangle = (1 - \mathbf{i} \frac{\mathbf{H}}{\hbar} dt) |\psi(t)\rangle$$

(this is the way you would write it, if you were going to implement this evolution numerically).⁵ I emphasize that specifying who is $\mathbf{H}$ is part of the specification of a quantum system.

The operator

$$\mathbf{U}(dt) \equiv 1 - \mathbf{i} \mathbf{H} dt$$

which generates time evolution by the step dt , is unitary, because $\mathbf{H}$ is self-adjoint (up to terms that are small like dt^2):

$$\mathbf{U}^\dagger \mathbf{U} = \mathbb{1} + \mathcal{O}(dt^2) .$$

Unitary is important because it means that it preserves the lengths of vectors.

Successive application of this operator (if $\mathbf{H}$ is time-independent) gives

$$\mathbf{U}(t) = e^{-it\mathbf{H}} .$$

⁵Notice that only the combination $\frac{\mathbf{H}}{\hbar}$ appears. Below I will just write $\mathbf{H}$. Similarly, instead of $\frac{\mathbf{p}}{\hbar}$ I will just write $\mathbf{p}$.

Notice that this solves the Schrödinger equation (2.2).

So: to evolve the state by a finite time interval, we just act with this unitary operator:

$$|\psi(t)\rangle = \mathbf{U}(t) |\psi(0)\rangle .$$

Notice that energy eigenstates $\mathbf{H}|\omega\rangle = \hbar\omega|\omega\rangle$ play a special role in that their time evolution is very simple:

$$\mathbf{U}(t)|\omega\rangle = e^{-i\frac{\mathbf{H}}{\hbar}t}|\omega\rangle = e^{-i\omega t}|\omega\rangle$$

they evolve just by a phase. Because of the equivalence relation (2.1), such states are appropriately called *stationary states*. The evolution operator is diagonal in the same basis as $\mathbf{H}$ (indeed: $[\mathbf{H}, e^{i\mathbf{H}t}] = 0$):

$$\mathbf{U} = \sum_{\omega} e^{-i\omega t} |\omega\rangle \langle \omega| .$$

Axiom 4: **Measurement**

This one has two parts:

- (a) In quantum mechanics, the answer you get when you measure an observable $\mathbf{A}$ is an eigenvalue of $\mathbf{A}$.
- (b) The measurement affects the state: right after the measurement, the system is in an eigenstate of $\mathbf{A}$ with the measured eigenvalue.

More quantitatively the two parts are: (a) if the quantum state just before the measurement were $|\psi\rangle$, then outcome a is obtained with probability

$$\mathbf{Prob}_{\psi}(a) = \|\mathbf{P}_a |\psi\rangle\|^2 = \langle \psi | \mathbf{P}_a | \psi \rangle .$$

In the case where the eigenvalue a is non-degenerate, we have $\mathbf{P}_a = |a\rangle \langle a|$ and

$$\mathbf{Prob}_{\psi}(a) = \langle \psi | a \rangle \langle a | \psi \rangle = |\langle a | \psi \rangle|^2 .$$

In words, the probability for outcome a_n for measurement of $\mathbf{A}$ is the expectation value of the projector onto the associated eigenspace $\mathbf{P}_n$.

(b) If we get the outcome a , the quantum state becomes not $\mathbf{P}_a |\psi\rangle$ (which is not normalized!) but

$$|\psi\rangle \xrightarrow{\text{measure } \mathbf{A}, \text{ get } a} \frac{\mathbf{P}_a |\psi\rangle}{(\langle \psi | \mathbf{P}_a | \psi \rangle)^{1/2}} \quad (2.3)$$

which is normalized. (Check that if a is non-degenerate, (2.3) reduces to the simpler statement: $|\psi\rangle \xrightarrow{\text{measure } \mathbf{A}, \text{ get } a} |a\rangle$.) Notice that if we do the measurement again right away, the rule tells us that we are going to get the same answer, with probability one.

Some comments:

- Notice that spectral decomposition of our observable $\mathbf{A}$ leads to the familiar expression for expectation values:

$$\langle \mathbf{A} \rangle_\psi \equiv \langle \psi | \mathbf{A} | \psi \rangle = \langle \psi | \sum_n a_n \mathbf{P}_n | \psi \rangle = \sum_n a_n \langle \psi | \mathbf{P}_n | \psi \rangle = \sum_n a_n \mathbf{Prob}_\psi(a_n) .$$

And notice that the fact that hermitian operators resolve the identity is crucial for the probabilistic interpretation: On the one hand

$$1 = \langle \mathbb{I} \rangle = \langle \psi | \psi \rangle = \| |\psi\rangle \|^2 .$$

On the other hand, for any $\mathbf{A} = \mathbf{A}^\dagger$, we can write this as

$$1 = \langle \psi | \psi \rangle = \langle \psi | \left(\sum_n \mathbf{P}_n \right) | \psi \rangle = \sum_n \mathbf{Prob}_\psi(a_n) .$$

Summing over all the probabilities has to give one.

- In light of the probabilistic interpretation in the measurement axiom, it makes sense that we want time evolution to happen via the action of a unitary operator, since the total probability (the probability that *something* will happen, including nothing as a possibility) had better always be 1, and this is equal to the magnitude of the state.
- Notice that while a sum of observables is an observable⁶, a product of observables $\mathbf{AB}$ is not necessarily itself an observable, since⁷

$$(\mathbf{AB})^\dagger = \mathbf{B}^\dagger \mathbf{A}^\dagger . \quad (2.4)$$

That is, $\mathbf{AB}$ is only self-adjoint if $\mathbf{A}$ and $\mathbf{B}$ commute, $[\mathbf{A}, \mathbf{B}] \equiv \mathbf{AB} - \mathbf{BA} = 0$. If the two operators may be simultaneously diagonalized, then by the measurement axiom, we can measure them simultaneously. So in that case, when the order of operation does not matter, it makes sense to think about the (unique) measurement of the product of the two.

⁶ $(\mathbf{A} + \mathbf{B})^\dagger = \mathbf{A}^\dagger + \mathbf{B}^\dagger = \mathbf{A} + \mathbf{B}$

⁷To see this: From the definition of adjoint, if $\mathbf{A} |u\rangle = |v\rangle$, then $\langle u | \mathbf{A}^\dagger = \langle v |$. So: $\mathbf{BA} |u\rangle = \mathbf{B} |v\rangle$, and so $\langle v | \mathbf{B}^\dagger = \langle u | \mathbf{A}^\dagger \mathbf{B}^\dagger$. But this is true for all $|u\rangle$, and we conclude (2.4).

- If the system is isolated (as suggested by the title of this section), how can we possibly make any measurements on it?
- You may notice a glaring difference in character of our Measurement Axiom – all the others, in particular time evolution, involve *linear* operations:

$$\hat{A}(|a\rangle + |b\rangle) = \hat{A}|a\rangle + \hat{A}|b\rangle.$$

Measurement, as described here (by the map labelled “ $\xrightarrow{\text{measure } \mathbf{A}, \text{ get } a}$ ” in (2.3) above) fails this property: it doesn’t play nicely with superpositions. On the other hand, we think that the devices that we use to measure things (*e.g.* our eyeballs) are governed by quantum mechanics, and evolve in time via the (linear!) Schrödinger equation!

We will revisit this issue when we discuss open quantum systems. Indeed, parts of Axiom 4 can be derived from a better understanding of how to implement measurements.

- We haven’t derived these axioms from something more basic, and (the previous comment aside) maybe they can’t be. In particular, the introduction of probability is an attempt to describe the results of experiments like particles going through a double slit, where interference patterns are observed. We don’t know any way to predict what any one particle will do, but we can use this machinery to construct a probability distribution which describes with exquisite accuracy what many of them will do. And this machinery seems to apply to any experiment anyone has ever done.

2.1 Illustrations with isolated two-state quantum systems

A one-dimensional Hilbert space is completely boring: because of the equivalence relation (2.1), there is a unique state of the system. So the simplest nontrivial case is when $\dim(\mathcal{H}) = 2$, a two-state system, aka a qbit.

You saw on the homework that a nice basis (over $\mathbb{C}$) for the linear operators on this Hilbert space (nice in the sense that it is ON with respect to the Hilbert-Schmidt inner product) was made up of the identity and the Pauli matrices

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (2.5)$$

(occasionally I will write $\sigma^x \equiv \sigma^1, \sigma^y \equiv \sigma^2, \sigma^z \equiv \sigma^3$). I have written their matrix representations in the basis where σ^z is diagonal. [\[End of Lecture 3\]](#)

A better definition of the Pauli operators would be: they are hermitian, square to one, and anticommute with each other

$$\sigma^a \sigma^b = -\sigma^b \sigma^a \quad (\text{if } b \neq a). \quad (2.6)$$

A slightly stronger statement which combines the previous two is

$$\sigma^a \sigma^b = \delta^{ab} \mathbb{1} + \mathbf{i} \epsilon^{abc} \sigma^c. \quad (2.7)$$

Here we used the Einstein summation convention on the RHS, and ϵ^{abc} is the completely antisymmetric tensor with $\epsilon^{xyz} = 1$. (I am not making any distinction between upper and lower indices.) Notice that the $\mathbf{i}$ in the second (antisymmetric) term on the RHS is required by hermiticity of the σ s.

The fact that $(\sigma^a)^2 = 1$ means their eigenvalues are all ± 1 . Often we will call the eigenstates of σ^z

$$\sigma^z |\pm_z\rangle = \pm |\pm_z\rangle$$

as $|+_z\rangle \equiv |\uparrow\rangle$, $|-_z\rangle \equiv |\downarrow\rangle$.

(2.6) has the consequence, for example, that

$$\sigma^x |\uparrow\rangle = |\downarrow\rangle, \quad \sigma^x |\downarrow\rangle = |\uparrow\rangle, \quad \sigma^z |\pm_x\rangle = |\mp_x\rangle.$$

They toggle each others' eigenstates.

The general Hamiltonian operator $\mathbf{h}$ on such a system has four real parameters. In the basis where σ^z is diagonal, they can be organized as

$$h = \begin{pmatrix} d_0 + d_3 & d_1 - \mathbf{i}d_2 \\ d_1 + \mathbf{i}d_2 & d_0 - d_3 \end{pmatrix} = d_0 \mathbb{1} + \vec{\sigma} \cdot \vec{d}. \quad (2.8)$$

Once we diagonalize this matrix, we've solved the dynamics for every (isolated) two-state quantum system. In each subfield of physics this solution has a different name. In atomic physics, it is named (like everything else in atomic physics) after Rabi.

Let's consider the special case that $d_0 = 0$ and $\vec{d} = \vec{n}$ is a unit vector. Then $\vec{n} \cdot \vec{\sigma}$ squares to one just like the Pauli matrices themselves. Its eigenvalues are therefore also ± 1 .

If we know the eigenvectors of $\vec{n} \cdot \vec{\sigma}$ we know those for $\mathbf{h}$ with arbitrary $d_{0,1,2,3}$. First, we add a multiple of the identity. But that doesn't change the eigenvectors at all: the identity commutes with everybody, and so can be simultaneously diagonalized with anybody. All this does is add d_0 to each eigenvalue. Second, $\vec{d}$ is not necessarily a unit vector; rather $\vec{d} = |d|\vec{n}$, where $|d| \equiv \sqrt{\vec{d} \cdot \vec{d}}$. But $\vec{d} \cdot \vec{\sigma}$ and $\vec{n} \cdot \vec{\sigma}$ also have the same eigenvectors, we are just multiplying the matrix by a number $|d|$. Combining

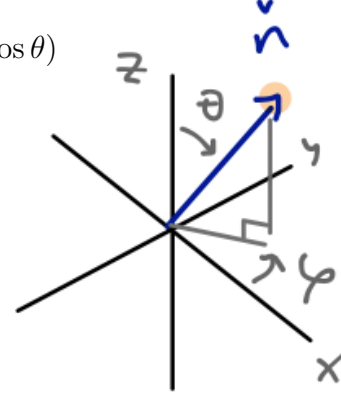
these facts, the eigenvectors of *any* 2-by-2 Hermitian matrix $\mathbf{h}$ are of the form (2.9) with eigenvalues

$$\epsilon_{\pm} = d_0 \pm \sqrt{|d|}.$$

If we choose the angles to be $\tilde{n} = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)$ (as in the figure), then the eigenstates are

$$\begin{aligned} |+\tilde{n}\rangle &= e^{-i\varphi/2} \cos \frac{\theta}{2} |\uparrow\rangle + e^{+i\varphi/2} \sin \frac{\theta}{2} |\downarrow\rangle, \\ |-\tilde{n}\rangle &= -e^{-i\varphi/2} \sin \frac{\theta}{2} |\uparrow\rangle + e^{+i\varphi/2} \cos \frac{\theta}{2} |\downarrow\rangle. \end{aligned} \quad (2.9)$$

A straightforward way to get (2.9) is by brute force. A good way to derive this result is to notice that $\tilde{n} \cdot \vec{\sigma}$ is related by a rotation to σ^z ; knowing how the rotation acts on the two-state system, we can just act on $|\uparrow\rangle$ and $|\downarrow\rangle$ to get (2.9).



2.1.1 Where to get your qbits.

Where can we find an isolated qbit? The demand that it is isolated is tricky and a subject of quite a bit of current research. If we are not too picky about this point, we can find lots of examples.

Photon polarization. [Baym, chapter 1] Begin with a coherent monochromatic laser beam, propagating in the z direction. The electric field looks something like (the real part of)

$$\vec{\mathcal{E}} = \vec{\mathcal{E}}_0 e^{i(kz - \omega t)} \quad (2.10)$$

with, in vacuum, $\omega = ck$. The magnetic field is $\tilde{z} \times \vec{\mathcal{E}}$. The polarization vector $\vec{\mathcal{E}}_0 = \mathcal{E}_x \tilde{x} + \mathcal{E}_y \tilde{y}$ is transverse to the direction of propagation $\tilde{k} = \tilde{z}$. It is an element of a complex 2d vector space over $\mathbb{C}$. Why complex? Taking the real part of $\vec{\mathcal{E}}$, we see that the *relative* phase $\arg \mathcal{E}_y / \mathcal{E}_x$ encodes a lag between the crests of the two polarizations. For example, a (right) circularly polarized wave has $\mathcal{E}_y = i\mathcal{E}_x = e^{i\pi/2} \mathcal{E}_x$, that is:

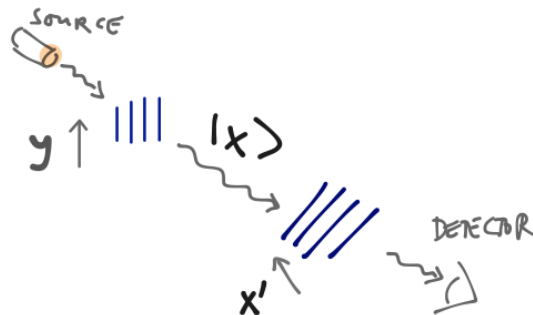
$$\text{Re} \left(\mathcal{E}_0 (\tilde{x} + i\tilde{y}) e^{i(kz - \omega t)} \right) = \tilde{x} \cos(kz - \omega t) - \tilde{y} \sin(kz - \omega t)$$

– the two components are out of phase by a quarter-period. On the other hand, the *overall* phase can be removed by restarting your watch. This should sound familiar.

The energy density in the wave is proportional to $|\vec{\mathcal{E}}|^2$. This moves at the speed of light and therefore the energy flux through a target normal to $\tilde{z}$ is also proportional to the norm of the polarization vector. This means that if we put a detector in front of the beam, we find a count rate proportional to the intensity of the beam: $I_1 \propto |\mathcal{E}_0|^2$.

I'm going to remind you about some basic optics phenomena. Then I'll tell you that basically the same statements hold if we attenuate the beam so much that only one photon at a time is passing through the apparatus. (But what's a photon? That question has a good answer, but for now I'm telling you an experimental fact: if you attenuate the beam enough, it becomes *particulate*, in the sense that the pattern on your detector is a discrete bunch of spots. Quantum mechanics predicts the probability distribution governing that pattern of spots.) For now, we're just going to be interested in the polarization state of the photons (the $\tilde{x}$ in (2.10)).

Recall that a *polarizer* along $\tilde{n}$ absorbs light with linear polarization along the direction transverse to $\tilde{n}$ (and the beam direction). If we send the beam through a $\tilde{x}$ -polarizer and then a $\tilde{y}$ -polarizer, nothing gets through. A polarizer is a projector in the sense that a second polarizer in the same direction does nothing to the intensity. If the second $\tilde{x}'$ -polarizer makes an angle θ with $\tilde{x}$, the count rate will be $I_\theta \propto |\mathcal{E}_0|^2 \cos^2 \theta$.



The polarization state is an observable. One the two state system with $\mathcal{H} = \text{span}\{|x\rangle, |y\rangle\}$, $\mathbf{P}_x \equiv |x\rangle\langle x|$ is a hermitian operator, and indeed a projector. The probability that a photon in a state $|\psi\rangle \in \mathcal{H}$ gets through an $\tilde{x}$ -polarizer (and is observed in a detector just behind it) is

$$P_x = \langle \psi | \mathbf{P}_x | \psi \rangle = |\langle x | \psi \rangle|^2.$$

Notice that this example does a lot to explain the (nonlinear) normalization step of the measurement axiom: *if* a photon gets through the $\tilde{x}$ -polarizer, it for sure has the right polarization. After this point, we are computing probabilities *conditioned* on the photon getting through the polarizer. The probabilities for the possible outcomes remaining need to sum up to one.

Why probabilities? You might accept based on previous experience with QM that the total energy in a wave of frequency ω has to be an integer multiple of $\hbar\omega$. The intensity therefore is of the form $N\hbar\omega$ for some integer N , called the *number of photons*. When a polarizer absorbs part of the beam, the transmitted beam a fraction of the energy, say half, also an integer multiple of $\hbar\omega$: only that many photons get through. Two questions for you: (1) *Which half??* This is a deep question. The photons are all the same and effectively do not interact with each other. How can they decide? The answer is that *each* photon has a probability (here, half) to get through and there is no way to know whether any given photon will get through.

(2) *Exactly half??* Certainly not. The number of photons in the initial beam certainly need not be an even number. Rather, half is the average number, the expected number, the fraction we'll get if we do the experiment over and over with lots of photons. Doing the experiment over and over is what happens in the classical limit. (This is called the *correspondence principle*.)

Another classical optics thing we can do is put a quarter-wave plate after the first polarizer; it will produce circularly polarized light. The plate is an optical element (of a certain thickness) through which right-handed and left-handed polarizations propagate with different speeds, c_L, c_R . A right(left)-handed polarized wave looks like:

$$\vec{\mathcal{E}}_{R/L} = \mathcal{E}_0 \left(\frac{\tilde{x} \pm \mathbf{i}\tilde{y}}{\sqrt{2}} \right) e^{\mathbf{i}(k_{R/L}z - \omega t)} .$$

(Notice I've allowed the wavenumber to depend on the polarization.) These other polarization states are linear combinations of the ones we've already discussed:

$$|R\rangle = \frac{1}{\sqrt{2}} (|x\rangle + \mathbf{i}|y\rangle), \quad |L\rangle = \frac{1}{\sqrt{2}} (|x\rangle - \mathbf{i}|y\rangle) .$$

Indeed any polarization state of a photon (with given $\vec{k} = \hat{z}k$) can be made as a superposition of either set of basis states:

$$|\psi\rangle = |R\rangle \langle R|\psi\rangle + |L\rangle \langle L|\psi\rangle = |x\rangle \langle x|\psi\rangle + |y\rangle \langle y|\psi\rangle .$$

Propagating at different speeds means their phases accrue at different rates: traversing a medium of thickness a , an input beam $|\psi(0)\rangle = c_R |R\rangle + c_L |L\rangle$ (for example, $|x\rangle = \frac{1}{\sqrt{2}} (|R\rangle + |L\rangle)$) will evolve into

$$|\psi(a)\rangle = e^{\mathbf{i}k_R a} c_R |R\rangle + e^{\mathbf{i}k_L a} c_L |L\rangle$$

Compare the time evolution of a superposition of energy eigenstates. (To make a quarter-wave plate, we can choose $e^{\mathbf{i}k_L a} c_L$ to be pure imaginary.)

Particle spin. A spin- $\frac{1}{2}$ particle has 'an inherent two-valuedness,' two orthogonal states for each position state. These two states can be split in energy by an external magnetic field $\vec{B}$:

$$\Delta H = -\mu \vec{S} \cdot \vec{B} .$$

Here $\vec{S} = \frac{\hbar}{2} \vec{\sigma}$ is the spin and μ is the magnetic moment of the particle (it depends on the particle). Sometimes it is convenient to define the magnetic moment as an operator: $\vec{\mu} \equiv \mu \vec{S}$. (Notice that we treating our moments as *neutral* particles, in the sense that we are ignoring any *orbital* coupling to the magnetic field; for charged particles like electrons this matters a lot.)

First, we can illustrate the time-evolution axiom. For an arbitrary state,

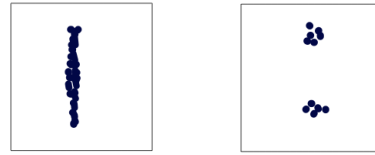
$$\frac{d}{dt} \langle \vec{S} \rangle = \langle \vec{\mu} \rangle \times \vec{B}.$$

The moment precesses about the field, just like in classical physics.

Second, we can give another illustration of the measurement axiom. How do we measure the spin? Classically, if we put a magnetic moment in an inhomogeneous magnetic field, $\vec{B} = \hat{z}B_z(z)$, it will exert a *force* on the moment

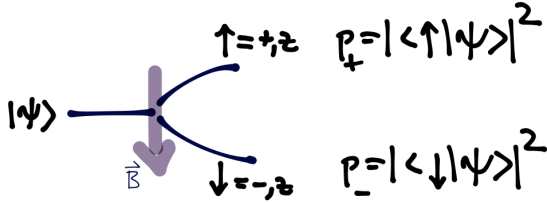
$$F_z = -\partial_z (\mu_z B_z(z)).$$

The same is true quantum mechanically, except that now $\mu_z \rightarrow \hat{\mu}_z$ is an operator, and it only takes values $\pm \frac{\hbar}{2}\mu$, where the $\pm$ is the eigenvalue of σ^z . So it makes two discrete spots on the wall behind the (Stern-Gerlach) apparatus, as in the right figure. Classically,

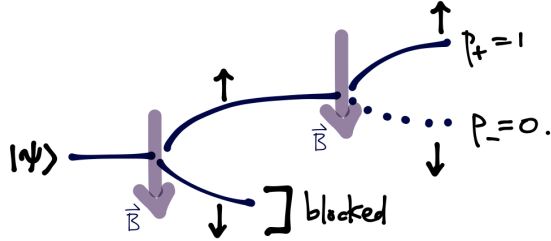


the z-component of the moment can vary continuously and the pattern is the left one. As with the photons and the polarizers, we can't predict what any one particle will do. Rather, we can compute probabilities, which tell us the relative intensities of the spots. (Here, unlike the polarizers, there is a further step to the classical limit which is to make the spin ($\propto \hbar$) small, so that the force difference between the two spin states goes away and the spots coalesce.) [\[End of Lecture 4\]](#)

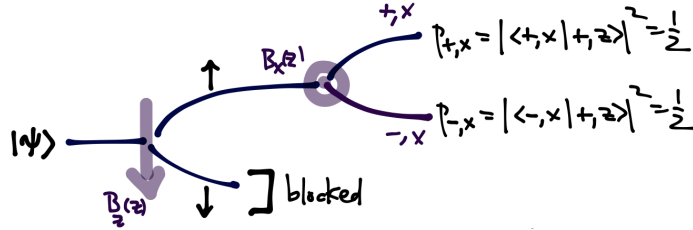
On the next page are some visualizations of Stern-Gerlach apparatuses. The thick dark blue lines represent possible trajectories of the spins, which are moving to the right. The big purple arrows represent the magnetic fields.



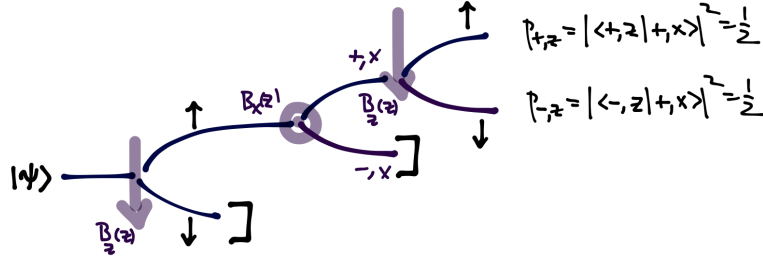
The Stern-Gerlach apparatus with one path blocked is a projector:



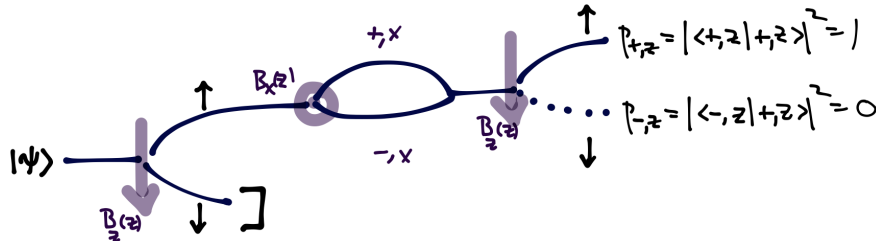
If we want to measure instead σ^x , we can rotate the magnetic field, so that it points in the x -direction and depends on x .⁸



What happens if we do both? If we measure σ^z and then σ^x and then σ^z , we can get different answers for the two σ^z measurements:



On the other hand, if we don't block either of the beams of the x -oriented S-G apparatus,



so that all particles enter the final z -oriented one, we reassemble the $|+_z\rangle = |\uparrow\rangle$ state. The 'loop' in the picture is an insertion of the identity operator in the form

$$\mathbf{P}(+_x) + \mathbf{P}(-_x) = |+_x\rangle \langle+_x| + |-_x\rangle \langle-_x| = \mathbb{1}.$$

⁸In the figures, I am imagining we can change the direction of the field (so that it is sensitive to a different component of the spin) while keeping its variation in the z direction (so that the force is still in the z direction).

Atomic levels. Another source of qbits is degenerate states of electrons in atoms. Such degeneracy requires some symmetry reason (like parity or spin or time-reversal) to be stable.

Particle flavor. Oscillations between neutrinos of various ‘flavor’ provide a vivid realization of a few-state system in high energy particle physics.

Particle position (interferometers). [Schumacher §2.1] Here is another possible realization of simple (few-state) quantum systems using photons. In free space, light can go all over the place; it’s easier to think about an *interferometer*, an apparatus wherein light is restricted to discrete *beams*, which can be guided or split apart and recombined. A beam is by definition a possible path for a photon. (It is useful to imagine a fiber-optics cable.)

Now: suppose we input a beam into an interferometer with two paths, which we’ll call upper and lower, and a detector at the end. Then each path will have an associated probability amplitude α , β and the probability that we would find the photon on the top path (if we were to put a detector there) is $P_{\text{upper}} = |\alpha|^2$. By superposition, can represent the intermediate state of the photon by

$$|\psi\rangle = \alpha |\text{upper}\rangle + \beta |\text{lower}\rangle .$$

If we know that we sent in a photon, then we must have

$$1 = P_{\text{upper}} + P_{\text{lower}} = |\alpha|^2 + |\beta|^2.$$

Notice that we are again encoding a qbit in a photon, but not in its polarization, but rather in its choice of path.

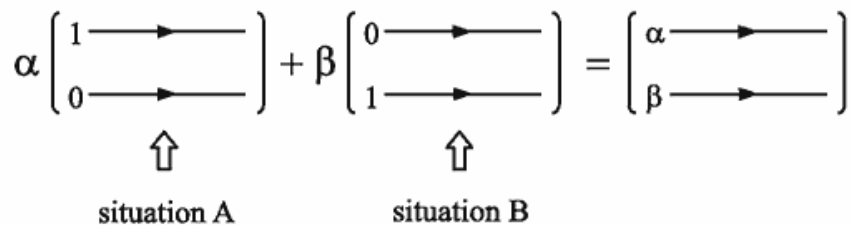


Figure 1: [From Schumacher] Here is an illustration of superposition.

There are a few things we can do to our photons while we’re sending them through our maze. One device we can send them through is a **phase shifter**. This could be a glass plate through which the light goes. In doing so, it gets slowed down, so its state

acquires a larger phase $\sim i\omega t$. We know the photon went in and will come out, so the probability $|\alpha|^2$ is preserved, but its phase changes:

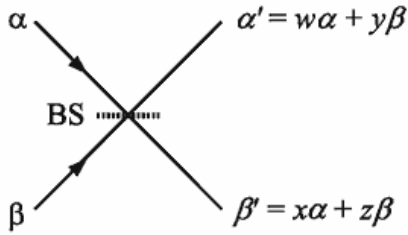
$$\alpha \mapsto e^{i\delta} \alpha$$

where δ is a property of the device. For example, it could be $\delta = \pi$, in which case this is $\alpha \mapsto -\alpha$. This phase can be important; for example, it can turn constructive interference into destructive interference. A phase shifter which acts only on the upper leg of the interferometer acts by the matrix (in the $\{|u\rangle, |d\rangle\}$ basis)

$$\mathbf{S}_{\text{upper}}(\delta) = \begin{pmatrix} e^{i\delta} & 0 \\ 0 & 1 \end{pmatrix}.$$

By another device, we can let the beams cross, so that we exchange the upper and lower amplitudes; this is represented by

$$\mathbf{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \boldsymbol{\sigma}^x.$$



Next, we can send our beam into a **beam-splitter**; this takes the input beam and splits it into two beams of lower intensity. A partially-silvered mirror will accomplish this by partially reflecting and partially transmitting the wave. Such a mirror has two possible basis inputs (see

fig): light from above or below, which we can denote respectively by the vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. By the principle of superposition (the assumption that the device is linear), we must have that the action of the beamsplitter is

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \mapsto \begin{pmatrix} \alpha' \\ \beta' \end{pmatrix} = \mathbf{U} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} w & y \\ x & z \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

If we know that all of the photons end up in one or the other of the two output beams, then the matrix $\mathbf{U}$ must be unitary – it must conserve probability.

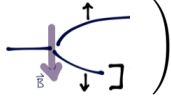
Photons can be annoying in that they can't stop moving. For our purposes of making a qbit we could also in principle use position degrees of freedom of a particle which just sits there in one of two wells, and we could imagine some similar construction of unitary gates ($\mathbf{S}(\delta)$, $\boldsymbol{\sigma}^x$, $\mathbf{U}$) acting on it as above.

An extreme case of an interferometer is if we let the particle be anywhere in the continuum, so its position Hilbert space is infinite dimensional. We'll treat this separately in §3.

2.1.2 Amplitudes add, not probabilities

[Baym] A few crucial points to notice from the discussion of measurements on a two-state system.

$$|\langle \uparrow | \psi \rangle|^2 = \mathbf{Prob}_\psi (\text{measure } \sigma^z \text{ get } +1)$$

$$|\langle \uparrow | \psi \rangle|^2 = \mathbf{Prob}_\psi \left(\text{particle passes through } \begin{array}{c} \uparrow \\ \downarrow \end{array} \right)$$


that is, this is the probability that the particle which is for sure in state $|\psi\rangle$ behaves as if it were in the state $|\uparrow\rangle$. Classically, this is something that does not happen: a particle with position and velocity (x, v) doesn't behave as if it is in the state (x', v') . (The exception is if these two states can't be resolved by our measuring devices. That kind of thing happens in QM, too, and we'll discuss it next, in §2.1.3. This is a different thing.)

This is made possible by superposition:

$$|\psi\rangle = |\uparrow\rangle \langle \uparrow | \psi \rangle + |\downarrow\rangle \langle \downarrow | \psi \rangle.$$

The probability of behaving like $\uparrow$ is only zero if $|\uparrow\rangle$ and $|\psi\rangle$ are orthogonal, *i.e.* if $0 = \langle \uparrow | \psi \rangle$. By the way, this overlap is called the *probability amplitude*, for a particle in state $|\psi\rangle$ to behave like $|\uparrow\rangle$.

Very important: why do I say this is a different thing from classical uncertainty? A participant in a Galilean dialogue (Simplicio) might ask: can't we just say that a particle is always in either the state $|\uparrow\rangle$ or the state $|\downarrow\rangle$ with some probability $p_{\uparrow/\downarrow} = |\langle \uparrow / \downarrow | \psi \rangle|^2$ and be done? No. Consider for example the states $|\pm, x\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle \pm |\downarrow\rangle)$. They both have $p_{\uparrow/\downarrow} = \frac{1}{2}$. But if $|\psi\rangle = |-, x\rangle$, the probability to measure σ^x and get $+$ is zero. If we used Simplicio's rules, we would say: if the particle is in state $|\uparrow\rangle$ (which happens half the time), we find $\sigma^x = +$ with probability $\frac{1}{2}$. if the particle is in state $|\downarrow\rangle$ (which happens half the time), we find $\sigma^x = -$ with probability $\frac{1}{2}$. There are two problems with this: these probabilities only add up to $\frac{1}{4} + \frac{1}{4} = \frac{1}{2}$, and also it's wrong. The thing that's missing is the *interference* terms in

$$\begin{aligned} \mathbf{Prob}_{|-x\rangle} (\sigma^x = +) &= |\langle -, x | +, x \rangle|^2 = |\langle +, x | (|\uparrow\rangle \langle \uparrow | -x\rangle + |\downarrow\rangle \langle \downarrow | -x\rangle)|^2 \\ &= |\langle + | \uparrow \rangle \langle \uparrow | - \rangle + \langle + | \downarrow \rangle \langle \downarrow | - \rangle|^2 \\ &= |\langle + | \uparrow \rangle|^2 |\langle \uparrow | - \rangle|^2 + |\langle + | \downarrow \rangle|^2 |\langle \downarrow | - \rangle|^2 \quad (\leftarrow \mathbf{Prob}_\uparrow(+) + \mathbf{Prob}_\downarrow(+)) \\ &\quad + \langle + | \uparrow \rangle \langle \uparrow | - \rangle (\langle + | \downarrow \rangle \langle \downarrow | - \rangle)^* + h.c. \quad (\leftarrow \text{interference!}) \end{aligned}$$

Of course this is obviously zero if we write it as

$$\mathbf{Prob}_{|-x\rangle} (\sigma^x = +) = |\langle -, x | +, x \rangle|^2 = 0.$$

The preceding combersome calculation was meant as an illustration of the weirdness of superposition to someone who has a preferred basis.

To emphasize further that it's the relative *phase* that matters, and not just the sign, consider *e.g.* the superposition $|+_y\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + \mathbf{i}|\downarrow\rangle)$ in which the probability to measure $\sigma^y = -$ is zero. (How do you measure σ^y with the S-G apparatus where the particle moves in the y direction? Just let $\vec{B} = \hat{y}B(z)$ point in the y direction and continue to depend on the z coordinate.)

2.1.3 Quantum and classical uncertainty

Unpolarized particles. So far, we've been speaking as if our qbits are in a definite, known quantum state, such as

$$|\psi\rangle = |+_x\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$$

(or as if our light was in a definite polarization state).

Consider instead a situation where our qbit dealer (*e.g.* our monochromatic laser) is composed of two sources $S_{1,2}$ which produce qbits in the states $|\psi_{1,2}\rangle$ respectively. A given qbit in our beam then is either from S_1 or S_2 , but we don't know which. Say the probability that it's from $S_{1,2}$ is $p_{1,2}$, so that $p_1 + p_2 = 1$.

This situation, where we aren't sure about what quantum state the system is in, is called a *mixed state*. In this mixed state, what is the expected value of the spin along z ,

$$\sigma^z = |\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|$$

(In the case where the 2-state system is photon polarization, the same discussion with different notation applies to the photon's angular momentum $\mathbf{L}_z \propto |R\rangle\langle R| - |L\rangle\langle L|$). With probability p_1 , the state is $|\psi_1\rangle$, in which case the expectation value is

$$\langle\sigma^z\rangle_1 \equiv \langle\psi_1|\sigma^z|\psi_1\rangle = |\langle\uparrow|\psi_1\rangle|^2 - |\langle\downarrow|\psi_1\rangle|^2. \quad (2.11)$$

With probability p_2 , the state is $|\psi_2\rangle$, in which case replace all the 1s with 2s in (2.11). Altogether then, we have

$$\langle\sigma^z\rangle_{\text{mixed}} = p_1 \langle\psi_1|\sigma^z|\psi_1\rangle + p_2 \langle\psi_2|\sigma^z|\psi_2\rangle. \quad (2.12)$$

That is, we just add the averages for each state, weighted by the probability for each state, simple. The tricky bit is the distinction with the case of superposition, that is if the qbit were in the definite ('pure') state

$$|\Phi\rangle \equiv a_1 |\psi_1\rangle + a_2 |\psi_2\rangle \quad (2.13)$$

with definite, known phases of a_1, a_2 . In that case, the probability of ‘behaving like it’s in state ψ_1 ’ is $|\langle \psi_1 | \Phi \rangle|^2 = |a_1|^2$ and we can make this look a bit like the mixed state above by setting $p_{1,2} = |a_{1,2}|^2$. In the mixed state, the particle is in the state $|\psi_1\rangle$ or the state $|\psi_2\rangle$ but we aren’t sure which. This means there is no possibility of interference between these outcomes.

To recap: in pure states, there is only quantum uncertainty. We are sure about the quantum state. In mixed states, there is both quantum uncertainty and also classical uncertainty about who is the quantum state. In particular, in the mixed state above there is complete uncertainty about the *phases* of $a_{1,2}$ in (2.13). If we compute the expectation of the z -spin in this state, we get

$$\begin{aligned} \langle \sigma^z \rangle_\Phi &\equiv \langle \Phi | \sigma^z | \Phi \rangle = (a_1^* \langle \psi_1 | + a_2^* \langle \psi_2 |) \sigma^z (a_1 |\psi_1\rangle + a_2 |\psi_2\rangle) \\ &= |a_1|^2 \langle \sigma^z \rangle_1 + |a_2|^2 \langle \sigma^z \rangle_2 \\ &\quad + (a_1^* a_2 \langle \psi_1 | \sigma^z | \psi_2 \rangle + c.c.). \end{aligned} \tag{2.14}$$

If we don’t know the phases $\phi_{1,2}$ in $a_{1,2} = |a_{1,2}| e^{i\phi_{1,2}}$ we should average the answer over them. But the interference terms (last line in (2.14)) are proportional to $a_1^* a_2 = |a_1 a_2| e^{i(\phi_2 - \phi_1)}$ which averages to zero.

2.1.4 Density operator

The projector onto a state vector $|\psi\rangle$ is

$$\mathbf{P}_\psi = |\psi\rangle \langle \psi| \quad .$$

If our system is in the state $|\psi\rangle$, we can compute the expectation value of any operator on $\mathcal{H}$ by

$$\langle \mathbf{A} \rangle_\psi = \langle \psi | \mathbf{A} | \psi \rangle = \text{tr}(\mathbf{A} \mathbf{P}_\psi) .$$

It is useful to generalize our machinery so we can use the rightmost expression here to define expectation values (and hence probabilities for any outcome) for mixed states. The role of the projector is played by the *state operator* or *density operator* or *density matrix* ρ :

$$\langle \mathbf{A} \rangle_\rho = \text{tr}(\mathbf{A} \rho) . \tag{2.15}$$

For a pure state, $\rho = \mathbf{P}_\psi$. For a mixed state where the probability to be in state $|\psi_i\rangle$ is p_i , then

$$\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i| ,$$

and (2.15) reproduces (2.12) above.

Here is an immediate virtue of this description of a quantum state: Notice that while the ray description of a state $|\psi\rangle$ involves a meaningless phase, its associated density matrix

$$\rho_\psi = |\psi\rangle\langle\psi|$$

only involves physically meaningful quantities.

What are the properties of the density operator?

In general the density operator

- 1: has unit trace (because $|\psi\rangle$ is a normalized state, and it had better because its trace is the sum of all the probabilities), $1 = \langle\mathbb{I}\rangle = \text{tr}(\rho_A)$
- 2: is self-adjoint, $\rho_A = \rho_A^\dagger$ (its eigenvalues are probabilities) and, moreover,
- 3: is positive (all of its eigenvalues are positive or zero (and in particular real), since they are probabilities). (In fact 3 implies 2.)

So: in general, the state of a subsystem is not a ray, it is a density operator. The density operator associated with a ray $|\psi\rangle$ (a pure state) is of the form $\rho_{\text{pure}} = |\psi\rangle\langle\psi|$, a projector of rank one, onto the space spanned by $|\psi\rangle$. $\rho_{\text{pure}}^2 = \rho_{\text{pure}}$. (Notice that although any projector satisfies this equation, it is only consistent with $\text{tr}\rho = 1$ if it is a projector of rank one.)

A more general (“mixed”) density matrix can be written (by spectral decomposition – it is positive and hence Hermitian) as

$$\rho_{\text{mixed}} = \sum_a p_a |\psi_a\rangle\langle\psi_a|$$

with $0 \leq p_a \leq 1$, $\sum_a p_a = 1$. You should think of p_a as the probability that the system is found in the state $|\psi_a\rangle$ (an eigenstate of the density operator).

Quantum vs Classical uncertainty. The density operator is useful for emphasizing the distinction between quantum superposition and a mere probabilistic distribution, even for a single qbit.

Returning to our example mixed state above, the associated density operator is

$$\rho_{\text{mixed}} = p_1 |\psi_1\rangle\langle\psi_1| + p_2 |\psi_2\rangle\langle\psi_2|.$$

In contrast, the density operator for the superposition $|\Phi\rangle = a_1 |\psi_1\rangle + a_2 |\psi_2\rangle$ with the same diagonal probabilities is

$$\rho_\Phi = |\Phi\rangle\langle\Phi| = (a_1 |\psi_1\rangle + a_2 |\psi_2\rangle)(a_1^* \langle\psi_1| + a_2^* \langle\psi_2|)$$

$$= |a_1|^2 |\psi_1\rangle \langle \psi_1| + |a_2|^2 |\psi_2\rangle \langle \psi_2| + \underbrace{(a_1 a_2^* |\psi_1\rangle \langle \psi_2| + h.c.)}_{\text{interference}} \quad (2.16)$$

(where *h.c.* stands for hermitian conjugate). By ‘diagonal probabilities’ I mean $p_{1,2} = |a_{1,2}|^2$. If we average the density matrix over the phases of $a_{1,2} = |a_{1,2}|e^{i\phi_{1,2}}$ the interference terms go away and we get the mixed density matrix.

An explicit example. In the state $(|\uparrow_z\rangle + |\downarrow_z\rangle)/\sqrt{2} = |\uparrow_x\rangle$, the value of σ^z is uncertain, but the measurement of σ^x gives +1 with probability one. The ensemble in which $|\uparrow_z\rangle$ and $|\downarrow_z\rangle$ occur with probability 1/2 is described by the density operator

$$\rho_{\text{mixed}} \equiv \frac{1}{2} |\uparrow_z\rangle \langle \uparrow_z| + \frac{1}{2} |\downarrow_z\rangle \langle \downarrow_z| = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2} \mathbb{1}.$$

This is a very different thing from the (*pure*) state $|\uparrow_x\rangle$, which has a density operator

$$\rho_{\text{pure}} = |\uparrow_x\rangle \langle \uparrow_x| = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Obviously these are different operators. To see how they give different predictions, consider the expectation value of the projection onto $|\uparrow_x\rangle$, $\mathbf{P}(\uparrow_x) = |\uparrow_x\rangle \langle \uparrow_x|$ in each of these two cases (this we interpret as the probability that we will get $\uparrow$ if we measure σ^x). In the state $|\uparrow_x\rangle$, this is

$$\langle \mathbf{P}(\uparrow_x) \rangle_{\text{pure}} = \text{tr} \rho_{\text{pure}} \mathbf{P}(\uparrow_x) = \langle \uparrow_x | \uparrow_x \rangle \langle \uparrow_x | \uparrow_x \rangle = 1,$$

that is, like I said, we are certain to get $\uparrow$ if we measure σ^x in the state $|\uparrow_x\rangle$. On the other hand, in the ensemble ρ_{mixed} , we get

$$\langle \mathbf{P}(\uparrow_x) \rangle_{\text{mixed}} = \text{tr} \rho_{\text{mixed}} \mathbf{P}(\uparrow_x) = \text{tr} \frac{1}{2} \mathbb{1} |\uparrow_x\rangle \langle \uparrow_x| = \frac{1}{2}.$$

In fact, as you can see from the fact that ρ_{mixed} is proportional to the identity on $\mathcal{H}_A$, if we measure the spin along *any* axis in this state, we get a completely random result.

So what I meant by an ‘unpolarized particle’ is really that we are ignorant of the polarization state, in which case the density matrix (for its polarization degree of freedom) is $\rho = \frac{1}{2} \mathbb{1}$. Notice that this is independent of which basis we use.

[End of Lecture 5]

2.2 Composite quantum systems

There was one more rule in my list above. Whether it needs to be added to the list of rules is arguable. Here is an argument against it. [Schumacher section 6.1, Le Bellac section 6.1]

Axiom 5: **Composition**

How do we combine quantum systems? Suppose we have two spin 1/2 particles. We will fix their locations, so that their position degree of freedom does not fluctuate, and so that they are distinguishable (quantum statistics of indistinguishable particles complicates the rule I am about to state). Each is described by a two-dimensional Hilbert space, a qbit:

$$\mathcal{H}_{a=1,2} = \text{span}\{|\uparrow\rangle_a, |\downarrow\rangle_a\} .$$

If the state of one spin imposes no hard constraint on the state of the other, we may specify their states independently. For example, if they are far apart and prepared completely independently, we would expect that the probabilities for outcomes of measurements on 1 and 2 separately should be uncorrelated: $P(1,2) = P(1)P(2)$. We can achieve this if the state of the combined system is of the form $|a\rangle_1 \otimes |b\rangle_2$ where the inner product behaves as

$$(\langle c|_1 \otimes \langle d|_2) (|a\rangle_1 \otimes |b\rangle_2) = \langle c|a\rangle_1 \langle d|b\rangle_2 .$$

But now Axiom 1 tells us that we must allow superpositions of vectors to also describe allowed states. The resulting combined vector space is the *tensor product* of the two Hilbert spaces:

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$$

a basis for which is

$$\{|\uparrow\rangle_1 \otimes |\uparrow\rangle_2, |\uparrow\rangle_1 \otimes |\downarrow\rangle_2, |\downarrow\rangle_1 \otimes |\uparrow\rangle_2, |\downarrow\rangle_1 \otimes |\downarrow\rangle_2\} \quad (2.17)$$

or more succinctly:

$$\mathcal{H} = \text{span}\{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\}.$$

You have already encountered the tensor product in describing the Hilbert space of a particle moving in more than one dimension: to indicate a position basis vector we must specify both its x -position and its y -position.

A few simple comments about tensor products.

- Scalars can be moved between the tensor factors – the state $(z|a\rangle) \otimes |b\rangle = |a\rangle \otimes (z|b\rangle)$.
- If $|a_1\rangle$ is orthogonal to $|a_2\rangle$ in $\mathcal{H}_1$ then $|a_1\rangle \otimes |b_1\rangle \perp |a_2\rangle \otimes |b_2\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$.

Note: this operation of taking the tensor product of vector spaces must be vigorously distinguished from the direct *sum*. A rough but useful way to remember the distinction is: To specify a state in the tensor product of $\mathcal{H}_a$ and $\mathcal{H}_b$ means we must specify the state of *both* a and b . To specify a state in the direct sum we must specify whether the vector is a state of a or b . So taking a direct sum of hilbert spaces is like making an addition to your house, but taking a direct product is like cloning yourself. More precisely, if $\mathcal{H}_a = \text{span}\{|i\rangle, i = 1..N\}$ and $\mathcal{H}_b = \text{span}\{|r\rangle, r = 1..M\}$ then

$$\mathcal{H}_a \otimes \mathcal{H}_b \equiv \text{span}\{|i\rangle \otimes |r\rangle, i = 1..N, r = 1..M\},$$

– the tensor product of an N -dimensional $\mathcal{H}$ and an M -dimensional one is NM -dimensional. In contrast, the direct sum

$$\mathcal{H}_a \oplus \mathcal{H}_b \equiv \text{span}\{|i\rangle, |r\rangle, i = 1..N, r = 1..M\}$$

is only $N + M$ dimensional.

Notice that generic vector in a tensor product $\mathcal{H}$ cannot be written as a product

$$|w\rangle = \sum_{i,m} w_{im} |i\rangle_1 \otimes |m\rangle_2 \neq |v^1\rangle_1 \otimes |v^2\rangle_2$$

for any $v^{1,2}$. This is only possible if the coefficient matrix factorizes as $w_{i,m} = v_i^1 v_m^2$. A matrix that can be written this way has rank 1 – only a one-dimensional eigenspace of nonzero eigenvalues. If $|w\rangle$ cannot be written this way, the two subsystems in the state $|w\rangle$ are said to be *entangled*. The rank of the matrix w is called the *Schmidt number* of the state $|w\rangle$; $|w\rangle$ is entangled if the Schmidt number is bigger than 1.

We may now allow the spins to interact with each other by introducing a Hamiltonian which acts on $\mathcal{H}$. So we have to think about operators acting on the tensor product. In general such an operator is a sum of operators of the form $\mathbf{A}_1 \otimes \mathbf{A}_2$ where $\mathbf{A}_{1,2}$ act on subsystems 1, 2.

A little more detail about matrix representation of tensor products

A few simple examples to illustrate the tensor product.

In the example with two qbits, the Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$ is four dimensional; a vector in it has four components. In the basis (2.17), we have the representation

$$|\uparrow\rangle_1 \otimes |\uparrow\rangle_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |\uparrow\rangle_1 \otimes |\downarrow\rangle_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad |\downarrow\rangle_1 \otimes |\uparrow\rangle_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad |\downarrow\rangle_1 \otimes |\downarrow\rangle_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Note that the order is a convention. In this basis, then, some famous operators look like (for example):

$$\mathbb{1}_1 \otimes \mathbb{1}_2 = \begin{pmatrix} \mathbb{1}_2 & 0 \\ 0 & \mathbb{1}_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbb{1}_1 \otimes \sigma_2^x = \begin{pmatrix} \sigma_2^x & 0 \\ 0 & \sigma_2^x \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

$$\sigma_1^x \otimes \mathbb{1}_2 = \begin{pmatrix} 0 & \mathbb{1}_2 \\ \mathbb{1}_2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \sigma_1^x \otimes \sigma_2^x = \begin{pmatrix} 0 & \sigma_2^x \\ \sigma_2^x & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Notice that if we re-ordered our basis as

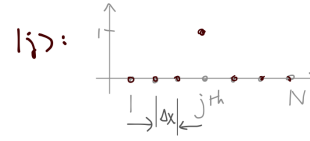
$$\{ |\uparrow\rangle_1 \otimes |\uparrow\rangle_2, |\downarrow\rangle_1 \otimes |\uparrow\rangle_2, |\uparrow\rangle_1 \otimes |\downarrow\rangle_2, |\downarrow\rangle_1 \otimes |\downarrow\rangle_2 \}$$

it would interchange matrix representations of the second and third examples above.

3 ∞ -dimensional Hilbert spaces, wave mechanics

Let's return to the discussion in §2.1.1 where we built our Hilbert space as

$$\mathcal{H} = \text{span}\{|j\rangle, j = 1..N\}$$



where $|j\rangle$ represents a particle sitting in a well (a little cup) at position $x = x_j$ on the line. Suppose the wells are evenly spaced: $x_j = j\Delta x$. These states are ON and form a complete basis:

$$\langle j|i\rangle = \delta_{ij}, \quad \mathbb{1} = \sum_j |j\rangle \langle j|. \quad (3.1)$$

We can think of the state $|j\rangle$ as in the figure at right.

A general state can be expanded (using the completeness relation in (3.1)) as

$$|\psi\rangle = \sum_j \psi(j) |j\rangle, \quad \psi(x_j) \equiv \psi_j = \langle j|\psi\rangle$$

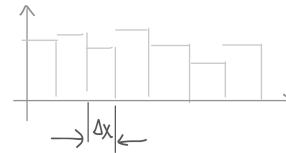
and the inner product between two states is

$$\langle \phi|\psi\rangle = \sum_j \phi^*(x_j)\psi(x_j). \quad (3.2)$$

So far there is nothing special coming from the notion of ‘locality’ introduced by the picture with the wells. Here it comes:

The notation (and the figures) encourage us to think of $\psi(x_j)$ as the values of a function $\psi(x)$ at discrete set of points.

We'd like to take a *continuum limit*, where $N \rightarrow \infty, \Delta x \rightarrow 0$, with $L = \Delta x N$ fixed. We'll have to replace the inner product in (3.2) with a (Riemann) integral:



$$\langle \phi|\psi\rangle \rightarrow \int dx \phi^*(x)\psi(x) = \lim_{N \rightarrow \infty, \Delta x \rightarrow 0} \Delta x \sum_{j=1}^N \phi^*(x_j)\psi(x_j).$$

To accomplish this we need a different inner product:

$$\langle \phi|\psi\rangle = \sum_j \phi_j^* \Delta x \psi_j = (\phi_1^*, \dots, \phi_N^*) \begin{pmatrix} \Delta x & 0 & \vdots \\ 0 & \Delta x & \vdots \\ \dots & \dots & \Delta x \end{pmatrix} \begin{pmatrix} \psi_1 \\ \dots \\ \psi_N \end{pmatrix}$$

This is pretty innocuous, in that it just multiplies everyone by a constant Δx , except that that constant is going to zero. A cost of this is that if we would like to keep

the relation $\phi(x) = \langle x|\phi\rangle$, then we need to fix the normalization of the basis states: $|j\rangle \rightarrow |x_j\rangle \equiv |j\rangle / \sqrt{\Delta x}$, so that

$$\langle x_i|x_j\rangle = \frac{1}{\Delta x}\delta_{ij}.$$

What becomes of the completeness relation?

Consider the continuum limit of

$$\psi(x_j) = \sum_j \Delta x \left(\frac{1}{\Delta x} \delta_{ij} \right) \psi(x_i) = \sum_j \Delta x \langle x_j|x_i\rangle \psi(x_i). \quad (3.3)$$

The first and last expressions become:

$$\psi(x') = \int dx \langle x'|x\rangle \psi(x).$$

That is, the inner product between x states is the Dirac delta function, defined by

$$\langle x'|x\rangle = \delta(x - x') \equiv \begin{cases} \infty, & x = x' \\ 0, & \text{else} \end{cases} \quad (3.4)$$

and

$$1 = \int_{\text{everyone}} \delta(x - x') dx. \quad (3.5)$$

Notice that the condition (3.5) means that the ∞ in (3.4) is a very specific one. There are many ways to define this object as a limit of finite things. One is the discrete way we've done it above, as a rectangle:

$$\delta_{\Delta x}(x) = \begin{cases} \frac{1}{\Delta x}, & x \in (0, \Delta x) \\ 0, & \text{else} \end{cases} \xrightarrow{\Delta x \rightarrow 0} \delta(x).$$

Alternatively, we can take a limit of smooth functions:

$$\delta_\sigma(x) \equiv \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \xrightarrow{\sigma \rightarrow 0} \delta(x).$$

Its integral from $-\infty$ up to x is

$$\theta(x) \equiv \int_{-\infty}^x dx' \delta(x') = \begin{cases} 1, & x > 0 \\ 0, & x < 0 \end{cases}$$

which by the fundamental theorem of calculus satisfies $\partial_x \theta(x) = \delta(x)$. We can also define its derivative $\delta'(x) \equiv \partial_x \delta(x)$ (and higher derivatives) by integration by parts:

$$\int_{-\epsilon}^{\epsilon} dx \delta'(x) f(x) = \delta(x) f(x) \Big|_{x=-\epsilon}^{x=\epsilon} - \int_{-\epsilon}^{\epsilon} dx \delta(x) \partial_x f(x) = -f'(0)$$

where in the last step we assumed that $f(x)$ was smooth enough that we could ignore the boundary term. This is a big hint about the significance of the continuum limit: we only allow functions where we can integrate by parts, which are the ones where you can draw a nice smooth picture (modulo the fact that they are complex functions), in the sense that the values at nearby points are nearby. This is a big reduction of the vector space as compared to the terrifying object $\text{span}\{|x\rangle\}$ about which we will not speak.

The same manipulation as in (3.3) shows

$$\langle x'|\psi\rangle = \int dx \langle x'|x\rangle \langle x|\psi\rangle$$

from which we infer the crucial completeness relation

$$\mathbb{1} = \int dx |x\rangle \langle x|.$$

(We will comment further below on the fact that this is a continuously infinite resolution of the identity.)

3.0.1 The position operator

The operator $\mathbf{x} = \int dx x |x\rangle \langle x|$ whose eigenstates are $|x\rangle$ is quite useful. However, it is not a *bounded* operator on the Hilbert space of a particle on the line. (So it is self-adjoint but not hermitian.) One way to see this is to notice that it has matrix elements which are not finite; another way to see this is to notice that if $\psi(x) = \langle x|\psi\rangle$ is a normalizable function, (so that we know $\int_{-\infty}^{\infty} dx |\psi(x)|^2 < \infty$) it is not necessarily the case that $x\psi(x) = \langle x|\mathbf{x}|\psi\rangle$ is as well (that is : $\int_{-\infty}^{\infty} dx x^2 |\psi(x)|^2 = ?$). Consider, for example, the case of $\psi(x) = \frac{1}{x}$.

This doesn't mean it isn't an observable. The position-space wavefunction is used to determine

$$dP_{\psi}(x) \equiv \mathbf{Prob}_{\psi} \left(\begin{array}{c} \text{position is measured and} \\ \text{found in the interval } (x, x + dx) \end{array} \right) = |\psi(x)|^2 dx.$$

You may have heard the statement “ $\delta(x)$ is not a function, it's a distribution.” This is a true statement, with two meanings: As a math statement, it means that it's well-defined under integrals against ‘nice’ functions: $f(x') = \int dx \delta(x - x') f(x)$. As a physics statement, its real meaning is that you'll never encounter a physical system which (*i.e.* whose hamiltonian) has $|x\rangle$ as a finite-energy state, *i.e.*, the state whose position-space wavefunction is

$$\psi_x(x') = \langle x'|x\rangle = \delta(x - x').$$

This is because of inertia, *i.e.* because of the kinetic terms in the Hamiltonian, the ones which depend on the momentum. So ‘nice’ means normalizable and finite energy, in particular, differentiable.

$|x\rangle$ is an eigenstate of $\mathbf{x}$, but is not a bounded operator on the actual Hilbert space, and $|x\rangle$ is not actually a possible physical state.

A good definition of the completeness property of a Hilbert space is that it is possible to resolve the identity by a completeness relation of the form

$$\mathbb{1} = \sum_n |n\rangle \langle n| \quad (3.6)$$

where, for the purposes of this statement, it actually matters for once that it is a sum and not an integral. The words for this are: $\mathcal{H}$ is *separable*. The fact that it is also possible to resolve the identity by a continuum is not a problem, and is very useful. We will see below whose eigenstates can give the discretuum in (3.6).

3.0.2 The derivative operator

Consider $\mathcal{H} = \text{span}\{|\psi\rangle, \psi(x) = \langle x|\psi\rangle \text{ differentiable, normalizable}\}$, with the inner product discussed above. A linear operator on this space is

$$\begin{aligned} \partial: \mathcal{H} &\rightarrow \mathcal{H} \\ |\psi\rangle &\mapsto |\partial\psi\rangle \quad \langle x|\partial\psi\rangle \equiv \partial_x \psi(x). \end{aligned}$$

[\[End of Lecture 6\]](#)

It is extremely useful to return to the lattice for a moment. Let’s make a (big!) column vector out of the components

$$\psi(x_j) = \langle x_j|\psi\rangle = \begin{pmatrix} \psi(x_1) \\ \vdots \\ \psi(x_N) \end{pmatrix}_j.$$

(In the last expression j is the index which runs up and down the column.) It is useful here to make a choice about boundary conditions; let’s take $\psi(x_0) = \psi(x_{N+1}) = 0$. In terms of these, the derivative is $\partial\psi(x_j) = \lim_{\Delta x \rightarrow 0} \frac{\psi(x_j + \Delta x) - \psi(x_j)}{\Delta x}$. The state resulting from the action of the derivative operator has components

$$\langle x_j|\partial|\psi\rangle = \begin{pmatrix} \psi'(x_1) \\ \vdots \\ \psi'(x_N) \end{pmatrix}_j = \frac{1}{\Delta x} \begin{pmatrix} \psi(x_2) - \psi(x_1) \\ \vdots \\ 0 - \psi(x_N) \end{pmatrix}_j = \frac{1}{\Delta x} \begin{pmatrix} -1 & 1 & 0 & \vdots \\ 0 & -1 & 1 & \vdots \\ 0 & \dots & 0 & -1 \end{pmatrix}_{jk} \begin{pmatrix} \psi(x_1) \\ \vdots \\ \psi(x_N) \end{pmatrix}_k.$$

This matrix expression makes it clear that it is a linear operator.

What are its matrix elements in the continuum?

$$\partial_x \psi(x) = \langle x | \boldsymbol{\partial} | \psi \rangle = \int_{-\infty}^{\infty} dx' \langle x | \boldsymbol{\partial} | x' \rangle \langle x' | \psi \rangle \stackrel{\text{IBP}}{=} \int dx' \delta(x - x') \partial_{x'} \psi(x').$$

In the third step, we used $\langle x | x' \rangle = \delta(x - x')$ and therefore

$$\langle x | \boldsymbol{\partial} | x' \rangle = \partial_x \delta(x - x') \stackrel{u \equiv x - x'}{=} \partial_u \delta(u) \stackrel{\text{chain rule}}{=} \underbrace{\frac{\partial x'}{\partial u}}_{=-1} \partial_{x'} \delta(u) = -\partial_{x'} \delta(x - x').$$

(Beware that sign from the chain rule.) Then we assumed we could ignore the boundary terms in the integration by parts (IBP) – this is an assumption about the properties of ψ which is required for $|\psi\rangle \in \mathcal{H}$.

More generally, given two states $|g\rangle, |f\rangle \in \mathcal{H}$,

$$\langle g | \boldsymbol{\partial} | f \rangle = \int dx g^*(x) \partial_x f(x)$$

whereas

$$\langle \partial g | f \rangle = \int dx (\partial_x g)^* f(x) = \int dx \partial_x g^*(x) f(x) \stackrel{\text{IBP}}{=} g^*(x) f(x) \Big| - \int dx g^*(x) \partial_x f(x) = -\langle g | \boldsymbol{\partial} | f \rangle.$$

Therefore

$$\boldsymbol{\partial}^\dagger = -\boldsymbol{\partial}$$

– $\boldsymbol{\partial}$ is antihermitian. Let

$$\mathbf{k} = -i\boldsymbol{\partial}.$$

$\mathbf{k}$ is hermitian.

Who are its eigenstates? If $\mathbf{k} |k\rangle = k |k\rangle$, then its position-basis representation is

$$\psi_k(x) \equiv \langle x | k \rangle$$

which satisfies

$$k \psi_k(x) = k \langle x | k \rangle = \langle x | \mathbf{k} | k \rangle = -i \langle x | \boldsymbol{\partial} | k \rangle = -i \partial_x \psi_k(x).$$

The equality of the first and last expressions give a 1st-order differential equation solved by

$$\psi_k(x) = A e^{ikx}. \quad (3.7)$$

We must discuss the normalization factor A . (Along the way, we will derive the Poisson summation formula and the basic theorem of Fourier analysis.)

First let's put the system in a periodic box, so we only admit functions which satisfy $\psi(x) = \psi(x + L)$. Normalizability then says $1 = \int_0^L dx |\psi(x)|^2$. For the states (3.7), this says $1 = |A|^2 L$. That is: $\psi_k(x) = \frac{1}{\sqrt{L}} e^{ikx}$. The periodicity condition gives

$$\psi(0) = \psi(L) \implies e^{ikL} = 1 \implies k = k_n = \frac{2\pi n}{L}, \quad n \in \mathbb{Z} = \{.. - 2, -1, 0, 1, 2, \dots\}.$$

These states form a complete basis⁹ :

$$\mathbb{1} = \sum_{n \in \mathbb{Z}} |k_n\rangle \langle k_n|$$

and we can expand any state in this basis

$$|f\rangle = \sum_n |k_n\rangle \underbrace{\langle k_n | f \rangle}_{\equiv f_n}. \quad (3.8)$$

This is the crucial step. Now

$$f_n = \langle k_n | f \rangle = \langle k_n | \underbrace{\mathbb{1}}_{=\int_0^L dx |x\rangle \langle x|} |f\rangle = \int_0^L dx \psi_{k_n}^*(x) f(x) = \int_0^L dx \frac{1}{\sqrt{L}} e^{-ik_n x} f(x). \quad (3.9)$$

So f_n is a Fourier coefficient of $f(x)$. Putting this back into (3.8), this gives the Poisson summation formula:

$$\begin{aligned} f(x) = \langle x | \mathbb{1} | f \rangle &= \langle x | \sum_n |k_n\rangle \langle k_n | f \rangle \\ &= \sum_n \psi_{k_n}(x) f_n \\ &= \sum_n \frac{1}{\sqrt{L}} e^{ik_n x} f_n \\ &\stackrel{(3.9)}{=} \sum_n \frac{1}{\sqrt{L}} e^{ik_n x} \int_0^L dy \frac{1}{\sqrt{L}} e^{-ik_n y} f(y) \\ &= \int_0^L dy \left(\frac{1}{L} \sum_n e^{ik_n(x-y)} \right) f(y). \end{aligned} \quad (3.10)$$

The object in parentheses is therefore

$$\frac{1}{L} \sum_{n \in \mathbb{Z}} e^{ik_n(x-y)} = \sum_{m \in \mathbb{Z}} \delta(x - y - mL)$$

a periodic delta function (sometimes called *Dirac's comb*).

⁹Although $\mathbf{k}$ itself is not a bounded operator on $\mathcal{H}$, e^{ik_a} is bounded and unitary and hence normal and shares its eigenvectors.

Now let's take the infinite-volume limit. Then k becomes continuous. The infinite-volume Fourier coefficients are normalized as

$$\tilde{f}(k_n) \equiv \int_0^L dy e^{-\mathbf{i}k_n y} f(y) = \sqrt{L} f_n \xrightarrow{\lim L \rightarrow \infty} \int_{-\infty}^{\infty} dy e^{-\mathbf{i}k y} f(y) \equiv \tilde{f}(k) \quad (3.11)$$

the Fourier transform of f . The expansion of f in fourier coefficients becomes

$$f(x) = \frac{1}{L} \sum_n e^{\mathbf{i}k_n x} \tilde{f}(k_n) = \frac{1}{2\pi} \underbrace{\frac{2\pi}{L}}_{=\Delta k} \sum_n e^{\mathbf{i}k_n x} \tilde{f}(k_n) \xrightarrow{\lim L \rightarrow \infty} \int_{-\infty}^{\infty} \mathrm{d}k e^{\mathbf{i}k x} \tilde{f}(k) = f(x). \quad (3.12)$$

This is the basic theorem of Fourier analysis. I defined

$$\mathrm{d}k \equiv \frac{dk}{2\pi}$$

(by analogy with $\hbar \equiv \frac{h}{2\pi}$).

Plugging (3.11) into (3.12) we learn that

$$f(x) = \int dy f(y) \underbrace{\int_{-\infty}^{\infty} \mathrm{d}k e^{\mathbf{i}k(x-y)}}_{=\delta(x-y)}.$$

We have derived the Fourier representation of the delta function:

$$2\pi\delta(k) = \int dx e^{\mathbf{i}kx}.$$

You may worry about the well-definedness of this kind of integral. Inserting a ‘convergence factor’ in the Fourier representation of the delta function, we reproduce the gaussian regulator:

$$\int \mathrm{d}k e^{\mathbf{i}kx} \sqrt{2\pi}\sigma e^{-\sigma^2 k^2/2} \propto \delta_\sigma(x).$$

To summarize: Position basis states satisfy

$$\mathbf{x} |x\rangle = x |x\rangle, \quad \langle x|x'\rangle = \delta(x-x'), \quad \mathbb{1} = \int dx |x\rangle \langle x|.$$

Momentum basis states satisfy $\mathbf{k} |k\rangle = k |k\rangle$ and they are normalized by

$$\langle k|k'\rangle = \int dx \langle k|x\rangle \langle x|k'\rangle = \int dx e^{\mathbf{i}x(k-k')} = 2\pi\delta(k-k').$$

We can similarly check the normalization of the momentum-space resolution of the identity

$$\mathbb{1} = \int \mathrm{d}k |k\rangle \langle k|$$

by taking matrix elements between position eigenstates:

$$\delta(x - x') = \langle x|x' \rangle = \int \mathrm{d}k \langle x|k \rangle \langle k|x' \rangle = \int \mathrm{d}k e^{\mathrm{i}k(x-x')} \quad \checkmark.$$

It is worth pausing a moment to think about why $2\pi\delta(k) = \int_{-\infty}^{\infty} dx e^{\mathrm{i}kx}$. For $k = 0$, the integrand is 1 for every x and so clearly we get infinity. For any other k , the integrand oscillates wildly.

One more theorem of Fourier analysis: The norm of a state is

$$\langle f|f \rangle = \int_{-\infty}^{\infty} |f(x)|^2 = \int_{-\infty}^{\infty} \mathrm{d}k |\tilde{f}(k)|^2.$$

(I think this is called Parseval's theorem.)

It shows us that $\mathbf{k}$ is not a bounded operator either, since the momentum-space representation of $\mathbf{k}|f\rangle$ is $\langle k|\mathbf{k}|f\rangle = k\tilde{f}(k)$ whose norm $\langle \mathbf{k}f|\mathbf{k}f \rangle = \int \mathrm{d}k k^2 |\tilde{f}(k)|^2$ may not be finite even if that of $|f\rangle$ is.

Finally, what are the matrix elements of $\mathbf{x}$ in the k basis:

$$\begin{aligned} \langle x|\mathbf{x}|f \rangle &= xf(x) = \int_{-\infty}^{\infty} \mathrm{d}k \underbrace{e^{\mathrm{i}kx} x}_{= -\mathrm{i}\partial_k e^{\mathrm{i}kx}} f(x) \\ &\stackrel{\text{IBP}}{=} \int \mathrm{d}k e^{\mathrm{i}kx} \left(+\mathrm{i}\partial_k \tilde{f}(k) \right). \end{aligned} \quad (3.13)$$

Therefore, in the position basis

$$\mathbf{x}f(x) = xf(x), \quad \mathbf{k}f(x) = -\mathrm{i}\partial_x f(x)$$

while in the momentum basis

$$\mathbf{x}\tilde{f}(k) = +\mathrm{i}\partial_k \tilde{f}(k), \quad \mathbf{k}\tilde{f}(k) = k\tilde{f}(k).$$

Let's consider the commutator of these operators:

$$[\mathbf{x}, \mathbf{k}]f(x) = \mathbf{x}(-\mathrm{i}\partial_x f(x)) - \mathbf{k}(xf(x)) = -\mathrm{i}x\partial_x f(x) + \mathrm{i}\partial_x(xf(x)) = \mathrm{i}f(x)$$

since this is true for all f ,

$$[\mathbf{x}, \mathbf{k}] = \mathrm{i}\mathbb{1}.$$

As you saw on the homework, this is not a possible commutation relation for operators on a finite-dimensional Hilbert space, or for bounded operators on any Hilbert space.

3.1 Uncertainty principle

(This subsection probably belongs in the mathematical equipment chapter, since it is also important for finite-dimensional Hilbert spaces, but it seems to have escaped.) On a general Hilbert space, consider an observable $\mathbf{A}$ which has no explicit dependence on time. Define the variance of $\mathbf{A}$ in a state ψ to be

$$\Delta A \equiv \sqrt{\left\langle \left(\mathbf{A} - \langle \mathbf{A} \rangle_\psi \right)^2 \right\rangle_\psi}$$

as usual, with $\langle \mathbf{A} \rangle_\psi \equiv \langle \psi | \mathbf{A} | \psi \rangle$. It will be useful here to speak about the operator $\delta \mathbf{A} \equiv \mathbf{A} - \langle \mathbf{A} \rangle_\psi \mathbb{1} = (\delta \mathbf{A})^\dagger$; notice that this is a weird object, since it depends explicitly on a state ψ .

We're going to quantify the statement that when two operators fail to commute, they can't be simultaneously measured. More specifically, we'll show

$$\Delta A \Delta B \geq \frac{1}{2} | \langle [\mathbf{A}, \mathbf{B}] \rangle |. \quad (3.14)$$

To prove (3.14), consider the states

$$|\alpha\rangle \equiv \delta \mathbf{A} |\psi\rangle, \quad |\beta\rangle \equiv \delta \mathbf{B} |\psi\rangle.$$

Apply the Cauchy-Schwarz inequality to their inner product:

$$\begin{aligned} |\langle \alpha | \beta \rangle|^2 &\leq \underbrace{\langle \alpha | \alpha \rangle}_{=\langle \psi | (\delta \mathbf{A})^\dagger \delta \mathbf{A} | \psi \rangle = \langle \delta \mathbf{A}^2 \rangle_\psi = \Delta A^2} \cdot \underbrace{\langle \beta | \beta \rangle}_{=\Delta B^2} \\ &= \Delta A^2 \Delta B^2 \end{aligned}$$

Now consider the LHS $\langle \alpha | \beta \rangle = \langle \delta \mathbf{A} \delta \mathbf{B} \rangle_\psi$.

$$\begin{aligned} \delta \mathbf{A} \delta \mathbf{B} &= \frac{1}{2} (\delta \mathbf{A} \delta \mathbf{B} + \delta \mathbf{B} \delta \mathbf{A}) + \frac{1}{2} (\delta \mathbf{A} \delta \mathbf{B} - \delta \mathbf{B} \delta \mathbf{A}) \\ &= \frac{1}{2} (\{ \delta \mathbf{A}, \delta \mathbf{B} \} + [\delta \mathbf{A}, \delta \mathbf{B}]) \\ &= \frac{1}{2} \left(\underbrace{\{ \delta \mathbf{A}, \delta \mathbf{B} \}}_{\text{hermitian}} + \underbrace{[\delta \mathbf{A}, \delta \mathbf{B}]}_{\text{anti-hermitian}} \right). \end{aligned} \quad (3.15)$$

In the last step, we used the fact that $\mathbb{1}$ commutes with everyone. In any state, the expectation value of a hermitian operator is real, and that of an antihermitian operator is pure imaginary. Therefore:

$$\Delta A^2 \Delta B^2 \geq | \langle \alpha | \beta \rangle |^2 = \frac{1}{4} |a + \mathbf{i}b|^2 \geq \frac{1}{4} |b|^2$$

with real a, b . Taking the square root gives (3.14).

[End of Lecture 7]

Here are two examples of its application: For a particle on a line,

$$[\mathbf{x}, \mathbf{p}] = i\hbar \implies \Delta x \Delta p \geq \frac{1}{2} |\langle i\hbar \rangle| = \frac{\hbar}{2}.$$

For a spin- $\frac{1}{2}$, the Pauli matrices (proportional to the components of the spin and hence the magnetic moment of the particle) satisfy

$$[\sigma^x, \sigma^y] = \frac{1}{2} i\sigma^z \implies \Delta\sigma^x \Delta\sigma^y \geq \frac{1}{2} \left| \left\langle \frac{i}{2} \sigma^z \right\rangle \right| = \frac{1}{4} |\langle \sigma^z \rangle|.$$

Notice that this implies that in an eigenstate of σ^x (hence $\Delta\sigma^x = 0$), we must have $\langle \sigma^z \rangle = 0$ (note that $\Delta\sigma^y$ is finite). A similar statement holds for the other permutations of the indices x, y, z .

I will pass on a useful complaint from Schumacher's book (page 96): the theorem (3.14) does not say that we are *uncertain* about the value of $\mathbf{A}$ or $\mathbf{B}$ in a state with nonzero $\langle [\mathbf{A}, \mathbf{B}] \rangle$. Rather, it says that both $\mathbf{A}$ and $\mathbf{B}$ *can not have definite values* in such a state. It's not that there's information about which we're ignorant – the information does not exist. Schumacher suggests the better name 'indeterminacy relation', but it's too late for that. This distinction may seem semantic but it is the essential point behind the violation by QM of Bell's inequalities and similar tricks of QM, to be discussed further in §5.

3.2 Wave mechanics

I need to comment on the identification between $\mathbf{p}$ and the momentum. For every observable, we are promised a hermitian operator. For a featureless particle on a line, we know we can measure the position and the momentum. What else could the momentum be? (Actually, by this argument it could be any (hermitian) operator like $\sum_{nm} a_{nm} \mathbf{x}^n \mathbf{k}^m$.) To make a closer identification, we can appeal to the satisfying connection with the canonical formalism for classical mechanics, where the Poisson bracket between canonically conjugate variables pairs them in a similar manner to the commutator:

$$\{x, p\}_{PB} = 1 \rightsquigarrow [\mathbf{x}, \mathbf{p}] = i\hbar.$$

This procedure of sticking an $i\hbar$ on the Poisson bracket, represented by a squiggly arrow above, is sometimes called 'quantization'. As useful as it is (*e.g.* it can be used to start from Maxwell's equations and make a theory of photons), this terminology is a

historical artifact of the fact that we learned about classical mechanics first. Classical mechanics is (just) a limit of quantum mechanics, not the other way around.

Sometimes a particle can move in more than one dimension, say d of them. In that case we need d $\mathbf{x}$ variables, $\mathbf{x}^j, j = 1..d$ and d $\mathbf{p}$ variables, and these have interesting commutators in pairs:

$$[\mathbf{x}^i, \mathbf{p}^j] = i\hbar\delta^{ij},$$

and $\mathbf{x}$ s and $\mathbf{p}$ s commute amongst themselves. The resulting Hilbert space is the tensor product of d copies of the particle in 1d¹⁰. Basically the only difference is we have to put some arrows on the position and momentum operators, which I will often omit.

Now we must choose a Hamiltonian for our particle. The time-evolution of a non-relativistic particle of mass m in an external potential $V(\mathbf{x}, t)$ is governed by the hamiltonian:

$$\mathbf{H} = \frac{\mathbf{p}^2}{2m} + V(\mathbf{x}, t).$$

(In $d > 1$, I mean $\mathbf{p}^2 \equiv \vec{\mathbf{p}} \cdot \vec{\mathbf{p}} \equiv \vec{\mathbf{p}}^2$.)

This is a special case of the Hamiltonian for a particle in an electromagnetic field:

$$\mathbf{H} = \frac{1}{2m} \left(\vec{\mathbf{p}} - \frac{q}{c} \vec{A}(\mathbf{x}, t) \right)^2 + q\Phi(\mathbf{x}, t).$$

Here I succumb to putting the arrows on the momentum operator since the vector structure matters for the vector potential $\vec{A}$. q is the electric charge of the particle. Let's retreat to the case with $A = 0$ for now.

If V has no explicit time dependence, the problem is completely solved by finding the energy eigenstates

$$\mathbf{H} |E\rangle = E |E\rangle.$$

In the position basis, $\langle x | \mathbf{H} | E \rangle = E \langle x | E \rangle$, using $\mathbf{p}\psi(x) = -i\hbar\partial_x\psi(x)$, this is the usual form of the Schrödinger equation

$$-\frac{\hbar^2}{2m}\partial_x^2\psi(x) + V(x)\psi(x) = E\psi(x).$$

¹⁰To see this more explicitly: It's clear that we can resolve the identity by *e.g.* simultaneous eigenstates of $\mathbf{x}^j$: $\mathbb{1} = \int d^d x |\vec{x}\rangle \langle \vec{x}|$. The norm is also the product norm we defined on the tensor product:

$$\langle f | g \rangle = \int d^d x f^*(\vec{x}) g(\vec{x}).$$

3.2.1 Probability current

I said earlier that for a region R

$$P(R) = \int_R dx |\psi(x)|^2 = \mathbf{Prob}_\psi (\text{particle found in region } R \text{ when measuring } \mathbf{x}).$$

That is, $\rho(x) \equiv |\psi(x)|^2$ is the *probability density*. Probability is a conserved quantity: it can't go away, it can only move around. This deep fact is built into our description, since from the general Schrodinger equation $-\mathbf{i}\hbar\partial_t\psi(x) = \mathbf{H}\psi(x)$,

$$\begin{aligned} \partial_t \rho(x) &= (\partial_t \psi^*(x)) \psi(x) + \psi^*(x) \partial_t \psi(x) \\ &= \left(\frac{\mathbf{i}}{\hbar} \mathbf{H} \psi \right)^* \psi + \psi^* \left(\frac{\mathbf{i}}{\hbar} \mathbf{H} \psi \right) \\ &= -\frac{\mathbf{i}}{\hbar} \left(-\frac{\hbar^2}{2m} \nabla^2 \psi^* + V \psi^* \right) \psi + \psi^* \left(\frac{\mathbf{i}}{\hbar} \left(-\frac{\hbar^2}{2m} \nabla^2 \psi + V \psi \right) \psi \right) \\ &= \frac{\mathbf{i}\hbar}{2m} ((\nabla^2 \psi^*) \psi - \psi^* \nabla^2 \psi) \\ &= \underbrace{\vec{\nabla} \cdot \left(\frac{\mathbf{i}\hbar}{2m} ((\vec{\nabla} \psi^*) \psi - \psi^* \vec{\nabla} \psi) \right)}_{\equiv -\vec{j}(x)} \end{aligned} \tag{3.16}$$

This $(\partial_t \rho + \vec{\nabla} \cdot \vec{j} = 0)$ is an equation of continuity. It has the consequence that the time rate of change of the probability of finding the particle in any region R comes only from probability leaking out of or into the boundaries ∂R of R :

$$\partial_t P(R) = - \int_R \vec{\nabla} \cdot \vec{j} = - \oint_{\partial R} \vec{j} \cdot d\vec{a}$$

where $d\vec{a}$ is a normal to the boundary of R . (The last step is Stokes' theorem.) In one dimension, for $R = [x_0, x_1]$, an interval, this is just

$$\partial_t P([x_0, x_1]) = - \int_{x_0}^{x_1} dx \frac{d}{dx} j_x = -j_x|_{x_0}^{x_1}.$$

So the probability current is

$$\vec{j} = \frac{\hbar}{2m\mathbf{i}} \left((\vec{\nabla} \psi^*) \psi - \psi^* \vec{\nabla} \psi \right).$$

Notice that if $\psi(x)$ is real, the probability does not move. Here's the simplest example where there is a nonzero current: Take

$$\psi(x) = ae^{\mathbf{i}\vec{k}\cdot\vec{x}} \quad \implies \quad \rho(x) = |a|^2, \quad \vec{j} = |a|^2 \frac{\hbar\vec{k}}{2m} \cdot 2 = |a|^2 \frac{\vec{p}}{m}.$$

This describes a uniform distribution, with a current moving with velocity $\vec{v} = \vec{p}/m$.

3.2.2 Free particle

Let's consider the case of a free particle, $V = 0$:

$$\mathbf{H} = \frac{\mathbf{p}^2}{2m}.$$

(We'll focus on 1d, I hope you can see where to put the arrows if $d > 1$.) Here $\mathbf{p} = \hbar \mathbf{k} = -i\hbar \mathbf{d}$. Since $\mathbf{H}$ depends only on $\mathbf{p}$ and not on $\mathbf{x}$, we have $[\mathbf{H}, \mathbf{p}] = 0$, which means that momentum eigenstates $\mathbf{k} |k\rangle = k |k\rangle$ are energy eigenstates,

$$\mathbf{H} |k\rangle = \frac{\hbar^2 k^2}{2m} |k\rangle \equiv E_k |k\rangle.$$

So a free (non-relativistic) particle has the *dispersion relation* $E_k = \frac{\hbar^2}{2m} k^2$.

So stationary states are plane waves $\psi_k(x) = \langle x | k \rangle = e^{ikx}$. Let us study the time evolution of some (non-stationary) states, where the particle is initially more localized. We are going to solve the Schrödinger equation in position space:

$$0 = \left(-i\hbar \partial_t - \frac{\hbar^2}{2m} \nabla^2 \right) \psi(x, t).$$

This equation describes diffusion in imaginary time, that is if we let $\tau = it$, it's the diffusion equation in τ . What does that do?

A useful step along the way to doing this (linear differential equations are eager to be solved using Green's functions) is to construct the time-evolution operator:

$$\mathbf{U}(t) = e^{-i\mathbf{H}t/\hbar} = \int dk |k\rangle \langle k| e^{-i\frac{\hbar^2 k^2}{2m} \frac{t}{\hbar}}$$

and find its position-space matrix elements:

$$K(x, x'; t) \equiv \langle x | \mathbf{U}(t) | x' \rangle = \int dk \underbrace{\langle x | k \rangle}_{=e^{ikx}} \underbrace{\langle k | x' \rangle}_{=e^{-ikx'}} e^{-i\frac{\hbar}{2m} t k^2} = \int dk e^{i(x-x')} e^{-i\frac{\hbar t}{2m} k^2}. \quad (3.17)$$

K is the amplitude to propagate from x' to x in time t , naturally, it is called the *propagator*.

The last expression in (3.17) is a gaussian integral.

$$\int_{-\infty}^{\infty} dy e^{-ay^2} = \sqrt{\frac{\pi}{a}}, \quad \text{Re}(a) > 0.$$

This formula can also be used to do integrals like

$$\int_{-\infty}^{\infty} dy e^{-ay^2+by+c}$$

(by completing the square) and

$$\int_{-\infty}^{\infty} dy y^2 e^{-ay^2} = -\partial_a \left(\int_{-\infty}^{\infty} dy e^{-ay^2} \right).$$

The integral we want to do has a pure imaginary. If we regard this as the limit of $\varphi \rightarrow \pi/2$ in $a = e^{i\varphi}$, we get

$$K(x, x'; t) = \sqrt{\frac{m}{2\pi i \hbar t}} e^{\frac{im}{2\hbar t} (x-x')^2}$$

with $\frac{1}{\sqrt{i}} = e^{i\pi/4}$.

How to use the propagator? If the initial state is $|\psi\rangle = |\psi(0)\rangle$, it evolves after a time t to

$$|\psi(t)\rangle = \mathbf{U}(t) |\psi\rangle.$$

Let's relate its wavefunction to $\psi(x, 0) \equiv \langle x | \psi \rangle$:

$$\psi(x, t) \equiv \langle x | \psi(t) \rangle = \langle x | \mathbf{U}(t) | \psi \rangle = \int dx' \underbrace{\langle x | \mathbf{U}(t) | x' \rangle}_{K(x, x'; t)} \underbrace{\langle x' | \psi(0) \rangle}_{=\psi(x', 0)}. \quad (3.18)$$

- Here is bad example of the application of this result. Start with $\psi(x', 0) = \delta(x' - x_0)$ – a position eigenstate. (It's a bad example because this state doesn't have finite energy, since it has $\Delta x = 0$ and hence $\infty = \Delta p^2 = \langle \mathbf{p}^2 \rangle \propto \langle \mathbf{H} \rangle$.) Then the state at time t is just the propagator itself: $\psi(x, t) = K(x, x_0; t)$. For any $t > 0$, the probability density is

$$\rho(x, t) = |\psi(x, t)|^2 = |K(x, x_0; t)|^2 = \frac{m}{2\pi \hbar t}$$

uniform in space. The position uncertainty instantly becomes infinite. (Isn't it a bit troubling that its amplitude decreases in time? Since we are starting with a state which is not in the Hilbert space, the probability can leak into other states that are not in the Hilbert space.)

[\[End of Lecture 8\]](#)

- Here's a better example:

$$\psi(x, 0) = (2\pi\Delta)^{-1/4} e^{-x^2/(4\Delta^2)} e^{ip_0 x/\hbar}.$$

This *Gaussian wavepacket* has probability density

$$\rho(x, 0) = |\psi(x, 0)|^2 = (2\pi\Delta)^{-1/2} e^{\frac{-x^2}{2\Delta^2}},$$

a gaussian with variance $\Delta x_0 = \Delta$, *i.e.* $\Delta x(0)^2 = \langle \mathbf{x}^2 \rangle_0 - \langle \mathbf{x} \rangle_0^2 = \Delta^2$. And the initial momentum is (peaked around) p_0 , in the sense that $\langle \mathbf{p} \rangle_0 = p_0$; also the probability current is

$$j_x = \rho p_0 / m.$$

(To see this last statement it's helpful to note that the gaussian envelope is real.) The initial spread in the momentum is

$$\Delta p(0)^2 = \langle \mathbf{p}^2 \rangle_0 - \langle \mathbf{p} \rangle_0^2 = \left(\frac{\hbar}{2\Delta} \right)^2.$$

Notice that this state saturates the uncertainty bound:

$$\Delta x(0)\Delta p(0) = \Delta \cdot \frac{\hbar}{2\Delta} = \frac{\hbar}{2}.$$

Using (3.18), the wavefunction at time t is another gaussian integral, since both K and ψ are gaussian in x' :

$$\psi(x, t) = \int dx' K(x, x'; t) \psi(x', 0) = e^{-(x-p_0 t/m)^2 / 4\delta^2} e^{i \frac{p_0}{\hbar} (x - \frac{p_0 t}{m})} (2\pi)^{-1/4} \delta(t)^{-1/2}$$

where $\delta \equiv \delta(t) \equiv \Delta + \frac{i\hbar t}{2m\Delta}$. This is a wavepacket whose center has $x_0(t) \sim p_0 t/m$ and whose width in position space is growing:

$$\Delta(t)^2 = |\delta(t)|^2 = \Delta^2 + \frac{\hbar^2 t^2}{m^2 \Delta^2}.$$

Why does the wavepacket spread? The initial distribution of momentum as particles whose momenta differ by as much as $\delta p \equiv \frac{\hbar}{2\Delta}$ from the central value (p_0); in a time t , they will be separated from the central particles by $\delta x(t) = t \delta v = t \frac{\delta p}{m} = t \frac{\hbar}{2\Delta m}$, which is just the answer we found. Notice that a narrower initial packet spreads faster.

When does this spreading of the wavepacket matter?

A grain of sand, according to answers.com, has a mass $m \sim 10^{-3}$ gram and $\Delta \sim 10^{-4}m$. In a time $t \sim 10^{10}$ years (the approximate age of the universe), $\Delta(t)$ is still of order Δ . If we cram a whole gram into an angstrom ($1\text{\AA} = 10^{-10}m$) (how?), and wait the age of the universe, we get $\Delta(t) \sim .5mm$.

On the other hand, if we start with an electron ($m \sim 10^{-31}kg$) in an atom (localized to say $1\text{\AA}$) and wait a nanosecond, we get $\Delta(t) \sim 1mm$.

In $d > 1$:

$$\mathbf{H} = \frac{\vec{\mathbf{p}} \cdot \vec{\mathbf{p}}}{2m} \equiv \frac{\vec{\mathbf{p}}^2}{2m} = \frac{\mathbf{p}_x^2}{2m} + \frac{\mathbf{p}_y^2}{2m} + \dots.$$

So the dispersion relation, instead of a parabola, describes a paraboloid: $E_k = \frac{\hbar^2 k^2}{2m}$. And the position-space propagator is

$$\langle \vec{x} | \mathbf{U}(t) | \vec{x}' \rangle = \int \underbrace{\underbrace{\mathrm{d}^d k}_{\equiv \mathrm{d}k_x \mathrm{d}k_y \dots}}_{= e^{i\vec{k} \cdot \vec{x}'}} \langle \vec{x}' | \vec{k}' \rangle \langle \vec{k}' | \vec{x} \rangle e^{-i\vec{k}'^2 \frac{\hbar}{2m} t} = \left(\frac{m}{2\pi i \hbar t} \right)^{d/2} e^{\frac{im}{2\hbar t} (\vec{x} - \vec{x}')^2}.$$

3.2.3 Particle in potential well

Now consider nonzero $V(x)$, independent of time, in one dimension. We solve the problem if we know the eigenstates of $\mathbf{H}$, $\mathbf{H} |E\rangle = E |E\rangle$. In position space, $\psi(x) \equiv \langle x | E \rangle$, this condition is:

$$-\frac{\hbar^2}{2m} \partial_x^2 \psi(x) + V(x) \psi(x) = E \psi(x).$$

$$\partial_x^2 \psi(x) + \frac{2m}{\hbar^2} (E - V(x)) \psi(x) = 0. \quad (3.19)$$

Notice that the complex conjugate of (3.19) says :

$$\partial_x^2 \psi^*(x) + \frac{2m}{\hbar^2} (E - V(x)) \psi^*(x) = 0, \quad (3.20)$$

i.e. $\psi^*(x)$ is also the wavefunction of an eigenvector of $\mathbf{H}$ with eigenvalue E . Is it the same one?

No degeneracy theorem and the Wronskian. Suppose we have another eigenfunction ϕ with the same eigenvalue E :

$$\partial_x^2 \phi(x) + \frac{2m}{\hbar^2} (E - V(x)) \phi(x) = 0 \quad (3.21)$$

(*e.g.* it could be ψ^*). Multiply (3.19) by ϕ^* and (3.21)* by ψ and subtract:

$$0 = \psi^* \partial_x^2 \phi - \partial_x^2 \phi^* \psi = \partial_x (\phi^* \partial_x \phi - \partial_x \psi^* \phi) \equiv \partial_x W(\phi, \psi).$$

This says that the *Wronskian* W of the two solutions is constant in space. (Notice how this generalizes our proof of probability current conservation, which is the special case with $\phi = \psi$.)

Here's why this is useful: Suppose that there is some point in space x where *both* ϕ and ψ vanish (and their derivatives are finite). For example, this could happen if they satisfy the same boundary condition, and the point x could be $\pm\infty$, where it may be necessary to have $\phi, \psi \rightarrow 0$ to be normalizable. Then $W = 0$ at x and therefore $W = 0$ everywhere. So

$$\frac{\partial_x \psi^*}{\psi^*} = \frac{\partial_x \phi}{\phi} \implies \psi(x) = c\phi(x)$$

which demanding normalized states means c is an unphysical phase. There is therefore no degeneracy under our assumption.

Notice that if V grows at both ends of the world, our hypothesis is met. If $V \rightarrow$ a constant at *both* $x = \pm\infty$ then states can be degenerate. (For a free particle $e^{\pm i k x}$ are two solutions with energy $E = \frac{\hbar^2 k^2}{2m}$.)

We can find real wavefunctions. For a nondegenerate level, the theorem above says

$$\psi(x) = c\psi^*(x)$$

and squaring the BHS gives $|\psi|^2 = |c|^2|\psi|^2 \implies c = e^{i\alpha}, \alpha \in \mathbb{R}$. We can rephase our wavefunction to make it real: $\psi_R \equiv e^{-i\alpha/2}\psi$ is real since

$$\psi_R^* = e^{+i\alpha/2}\psi^* = e^{i\alpha/2}(e^{-i\alpha/2}\psi^*) = \psi_R^*.$$

Forbidden and allowed regions. Consider a real wavefunction $\psi(x)$ with energy E .

- If $V(x) > E$, then

$$0 < \kappa^2 \equiv \frac{2m}{\hbar^2} (V(x) - E) \stackrel{(3.19)}{=} \frac{\psi''(x)}{\psi(x)}.$$

If κ were constant, the solutions would be $e^{\pm\kappa x}$, but more generally such ψ exhibit exponential growth or decay.

- If $V(x) < E$, then

$$0 > -K^2 \equiv \frac{2m}{\hbar^2} (V(x) - E) \stackrel{(3.19)}{=} \frac{\psi''(x)}{\psi(x)}.$$

If K were constant, the solutions would be $e^{\pm iKx}$, but more generally such ψ exhibit oscillation.

The behavior of the potential at ∞ leads to an important dichotomy between *bound-states* and *scattering states*. In the case of the free particle ($V = 0$) we only encounter the latter. A boundstate an energy eigenstate where the wavefunction is actually normalizable, unlike a plane wave. Such states will occur if the potential is a well which can hold a particle from leaking off to infinity, for example if $V(x)$ grows as $|x| \rightarrow \infty$. Even if $V(|x| \rightarrow \infty)$ is a constant V_0 , there can be boundstates if $V(x) < V_0$ for some finite x .

No nodes in the groundstate theorem. Suppose V is such that boundstates are nondegenerate. Then the groundstate wavefunction has no nodes.

Here's the proof [Feynman, *Statistical Mechanics*, p. 322]: Take ψ real WLOG. Suppose $\psi(x)$ passes through zero at $x = x_0$. We'll use this assumption to construct a $\tilde{\psi}(x)$

with lower energy expectation $\langle \mathbf{H} \rangle_{\tilde{\psi}} < \langle \mathbf{H} \rangle_{\psi}$; this means ψ can't be the groundstate by the variational theorem (which you proved on the homework).

To do this, first consider $\phi(x) \equiv |\psi(x)|$. Although the sign of its derivative jumps at x_0 , this state has the same values of $(\partial_x \phi)^2 = (\partial_x \psi)^2$ and $\phi^2 = \psi^2$. Therefore

$$\langle \mathbf{H} \rangle_{\phi} = \frac{\int (\phi(-a\partial^2)\phi + V\phi^2)}{\int \phi^2} = \frac{\int (a(\partial\phi)^2\phi + V\phi^2)}{\int \phi^2} = \frac{\int (a(\partial\psi)^2\phi + V\psi^2)}{\int \psi^2} = \langle \mathbf{H} \rangle_{\psi}.$$

But now we can lower the kinetic energy term by smoothing out the cusp, where $\partial\phi$ jumps. The resulting state will have $\langle \mathbf{H} \rangle_{\tilde{\psi}} < \langle \mathbf{H} \rangle_{\psi}$ as promised.

Note that this result implies that the $n = 0$ state is non-degenerate, since if there were a two-dimensional subspace of $n = 0$ states, we could choose $a_{1,2}$ so that $a_1\psi_1(x) + a_2\psi_2(x)$ vanishes for some x , and it therefore can't be a groundstate.

This statement is actually much more general than 1d single-particle potential problems. Consider several, say N , particles in d dimensions. (Ignoring identical-particle weirdness for now) their Hilbert space is $\mathcal{H}_d \otimes \mathcal{H}_d \otimes \cdots \otimes \mathcal{H}_d$ a tensor product of N copies of that of a single particle. We can resolve the identity by

$$\mathbb{1} = \int d^d x_1 d^d x_2 \cdots d^d x_N |\vec{x}_1, \cdots \vec{x}_N\rangle \langle \vec{x}_1, \cdots \vec{x}_N|.$$

The wavefunction of a state Ψ is

$$\Psi(\vec{x}_1, \cdots \vec{x}_N) \equiv \langle \vec{x}_1, \cdots \vec{x}_N | \Psi \rangle.$$

A useful hamiltonian for non-relativistic particles is

$$\mathbf{H} = \sum_{a=1}^N \frac{\vec{p}_a^2}{2m} + V(\{x\}).$$

If V is not just a sum of functions of one variable at a time V describes interactions among the particles. This Hamiltonian gives a useful description of many real and interesting physical systems.

What's the groundstate? In general this is a very difficult problem. One thing we can say however, is that it has no nodes. To see this, let

$$\psi(\xi) \equiv \Psi(\vec{x}_1, \cdots \vec{x}_N)$$

where ξ is one of the dN coordinates. Now apply the argument above to the function $\psi(\xi)$.

Theorem on the interleaving of the nodes. Here's a generalization of this theorem; the previous result was the $n = 0$ bit of this statement. Claim: the wavefunction for the n th level $\psi_n(x)$ has n nodes (*i.e.* vanishes at n values of x). Each node of $\psi_n(x)$ lies between two nodes of ψ_{n+1} .

For the proof see the book by Morse and Feshbach.

[End of Lecture 9]

Here's another application of the no-nodes theorem. Consider a symmetric double well with a large barrier. A wavefunction localized in the left well is nearly a groundstate, as is the right-well groundstate. But by our result above, they cannot be degenerate if the barrier is finite.

Let $\mathbf{P} \equiv \int dx |x\rangle \langle -x|$ be the *parity* operator. Since our potential is symmetric, $V(x) = V(-x)$, we have $[\mathbf{H}, \mathbf{P}] = 0$, and parity eigenstates are energy eigenstates, if they are nondegenerate.

Claim: $\psi_{\pm} = \frac{1}{\sqrt{2}}(\psi_L \pm \psi_R)$ are eigenstates of $\mathbf{P}$.

Who is the groundstate? ψ_+ has no node, ψ_- has a node. So ψ_- cannot be the groundstate.

Galilean invariance of the Schrödinger equation.

If you watch water waves while driving past the ocean, their wavelength is the same as if you were sitting on the beach. Hydrodynamics is invariant under a Galilean boost, $x \rightarrow x' = x + vt, t' = t$, and the wavenumber is invariant. (Note that I am assuming $v \ll c$; we are talking about Galilean relativity.) But de Broglie waves are weird. They have wavelength $\lambda = \frac{h}{p}$. Under a Galilean boost, $p \rightarrow p' = p + mv$, so $\lambda \rightarrow \lambda' = \frac{h}{p'}$, not invariant. Nevertheless the Schrödinger equation for a free particle is Galilean invariant. Exactly how this works out is a problem on the homework.

3.2.4 Relative and center-of-mass coordinates

Suppose we have two particles, interacting via a potential that depends only on their separation:

$$\mathbf{H} = \frac{\vec{\mathbf{p}}_1^2}{2m_1} + \frac{\vec{\mathbf{p}}_2^2}{2m_2} + V(\vec{x}_1 - \vec{x}_2).$$

As in classical mechanics, we can eliminate the center of mass degree of freedom, which is free, by introducing:

$$x \equiv x_1 - x_2, \quad X \equiv \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2}$$

in terms of which the kinetic terms in the position-space Schrödinger equation are

$$-\frac{\hbar^2}{2m_1}\partial_{x_1}^2 - \frac{\hbar^2}{2m_2}\partial_{x_2}^2 = -\frac{\hbar^2}{2M}\partial_X^2 - \frac{\hbar^2}{2\mu}\partial_x^2$$

with

$$M \equiv m_1 + m_2, \quad \mu \equiv \frac{m_1 m_2}{m_1 + m_2}$$

is the reduced mass. A basis of solutions of the Schrödinger equation are of the factorized form

$$\Psi(x_1, x_2) = \langle x_1, x_2 | \Psi \rangle = \phi(X)\psi(x)$$

where

$$\mathbf{i}\partial_t\psi(x) = -\frac{\hbar^2}{2\mu}\partial_x^2\psi(x) + V(x)\psi(x), \quad \mathbf{i}\partial_t\phi(X) = -\frac{\hbar^2}{2M}\partial_X^2\phi(X).$$

So the center of mass degree of freedom is a free particle which factors out of the problem.

3.2.5 Scattering and bound states: piecewise-constant potentials

Example 1: $V(x) = W_0\delta(x)$. Note that W_0 has units of work. Let's look for a boundstate. Away from $x = 0$, the Schrödinger equation is

$$\psi''(x) + \underbrace{\frac{2m}{\hbar^2}E}_{\equiv -\kappa^2} \psi = 0 \tag{3.22}$$

which is solved by $\psi \sim e^{\pm\kappa x}$. We are going to glue solutions from different regions. A normalizable state must decay far from the origin, so we must have

$$\psi_{<}(x) = Ae^{+\kappa x}, \quad \psi_{>}(x) = Be^{-\kappa x}$$

where $\psi_{\leq}$ is the solution for $x \leq 0$.

How to match the two? If ψ' jumps at $x = 0$ then

$$\psi'(\epsilon) - \psi'(-\epsilon) = \int_{-\epsilon}^{\epsilon} dx \psi''(x) \stackrel{\text{Schröd}}{=} \int_{-\epsilon}^{\epsilon} dx \psi(V-E) \frac{2m}{\hbar^2} = \frac{2m}{\hbar^2} (W_0\psi(0) - E\psi(0)\epsilon) = \frac{2m}{\hbar^2} W_0\psi(0).$$

(Here we took $\epsilon \ll 1$ tiny.) So the junction conditions are:

$$\begin{aligned} \psi_{>}(\epsilon) &= \psi_{<}(-\epsilon) = \psi(0) \\ \psi'_{>}(\epsilon) - \psi'_{<}(-\epsilon) &= \frac{2m}{\hbar^2} W_0\psi(0). \end{aligned} \tag{3.23}$$

So $A = B$ and $-\kappa 2A = \frac{2m}{\hbar^2} W_0 A$. The second equation is only solvable if $W_0 < 0$, that is, if the potential is *attractive*. The solution in that case is $\kappa = \frac{m|W_0|}{\hbar^2}$. The boundstate energy is then

$$E_b = -\frac{\hbar^2 \kappa^2}{2m} = -\frac{m|W_0|^2}{2\hbar^2}.$$

There is a single boundstate.

Now let's think about scattering off of this potential. This is the whole story for $W_0 > 0$. $E > 0$ solutions of (3.22) are obtained by $k^2 \equiv \frac{2mE}{\hbar^2} > 0$

$$\psi_{<} = c_1 e^{ikx} + c_2 e^{-ikx}, \quad \psi_{>} = c_3 e^{ikx} + c_4 e^{-ikx}.$$

Let's suppose our particle beam is sending particles in from the left, so $c_4 = 0, c_1 = 1$. $c_3 \equiv T$ is the transmission amplitude, and $c_2 = R$ is the reflection amplitude. The junction conditions at $x = 0$ are then

$$1 + R = T = \psi(0), \quad \psi'_{<}(-\epsilon) - \psi'_{>}(\epsilon) = ikT - (ik - ikR) = W_0 \frac{2m}{\hbar^2} T.$$

Therefore

$$T = \frac{ik}{ik - mW_0/\hbar^2}, \quad R = 1 - T = \frac{mW_0/\hbar^2}{ik - mW_0/\hbar^2}.$$

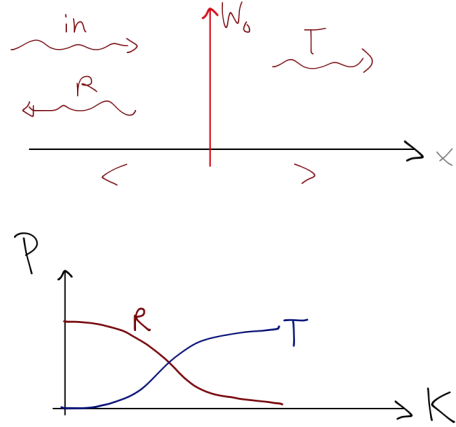
This solution is valid for either sign of W_0 .

Let's do some accounting of the scattering of probability-amplitude waves. Recall that for a plane wave $\psi = ae^{ikx}$, the probability density and current are $\rho = |a|^2, j = |a|^2 \hbar k/m$. We identify three different currents, associated to the incident, reflected and transmitted waves: $j_{in} = \hbar k/m, j_R = -|R|^2 \hbar k/m, j_T = |T|^2 \hbar k/m$. The probability of transmission is then the ratio of the left-going probability fluxes:

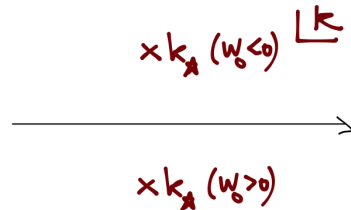
$$P_T = \frac{j_T}{j_{in}} = |T|^2 = \frac{k^2}{k^2 + \left(\frac{mW_0^2}{\hbar^2}\right)^2}.$$

$$P_R = \frac{j_R}{j_{in}} = |R|^2 = \frac{\left(\frac{mW_0^2}{\hbar^2}\right)^2}{k^2 + \left(\frac{mW_0^2}{\hbar^2}\right)^2}.$$

Notice that $P_R + P_T = 1$ accounts for all the possibilities. This probability doesn't care about the sign of W_0 .



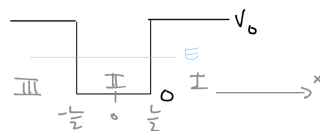
To tell the difference between attractive and repulsive potentials by scattering requires *phase* information. The transmission amplitude has a *pole* at $k_* = -i\frac{mW_0}{\hbar^2}$. This value of k is what determines the bound state energy: $E_b = \frac{\hbar^2 k_*^2}{2m}$. How could you measure this information? The phase of the reflected wavefunction can be measured by looking at its interference with the incident wave.



The kind of function P_R and P_T is called a *Lorentzian* and is the absolute value of a complex function with a pole. It has a bump, called a resonance, whose width is determined by the distance of the pole from the real axis.

Example 2: Square well, $V(x) = \begin{cases} V_0, & |x| > L/2, \\ 0, & |x| < L/2 \end{cases}$.

Now we have three regions to match, but the derivatives should all be continuous. We will think of V_0 as really big, so let's consider E in between 0 and V_0 . In region I ($x > L/2$), then, we are below the barrier, and the solution is



$$\psi_I(x) = Ae^{\kappa x} + Be^{-\kappa x}, \quad \kappa^2 \equiv \frac{2m}{\hbar^2}(V_0 - E) > 0.$$

Normalizability requires $A = 0$.

Similarly, in region III ($x < -L/2$),

$$\psi_{III}(x) = Fe^{\kappa x} + Ge^{-\kappa x}, \quad \kappa^2 \equiv \frac{2m}{\hbar^2}(V_0 - E) > 0.$$

Normalizability requires $G = 0$.

In region II ($x \in (-L/2, L/2)$),

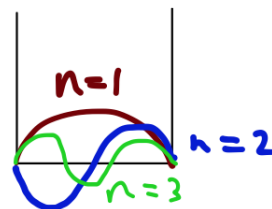
$$\psi_{II}(x) = Ce^{ikx} + De^{-ikx}, \quad k^2 = \frac{2mE}{\hbar^2}.$$

Taking $V_0 \rightarrow \infty$ means that B and F don't matter: $e^{-\infty} = 0$ anyway, and the particle is localized to the well. In that case (on which we focus to avoid horrible algebra), the boundary conditions say

$$0 = \psi(-L/2) = \psi(L/2) = Ce^{ikL/2} + De^{-ikL/2} = Ce^{-ikL/2} + De^{ikL/2}.$$

These equations are only solvable for nonzero C, D if

$$0 = \det \begin{pmatrix} e^{ikL/2} & e^{-ikL/2} \\ e^{-ikL/2} & e^{ikL/2} \end{pmatrix} = e^{ikL} - e^{-ikL} = 2i \sin kL.$$



Therefore nonzero solutions only exist for $k = k_n \equiv \frac{\pi n}{L}$.

The solution is nonzero in region II and looks like:

$$\psi_n(x) = C_n e^{ik_n x} + D_n e^{-ik_n x} = \begin{cases} \sqrt{\frac{2}{L}} \cos k_n x, & n = 1, 3, 5, \dots \\ \sqrt{\frac{2}{L}} \sin k_n x, & n = 2, 4, 6, \dots \end{cases}$$

Notice that the parity operator $\mathbf{P} = \int dx |x\rangle \langle -x|$ commutes with $\mathbf{H}$, since the potential is even, $V(-x) = V(x)$. The states we've found $\psi_n(x) = \langle x|n\rangle$ are also parity eigenstates:

$$\mathbf{P} |n\rangle = (-1)^{n+1} |n\rangle.$$

The odd functions have a node at $x = 0$; the even functions have $\psi'(0) = 0$. The energies are

$$E_n = \frac{\hbar^2}{2m} k_n^2 = \frac{\hbar^2}{2m} \frac{\pi^2}{L^2} n^2.$$

Finite V_0 . Returning to finite V_0 , a point worth noting from this example is that an attractive potential does not automatically mean a boundstate.

Focus on states with $0 < E < V_0$, those that are trapped in the well. They are non-degenerate parity eigenstates, and the groundstate must be even, since odd functions have a zero at the origin. So we need only consider what happens for $x > 0$. First even states:

$$\psi_{II} = A \cos k_1 x, \quad \psi_{III} = B e^{-\kappa_2 x}$$

with $k_1^2 = \frac{2m}{\hbar^2} E$ and $\kappa_2^2 = \frac{2m}{\hbar^2} (V_0 - E)$.

$$A \cos k_1 L/2 = B e^{-\kappa_2 L/2}, \quad -k_1 A \sin k_1 L/2 = -\kappa_2 B e^{-\kappa_2 L/2}.$$

Dividing the first equation by the second this gives an awful transcendental condition on k_1 and hence the boundstate energy:

$$\tan \frac{k_1 L}{2} = \frac{\kappa_2}{k_1}.$$

In terms of $\omega \equiv \sqrt{\frac{2m}{\hbar^2} E} L/2$, this is

$$\tan \omega = f(\omega) \equiv \sqrt{\left(\frac{V_0 L^2}{4\omega^2}\right)^2 - 1}.$$

This has a solution for any attractive potential.

Now consider the odd states:

$$\psi_{II} = A \sin k_1 x, \quad \psi_{III} = B e^{-\kappa_2 x}.$$

The same analysis gives instead $-\cot \omega = f(\omega)$. This equation only has a solution if $V_0 > \left(\frac{\pi}{L}\right)^2$, since $f(\omega)$ is only real for $\omega^2 < V_0 L^2/4$ (and it is positive) but $-\cot \omega$ is negative until $\omega = \pi/2$.

Why do we care about the odd states? Suppose we consider instead the potential

$$V(x) = \begin{cases} V_0, & x > L/2, \\ 0, & 0 < x < L/2, \\ \infty, & x < 0 \end{cases}$$

Eigenfunctions for this potential have to have $\psi(0) = 0$, but are otherwise the same as those above: this is exactly the odd ones. So *this* potential only has a boundstate when the potential is attractive *enough*.

We can also discern the existence of boundstates from poles in the scattering amplitude (with the appropriate sign of $\text{Im} k$). The transmission amplitude is (I am not going to do the algebra at you)

$$T = \frac{2e^{-ik_1 L}}{2 \cos(k_1 L) - \frac{i(k_1^2 + k_2^2)^2}{k_1 k_2} \sin(k_1 L)},$$

with $k_1^2 = \frac{2m}{\hbar^2} E$ and $k_2^2 = \frac{2m}{\hbar^2} (E - V_0)$.

Notice that for $k_2 L = 2n\pi$ we have $|T|^2 = 1$ – perfect transmission; this is just like in optics where the reflected wave from the backside of a piece of glass interferes destructively with the reflected with from the front side.

Example 3: Potential step. Consider scattering ($E > 0$) from the left off of $V(x) = \begin{cases} 0, & x > 0, \\ V_0, & x < 0 \end{cases}$. This is like water waves hitting the shore, where their wavenumber changes. On the left,

$$\psi_{<}(x) = e^{ik_{<}x} + R e^{-ik_{<}x}, \quad k_{<}^2 = \frac{2mE}{\hbar^2}.$$

On the right:

$$\psi_{>}(x) = T e^{ik_{>}x}, \quad k_{>}^2 = \frac{2m(E - V_0)}{\hbar^2}.$$

Matching conditions are:

$$1 + R = T, \quad ik_{<}(1 - R) = ik_{>}T \quad \implies \quad R = \frac{k_{<} - k_{>}}{k_{<} + k_{>}}, \quad T = \frac{2k_{<}}{k_{<} + k_{>}}.$$

The formulae for the various currents are as above, except now the left and right wavenumbers are different: $j_{in} = \hbar k_{<}/m$, $j_R = -|R|^2 \hbar k_{<}/m$, $j_T = |T|^2 \hbar k_{>}/m$. Reflection and transmission probabilities are then

$$P_R = \frac{j_R}{j_{in}} = |R|^2 = \left| \frac{k_{<} - k_{>}}{k_{<} + k_{>}} \right|^2.$$

$$P_T = \frac{j_T}{j_{in}} = |T|^2 \frac{k_{>}}{k_{<}} = \frac{4k_{<}k_{>}}{(k_{<} + k_{>})^2}.$$

Notice the highlighted step: the particle flux depends on the wavenumber. Happily, $P_R + P_T = 1$.

Notice that for $0 < E < V_0$, the wavefunction on the right is $e^{-\sqrt{\frac{2m(V_0-E)}{\hbar^2}}x}$, which is real. Therefore $j_T = 0$, and $P_T = 0, P_R = 1$. Our formulae above smoothly approach this answer as $E \rightarrow V_0$ from below.

[End of Lecture 10]

3.2.6 Index theorem

Consider an operator

$$\alpha : V \rightarrow W, \quad \alpha^\dagger : W \rightarrow V.$$

Then $\alpha^\dagger \alpha : V \rightarrow V$ is self-adjoint: $(\alpha^\dagger \alpha)^\dagger = \alpha^\dagger (\alpha^\dagger)^\dagger = \alpha^\dagger \alpha$. Similarly, $\alpha \alpha^\dagger : W \rightarrow W$ is self-adjoint. Further, for any state in V ,

$$\langle v | \alpha^\dagger \alpha | v \rangle = \| \alpha | v \rangle \|^2 \geq 0.$$

In particular, for a (normalized) eigenstate of $\alpha^\dagger \alpha$, $\alpha^\dagger \alpha |\lambda\rangle_v = \lambda |\lambda\rangle_v$,

$$\lambda = \langle \lambda | \alpha^\dagger \alpha | \lambda \rangle_v = \| \alpha | \lambda \rangle_\lambda \|^2 \geq 0.$$

A similar statement holds for eigenvalues of $\alpha \alpha^\dagger$.

Now consider a state $\alpha |\lambda\rangle_v \in W$:

$$\alpha \alpha^\dagger (\alpha |\lambda\rangle_v) = \alpha \left(\underbrace{\alpha^\dagger \alpha | \lambda \rangle_v}_{= \lambda | \lambda \rangle_v} \right) = \lambda \alpha |\lambda\rangle_v.$$

Therefore, if $\lambda > 0$, $\frac{1}{\sqrt{\lambda}} \alpha |\lambda\rangle_v$ is a (normalized) eigenvector of $\alpha \alpha^\dagger$ with the same eigenvalue λ . On the other hand, if $\lambda = 0$, then this is the zero state: $\alpha |\lambda = 0\rangle_v = 0$, since its norm is $\lambda = 0$. This is called a ‘zeromode’. A similar statement applies going from W to V . This means that eigenvectors with nonzero eigenvalues of $\alpha^\dagger \alpha$ and $\alpha \alpha^\dagger$ come in pairs, one in V and one in W . Only zeromodes don’t have partners.

This means that we can define the

$$\text{index}(\alpha) \equiv \# \text{ of zeromodes of } \alpha - \# \text{ of zeromodes of } \alpha^\dagger$$

and it is *topological*, in the sense that it cannot change under continuous deformations. For example, by smoothly changing the nonzero eigenvalues $\lambda_i > 0$, taking some of them to zero, we can only add or remove zeromodes in pairs, and the index changes by $n - n = 0$.

3.2.7 Harmonic oscillator

The simple harmonic oscillator is ubiquitous in physics, not just because it is exactly solvable both classically and quantum mechanically, but because it arises as the leading approximation to any system near a stable equilibrium. Expanding a potential in one dimension about its minimum, Taylor's theorem says

$$V(x) = V(x_0) + 0 + \frac{1}{2}V''(x_0)(x - x_0)^2 + \dots$$

and we can often ignore the ... (aka anharmonic terms) for systems that are near their equilibrium configuration.

$$\mathbf{H} = \frac{\mathbf{p}^2}{2m} + \frac{1}{2}m\omega^2\mathbf{x}^2 = \frac{\hbar\omega}{2}(\mathbf{P}^2 + \mathbf{Q}^2) = \hbar\omega\left(\mathbf{a}^\dagger\mathbf{a} + \frac{1}{2}\right)$$

with

$$\mathbf{a} \equiv \frac{1}{\sqrt{2}}(\mathbf{Q} + \mathbf{iP}), \quad \mathbf{a}^\dagger \equiv \frac{1}{\sqrt{2}}(\mathbf{Q} - \mathbf{iP}).$$

Here I've defined these new operators to hide the annoying factors:

$$\mathbf{Q} \equiv \left(\frac{m\omega}{\hbar}\right)^{1/2}\mathbf{x}, \quad \mathbf{P} \equiv \left(\frac{1}{m\hbar\omega}\right)^{1/2}\mathbf{p}.$$

$$[\mathbf{x}, \mathbf{p}] = \mathbf{i}\hbar\mathbb{1} \quad \implies \quad [\mathbf{a}, \mathbf{a}^\dagger] = \mathbb{1}.$$

(Note that $[\mathbf{Q}, \mathbf{P}] = \mathbf{i}$.)

The *number operator* $\mathbf{N} \equiv \mathbf{a}^\dagger\mathbf{a}$ satisfies

$$[\mathbf{N}, \mathbf{a}] = -\mathbf{a}, \quad [\mathbf{N}, \mathbf{a}^\dagger] = +\mathbf{a}^\dagger.$$

So $\mathbf{a}$ and $\mathbf{a}^\dagger$ are lowering and raising operators for the number operator, in the sense that if $\mathbf{N}|n\rangle = n|n\rangle$

$$\mathbf{N}(\mathbf{a}^\dagger|n\rangle) = (\mathbf{a}^\dagger\mathbf{N} + \mathbf{a}^\dagger)|n\rangle = (n+1)(\mathbf{a}^\dagger|n\rangle), \quad \mathbf{N}(\mathbf{a}|n\rangle) = (n-1)(\mathbf{a}|n\rangle).$$

This means the eigenvalues n of the number operator are spaced by 1. The eigenvalues of the number operator have to be positive, since

$$0 \leq \|\mathbf{a}|n\rangle\|^2 = \langle n|\mathbf{a}^\dagger\mathbf{a}|n\rangle = \langle n|\mathbf{N}|n\rangle = n\langle n|n\rangle.$$

So the smallest eigenvalue of $\mathbf{N}$ is between 0 and 1. But if it isn't zero $\mathbf{a}|n\rangle$ is also an eigenvector of $\mathbf{N}$ with eigenvalue $n-1$, which would be negative. This means that for

$n = 0$ we must have $\mathbf{a}|n=0\rangle = 0$. It has to stop somewhere, and the only way it can stop is to have a zeromode. So the eigenstates of $\mathbf{N}$ (and hence of $\mathbf{H} = \hbar\omega (\mathbf{N} + \frac{1}{2})$) are

$$|0\rangle, \quad |1\rangle \equiv \mathbf{a}^\dagger |0\rangle, \quad \dots, |n\rangle = c_n (\mathbf{a}^\dagger)^n |0\rangle \dots$$

where we must choose c_n to normalize these states.

The answer which gives $\langle n|n\rangle = 1$ is $c_n = \frac{1}{\sqrt{n!}}$. To see this, start with $\langle 0|0\rangle = 1$. Then $|1\rangle = \mathbf{a}^\dagger |0\rangle$ is normalized since

$$\langle 1|1\rangle = \langle 0|\mathbf{a}\mathbf{a}^\dagger|0\rangle = \langle 0|(1 + \mathbf{a}^\dagger\mathbf{a})|0\rangle = \langle 0|0\rangle = 1.$$

Now the induction step. Suppose we've normalized $n-1$. Let $|n\rangle = C_n \mathbf{a}^\dagger |n-1\rangle$, so

$$\begin{aligned} \langle n|n\rangle &= |C_n|^2 \langle n-1|\mathbf{a}\mathbf{a}^\dagger|n-1\rangle = |C_n|^2 \langle n-1|\left(1 + \underbrace{\mathbf{a}^\dagger\mathbf{a}}_{=\mathbf{N}=n-1}\right)|n-1\rangle \\ &= |C_n|^2 \langle n-1|n-1\rangle (1 + (n-1)) = |C_n|^2 n. \end{aligned} \quad (3.24)$$

So $|n\rangle = \frac{1}{\sqrt{n}} \mathbf{a}^\dagger |n-1\rangle$, and we verify $c_n = \frac{1}{\sqrt{n!}}$.

Index. So the SHO is a special case of our previous discussion where we further specify that $[\mathbf{a}, \mathbf{a}^\dagger] = 1$. In that case, we have the matrix representations

$$\mathbf{a} = \begin{pmatrix} 0 & 1 & & \\ 0 & 0 & \sqrt{2} & \\ 0 & 0 & \sqrt{3} & \\ 0 & 0 & 0 & \sqrt{4} \end{pmatrix}, \quad \mathbf{a}^\dagger = \begin{pmatrix} 0 & 0 & & \\ 1 & 0 & 0 & \\ \sqrt{2} & 0 & 0 & \\ \sqrt{3} & 0 & 0 & \end{pmatrix}$$

and

$$\mathbf{a}^\dagger \mathbf{a} = \begin{pmatrix} 0 & & & \\ 1 & & & \\ & 2 & & \\ & & 3 & \end{pmatrix} = \mathbf{N} \quad \mathbf{a}\mathbf{a}^\dagger = \mathbf{N} + 1 = \begin{pmatrix} 1 & & & \\ & 2 & & \\ & & 3 & \\ & & & 4 \end{pmatrix}$$

So that only $\mathbf{a}$ has a single zeromode, and $\mathbf{a}^\dagger$ has no zeromode; all other states are paired, as promised. The index of $\mathbf{a}$ is therefore 1.

Oscillator wavefunctions. Wavefunctions for these states are now easy to obtain. The condition $0 = \mathbf{a}|0\rangle$ in position space is a first-order differential equation:

$$0 = \langle x|\mathbf{a}|0\rangle \propto \langle x|\left(\mathbf{x} + \frac{\mathbf{i}}{m\omega}\mathbf{p}\right)|0\rangle = \left(x - \frac{\hbar}{m\omega}\partial_x\right)\phi_0(x).$$

In terms of $\mathbf{Q}$ and $\mathbf{P}$ ($\mathbf{Q}|q\rangle = q|q\rangle$, $\psi(q) = \langle q|\psi\rangle$) the condition is

$$0 = \langle q|(\mathbf{Q} + i\mathbf{P})|0\rangle \implies \psi(q) = \pi^{-1/4} e^{-q^2/2}, \quad 1 = \int dq |\psi_0(q)|^2.$$

To find $\phi_n(x) \equiv \langle x|n\rangle$ now, we can just act with $\mathbf{a}^\dagger$ n times (and normalize).

Parity is a symmetry of the harmonic oscillator; it acts by

$$\mathbf{x} \rightarrow \mathbf{P}\mathbf{x}\mathbf{P} = -\mathbf{x}, \mathbf{p} \rightarrow -\mathbf{p}, \mathbf{a} \rightarrow -\mathbf{a}, \mathbf{a}^\dagger \rightarrow -\mathbf{a}^\dagger$$

which means that the number eigenstates have

$$\mathbf{P}|n\rangle = (-1)^n |n\rangle.$$

Notice that again the groundstate is parity-even.

What are the position-space properties of the energy eigenstates?

$$\langle n|\mathbf{x}|n\rangle = \sqrt{\frac{\hbar}{2m\omega}} \langle n|(\mathbf{a} + \mathbf{a}^\dagger)|n\rangle = \sqrt{\frac{\hbar}{2m\omega}} (\langle n|n-1\rangle \sqrt{n} + \langle n|n+1\rangle \sqrt{n+1}) = 0.$$

$$\begin{aligned} \langle n|\mathbf{x}^2|n\rangle &= \frac{\hbar}{2m\omega} \langle n|(\mathbf{a} + \mathbf{a}^\dagger)^2|n\rangle = \frac{\hbar}{2m\omega} \langle n|(\mathbf{a}^2 + (\mathbf{a}^\dagger)^2 + \mathbf{a}\mathbf{a}^\dagger + \mathbf{a}^\dagger\mathbf{a})|n\rangle \\ &= \frac{\hbar}{2m\omega} (\langle n|n\rangle n + \langle n+1|n+1\rangle (n+1)) = \frac{\hbar}{2m\omega} (2n+1). \end{aligned} \quad (3.25)$$

$$\Delta x_n = \sqrt{\langle \mathbf{x}^2 \rangle_n - \langle \mathbf{x} \rangle_n^2} = \sqrt{\frac{\hbar}{2m\omega}} \sqrt{2n+1}.$$

Similarly, $\langle \mathbf{p} \rangle_n = 0$ and $\Delta p_n = \sqrt{\frac{m\omega\hbar}{2}} \sqrt{2n+1}$. So $\Delta x_n \Delta p_n = \frac{\hbar}{2} (2n+1) \geq \frac{\hbar}{2}$. The groundstate $n=0$ is a minimum-uncertainty state.

What's oscillating? In any of these states $\langle \mathbf{Q} \rangle_n = \langle \mathbf{P} \rangle_n = 0$. A solution in which something oscillates must be a superposition of energy eigenstates. How do you make a spring oscillate? Displace it from equilibrium and let it go. To effect such a displacement, we act with the *displacement operator*

$$\mathbf{T}(s) \equiv e^{-i\mathbf{P}s}.$$

Why is this the displacement operator: Taylor's theorem. If $\langle q|\psi\rangle = \psi(q)$, then

$$\langle q|\mathbf{T}(s)|\psi\rangle = e^{+s\partial_q} \psi(q) = \psi(q+s).$$

The wavefunction gets translated by s .

In terms of the creation and annihilation operators, $\mathbf{P} = \frac{i}{\sqrt{2}} (\mathbf{a}^\dagger - \mathbf{a})$. Therefore

$$\mathbf{T}(s) = e^{-i\mathbf{P}s} = e^{+\frac{s}{\sqrt{2}}(\mathbf{a}^\dagger - \mathbf{a})}$$

and we can use

$$e^{\mathbf{A}+\mathbf{B}} = e^{\mathbf{A}}e^{\mathbf{B}}e^{-\frac{1}{2}c}, \quad \text{if } [\mathbf{A}, \mathbf{B}] = c, \text{ a c-number,}$$

which you proved on the homework. Here $\mathbf{A} = \frac{s}{\sqrt{2}}\mathbf{a}^\dagger$, $\mathbf{B} = -\frac{s}{\sqrt{2}}\mathbf{a}$, $c = s^2/2$. Therefore, if we translate the vacuum state $|0\rangle$ by s , we have

$$\begin{aligned} |\Psi(0)\rangle &= e^{-i\mathbf{P}s} |0\rangle = e^{-s^2/4} e^{\frac{s}{\sqrt{2}}\mathbf{a}^\dagger} \underbrace{e^{-\frac{s}{\sqrt{2}}\mathbf{a}} |0\rangle}_{=|0\rangle} \\ &= e^{-s^2/4} e^{\frac{s}{\sqrt{2}}\mathbf{a}^\dagger} |0\rangle = e^{-s^2/4} \sum_{n=0}^{\infty} \frac{(s/\sqrt{2})^n}{\sqrt{n!}} |n\rangle. \end{aligned} \quad (3.26)$$

So all levels appear. What's the probability of finding level m if we measure $\mathbf{N}$?¹¹

$$P_m = |\langle m|\Psi(0)\rangle|^2 = e^{-s^2/2} \left(\frac{s^2}{2}\right)^m \frac{1}{m!}$$

– this is a Poisson distribution.

Time evolution. Now take this displaced wavefunction as the initial state and time evolve it with the SHO Hamiltonian.

$$\begin{aligned} |\psi(t)\rangle &= \mathbf{U}(t) |\psi(0)\rangle \\ &= e^{-s^2/4} \left(\mathbf{U} e^{\frac{s}{\sqrt{2}}\mathbf{a}^\dagger} \mathbf{U}^\dagger \right) \left(\mathbf{U} e^{-\frac{s}{\sqrt{2}}\mathbf{a}} \mathbf{U}^\dagger \right) \underbrace{\mathbf{U} |0\rangle}_{=e^{-i\omega t}|0\rangle}. \end{aligned} \quad (3.27)$$

Considering first just the conjugation of $e^{\mathbf{a}^\dagger}$:

$$\left(\mathbf{U} e^{-\frac{s}{\sqrt{2}}\mathbf{a}^\dagger} \mathbf{U}^\dagger \right) = \sum_{k=0}^{\infty} \frac{1}{k!2^k} s^k \underbrace{\mathbf{U} \mathbf{a}^\dagger \mathbf{U}^\dagger}_{=(\mathbf{U} \mathbf{a}^\dagger \mathbf{U}^\dagger)^k} = e^{\frac{s}{\sqrt{2}}\mathbf{U} \mathbf{a}^\dagger \mathbf{U}^\dagger}. \quad (3.28)$$

What is $\mathbf{U} \mathbf{a}^\dagger \mathbf{U}^\dagger$? For any (time-independent) operator $\mathcal{O}$, it is sometimes useful to think of such an object as

$$\mathcal{O}(t) \equiv \mathbf{U}(t) \mathcal{O} \mathbf{U}^\dagger(t)$$

¹¹How do you measure $\mathbf{N}$? For a ball on a spring it's not so obvious; maybe you could use the fact that $\mathbf{N}$ determines the energy. In other contexts described by an SHO it is clearer. For example, a single mode (given wavenumber and polarization) of the electromagnetic field is described by such an oscillator. In that context, $\mathbf{N}$ is the number of photons in that mode.

(Such an organization of the time-dependence is sometimes grandiosely called ‘Heisenberg picture’). It satisfies the differential equation

$$-i\hbar\partial_t\mathcal{O}(t) = [\mathbf{H}, \mathcal{O}(t)] .$$

For the case of $\mathcal{O} = \mathbf{a}^\dagger$ this is

$$-i\partial_t\mathbf{a}^\dagger(t) = \hbar^{-1}[\mathbf{H}, \mathbf{a}^\dagger(t)] = \omega\mathbf{a}^\dagger(t)$$

which is solved by

$$\mathbf{a}^\dagger(t) = e^{-i\omega t}\mathbf{a}^\dagger.$$

Similarly, $\mathbf{a}(t) = e^{+i\omega t}\mathbf{a}$. Returning to (3.28) we have

$$\left(\mathbf{U}e^{-\frac{s}{\sqrt{2}}\mathbf{a}}\mathbf{U}^\dagger\right) = e^{\frac{s}{\sqrt{2}}\mathbf{U}\mathbf{a}^\dagger\mathbf{U}^\dagger} = e^{\frac{s}{\sqrt{2}}e^{-i\omega t}\mathbf{a}^\dagger}$$

and finally

$$\begin{aligned} |\psi(t)\rangle &= \mathbf{U}(t) |\psi(0)\rangle \\ &= e^{-i\omega t/2} e^{-s^2/4} e^{\frac{s}{\sqrt{2}}\mathbf{a}^\dagger} e^{-i\omega t} e^{-\frac{s}{\sqrt{2}}e^{i\omega t}\mathbf{a}} |0\rangle \\ &= e^{-i\omega t/2} \exp \left[\underbrace{\frac{s}{\sqrt{2}} (\mathbf{a}^\dagger e^{-i\omega t} - \mathbf{a} e^{+i\omega t})}_{= \frac{s}{\sqrt{2}} ((\mathbf{a}^\dagger - \mathbf{a}) \cos \omega t - (\mathbf{a}^\dagger + \mathbf{a}) \sin \omega t)} \right] |0\rangle \\ &= e^{-i\omega t/2} \exp [s (-i\mathbf{P} \cos \omega t + \mathbf{Q} \sin \omega t)] |0\rangle \\ &= e^{-i\omega t/2} e^{i\frac{s^2}{4} \sin 2\omega t} e^{-is\mathbf{Q} \sin \omega t} e^{-is\mathbf{P} \cos \omega t} |0\rangle \end{aligned} \tag{3.29}$$

So the wavefunction at time t is

$$\psi(Q, t) \equiv \langle Q | \psi(t) \rangle = \underbrace{e^{-i\omega t/2} e^{i\frac{s^2}{4} \sin 2\omega t} e^{-isQ \sin \omega t}}_{\text{phase}} \langle Q | e^{-is\mathbf{P} \cos \omega t} |0\rangle$$

and the probability density is

$$\rho(Q, t) = |\psi(Q, t)|^2 = |\langle Q | e^{-is\mathbf{P} \cos \omega t} |0\rangle|^2 = |\langle Q - s \cos \omega t | 0 \rangle|^2 = \rho(Q - s \cos \omega t, 0).$$

The shape of the distribution stays fixed, and it sloshes back and forth. This is special to the *harmonic* oscillator.

[End of Lecture 11]

Classical limit. Let's think about the classical limit of the previous discussion. The biggest contribution to the state comes from $n = \bar{n} = s^2/2$, which has

$$E_{\bar{n}} = \hbar\omega \left(\bar{n} + \frac{1}{2} \right) = \hbar\omega \left(\frac{s^2}{2} + \frac{1}{2} \right).$$

Let's consider the case of a large initial displacement, $s \gg 1$. In terms of $\mathbf{Q} = \sqrt{\frac{m\omega}{\hbar}}\mathbf{x}$, this means the initial amplitude of displacement in x is $d = \sqrt{\frac{\hbar}{m\omega}}s$, and the initial average energy is

$$E_{\bar{n}} = \hbar\omega \left(\frac{s^2}{2} + \frac{1}{2} \right) \approx \hbar\omega \frac{s^2}{2} = \hbar\omega^2 d^2 = E_{cl}$$

which is the energy of a classical harmonic oscillator at rest at distance d from its equilibrium. Let's put in some numbers for a macroscopic object: if $m = 1g, d = 1cm, \omega = 1$ radian/second then $s \sim 10^{13}, \bar{n} = s^2/2 \sim 10^{26}, \Delta n = \sqrt{\bar{n}} \sim 10^{13}$, and the fractional variance is

$$\frac{\Delta n}{\bar{n}} \sim 10^{-13}$$

quite negligible.

Coherent states. The displaced state we've been studying is a special case of *coherent states*, which are defined to be eigenstates of the annihilation operator

$$\mathbf{a}|\alpha\rangle = \alpha|\alpha\rangle.$$

Notice that since $[\mathbf{a}, \mathbf{a}^\dagger] = 1$, $\mathbf{a}$ is not normal and so α can be complex and they need not be orthogonal for different α .

Also, eigenstates of the creation operator $\mathbf{a}^\dagger|b\rangle = b|b\rangle$ are boring. To see this consider expanding them in number-eigenstates: $|b\rangle = \sum_n c_n|n\rangle$, but then $\mathbf{a}^\dagger|b\rangle = \sum_{n=0} c_n \sqrt{n+1}|n+1\rangle \stackrel{!}{=} b \sum_n c_n|n\rangle$ and we learn that $c_0 = 0$ unless $b = 0$; acting again tells us $c_1 = 0$ etc.

Let's solve the coherent state definition in the number basis: $|\alpha\rangle \equiv \sum_n a_n|n\rangle$. Acting on both sides with $\mathbf{a}$ gives

$$\mathbf{a}|\alpha\rangle = \sum_n \underbrace{\mathbf{a}|n\rangle}_{=\sqrt{n}|n-1\rangle} a_n = \sum_n a_n \alpha|n-1\rangle$$

and equating the coefficients of $|n\rangle$ on both sides gives

$$a_{n+1}\sqrt{n+1} = \alpha a_n \implies a_{n+1} = \frac{\alpha a_n}{\sqrt{n+1}}$$

and therefore

$$|\alpha\rangle = a_0 \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle.$$

Normalizing requires

$$1 = \langle\alpha|\alpha\rangle = |a_0|^2 \sum_n \frac{|\alpha|^{2n}}{n!} = |a_0|^2 e^{|\alpha|^2} \implies a_0 = e^{-|\alpha|^2/2}.$$

So

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_n \frac{\alpha^n}{\sqrt{n!}} \underbrace{|n\rangle}_{=\frac{(\mathbf{a}^\dagger)^n}{\sqrt{n!}}|0\rangle} = e^{-|\alpha|^2/2} \sum_n \frac{\alpha^n (\mathbf{a}^\dagger)^n}{n!} |0\rangle = e^{-|\alpha|^2/2} e^{\alpha \mathbf{a}^\dagger} |0\rangle.$$

Note that α can be complex. The ‘displaced’ groundstate we studied above was a special case with α real, but we saw that it evolved in time to complex α . We could have directly made it complex by acting with

$$e^{\alpha \mathbf{x} + b \mathbf{p}} = e^{\beta \mathbf{a}^\dagger - \beta^* \mathbf{a}} = e^{-|\beta|^2/2} e^{\beta \mathbf{a}^\dagger} e^{-\beta^* \mathbf{a}}.$$

More generally, coherent states behave nicely under time evolution, in the sense that they remain coherent states:

$$|\alpha(t)\rangle \equiv \mathbf{U}(t) |\alpha\rangle = e^{-i\omega t/2} |\alpha e^{-i\omega t}\rangle$$

In this last expression we have the coherent state where the eigenvalue of $\mathbf{a}$ is $\alpha e^{-i\omega t}$: the value of α goes in a circle around the origin in the complex plane as time evolves.

One more comment about coherent states. Since they are eigenstates of a non-normal operator, they needn’t be orthogonal, and they aren’t:

$$|\langle\beta|\alpha\rangle|^2 = e^{-|\alpha-\beta|^2}$$

coherent states with nearby eigenvalues have more overlap. Moreover, it isn’t obvious that they form a basis, but they do. They are sometimes called *overcomplete* in the sense that

$$\mathbb{1} = \int \frac{d^2\alpha}{2\pi} |\alpha\rangle \langle\alpha|.$$

3.2.8 Shape-invariance

Solvable Schrödinger problems are rare. Here is a machine for making more. Suppose we have a 1d Hamiltonian $\mathbf{H}_- \equiv \frac{\mathbf{p}^2}{2m} + V_-(\mathbf{x})$ with a collection of boundstates $\mathbf{H}_- |u_n\rangle =$

$E_n |u_n\rangle$, $n = 0, 1, 2, \dots$, and we've subtract a constant so that $E_0 = 0$. Let's also set $\hbar = m = 1$ for this discussion. So in terms of the position-space representation $u_n(x) = \langle x | u_n \rangle$, we have

$$0 = \left(-\frac{1}{2} \partial_x^2 + V_-(x) \right) u_0(x) \quad \leftrightarrow \quad V_-(x) = \frac{u_0''}{2u_0}$$

(note that dividing by u_0 is ok everywhere, since the groundstate has no nodes!) so we can write the (position-space representation of the) hamiltonian in terms of the groundstate wavefunction as

$$H_- = \frac{1}{2} \left(-\partial_x^2 + \frac{u_0''}{u_0} \right).$$

Next, introduce the *superpotential* $W \equiv -\frac{1}{\sqrt{2}} \frac{u_0'}{u_0}$ and

$$\alpha \equiv \frac{1}{\sqrt{2}} \left(\partial_x - \frac{u_0'}{u_0} \right) = \frac{1}{\sqrt{2}} \partial_x + W, \quad \alpha^\dagger \equiv \frac{1}{\sqrt{2}} (-\partial_x) + W$$

in terms of which

$$\mathbf{H}_- = \alpha^\dagger \alpha.$$

Notice that this means that the groundstate satisfies

$$\alpha |u_0\rangle = 0 \quad \implies \quad \left(\frac{1}{\sqrt{2}} \partial_x + W \right) u_0 = 0 \quad \implies \quad u_0 \propto e^{-\sqrt{2} \int^x W(y) dy}.$$

Now you can maybe guess what's coming next, given our discussion of the index of an operator. Consider now

$$\mathbf{H}_+ \equiv \alpha \alpha^\dagger = \mathbf{H}_- - \partial_x \left(\frac{u_0'}{u_0} \right) = \frac{\mathbf{p}^2}{2} + V_+.$$

The two potentials at play here are $V_\pm = W^2 \pm \frac{1}{\sqrt{2}} \partial_x W$. Except for the zeromode, with $E_0 = 0$, all of the boundstate energies E_n of the two potentials are the same!

$$H_- u_n = E_n u_n, \quad v_n \equiv \alpha u_n \text{ satisfies } H_+ v_n = E_n v_n.$$

Similarly, $H_- (\alpha^\dagger v_n) = E_n (\alpha^\dagger v_n)$.

Here's an example. Take the infinite square well, $V(x) = \begin{cases} \infty, & |x| > \pi/2 \\ -\epsilon_0, & |x| < \pi/2. \end{cases}$ I have chosen the width for convenience, and the depth $-\epsilon_0 \equiv -\frac{1}{2}$ so that the groundstate energy is zero. So the groundstate is $u_0 \propto \cos x$ inside the well and zero for x outside

the well; we will not speak of what happens outside the well any further. So $W = -\frac{1}{\sqrt{2}} \frac{u'_0}{u_0} = \frac{1}{\sqrt{2}} \tan x$ and

$$V_+ = W^2 + \frac{1}{\sqrt{2}} W' = \frac{1}{2} (\tan^2 x + \sec^2 x) = \frac{1}{2} + \tan^2 x.$$

The expression for $\alpha = \frac{1}{\sqrt{2}} (\partial_x + \tan x)$ allows us to construct the eigenstates of H_+ .

level	u_n	E_n	$v_n = \alpha u_n$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
2	$\cos 3x$	8/2	$\sin x \cos^2 x$
1	$\sin 2x$	3/2	$\cos^2 x$
0	$\cos x$	0	—

(The wavefunctions in the table are not normalized.)

It is sometimes useful to keep going: shift the energy of H_+ so that the groundstate is at zero, and make it the new H_- .

A second example to try is the SHO:

$$V_- = \frac{1}{2} \omega^2 x^2 - \omega/2 = W^2 - \frac{1}{\sqrt{2}} \partial_x W$$

with $W = \omega x / \sqrt{2}$. The groundstate is

$$u_0(x) = c e^{-\sqrt{2} \int^x W(y) dy} = c e^{-\int^x \omega y dy} = c e^{-\omega x^2 / 2}.$$

The potential $V_+ = W^2 + \partial_x W$ is also quadratic, also an SHO. And the α operator is just the usual lowering operator. So that wasn't so useful, but at least it's self-consistent.

If you are interested in reading more about this kind of thing, look at [this review article by Cooper et al.](#)

3.2.9 Particle in a uniform force field

Consider a particle in a constant (in time and space) force field, such as gravity on the Earth, or a constant electric field. This means a linear potential $V(y) = mgy$. This is important both as a simple, real, solvable example, as well as as an ingredient in the WKB approximation. The position-space Schrödinger equation for an energy eigenstate is

$$-\frac{\hbar^2}{2m} \partial_y^2 \psi + mgy \psi = E \psi.$$

In terms of $\alpha \equiv \frac{2mE}{\hbar^2}$, $\beta \equiv \frac{2gm^2}{\hbar^2}$, $z \equiv \beta^{-2/3}(\beta y - \alpha)$, Schröd becomes

$$0 = \partial_y^2 \psi + (\alpha - \beta y) \psi \quad \leftrightarrow \quad \partial_z^2 \psi - z \psi = 0.$$

This last is the Airy differential equation. The fact that it has a person's name means that there isn't a solution in terms of elementary functions. However, we can learn quite a bit about the solutions.

To make progress, let's go to momentum space, $\phi(k) \equiv \langle k|E \rangle$. There the Schröd equation is a first order ODE, because the potential is linear in $x \sim +i\hbar\partial_k$!

$$\begin{aligned} 0 &= \langle k|(\mathbf{H} - E)|E \rangle = -\frac{\hbar^2 k^2}{2m} \phi(k) + mg(+i\hbar\partial_k) \phi(k) \\ \Rightarrow \partial_k \phi(k) &= -\frac{i}{\hbar mg} \left(E - \frac{\hbar^2 k^2}{2m} \right) \phi(k) \quad \Rightarrow \quad -\frac{i}{\hbar mg} \left(E - \frac{\hbar^2 k^2}{2m} \right) = \frac{\partial_k \phi(k)}{\phi(k)} = \partial_k \log \phi(k) \\ \Rightarrow \log \phi(k) &= -\int^k \left(dk \frac{i}{\hbar mg} \left(E - \frac{\hbar^2 k^2}{2m} \right) \right) = -\frac{i}{\hbar mg} \left(Ek - \frac{\hbar^2 k^3}{6m} \right) + c \\ \Rightarrow \phi(k) &= ae^{-\frac{i}{\hbar mg} \left(Ek - \frac{\hbar^2 k^3}{6m} \right)}. \end{aligned}$$

Then the position-space wavefunction is

$$\psi(x) = \langle x|E \rangle = \int dk e^{ikx} \phi(k) = a \int dk e^{i\left(x - \frac{E}{\hbar mg}\right)k + i\frac{\hbar k^3}{6m^2 g}}.$$

This is the integral representation of the Airy function. You can check by direct differentiation that ∂_x^2 acting on it gives something proportional to x plus a constant.

12

[End of Lecture 12]

Here's one application of this integral representation: an integral of this form $\int dk e^{iS(k)}$ can be approximated by *stationary phase* (or *saddle point*) if the function $S(k)$ is rapidly varying. This happens when the argument $|x|$ is large. The approximation proceeds by Taylor expanding $S(k)$ about its extremum: $S(k) = S(k_0) + S''(k_0)(k -$

¹²Actually, here's a better way to do it: since there's no boundary to the region of integration, we can write

$$0 = \int dk \frac{\partial}{\partial k} \left(e^{iS(k)} \right)$$

with $S(k) \equiv k \left(x - \frac{E}{\hbar mg} \right) + \frac{\hbar k^3}{6m^2 g}$. This gives exactly the Schrödinger equation:

$$0 = \int dk \partial_k S \left(e^{iS(k)} \right) = \int dk \left(x - \frac{E}{\hbar mg} + \frac{\hbar k^2}{2m^2 g} \right) e^{ikx} \phi(k) = \frac{1}{mg} \left(mgx - E + \frac{\mathbf{p}^2}{2m} \right) \psi(x).$$

(Generalizations of this trick are called *Schwinger-Dyson equations* in the context of quantum field theory.)

$k_0) + \dots$. If we keep only these terms, the k integral becomes Gaussian. The leading approximation is just $\int dk e^{iS(k)} \sim e^{iS(k_0)}|_{S'(k_0)=0}$. For the Airy integral,

$$\text{Ai}(x) = \int_{-\infty}^{\infty} dk e^{iS(k)}, \quad S(k) = \frac{k^3}{3} + ky$$

this gives $k_0^2 = -y$ and

$$\text{Ai}(x) \begin{cases} x \rightarrow +\infty & e^{-\frac{2}{3}x^{3/2}} \\ x \rightarrow -\infty & e^{\pm i(\frac{2}{3}x^{3/2})}. \end{cases}$$

A more careful accounting of the prefactors (in particular including the Gaussian integral bit) gives

$$\text{Ai}(x) \begin{cases} x \rightarrow +\infty & \frac{1}{2\sqrt{\pi x^{1/4}}} e^{-\frac{2}{3}x^{3/2}} \\ x \rightarrow -\infty & \frac{1}{\sqrt{\pi x^{1/4}}} \sin\left(\frac{2}{3}x^{3/2}\right). \end{cases}$$

It is sometimes useful to note that the Airy function can be written in terms of Bessel functions

$$\text{Ai}(x) = \frac{\sqrt{x}}{3} \left(I_{-1/3} \left(\frac{2}{3}x^{3/2} \right) - I_{1/3} \left(\frac{2}{3}x^{3/2} \right) \right) \quad (\text{for } x > 0).$$

(For negative argument, we must replace BesselI with BesselJ (in Mathematica's notation), where $I_\nu(z) = e^{-i\pi\nu/2} J_\nu(iz)$. The information encoded in this extra phase will be important below.)

So the time-evolution of an energy eigenstate for a particle in a constant force field is

$$\psi(y, t) = \text{Ai} \left(\left(\frac{2m^2g}{\hbar^2} \right)^{1/3} \left(y - \frac{E}{mg} \right) \right) e^{-iEt/\hbar}. \quad (3.30)$$

The Einstein equivalence principle says that in a frame with constant acceleration g , we can get rid of the uniform force field. This frame has $y' = y - y_0 + \frac{1}{2}gt^2$, $t' = t$. In this frame, the (time-dependent) Schrödinger equation is

$$\left(-\frac{\hbar^2}{2m} \partial_{y'}^2 + mg \left(y' + y_0 - \frac{1}{2}gt^2 \right) \right) \psi = i\hbar (g + \partial_{y'} + \partial_t) \psi.$$

Since we expect the particle in this frame to feel no force, we try an ansatz involving the free wave equation in y' :

$$\psi(y', t) = u(y', t) e^{i\lambda(y', t)}, \quad \frac{-\hbar^2}{2m} \partial_{y'}^2 u = i\hbar \partial_t u.$$

Setting $y_0 = E/mg$, the solution of this equation is (this takes some work to check)

$$u(y', t) = \text{Ai} \left(\beta^{1/3} (y' - gt^2/2) \right) e^{i \left(\frac{mgt}{\hbar} y' - \frac{m^2 g^2 t^3}{3\hbar} \right)}.$$

In frame F , we found a solution with definite E which was stationary, ρ a constant in time, and $\vec{j} = 0$. In frame F' , the classical particle is stationary, but the solution we find describes an accelerating wavepacket. The resolution of this apparent paradox is that the wavepacket in F which is approximated by a classical particle is a superposition of many energy eigenstates, $\psi^F = \sum_E c_E \psi_E$. And furthermore, ψ^F boosts to a square-integrable, non-accelerating packet in F' , which disperses in time (like $\sqrt{t}$), unlike the solution u above.

3.3 WKB approximation

When we first discussed a potential problem, we described the physics in terms of

$$k(x) \equiv \sqrt{\frac{2m}{\hbar} (E - V(x))}$$

and noted that if k were constant, then we have plane waves or exponentials. When $k(x)$ varies slowly, this is the basis for a systematic approximation. Slowly compared to what? The condition with the right units is $\partial_x V \ll k(x)$; another way to state it is: the potential varies little over one wavelength of the de Broglie wave, $\lambda = \frac{2\pi}{k}$.

In terms of $k(x)$, the Schrödinger equation is

$$\partial_x^2 \psi(x) + k(x)^2 \psi(x) = 0.$$

If we write $\psi(x) \equiv e^{is(x)/\hbar}$ this is

$$\frac{\mathbf{i}}{\hbar} s'' - \frac{(s')^2}{\hbar^2} + k^2(x) = 0.$$

For constant k , a solution is $s' = \pm \hbar k$ and $s'' = 0$, and $\psi = e^{\pm \mathbf{i} k x}$. For slowly-varying k , we can treat the first $s''/\hbar$ term as a perturbation. Heuristically, think of $\hbar$ as a small quantity, so that the term with two powers of $\hbar$ in the denominator is bigger. So $s' \simeq \pm \hbar k$ to leading order, so that $s'' \simeq \pm \hbar k'$ to leading order; plugging back gives

$$\begin{aligned} \frac{(s')^2}{\hbar^2} &= k(x)^2 \pm \mathbf{i} k'(x) = k^2 \left(1 \pm \frac{\mathbf{i} k'}{k^2} \right) \\ \implies s' &= \pm \hbar k \sqrt{1 \pm \mathbf{i} \frac{k'}{k^2}} \simeq \pm \hbar k + \mathbf{i} \frac{\hbar k'}{2k} \\ \implies \frac{\mathbf{i} s(x)}{\hbar} &\simeq \pm \mathbf{i} \int^x k(y) dy - \frac{1}{2} \ln k(x). \end{aligned}$$

Here we've ignored terms which would come with more powers of $\hbar$. In this approximation, the two solutions are

$$\psi_{\pm}(x) \simeq \frac{A_{\pm}}{|2m(E - V(x))|^{1/4}} e^{\pm i \int^x \sqrt{2m(E - V(y))} dy}.$$

It is possible to iterate this procedure further, expanding $s = s_0 + \hbar s_1 + \hbar^2 s_2 + \dots$, but I don't know an example where it is useful to do so.

We can learn some thing already from this expression. In this approximation, assuming we are in the allowed region, the probability density is $\rho(x) = \psi_{\text{WKB}}(x)^2 \propto \frac{1}{|p(x)|}$. What's this? Consider a classical particle in a bound orbit of some potential with period T and energy $E = \frac{p^2}{2m} + V(x)$. Let $\rho_{\text{cl}}(x)dx \equiv$ the fraction of time it spends in the interval $(x, x + dx)$. This is

$$\rho_{\text{cl}}(x)dx = \frac{2}{T}dt$$

where where the 2 is because it passes through the region twice a period and dt is the time it takes to pass through the interval on each pass. Then

$$\rho_{\text{cl}}(x)dx = \frac{2}{T}dt = \frac{2}{T}dx \frac{dt}{dx} = dx \frac{2}{T} \frac{1}{v(x)} = dx \frac{2m}{T} \frac{1}{p(x)}.$$

So the leading WKB approximation reproduces the density you would get if you sent a swarm of classical particles and took some long-exposure photograph. It peaks at the turning points because the particles have to stop to turn around.

What is the approximation we are making here? To see this more clearly, it's useful to think briefly about $d > 1$. Then we put arrows on stuff in the Schrödinger equation

$$-\nabla^2 \psi(\vec{x}) + \sqrt{\frac{2m}{\hbar^2}} (V(\vec{x}) - E) \psi(\vec{x}) = 0$$

and the same redefinition $\psi(\vec{x}) \equiv e^{\frac{i}{\hbar}s(\vec{x})}$ gives

$$0 = +\frac{i}{\hbar} \nabla^2 s - \frac{1}{\hbar^2} \vec{\nabla} s \cdot \vec{\nabla} s + \sqrt{\frac{2m}{\hbar^2}} (V(\vec{x}) - E).$$

A plane wave $\psi(\vec{x}) = e^{i\vec{s} \cdot \vec{x}} = e^{i\vec{p} \cdot \vec{x}}$ is called this because the wave-fronts, the surfaces of constant phase, are planes, $\vec{p} \cdot \vec{x} = \text{const}$. $\vec{p}$ is the normal to these planes. The WKB approximation rests on the idea that the surfaces of constant phase are nearly planes, are varying slowly, at the same length scale as the potential. $\vec{p} \equiv \vec{\nabla} s(x)$ is now the normal to the plane at $\vec{x}$. (Notice that in $d > 1$ this definition replaces the less general expression in terms of the sqrt of the potential.) Two wavefronts whose phase differs by a whole period $\Delta s = 2\pi\hbar$, are separated by a wavelength, $\lambda = \frac{2\pi\hbar}{|\vec{\nabla} s|} = \frac{2\pi\hbar}{|p|}$. The condition for validity is that the potential can't vary too much over this distance.

Back to the 1d case, two more steps before we are done:

- 1: To specify the correct linear combination of $\psi_{\pm}$ in a given problem, we need to impose boundary conditions.
- 2: The approximation fails near *turning points*, where $k(x) = 0$, the boundaries between classically allowed and forbidden regions.

To see how these problems solve each other, consider the case where the potential rises to the left, *i.e.* $\psi(x \rightarrow -\infty) = 0$. Choose an x coordinate to put the turning point at $x = 0$. Then the unique solution for $x = x_0 > 0$, in WKB approximation, is

$$\psi(x_0) = \frac{A}{\sqrt{p}} \sin \left(\frac{1}{\hbar} \int_0^{x_0} dx p(x) + \alpha \right) \quad (3.31)$$

where $p(x) \equiv \sqrt{2m(E - V(x))}$, and $\alpha \in \mathbb{R}$ is to be determined. How? Compare (3.31) with the solution for a linear potential in the neighborhood of the turning point:

$$V(x) \simeq V_0(x) = E + \partial_x V|_{x=0} x \equiv E - F_0 x$$

this describes a uniform force $F_0 \equiv \partial_x V|_{x=0}$. More specifically, we will compare with the *asymptotic* solution of this problem at x_0 big enough that the asymptotic behavior is reached, but small enough that $V(x)$ is still well-approximated by the first two terms of its Taylor expansion. This will only work if F_0 is small enough compared to higher-order terms in the Taylor expansion of V .

V_0 : $0 = \psi'' + \frac{2m}{\hbar^2} (E - V_0(x)) \psi = \psi'' + \frac{2m}{\hbar^2} F_0 x \psi$ In terms of $z \equiv (2mF_0/\hbar^2)^{1/3} x \equiv \beta^{1/3} x$ this is Airy equation: $\partial_z^2 \psi + z \psi = 0$. The asymptotic solution for $z \gg 0$ is

$$\psi = \frac{c}{z^{1/4}} \sin \left(\frac{2}{3} z^{3/2} + \frac{\pi}{4} \right)$$

Setting this equal to (3.31) and noticing that

$$\frac{1}{\hbar} \int_0^x dy \sqrt{2m(E - V_0(y))} = \frac{2}{3} z^{3/2}$$

we learn that $\alpha = \pi/4$.

Similarly, if the potential rises to the right, the ‘connection’ formulae, for $x_1 < 0 < x_2$, are

$$\begin{aligned} \psi(x_1) &= \frac{A}{\sqrt{p(x_1)}} \sin \left(\frac{1}{\hbar} \int_{x_1}^0 dx p(x) + \frac{\pi}{4} \right) \\ \psi(x_2) &= \frac{A}{2\sqrt{|p(x_2)|}} e^{-\frac{1}{\hbar} \int_0^{x_2} dx |p(x)|} \end{aligned}$$

where, crucially, it is the same A in both of these formulae.

Application 1: (Bohr-Sommerfeld quantization.) The WKB approximation then implies the old-school Bohr-Sommerfeld quantization formula for bound orbits, since if we have a potential which rises on both left and right then all the above formulae apply. Take the case where the two turning points are $x_- < x_+$. Then for $x \in (x_-, x_+)$, both formula can only be true with real coefficients A, A' if

$$\psi(x) = \frac{A}{\sqrt{p(x)}} \sin \left(\frac{1}{\hbar} \int_{x_-}^x dy p(y) + \frac{\pi}{4} \right) = \frac{A'}{\sqrt{p(x)}} \sin \left(\frac{1}{\hbar} \int_x^{x_+} dy p(y) + \frac{\pi}{4} \right)$$

where, since $\int_x^{x_+} = \int_{x_-}^{x_+} - \int_{x_-}^x$, we have

$$\int_{x_-}^{x_+} dy p(y) = (m + \frac{1}{2})\pi\hbar, \quad m = 0, 1, 2, \dots, \quad A' = (-1)^m A.$$

The LHS here is related to the area enclosed by the orbit in phase space, by $\int_{x_-}^{x_+} dy p(y) = \frac{1}{2} \oint p(y) dy$. So there is some truth to the idea that areas of phase space $\oint p(x) dx$ come in Planck-sized chunks $h = 2\pi\hbar$.

[\[End of Lecture 13\]](#)

Application 2: (Potential barrier penetration by tunneling.) Assume that the potential looks like a big hump separating two allowed regions. Assuming that the rate for penetrating the barrier is small, we can calculate it.

A wave incident from the left in I (or from the right in II) is nearly perfectly reflected, hence nearly a standing wave, *e.g.*

$$v_{II}(x > x_2) = \frac{1}{\sqrt{p}} \sin \left(\frac{1}{\hbar} \int_{x_2}^{x > x_2} dy p(y) + \frac{\pi}{4} \right)$$

where we chose a normalization.

The transmitted wave in II is

$$u_{II}(x > x_2) = \frac{C}{\sqrt{p}} e^{\frac{i}{\hbar} \int_{x_2}^{x > x_2} dy p(y) - i\pi/4}$$

where since C is complex, the phase in the exponent is just a convenience.

Since these are two (approximate) solutions of the same Schrödinger problem, their Wronskian is a constant

$$W(u_{II}, v_{II}) = uv' - vu' = -\frac{iC}{\hbar}.$$

The connection formula connects v_{II} to

$$v_{III} = \frac{1}{2\sqrt{|p|}} e^{-\frac{1}{\hbar} \int_{x < x_2}^{x_2} |p(y)| dy}$$

– this solution decreases as x decreases. The other linearly-independent solution in (the forbidden) region III is

$$u_{III} = \frac{C'}{\sqrt{|p|}} e^{+\frac{1}{\hbar} \int_{x < x_2}^{x_2} |p(y)| dy}$$

which increases as x decreases. Their Wronskian should have the same value as the Wronskian for region II:

$$W(u_{III}, v_{III}) = \frac{C'}{\hbar} \stackrel{!}{=} W(u_{II}, v_{II}) \quad \Longrightarrow \quad C' = -\mathbf{i}C.$$

The standing wave on the left connects via

$$u_I = \frac{A}{\sqrt{p}} \sin \left(\frac{1}{\hbar} \int_{x < x_1}^{x_1} dy p(y) + \frac{\pi}{4} \right) \rightsquigarrow u_{III} = \frac{A}{2\sqrt{|p|}} e^{-\frac{1}{\hbar} \int_{x_1}^{x > x_1} |p(y)| dy}$$

which using $\int_{x_1}^{x > x_1} = \int_{x_1}^{x_2} - \int_{x < x_2}^{x_2}$ gives

$$C = \mathbf{i} \frac{A}{2} e^{-\frac{1}{\hbar} \int_{x_1}^{x_2} |p(y)| dy}.$$

Hence, the transmission probability is

$$\left| \frac{C}{A} \right|^2 = e^{-2 \frac{1}{\hbar} \int_{x_1}^{x_2} |p(y)| dy}.$$

3.4 Wave mechanics of charged particles

When we discussed the question of choosing a Hamiltonian for our quantum particle, I mentioned the possibility of considering a charged particle, and putting it in an electromagnetic field $(\vec{E}, \vec{B})$. In order to do so, we must introduce the electromagnetic potentials, $\Phi, \vec{A}$, such that

$$\vec{E} = -\vec{\nabla}\Phi - \frac{1}{c}\partial_t\vec{A}, \quad \vec{B} = \vec{\nabla} \times \vec{A}.$$

You will recall that the form of E, B is preserved by a *gauge transformation*,

$$\vec{A} \rightarrow \vec{A} - \vec{\nabla}\lambda, \quad \Phi \rightarrow \Phi - \frac{1}{c}\partial_t\lambda. \quad (3.32)$$

These expressions are nicely packaged in terms of the four-vector potential $A_\mu \equiv (\vec{A}, c\Phi)_\mu \rightarrow A_\mu - \partial_\mu\lambda$, but will not consider any relativistic situation here. This ambiguity may make it seem like the vector potential is not real. Some of the vector potential is indeed an artifice, but notice that $\oint_C \vec{A} \cdot d\vec{x}$ is not changed by the gauge transformation (3.32). We will now see how this quantity arises physically, because of quantum mechanics.

It is useful to think also about the Lagrangian description of the charged particle, which has

$$L = \frac{1}{2}m\dot{x}^2 - e\Phi - \frac{e}{c}\dot{\vec{x}} \cdot \vec{A}.$$

Notice that the extra term in the action $S = \int dt L$ from the vector potential is *geometric*, in the sense that $\int dt \dot{\vec{x}} \cdot \vec{A} = \int d\vec{x} \cdot \vec{A}$ depends only on the path of the particle, not on how quickly it's traversed.

The canonical momentum deduced from the Lagrangian L is

$$p = \frac{\partial L}{\partial \dot{x}} = m\dot{x} - \frac{e}{c}A$$

and now differs from the *mechanical momentum* $m\dot{x}$. This gives back the Hamiltonian I wrote above:

$$H = \frac{1}{2m} \left(p - \frac{e}{c}A \right)^2 + e\Phi.$$

Let's consider time-independent EM fields for simplicity. The Schrödinger equation in position space is then

$$-\mathbf{i}\hbar\partial_t\psi(x, t) = \langle x | \mathbf{H} | \psi(t) \rangle = \frac{1}{2m} \left(-\mathbf{i}\hbar\vec{\nabla} - \frac{e}{c}\vec{A}(x) \right)^2 \psi(x, t) + e\Phi(x)\psi(x, t). \quad (3.33)$$

The probability density is still $\rho(x, t) = |\psi(x, t)|^2$. Is it still conserved? The calculation we did above shows that $\partial_t\rho + \vec{\nabla} \cdot \vec{j}_A = 0$ but where

$$\vec{j}_A \equiv \frac{\hbar}{2m\mathbf{i}} \left(\psi^*\vec{\nabla}\psi - \psi\vec{\nabla}\psi^* \right) - \frac{e}{mc}\psi^*\psi\vec{A}$$

has an extra term. This depends very explicitly on $\vec{A}$! Is it gauge invariant? Only if the gauge transformation acts on the wavefunction:

$$\vec{A} \rightarrow \vec{A} + \vec{\nabla}\lambda, \quad \psi \rightarrow e^{-i\frac{e\lambda}{\hbar c}}\psi.$$

Now we turn to showing that the vector potential (or at least its integral around closed curves) is real, via the Aharonov-Bohm effect. Two bits of warmup.

Here is a useful mathematical nugget: Suppose we know the solution $u_0(t)$ in $i\hbar\partial_t u_0 = H_0 u_0$ for some H_0 . Suppose we want to solve $i\hbar\dot{\psi} = H\psi$ with $H \equiv H_0 + V(t)$, where $V(t)$ is a c-number function when acting on u_0 . Then let $\psi = f(t)u_0$ and

$$\begin{aligned} H u_0 f = i\hbar \left(u_0 \dot{f} + \dot{u}_0 f \right) &\implies \frac{\dot{f}}{f} = -\frac{i}{\hbar} V(t) \\ \implies f(t) &= f(0) e^{-\frac{i}{\hbar} \int_0^t V(s) ds}. \end{aligned} \tag{3.34}$$

Next, let's consider a double-slit experiment, where a particle may follow two possible paths from a source to a detector, which we'll call u(pper) and l(ower). The intensity at position x on the detector is

$$I_x = |\psi_u(x) + \psi_l(x)|^2$$

where $\psi_{u/l}(x)$ is the amplitude for the case where the l/u path is blocked.

Now here's the real thing. Consider a charged particle traversing the same double-slit experiment, where now the (completely impenetrable) wall between the slits contains a (perfect) solenoid, with

$$\vec{B} = \begin{cases} B\hat{z}, & \text{inside} \\ 0, & \text{outside} \end{cases}.$$

The Schrödinger equation is

$$-i\hbar\partial_t\psi(x,t) = \frac{1}{2m} \left(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}(x) \right)^2 \psi(x,t) + V(x)\psi(x,t) \tag{3.35}$$

where V includes the walls of the barrier, and $\vec{\nabla} \times \vec{A} = \vec{B}$. Notice that the wavefunction vanishes wherever the magnetic field is nonzero (since the walls are impenetrable ($V = \infty$ inside the walls)) and the magnetic field vanishes wherever the wavefunction is nonzero (since the solenoid is perfect ($B = 0$ outside the solenoid)).

Now, suppose we know the solution u_0 for $\vec{A} = 0$:

$$-i\hbar\partial_t u_0(x,t) = \frac{1}{2m} \left(-i\hbar\vec{\nabla} + 0 \right)^2 u_0(x,t) + V(x)u_0(x,t). \tag{3.36}$$

This might be quite difficult to find, but we don't need to find it. I claim that the solution to (3.35) is

$$\psi = e^{\mathbf{i}\frac{e}{\hbar c} \int_{x_0}^x d\vec{\ell} \vec{A}(\ell)} u_0 \equiv e^g u_0.$$

Notice the similarity with (3.34). To check this, let's differentiate:

$$\left(-\mathbf{i}\hbar\vec{\nabla} - \frac{e}{c}\vec{A}\right) \psi(x) \stackrel{\text{FTC}}{=} -\mathbf{i}\hbar\left(\vec{\nabla} u_0\right) e^g - \mathbf{i}\hbar u_0 \left(\frac{\mathbf{i}e}{\hbar c}\vec{A}\right) e^g - \frac{e}{c}\vec{A} u_0 e^g.$$

Doing it again:

$$\frac{1}{2m} \left(-\mathbf{i}\hbar\vec{\nabla} - \frac{e}{c}\vec{A}\right)^2 (u_0 e^g) = \frac{1}{2m} \left(\left(-\mathbf{i}\hbar\vec{\nabla}\right)^2 u_0\right) e^g.$$

Notice that we didn't have to choose a path in $\int^x A$.

The combination $D \equiv \left(-\mathbf{i}\hbar\vec{\nabla} - \frac{e}{c}\vec{A}\right)$ is sometimes called a *covariant derivative*, because it preserves the gauge transformation properties of the function it acts on:

$$\psi \rightarrow e^{\mathbf{i}\lambda} \psi \implies (D\psi) \rightarrow e^{\mathbf{i}\lambda} (D\psi).$$

3.4.1 Aharonov-Bohm effect

Without $\vec{B}$, the probability amplitude at the location x on the detector is $u_0^u(x) + u_0^l(x)$. With $\vec{B}$, it is

$$\begin{aligned} \psi(x) &= \psi^u(x) + \psi^l(x) \\ &= u_0^u(x) e^{\mathbf{i}\frac{e}{\hbar c} \int_{C_u} \vec{A}(\ell) \cdot d\vec{\ell}} + u_0^l(x) e^{\mathbf{i}\frac{e}{\hbar c} \int_{C_l} \vec{A}(\ell) \cdot d\vec{\ell}} \\ &= e^{\mathbf{i}\frac{e}{\hbar c} \int_{C_u} \vec{A}(\ell) \cdot d\vec{\ell}} \left(u_0^u(x) + u_0^l(x) e^{\mathbf{i}\frac{e}{\hbar c} \oint_{C_l - C_u} \vec{A}(\ell) \cdot d\vec{\ell}} \right). \end{aligned} \quad (3.37)$$

The relative phase between the two terms is changed by the integral of $\vec{A}$ over a *closed* path:

$$\frac{e}{\hbar c} \oint_{C_l - C_u} \vec{A}(\ell) \cdot d\vec{\ell} \stackrel{\text{Stokes}}{=} \frac{e}{\hbar c} \int_S \underbrace{\vec{\nabla} \times \vec{A}}_{=\vec{B}} \cdot \underbrace{d\vec{a}}_{\equiv \vec{n} d^2x} \equiv \frac{e}{\hbar c} \Phi_B(S).$$

Here S is any surface bounded by $\partial S = C_l - C_u$, and the magnetic flux through the surface is $\Phi_B(S) = BA$ in this case.

Notice that Φ_B is gauge invariant. And this is good because the measured result depends on it: The probability of seeing the particle hit the location x is

$$P_x = |\psi(x)|^2 = \left| u_0^u + u_0^l e^{\mathbf{i}\frac{e}{\hbar c} \Phi_B} \right|^2$$

– if we turn the knob controlling the current through the solenoid, the interference pattern will shift! (Notice that the pattern returns to itself when $\frac{e}{\hbar c}\Phi_B$ changes by 2π .)

The observability of this kind of phase was used brilliantly by Dirac to show that if magnetic monopoles exist with magnetic charge g , then the combination eg is quantized.

3.4.2 Magnetic monopoles and quantization of charge and angular momentum

[from Commins] A magnetic monopole is a point source of magnetic field, around which

$$\vec{B} = \frac{g}{r^2} \vec{r}$$

where g is the magnetic charge, $\oint \vec{B} \cdot d\vec{a} = 4\pi g$. We can learn quite a lot by supposing one of these exists.

Consider also an electric charge, separated by $\vec{a}$ from the origin where the monopole sits, so

$$\vec{E} = \frac{e}{R^2} \vec{R}, \quad \vec{R} \equiv \vec{r} - \vec{a}.$$

If neither moves, there is no force on either. But notice that $\vec{E}$ is not parallel to $\vec{B}$, so there is energy and momentum in the fields; the momentum density (in cgs) is

$$\begin{aligned} \vec{\wp} &= \frac{1}{4\pi c} \vec{E} \times \vec{B} \\ &= \frac{1}{4\pi c} \frac{ge}{R^3 r^3} \vec{r} \times \underbrace{\vec{R}}_{\vec{r}-\vec{a}} \\ &= -\frac{1}{4\pi c} \frac{ge}{R^3 r^3} \vec{r} \times \vec{a} \end{aligned} \tag{3.38}$$

The angular momentum density about the origin is

$$\vec{j} = \vec{r} \times \vec{\wp} = \frac{1}{4\pi c} \frac{ge}{r^3 R^3} \vec{r} \times (\vec{a} \times \vec{r}).$$

The total angular momentum is

$$\vec{S} = \int \vec{j} d^3v ;$$

this must be interpreted as the spin of the charge-monopole pair.

[End of Lecture 14]

What is $\vec{S}$? It must point along $\vec{a}$, since there is azimuthal symmetry about the $\vec{a}$ axis. And if $\vec{S}$ points along $\vec{a}$, then a configuration where we put charges $e/2$ at

$\vec{a}$ and $-\vec{a}$ must have zero angular momentum. Starting from that configuration, we can find $\vec{S}$ by integrating the torque $\vec{\tau} = \partial_t \vec{S}$ necessary to put the charges together: $\vec{S} = \Delta \vec{S} = \int dt \vec{\tau}$. Move the charge along an arc of radius $|\vec{a}|$. The Lorentz force on the charge is $\vec{F} = \frac{e}{2} \frac{\vec{v}}{c} \times \vec{B} = \frac{e}{2c} v B \hat{z}$ (where $\hat{z}$ is out of the board). We must exert $-\vec{F}$ to keep the charge moving along the circle (other forces such as Coulomb repulsion don't contribute to the torque along $\vec{a}$). The component of the torque along $\vec{a}$ is

$$\vec{\tau} \cdot \vec{a} = \left(\vec{r} \times \vec{F} \right) \cdot \vec{a} = a \sin \theta \frac{v B e}{c} \frac{1}{2}.$$

Along the arc, $v dt = d\ell = a d\theta$, so $dt = \frac{a d\theta}{v}$ and

$$S \equiv |\vec{S}| = \int dt \tau = \int_0^\pi \frac{d\theta a^2}{v} \frac{v B e}{c} \frac{1}{2} = a^2 \frac{g}{a^2} \frac{e}{2c} \underbrace{\int_0^\pi \sin \theta d\theta}_{=2} = \frac{eg}{c}.$$

I claim that angular momentum is quantized in units of $\frac{\hbar}{2}$. This means that $S = \frac{eg}{c} = \frac{\hbar}{2} n, n \in \mathbb{Z}$. And it means that electric charges only come in integer multiples of $\frac{\hbar c}{g}$. We'll say more about angular momentum quantization anon, but here is some evidence.

Since $\vec{\nabla} \cdot \vec{B} = 0$ away from $r = 0$, we can still find a vector potential, though it will be singular at $r = 0$.

$$\begin{aligned} \vec{B} = \frac{g}{r^2} \vec{r} &= \vec{\nabla} \times \vec{A} = \frac{1}{r \sin \theta} (\partial_\theta (\sin \theta A_\varphi) - \partial_\varphi A_\theta) \vec{r} + \dots \\ \implies \frac{g}{r^2} &= \frac{1}{r \sin \theta} (\partial_\theta (\sin \theta A_\varphi) - \partial_\varphi A_\theta) \end{aligned}$$

We can set $A_\theta = 0$ by azimuthal symmetry, so we need

$$\frac{g}{r} \sin \theta = \partial_\theta (\sin \theta A_\varphi(r, \theta))$$

Integrating gives

$$-g \frac{\cos \theta}{r} + \frac{K(r)g}{r} = \sin \theta A_\varphi \implies A_\varphi = \frac{g}{r \sin \theta} (K(r) - \cos \theta).$$

(Vanishing of the other components of $\vec{\nabla} \times \vec{A}$ will tell us that $K(r) = k$ is constant.)

Even for $r \neq 0$, this solution A_φ will be singular when $\sin \theta = 0$, *i.e.* at the poles $\theta = 0, \pi$. We can avoid this by choosing k :

$$k = +1 : A_\varphi^{(+)} = \frac{g}{r} \left(\frac{1 - \cos \theta}{\sin \theta} \right) = \frac{g}{r} \frac{2 \sin^2 \frac{\theta}{2}}{2 \sin \frac{\theta}{2} \cos \frac{\theta}{2}} = \frac{g}{r} \tan \frac{\theta}{2}$$

is singular at $\theta = \pi$, but not $\theta = 0$.

$$k = -1 : A_{\phi}^{(-)} = \frac{g}{r} \left(\frac{-1 - \cos \theta}{\sin \theta} \right) = -\frac{g}{r} \frac{2 \cos^2 \frac{\theta}{2}}{2 \sin \frac{\theta}{2} \cos \frac{\theta}{2}} = \frac{g}{r} \cot \frac{\theta}{2}$$

is singular at $\theta = 0$, but not $\theta = \pi$. So we can make a vector potential which is nonsingular everywhere (except the origin) by patching together solutions: use $k = 1$ in the northern hemisphere, and $k = -1$ in the southern hemisphere.

For this to work, answers for physics at the equator must agree. (Note that the two solutions are related by a gauge transformation $A^{(+)} = A^{(-)} + d\lambda$ where the gauge function λ is well-defined around the equator.) Let's do the A-B experiment, sending two charged-particle beams in circular paths around the equator of a sphere surrounding a magnetic monopole. The relative phase shift between the two paths is $e^{i\delta^{\pm}}$ with

$$\delta^{\pm} = \frac{e}{\hbar c} \oint \vec{A}^{(\pm)} \cdot d\vec{\ell} = \frac{e}{\hbar c} \cdot 2\pi r \cdot \frac{\pm g}{r} = \pm \frac{2\pi e g}{\hbar c},$$

where the $\pm$ indicates whether we use the northern gauge or the southern gauge. The difference is unobservable only if it's an integer multiple of 2π :

$$\frac{2\pi e g}{\hbar c} = 2\pi n - \frac{2\pi e g}{\hbar c} \quad \leftrightarrow \quad \frac{e g}{\hbar c} 4\pi = 2\pi n \quad \leftrightarrow \quad e g = \frac{\hbar c}{2} n$$

where $n \in \mathbb{Z}$. This is the condition for angular momentum quantization.

3.5 Path integrals, briefly

Some of the things I've said above are unified by the description of quantum mechanics in terms of path integral. It is the quantum mechanical analog of Lagrangian mechanics. It describes, for example, the propagator, as a sum of phases, each associated with a possible path. The phase for a path is $e^{\frac{i}{\hbar}S[\text{path}]}$.

The derivation of this statement is straightforward, in that it involves inserting lots of resolutions of the identity, but it is a little annoying and takes most of a lecture. Here are some payoffs:

- Classical mechanics is the stationary phase approximation.
- The fact that the vector potential term in the action is geometric makes the A-B effect very natural.
- It gives a straightforward way to derive the propagator for the harmonic oscillator.
- In quantum systems with many degrees of freedom, such as quantum field theories, the path integral is more useful, especially as an organizing device.
- Path integrals, with fewer factors of i , are also useful in statistical physics. The derivation from quantum mechanics gives a useful perspective on them.

Propagator. Recall that the time evolution operator is the unitary $\mathbf{U}(t_b, t_a)$ which satisfies $\mathbf{U}(t_a, t_a) = \mathbb{1}$ and

$$i\hbar\partial_{t_b}\mathbf{U}(t_b, t_a) = \mathbf{H}(t_b)\mathbf{U}(t_b, t_a).$$

If $\mathbf{H}(t) = \mathbf{H}$ is time-independent, then

$$\mathbf{U}(t_b, t_a) = \mathbf{U}(t_b - t_a) = e^{-i\mathbf{H}(t_b - t_a)/\hbar}.$$

(More generally, we need to specify an ordering in the exponential.) The (position-space) propagator is its matrix elements in position-space:

$$K(x, t; x_a, t_a) = \langle x | \mathbf{U}(t, t_a) | x_a \rangle,$$

the amplitude to propagate from x_a and time t_a to x at time t . It satisfies the Schrödinger equation:

$$i\hbar\partial_t K(x, t; x_a, t_a) = \langle x | i\hbar\partial_t \mathbf{U}(t - t_a) | x_a \rangle = \langle x | \mathbf{H} \mathbf{U}(t - t_a) | x_a \rangle.$$

For example, for a point particle, with $\mathbf{H}(\mathbf{x}, \mathbf{p}) \stackrel{e.g.}{=} \frac{\mathbf{p}^2}{2m} + V(\mathbf{x})$, we may replace $\langle x | \mathbf{H}(\mathbf{x}, \mathbf{p}) = \langle x | H(x, -i\hbar\partial_x)$, so

$$i\hbar\partial_t K(x, t; x_a, t_a) = H(x, -i\hbar\partial_x) K(x, t; x_a, t_a).$$

The initial condition on this differential equation is

$$K(x, t_a; x_a, t_a) = \langle x | \mathbf{U}(t_a, t_a) | x_a \rangle = \delta(x - x_a).$$

For a free particle, $V = 0$, we solved this above by Fourier transform:

$$i\hbar\partial_t K(x, t; x_a, t_a) = -\frac{\hbar^2\partial_x^2}{2m} K(x, t; x_a, t_a) \implies K(x, t; x_a, t_a) = \sqrt{\frac{m}{2\pi i t}} e^{\frac{im(x-x_a)^2}{2t}}$$

(I set $\hbar = 1$ in the last expression.)

Since waiting a bit and waiting some more is the same as waiting a lot, the time evolution operators satisfy a nice composition property:

$$\mathbf{U}(t_b, t_a) = \mathbf{U}(t_b, t) \mathbf{U}(t, t_a)$$

for any t (which need not even be in between t_b, t_a). If $\mathbf{H}$ is time independent, this just comes from adding the exponents. But we can stick a resolution of the identity in between the two factors on the RHS:

$$\mathbf{U}(t_b, t_a) = \int dx \mathbf{U}(t_b, t) |x\rangle \langle x| \mathbf{U}(t, t_a)$$

and make a sandwich of the whole thing:

$$K(x_b, t_b; x_a, t_a) = \int dx K(x_b, t_b; x, t) K(x, t; x_a, t_a).$$

This is like sticking a wall full of holes along the path of the particle.

Path integral representation of propagator. Now do this $N - 1$ times: divide the interval into N equal segments of (tiny) size $\Delta t = \frac{t_b - t_a}{N}$:

$$K(x, t; x_a, t_a) = \int dx_1 \cdots dx_{N-1} K(x_b, t_b; x_{N-1}, t_{N-1}) \cdots K(x_1, t_1; x_a, t_a).$$

Writing $\mathbf{H} = \mathbf{T}(\mathbf{p}) + \mathbf{V}(\mathbf{x})$, the propagator for a tiny step is

$$\begin{aligned} K(x_j, t_j; x_{j-1}, t_{j-1}) &= \langle x_j | e^{-i\Delta t \mathbf{T} - i\Delta t \mathbf{V}} | x_{j-1} \rangle \\ &= \langle x_j | e^{-i\Delta t \mathbf{T}} e^{-i\Delta t \mathbf{V}} | x_{j-1} \rangle + \mathcal{O}(\Delta t^2) \\ &= \int dp_j \langle x_j | e^{-i\Delta t \mathbf{T}} | p_j \rangle \langle p_j | e^{-i\Delta t \mathbf{V}} | x_{j-1} \rangle + \mathcal{O}(\Delta t^2) \end{aligned}$$

$$\begin{aligned}
&= \int \tilde{d}p_j e^{-i T(p_j) \Delta t - i V(x_{j-1}) \Delta t} e^{i p_j (x_j - x_{j-1})} + \mathcal{O}(\Delta t^2). \\
&= \sqrt{\frac{m}{2\pi i \Delta t}} \exp \left(i \Delta t \left(\frac{m}{2} \frac{(x_j - x_{j-1})^2}{\Delta t^2} - V \left(\frac{x_j + x_{j-1}}{2} \right) \right) \right) + \mathcal{O}(\Delta t^2).
\end{aligned}$$

In the last step we used $\int_{-\infty}^{\infty} dp e^{-i a p^2} = \sqrt{\frac{\pi}{i a}}$. Assembling the whole thing,

$$\begin{aligned}
K(x, t; x_a, t_a) &= A^N \int \prod_{j=1}^{N-1} dx_j \exp \left(i \sum_{j=1}^{N-1} \frac{\Delta t}{\hbar} L \left(t_j, \frac{x_j + x_{j-1}}{2}, \frac{x_j - x_{j-1}}{\Delta t} \right) \right) \\
&\equiv \int [dx]_{x(t_a/b)=x_{a,b}} e^{\frac{i}{\hbar} \int_{t_a}^{t_b} dt L(t, x, \dot{x})}. \tag{3.39}
\end{aligned}$$

with $L(t, x, \dot{x}) = \frac{1}{2} m \dot{x}^2 - V(x)$. We see that the true role in life of $\hbar$ is to make up the units of action in this expression. The first expression here defines what we mean by the path integral $\int [dx]$. It includes horrible things which are not differentiable at all (they have $\Delta x \sim \sqrt{\Delta t}$, not $\Delta x \sim \Delta t$); when this matters, you need to use the discretization to make it well-defined.

[End of Lecture 15]

Stationary phase and classical mechanics. Recall that

$$I = \int dx e^{-N f(x)} = e^{-N f(x_0)} \int dy e^{-\frac{N}{2} f''(x_0) y^2 + \dots} \simeq e^{-N f(x_0)} \sqrt{\frac{2\pi}{N f''(x_0)}}$$

where $f'(x_0) = 0$ and $y \equiv x - x_0$. This is called saddle point approximation. If instead we have

$$\tilde{I} = \int dx e^{-i N f(x)} = e^{-i N f(x_0)} \int dy e^{-i \frac{N}{2} f''(x_0) y^2 + \dots} \simeq e^{-i N f(x_0)} \sqrt{\frac{2\pi i}{N f''(x_0)}}$$

where $f'(x_0) = 0$ and $y \equiv x - x_0$, this is called stationary phase. With more variables:

$$\int dx_1 \cdots dx_N e^{-i N f(x_1 \cdots x_N)} \simeq e^{-i N f(\underline{x})} \sqrt{\det_{ij} \left(\frac{\pi}{i N \partial_i \partial_j f} \right)}.$$

where

$$0 = \partial_{x_j} f(x)|_{x=\underline{x}} \forall j. \tag{3.40}$$

The path integral is just of the previous form with $\hbar = \frac{1}{N}$ and $f(x) = S[x] = \int dt L(t, x, \dot{x})$, the action functional. The square brackets indicate that the argument x is a whole function, rather than just a finite collection of numbers, but basically it just

means that N is big. The saddle point condition (3.40) can be rewritten in the fancy notation:

$$0 = \frac{\delta}{\delta x(t)}[\underline{x}] \quad , \forall t.$$

What is this when $S[x] = \int ds \left(\frac{1}{2} m \dot{x}^2(s) - V(x(s)) \right)$? The key ingredient is the ‘fundamental rule of variational calculus’, which in words is different independent variables are independent:

$$\frac{\partial x_j}{\partial x_i} = \delta_{ij} \quad \leftrightarrow \quad \frac{\delta x(s)}{\delta x(t)} = \delta(t-s).$$

$$\begin{aligned} \frac{\delta S}{\delta x(t)} &= \int ds \left(m \partial_s x(s) \partial_s \delta(s-t) - \frac{\partial V(x(s))}{\partial x(s)} \delta(t-s) \right) \\ &\stackrel{\text{IBP}}{=} \int ds \left(-m \partial_s^2 x(s) - \partial_{x(s)} V(x(s)) \right) \delta(t-s) \\ &= -m \ddot{x} - \partial_x V \end{aligned} \tag{3.41}$$

which is an equation due to Newton.

Evaluate for free particle (in a way which is useful for non-free particles). For $V = 0$, the classical path is a straight line:

$$\ddot{x} = 0 \quad \implies \quad \underline{x}(t) = x_a + (t - t_a) \frac{x_b - x_a}{t_b - t_a}.$$

$$S[\underline{x}] = \int_{t_a}^{t_b} L dt = \int \frac{m}{2} \dot{x}^2 = \frac{m}{2} \frac{(x_b - x_a)^2}{t_b - t_a}.$$

The fluctuations about the classical path are

$$y_j \equiv \delta x_j = x_j - \underline{x}(t_j).$$

$x_0 = x_a, x_N = x_b$ don’t fluctuate. The path integral is actually gaussian:

$$K(x_b, t_b; x_a, t_a) = \left(\frac{m}{2\pi i \Delta t} \right)^{N/2} \int dx_1 \cdots dx_{N-1} \prod_{j=1}^N \exp \left(i \frac{m}{2} \frac{(x_j - x_{j-1})^2}{\Delta t} \right)$$

Now plugging in $x_j = \underline{x}_j + y_j$,

$$(x_j - x_{j-1})^2 = \underbrace{(\underline{x}(t_j) - \underline{x}(t_{j-1}))^2}_{\rightarrow S[\underline{x}]} + (y_j - y_{j-1})^2 + \text{terms linear in } \delta x_j$$

we can ignore the terms linear in δx because they are zero by definition of the classical path.

$$K(x_b, t_b; x_a, t_a) = \left(\frac{m}{2\pi i \Delta t} \right)^{N/2} e^{\frac{i}{\hbar} S[\underline{x}]} \underbrace{\int dy_1 \cdots dy_{N-1} e^{i \sum_{jk} y_j M_{jk} y_k}}_{= \frac{\pi^{\frac{N-1}{2}}}{\sqrt{\det(-iM)}}}$$

where the $N - 1 \times N - 1$ matrix M is the discrete laplacian (recall the homework)

$$M_{jk} = \frac{m}{2\Delta t} \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix}_{jk}$$

We need to compute its determinant. On the homework you saw that $[M, T] = 0$ where $T = \begin{pmatrix} 0 & 1 & 0 & & \\ 0 & 0 & 1 & & \\ & & & \ddots & \\ & & & & 0 & 1 & 0 \end{pmatrix}$ is the discrete translation operator. Its eigenstates are plane waves: $v_j(p) = ae^{ipj} + be^{-ipj}$; the boundary conditions $y_0 = y_N = 0$ determine $v_j(p) = \sin pj$, $p = p_n = \frac{2\pi n}{N}$. So we could try to compute $\det M = \prod_{n=1..N-1} (2 - \cos p_n)$. Instead, here's a trick: let

$$M_N(u) \equiv \begin{pmatrix} 2 \cosh u & -1 & & & \\ -1 & 2 \cosh u & -1 & & \\ & -1 & 2 \cosh u & -1 & \\ & & & \ddots & \\ & & & -1 & 2 \cosh u & -1 \\ & & & & -1 & 2 \cosh u \end{pmatrix}.$$

and $d_N \equiv \det M_N$. The recursion relation

$$\det M_N = 2 \cosh u \det M_{N-1} - \det M_{N-2} (-1)^2$$

follows from the block structure of the matrix. For $u = 0$, it's just $d_N = 2d_{N-1} - d_{N-2}$. We need two initial conditions to solve this two-step recursion: $\det M_1 = 2 \cosh u \rightarrow 2$ for $u \rightarrow 0$. $\det M_2 = 4 \cosh^2 u - 1 \rightarrow 3$ for $u \rightarrow 0$. The solution for $u = 0$ is simple: $d_{N-1} = d_{N-2} + 1$, hence $d_{N-1} = N$. (For general u , it is $d_N = \frac{\sinh(N+1)u}{\sinh u}$.)

$$\begin{aligned} K(x_b, t_b; x_a, t_a) &= \left(\frac{m}{2\pi i \Delta t} \right)^{N/2} e^{\frac{i}{\hbar} S[x]} \frac{\pi^{\frac{N-1}{2}}}{\sqrt{\det(-iM)}} \\ &= \left(\frac{m}{2\pi i \Delta t} \right)^{N/2} e^{\frac{i}{\hbar} S[x]} \left(\frac{m}{2\pi i \Delta t} \right)^{-(N-1)/2} N^{-1/2} \\ &= \sqrt{\frac{m}{2\pi(t_b - t_a)i}} e^{\frac{i}{\hbar} S[x]}. \end{aligned} \tag{3.42}$$

Coherent state path integral. The position basis resolution on the identity does not have a monopole here; we can use whatever resolution we like. We can even use coherent states! This representation is very useful in quantum many-body physics.

So suppose $\mathcal{H} = \mathcal{H}_{\text{SHO}} = \text{span}\{|0\rangle, |1\rangle, \dots\}$ where $\mathbf{a}^\dagger \mathbf{a} |n\rangle = n |n\rangle$, $[\mathbf{a}, \mathbf{a}^\dagger] = 1$. We need not take $\mathbf{H} = \hbar\omega (\mathbf{a}^\dagger \mathbf{a} + \frac{1}{2})$, but we will.

Recall that coherent states have $\mathbf{a} |\alpha\rangle = \alpha |\alpha\rangle$. The coherent-state representation of the propagator is

$$K(\alpha_b, t_b; \alpha_a, t_a) \equiv \langle \alpha_b | \mathbf{U}(t_b, t_a) | \alpha_a \rangle.$$

They resolve the identity by

$$\mathbb{1} = \int \frac{d\text{Re}\alpha d\text{Im}\alpha}{\pi} |\alpha\rangle \langle \alpha|$$

and their overlaps are

$$\langle \alpha | \alpha' \rangle = e^{-\frac{1}{2}(|\alpha|^2 + |\alpha'|^2) + \alpha^* \alpha'}.$$

So sticking $N - 1$ walls full of holes in the coherent state basis:

$$K(\alpha_b, t_b; \alpha_a, t_a) = \int \frac{d^2\alpha_1 \cdots d^2\alpha_{N-1}}{\pi^{N-1}} K(\alpha_b, t_b; \alpha_{N-1}, t_{N-1}) \cdots K(\alpha_1, t_1; \alpha_a, t_a)$$

where the small step is

$$K(\alpha_j, t_j; \alpha_{j-1}, t_{j-1}) = \langle \alpha_j | e^{-\frac{i}{\hbar} \Delta t \mathbf{H}} | \alpha_{j-1} \rangle.$$

This is useful if $\mathbf{H}$ acts nicely on the coherent states, as is the case when it can be written in terms of creation and destruction operators. For simplicity consider the (cop out!) case of the SHO:

$$\begin{aligned} K(\alpha_j, t_j; \alpha_{j-1}, t_{j-1}) &= \langle \alpha_j | e^{-i\Delta t \omega \mathbf{a}^\dagger \mathbf{a}} | \alpha_{j-1} \rangle \\ &= e^{-i\Delta t \omega \alpha_j^* \alpha_{j-1}} \langle \alpha_j | \alpha_{j-1} \rangle + \mathcal{O}(\Delta t^2) \\ &= e^{-i\Delta t \omega \alpha_j^* \alpha_{j-1} + \mathcal{O}(\Delta t^2)} \exp \left(\underbrace{-\frac{1}{2}|\alpha_j|^2 - \frac{1}{2}|\alpha_{j-1}|^2 + \alpha_j^* \alpha_{j-1}}_{=\frac{1}{2}\alpha_j^* (\alpha_j - \alpha_{j-1}) + \frac{1}{2}\alpha_{j-1} (-\alpha_{j-1}^* + \alpha_j^*)} \right) \end{aligned} \quad (3.43)$$

$$\begin{aligned} K(\alpha_b, t_b; \alpha_a, t_a) &= \int \prod_{j=1}^{N-1} \frac{d^2\alpha_j}{\pi} e^{i \int_{t_a}^{t_b} dt \left(\frac{1}{2}(\alpha^* \dot{\alpha} - \alpha \dot{\alpha}^*) - \omega \alpha^* \alpha \right)} \\ &= \int [d\alpha] e^{i \int dt L(\alpha, \dot{\alpha})} \end{aligned} \quad (3.44)$$

with $L = p\dot{x} - H$.

For the special case of the SHO, recall that we know the answer since coherent states remain coherent under time evolution by the SHO hamiltonian:

$$\mathbf{U}(t) |\alpha\rangle = \mathcal{N}_\alpha \underbrace{\mathbf{U}(t) e^{\mathbf{a}^\dagger \alpha} \mathbf{U}^\dagger(t)}_{=e^{\mathbf{a}^\dagger \alpha e^{-i\omega t}}} \underbrace{\mathbf{U}(t) |0\rangle}_{=e^{-i\omega t/2} |0\rangle} = e^{-i\omega t/2} |\alpha e^{-i\omega t}\rangle.$$

So

$$K(\alpha_b, t_b; \alpha, 0) = \langle \alpha_b | \mathbf{U}(t) |\alpha\rangle = e^{-i\omega t/2} \langle \alpha_b | \alpha e^{-i\omega t} \rangle = e^{-i\omega t/2} e^{-|\alpha_b|^2/2 - |\alpha|^2/2 + \alpha_b^* \alpha e^{-i\omega t}}.$$

The situation when this comes into its own is when we have a collection of SHOs:

$$\mathbf{H} = \sum_k \hbar \omega_k \left(\mathbf{a}_k^\dagger \mathbf{a}_k \right) + \dots$$

where the $\dots$ represents some interactions which may be treated as small.

4 Theory of Angular Momentum

A natural next step would be to study wave mechanics in $d > 1$. However, an essential tool in doing so is the use of rotation invariance. So we must first develop an understanding of the consequences of symmetries in quantum mechanics.

4.1 Continuous symmetries and the quantum Noether theorem

As in classical mechanics, in QM there is direct connection between conservation laws (the existence of quantities which don't change under time evolution) and continuous symmetries. A symmetry in quantum mechanics is a map on $\mathcal{H}$ which doesn't change things we can observe about the system. So imagine we can do something to the system, call it T for 'transformation,' it could be a rotation or translation or Lorentz boost. This maps each state $|\psi\rangle$ of the system to another state $|\psi_T\rangle$. (The perspective we are taking here, of actually acting upon the system, is called the *active* viewpoint.)

Given some observable $\mathbf{A} = \mathbf{A}^\dagger = \sum_a a |a\rangle \langle a|$ acting on $\mathcal{H}$, the transformation acts on its eigenstates

$$|a\rangle \xrightarrow{T} |a_T\rangle.$$

We care about this because we can measure

$$\text{Prob}_{|\psi\rangle}(a) = |\langle a|\psi\rangle|^2$$

If T is a symmetry, we require it to preserve our accounting of probabilities for the outcomes for measuring $\mathbf{A}$:

$$\text{Prob}_{|\psi\rangle}(a) \stackrel{!}{=} \text{Prob}_{|\psi_T\rangle}(a) \quad \leftrightarrow \quad |\langle a|\psi\rangle|^2 = |\langle a_T|\psi_T\rangle|^2.$$

Now, a non-trivial fact is that for any such transformation, we can write its effects as

$$|\psi_T\rangle = \mathbf{U}_T |\psi\rangle$$

where $\mathbf{U}_T$ is either a linear operator or an anti-linear operator. This is a theorem of Wigner whose proof I don't find very instructive; it can be found in appendix A of Le Bellac's book. More precisely, it says: suppose that are given a correspondence between *rays* on $\mathcal{H}$:

$$\tilde{\psi} \mapsto \tilde{\psi}_T$$

(here I am denoting the ray by a tilde: $\tilde{\psi} = \widetilde{e^{i\phi}\psi}$) that preserves inner products

$$|\langle \tilde{\psi} | \tilde{\chi} \rangle| = |\langle \tilde{\psi}_T | \tilde{\chi}_T \rangle| \quad \forall \psi, \chi.$$

then we can choose the phases of the vectors so that

$$|\phi\rangle \mapsto |\phi_T\rangle = \mathbf{U}_T |\phi\rangle.$$

Then the inner-product preservation says

$$|\langle\phi|\chi\rangle| = |\langle\mathbf{U}_T\phi|\mathbf{U}_T|\chi\rangle| = |\langle\phi|\mathbf{U}_T^\dagger\mathbf{U}_T|\chi\rangle|.$$

Since this is true for all states, we have $\mathbf{U}_T^\dagger\mathbf{U}_T = \pm 1$. $\mathbf{U}_T$ is either linear and unitary or anti-linear and unitary (which is called anti-unitary) and means

$$\langle\mathbf{U}_T\phi|\mathbf{U}_T\chi\rangle = \langle\chi|\phi\rangle = \langle\phi|\phi\rangle^*.$$

In order for T to be a symmetry, we must further demand that this situation does not change under time evolution. That is, a symmetry is represented by a map $\mathbf{U}_T$ such that

$$\mathbf{U}_T\mathbf{U}(t) = \mathbf{U}(t)\mathbf{U}_T$$

where $\mathbf{U}(t)$ is the time evolution operator. Since this should be true for all t , we can expand in small t and conclude that

$$\mathbf{U}_T\mathbf{H} = \mathbf{H}\mathbf{U}_T \quad i.e. \quad [\mathbf{U}_T, \mathbf{H}] = 0. \quad (4.1)$$

Symmetries form a group: if you can do something T_1 and something else T_2 , then you can do something and then something else $T_2 \circ T_1$. One thing you can always do is nothing: $T = \mathbb{1}$. Moreover, you can undo each of them (once you believe the Wigner theorem, you certainly can, since unitary matrices are invertible).

The multiplication law of the abstract group of transformations has to be respected by its representation on the quantum states:

$$\mathbf{U}(T_2)\mathbf{U}(T_1) \stackrel{?}{=} \mathbf{U}(T_2 \circ T_1).$$

Actually, we can allow the slight generalization of this representation law:

$$\mathbf{U}(T_2)\mathbf{U}(T_1) = u(T_2, T_1)\mathbf{U}(T_2 \circ T_1).$$

Here the object u is a (sometimes important) loophole: states are only defined up to an overall phase, so the group law only needs to be true up to such a phase $u(T_2, T_1)$ (this is called a *projective representation*). We will see an important example below with the action of rotations on a half-integer spin.

For a continuous symmetry, where the transformation is labelled by a continuously-variable parameter (such as rotations or translations), we can say more. We can choose

T to be very close to doing nothing, in which case $\mathbf{U}$ must be close to the identity on the Hilbert space

$$\mathbf{U} = \mathbb{1} - \mathbf{i}\epsilon\mathbf{Q} + \mathcal{O}(\epsilon^2) \quad .$$

$\mathbf{U}$ is unitary; this implies that $\mathbf{Q}^\dagger = \mathbf{Q}$, so $\mathbf{Q}$ is an observable. Expanding (4.1) to first order in ϵ , we find

$$[\mathbf{Q}, \mathbf{H}] = 0 \quad . \quad (4.2)$$

This is a *conservation law* in the following sense: if the system is in an eigenstate of $\mathbf{Q}$, the time evolution by the Schrödinger equation doesn't change that situation. Symmetries imply conservation laws.

The converse is also true: an observable $\mathbf{Q}$ which commutes with $\mathbf{H}$ can be used to make a family of unitaries which commute with $\mathbf{H}$, by making many infinitesimal transformations:

$$(\mathbb{1} - \mathbf{i}\epsilon\mathbf{Q})^N \xrightarrow{N \rightarrow \infty, \epsilon \rightarrow 0, a \equiv \epsilon N} e^{-\mathbf{i}\mathbf{Q}a} \equiv \mathbf{U}_a.$$

Example: translations. As an example, let's think about a free particle that lives on a line, with

$$\mathbf{H} = \frac{\hat{p}^2}{2m} = -\frac{\hbar^2}{2m} \partial_x^2 \quad .$$

Notice that shifting the origin of the line doesn't change the Hamiltonian; formally, this is the statement that

$$[\hat{p}, \mathbf{H}] = 0 \quad .$$

The fact that the Hamiltonian doesn't depend on x means that momentum is a conserved quantity. What is the finite transformation generated by the conserved quantity $\hat{p}$? It is just $\mathbf{U}(a) = e^{-ia\hat{p}/\hbar}$ which acts by $|\psi\rangle \rightarrow \mathbf{U}(a) |\psi\rangle \equiv |\psi_a\rangle$ and on a position-basis wavefunction $\psi(x) = \langle x|\psi\rangle$ by

$$\mathbf{U}(a)\psi(x) = e^{-ia\hat{p}/\hbar}\psi(x) = e^{-a\partial_x}\psi(x) = \psi(x) - a\partial_x\psi(x) + \frac{1}{2!}a^2\partial_x^2\psi(x) + \dots = \psi(x-a) = \psi_a(x) \quad , \quad (4.3)$$

a translation. Here I used Taylor's theorem.

Directly using the canonical commutation relation, you can see that

$$\mathbf{U}(a) |x\rangle = |x+a\rangle \quad .$$

More generally, it acts by

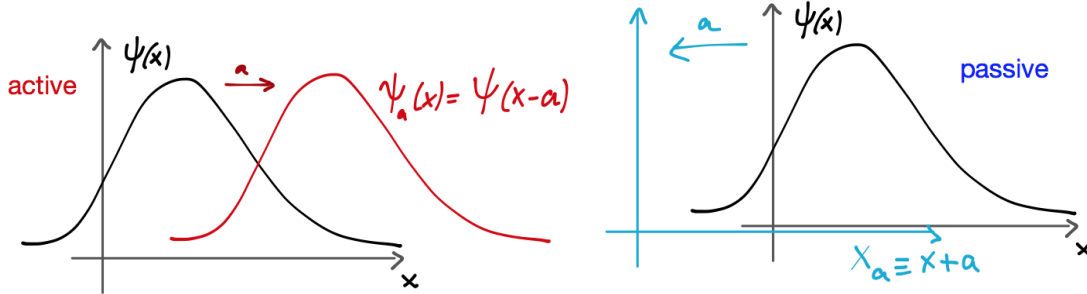
$$\mathbf{U}(a) |\psi\rangle = |\psi_a\rangle$$

whose wavefunction is shifted to the right by a .

Active versus passive:

$$\langle \psi | \mathcal{O} | \psi \rangle \xrightarrow{a} \langle \psi_a | \mathcal{O} | \psi_a \rangle \stackrel{\text{active}}{=} \underbrace{\langle \psi | \mathbf{U}(a)^\dagger}_{= \langle \psi_a |} \mathcal{O} \underbrace{\mathbf{U}(a) | \psi \rangle}_{= | \psi_a \rangle} \stackrel{\text{passive}}{=} \langle \psi | \underbrace{\mathbf{U}(a)^\dagger \mathcal{O} \mathbf{U}(a)}_{\equiv \mathcal{O}_a} | \psi \rangle.$$

So we can think of the transformation as $\mathcal{O} \rightarrow \mathcal{O}_a \equiv \mathbf{U}(a)^\dagger \mathcal{O} \mathbf{U}(a)$, and keep the states the same. For translations, this acts on the position operator as $\mathbf{x} \rightarrow \mathbf{x}_a \equiv \mathbf{x} + a\mathbb{1}$, and hence shifts its eigenvalues to the left.



The group law in this example is very simple: $\mathbf{U}(a)\mathbf{U}(b) = \mathbf{U}(a+b) = e^{-i(a+b)\hat{p}/\hbar} = \mathbf{U}(b)\mathbf{U}(a)$. This group is *abelian*: all the elements commute.

Another example of a continuous symmetry is time translations. If $\partial_t \mathbf{H} = 0$ then $[\mathbf{H}, \mathbf{H}] = 0$ at all times, and $[\mathbf{U}(t), \mathbf{H}] = 0$, so $\mathbf{H}$ is a conserved charge: energy is conserved. The finite transformation is unsurprisingly implemented by exponentiating the conserved charge $\mathbf{U}(t) = e^{-it\mathbf{H}/\hbar}$.

Our main focus here will be an example where the symmetry group (namely, the group of rotations) is non-abelian, and where the projective loophole is exploited. (The associated conserved quantity which generates rotations is the angular momentum.)

[End of Lecture 16]

4.2 Rotations in ordinary 3-space

We'll take the active point of view that a rotation acts by moving a point $R : P \rightarrow P' = R(P)$. A *rotation* is such a transformation which preserves an origin, and which preserves distances (and hence angles).

If we choose coordinates on our space, it acts on them by $\vec{r} \rightarrow \vec{r}' = R(\vec{r})$. The definition implies that the transformation is linear and hence invertible (choosing the origin of coordinates to coincide with the origin of the rotation,



it, the zero vector, is the only element of the kernel of R).

In a basis, $\{\check{e}_1 \equiv \check{x}, \check{e}_2 \equiv \check{y}, \check{e}_3 \equiv \check{z}\}$, a point is labelled by the components $\vec{r} = \sum_i \check{e}_i r_i$, $\vec{r}' = \sum_i \check{e}'_i r'_i$, the rotation operator has matrix elements

$$R_{ij} \equiv \check{e}_i \cdot (R\check{e}_j)$$

which determine the components of $\vec{r}'$: $r'_i = \sum_j R_{ij} r_j$. The condition $|\vec{r}'| = |\vec{r}|$ implies

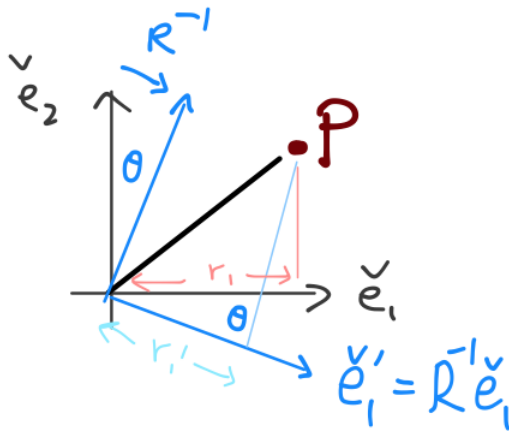
$$R^t R = \mathbb{1}$$

where t is for ‘transpose’. This in turn implies $(\det R)^2 = 1$.

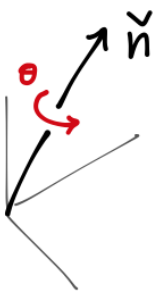
The set of rotations of $\mathbb{R}^n$ is a group called $O(n)$. The subgroup with $\det R = 1$ is continuously connected to $\mathbb{1}$ (these are sometimes called ‘proper rotations’) and is called $SO(n)$.

One more comment about the alternative passive viewpoint. In the active viewpoint, the transformation takes one point P to another $R(P)$, and we use the same basis to describe both. In the passive viewpoint, we think about a single point P , and consider the action of a symmetry on the basis: one point, two frames of reference. To make the same transformation on the components in the passive viewpoint, we must act with the inverse transformation:

$$\check{e}'_i \equiv R^{-1} \check{e}_i.$$



Axis and angle. A slightly nontrivial fact is that a proper rotation in 3d is specified by a rotation axis, and an angle. The rotation axis is preserved by the rotation $R(\check{n}, \theta)\check{n} = \check{n}$ for all angles: it is an eigenvector with eigenvalue one. The content of this ‘axis-angle theorem’ is that such an eigenvector always exists; we’ll see why just below.



Rotations about different axes do not commute, as you can see by doing an appropriate dance. The rotations about a fixed axis form a one-parameter abelian subgroup:

$$R(\check{n}, \theta_1)R(\check{n}, \theta_2) = R(\check{n}, \theta_1 + \theta_2). \quad (4.4)$$

It will be useful to make explicit the matrices for rotations about the basis vectors:

$$R(\hat{x}, \theta) = \begin{pmatrix} 1 & & \\ & \cos \theta & -\sin \theta \\ & \sin \theta & \cos \theta \end{pmatrix}, \quad R(\hat{y}, \theta) = \begin{pmatrix} \cos \theta & -\sin \theta & \\ & 1 & \\ \sin \theta & & \cos \theta \end{pmatrix}, \quad R(\hat{z}, \theta) = \begin{pmatrix} \cos \theta & -\sin \theta & \\ \sin \theta & \cos \theta & \\ & & 1 \end{pmatrix}.$$

Now consider an infinitesimal rotation,

$$R = \mathbb{1} + \epsilon A \tag{4.5}$$

(plus terms of order ϵ^2 which we regard as zero unless they are the leading term). The orthogonality condition says

$$\mathbb{1} = R^t R \implies A + A^t = 0$$

Now we can prove the axis-angle theorem: The nonzero eigenvectors of a real antisymmetric matrix are zero or come in complex-conjugate pairs of the form $\pm i a$ where a is real. A 3×3 such matrix must then have one eigenvector with eigenvalue zero; this is the rotation axis.

It is useful to make a basis for such matrices A :

$$A = \begin{pmatrix} 0 & -a_3 & a_2 \\ & 0 & -a_1 \\ & & 0 \end{pmatrix} \equiv \sum_i a_i \mathcal{J}_i = \vec{a} \cdot \vec{\mathcal{J}}.$$

$\vec{\mathcal{J}}$ is a 3-vector of matrices, just like the Paulis. Explicitly,

$$\mathcal{J}_1 = \begin{pmatrix} 0 & & \\ & 0 & -1 \\ & 1 & 0 \end{pmatrix}, \quad \mathcal{J}_2 = \begin{pmatrix} 0 & 1 & \\ & 0 & \\ -1 & & 0 \end{pmatrix}, \quad \mathcal{J}_3 = \begin{pmatrix} 0 & -1 & \\ 1 & 0 & \\ & & 0 \end{pmatrix}.$$

A nice compact formula for the matrix elements is

$$(\mathcal{J}_i)_{jk} = -\epsilon_{ijk}.$$

By direct calculation, or using epsilon symbol identities, these matrices satisfy

$$(\mathcal{J}_i \mathcal{J}_j)_{kl} = \delta_{il} \delta_{kj} - \delta_{ij} \delta_{kl}$$

and hence

$$[\mathcal{J}_i, \mathcal{J}_j] = \epsilon_{ijk} \mathcal{J}_k$$

or $[\vec{a} \cdot \vec{\mathcal{J}}, \vec{b} \cdot \vec{\mathcal{J}}] = (\vec{a} \times \vec{b}) \cdot \vec{\mathcal{J}}$. The last step is familiar from your study of Paulis.

Now let's reconstruct a finite rotation. Here's another way to see what the right thing to do is:

$$\partial_\theta R(\tilde{n}, \theta) = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (R(\tilde{n}, \theta + \epsilon) - R(\tilde{n}, \theta)) \stackrel{(4.4)}{=} \underbrace{\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (R(\tilde{n}, \epsilon) - \mathbb{1})}_{=A} R(\tilde{n}, \theta)$$

where $A = \vec{a} \cdot \vec{\mathcal{J}}$ is as in (4.5) and is proportional to θ . With the initial condition $R(\tilde{n}, 0) = \mathbb{1}$, the solution of this differential equation is

$$R(\tilde{n}, \theta) = e^{\theta \tilde{n} \cdot \mathcal{J}}$$

where θ is the angle and $\tilde{n}$ is a unit vector. Let's check the normalization by evaluating

$$R(\tilde{z}, \theta) = e^{\theta \mathcal{J}_3}.$$

Using

$$\mathcal{J}_3^2 = -\mathbb{1} + |z\rangle \langle z|, \mathcal{J}_3^3 = -\mathcal{J}_3, \mathcal{J}_3^4 = -\mathcal{J}_3^2 \dots$$

we have

$$R(\tilde{z}, \theta) = e^{\theta \mathcal{J}_3} = \sum_{n=0}^{\infty} \frac{\theta^n}{n!} \mathcal{J}_3^n = \mathbb{1} + \mathbf{i} \mathcal{J}_3 \sum_{n \text{ odd}} \frac{\mathbf{i}^n \theta^n}{n!} + (\mathbb{1} - |z\rangle \langle z|) \sum_{n \text{ even}} \frac{\mathbf{i}^n \theta^n}{n!} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \\ & & 1 \end{pmatrix}.$$

(Alternatively, we could just exponentiate the upper 2×2 block, noticing that that part of $\mathcal{J}_3$ is just $\mathbf{i}\sigma^2$.)

4.3 Spectra of the angular momentum operators

It is conventional to write the relationship between finite transformation and generator as

$$\mathbf{U}(\theta) = e^{-\mathbf{i} \mathbf{G} \theta / \hbar}$$

so that the angular momentum operators are

$$\mathbf{J}_j \equiv -\mathbf{i} \hbar \mathcal{J}_j$$

and have units of $\hbar$, action, angular momentum. The factor of $\mathbf{i}$ is significant because it makes the generators $\mathbf{J}_j$ Hermitian. They satisfy the algebra

$$[\mathbf{J}_i, \mathbf{J}_j] = \mathbf{i} \hbar \epsilon_{ijk} \mathbf{J}_k \tag{4.6}$$

(note that an $\mathbf{i}^2 = -1$ appeared at an intermediate step). This is the (Lie) algebra called $\mathfrak{su}(2)$ or $\mathfrak{so}(3)$. We have seen a close relative when studying Pauli matrices. We

are going to suppose this algebra is realized on our Hilbert space $\mathcal{H}$ and understand what that means for the structure of $\mathcal{H}$. The thing we are doing is the essence of the representation theory of Lie algebras.

As a first step notice that algebra (4.6) implies that the *total angular momentum*

$$\mathbf{J}^2 \equiv \vec{\mathbf{J}} \cdot \vec{\mathbf{J}} = \mathbf{J}_1^2 + \mathbf{J}_2^2 + \mathbf{J}_3^2$$

commutes with each component:

$$[\mathbf{J}^2, \vec{\mathbf{J}}] = 0.$$

This means further that $[\mathbf{J}^2, \mathbf{U}(\vec{n}, \theta)] = 0$ for any $\vec{n}, \theta$. So it doesn't change under a rotation. It is called a *Casimir* operator. Because $\mathbf{J}_i$ are hermitian, $\mathbf{J}^2 \geq 0$ in the sense that its eigenvalues $a \geq 0$.

So it is possible that a complete set of commuting operators on $\mathcal{H}$ is $\{\mathbf{J}^2, \mathbf{J}_3 \equiv \mathbf{J}_z\}$, with eigenvalues which we'll call $\hbar^2 a, \hbar m$ and eigenvectors $|a, m\rangle$. (notice that a, m are dimensionless). There could be other operators which commute with these; ignore them for now. Assume we've normalized $\langle a, m | a, m \rangle = 1$. What's special about $\mathbf{J}_3$ here? Nothing.

Define raising and lowering operators

$$\mathbf{J}_{\pm} \equiv \mathbf{J}_1 \pm i\mathbf{J}_2,$$

which are each others adjoints: $\mathbf{J}_{\pm}^{\dagger} = \mathbf{J}_{\mp}$. They are called raising and lowering because

$$[\mathbf{J}_3, \mathbf{J}_{\pm}] = \pm \hbar \mathbf{J}_{\pm}, \quad [\mathbf{J}_+, \mathbf{J}_-] = 2\hbar \mathbf{J}_3, \quad [\mathbf{J}^2, \mathbf{J}_{\pm}] = 0.$$

The first equation says that $\mathbf{J}_{\pm} |am\rangle = c_{am} |a, m \pm 1\rangle$ for some constant c_{am} (just like with the SHO: $[\mathbf{n}, \mathbf{a}^{\dagger}] = \mathbf{a}^{\dagger}$). It will be useful to note that

$$\mathbf{J}^2 = \frac{1}{2} (\mathbf{J}_+ \mathbf{J}_- + \mathbf{J}_- \mathbf{J}_+) + \mathbf{J}_3^2 = \mathbf{J}_+ \mathbf{J}_- + \mathbf{J}_3 (\mathbf{J}_3 + \hbar) = \mathbf{J}_- \mathbf{J}_+ + \mathbf{J}_3 (\mathbf{J}_3 - \hbar). \quad (4.7)$$

Now make a sandwich of $\mathbf{J}_- \mathbf{J}_+$:

$$0 \leq \|\mathbf{J}_+ |am\rangle\|^2 = \langle am | \mathbf{J}_- \mathbf{J}_+ |am\rangle \stackrel{(4.7)}{=} \hbar^2 (a - m(m+1)). \quad (4.8)$$

Similarly,

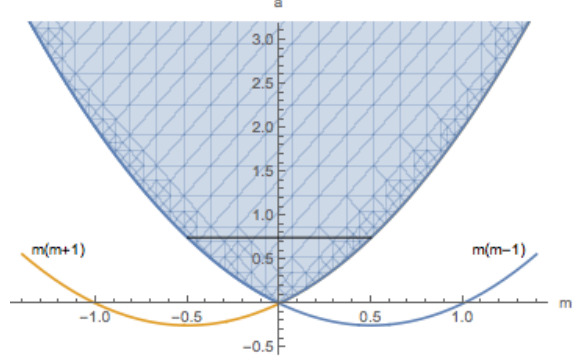
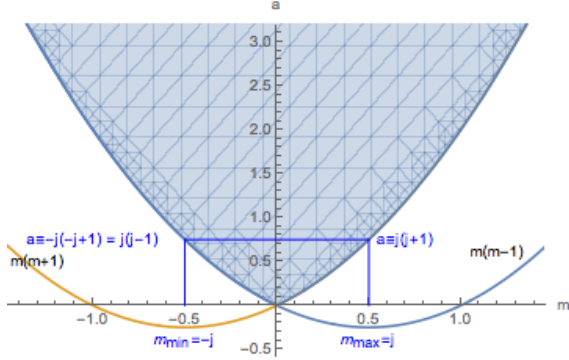
$$0 \leq \|\mathbf{J}_- |am\rangle\|^2 = \langle am | \mathbf{J}_+ \mathbf{J}_- |am\rangle \stackrel{(4.7)}{=} \hbar^2 (a - m(m-1)). \quad (4.9)$$

We learn that

$$a \geq \max(m(m+1), m(m-1)).$$

For a given value of a , we therefore have a biggest possible value of $|m|$, which we call j :

$$m \in [-j, j], \quad a \equiv j(j+1).$$



This defines the quantity j ; we'll use this to label the eigenvalue of $\mathbf{J}^2$ from now on and denote the eigenstate

$$|a, m\rangle \equiv |j, m\rangle,$$

$$\mathbf{J}^2 |j, m\rangle = \hbar^2 j(j+1) |j, m\rangle, \quad \mathbf{J}_3 |j, m\rangle = \hbar m |j, m\rangle.$$

[End of Lecture 17]

The norm conditions (4.8), (4.9) can be written (just using algebra) as

$$0 \leq \|\mathbf{J}_+ |am\rangle\|^2 = \hbar^2 (j(j+1) - m(m+1)) = \hbar^2 (j-m)(j+m+1)$$

$$0 \leq \|\mathbf{J}_- |am\rangle\|^2 = \hbar^2 (j(j+1) - m(m-1)) = \hbar^2 (j+m)(j-m+1).$$

When are these inequalities saturated? The states which do so are called highest-weight (sometimes lowest-weight) states:

$$\mathbf{J}_+ |jm\rangle = 0 \quad \Leftrightarrow \quad j-m=0 \quad \text{or} \quad j+m+1=0$$

The second equality is not possible since $|m| \leq j$. So $\mathbf{J}_+$ annihilates states which can't be raised anymore:

$$\mathbf{J}_+ |j, m=j\rangle = 0, \quad \mathbf{J}_- |j, m=-j\rangle = 0.$$

Similarly, $\mathbf{J}_-$ annihilates states which can't be lowered anymore. In between these two extremes, $\mathbf{J}_\pm$ raise and lower the (z -component of) angular momentum by integer multiples of $\hbar$: If the state $\mathbf{J}_\pm |jm\rangle$ is not zero then

$$\mathbf{J}_3 (\mathbf{J}_\pm |jm\rangle) = \mathbf{J}_\pm (\underbrace{\mathbf{J}_3 \pm \hbar}_{=\hbar(m\pm 1)} |jm\rangle) = \hbar(m \pm 1) (\mathbf{J}_\pm |jm\rangle). \quad (4.10)$$

And by the way, all these states have the same $\mathbf{J}^2$:

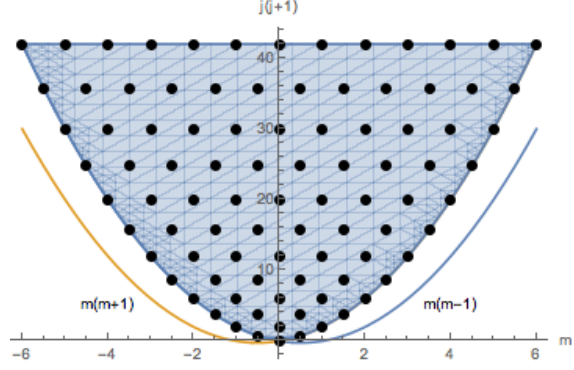
$$\mathbf{J}^2 (\mathbf{J}_+ |jm\rangle) = \mathbf{J}_+ \mathbf{J}^2 |jm\rangle = \hbar^2 j(j+1) (\mathbf{J}_+ |jm\rangle).$$

The + equation of (4.10) says that m must be of the form $m = j - n_1$ where $n_1 \in \mathbb{Z}_{\geq 0}$, lest we be able to raise forever without ever getting a null vector.

The – equation of (4.10) says that $m = -j + n_2$ where $n_2 \in \mathbb{Z}_{\geq 0}$ is a non-negative integer, lest we be able to lower forever. Taking the difference of these two equations we learn

$$2j = n_1 + n_2 \in \{0, 1, 2, 3, \dots\}$$

$$\implies j \in \{0, \frac{1}{2}, 1, \frac{3}{2}, \dots\}.$$



Unitary representations of the algebra (4.6) have only these values of j . For a given value of j , m ranges over the $2j + 1$ values of the form

$$m \in \{-j, -j + 1, \dots, j - 1, j\} \quad .$$

Angular momentum is quantized! This followed just from demanding that the algebra (4.6) is realized on a Hilbert space (where all the states of positive norm).

4.4 Matrix elements of the angular momentum operators

We noticed earlier that $\mathbf{J}_{\pm} |jm\rangle = c_{\pm} |j, m \pm 1\rangle$. Demanding $\| |j, m \pm 1\rangle \|^2 = 1$,

$$|c_{\pm}|^2 = \langle jm | \mathbf{J}_{\mp} \mathbf{J}_{\pm} |jm\rangle \stackrel{(4.8), (4.9)}{=} \hbar^2 (j \mp m)(j \pm m + 1).$$

Hence

$$\mathbf{J}_{\pm} |jm\rangle = \hbar \sqrt{(j \mp m)(j \pm m + 1)} |j, m \pm 1\rangle \quad (4.11)$$

up to a phase. We can (partially) fix this phase by the convention of demanding that $|jj\rangle$ is real. (With this convention $\mathbf{J}_1 = \frac{1}{2}(\mathbf{J}_+ + \mathbf{J}_-)$ and $\mathbf{J}_2 = \frac{1}{2i}(\mathbf{J}_+ - \mathbf{J}_-)$ have real and imaginary matrix elements respectively. Like σ^1 and σ^2 in the σ^3 basis.)

Induction implies that

$$|jm\rangle = \left(\frac{\mathbf{J}_-}{\hbar} \right)^{j-m} |jj\rangle \sqrt{\frac{(j+m)!}{(2j)!(j-m)!}}.$$

This is just like the SHO, except that it stops when $m = -j$, in which case we get zero if we do it again. It's like an SHO with a vacuum at both ends.

So far, we've assumed that j, m are complete set of labels on a state in $\mathcal{H}$, so $\mathcal{H} = \mathcal{H}_j \equiv \text{span}\{|jm\rangle, m = -j, -j+1, \dots, j-1, j\}$ of dimension $2j+1$ for some particular j . Such a thing with one particular j is called an *irreducible* representation of $\mathfrak{su}(2)$. Here irreducible means that you can make any state in the representation from any

other by acting just with generators of the algebra. This is not always true. In general, we can have a reducible representation where moreover the same j can appear more than once:

$$\mathcal{H} = \oplus_{j_a \in S} \mathcal{H}_{j_a}$$

where S is some set of labels chosen from $\mathbb{Z}_{\geq 0}/2$ possibly with repetitions. The full identity on this space is

$$\mathbb{1} = \sum_{j_a \in S} \sum_{m=-j_a, -j_a+1, \dots, j_a-1, j_a} |j_a m, a\rangle \langle j_a m, a|.$$

The $\vec{\mathbf{J}}$ s can't change j_a or the a label. Acting on this space, $\mathbf{J}^2$ is of the form

$$\mathbf{J}^2 = \sum_{j_a \in S} \mathbb{1}_{2j_a+1 \times 2j_a+1} \hbar^2 j_a(j_a + 1)$$

where a is an extra label (some eigenvalue of some other observable), and the identity operator acting on the space of values of m_j .

Examples, starting from the simplest case:

Spin zero: If $j = 0$, the irrep has only one state $|0, 0\rangle$, and all the generators annihilate it: $\mathbf{J}^2 |0, 0\rangle = 0$, $\vec{\mathbf{J}}^z |0, 0\rangle = 0$, $\vec{\mathbf{J}}^\pm |0, 0\rangle = 0$. They are all given by a 1×1 matrix of zeros. Done.

Spin- $\frac{1}{2}$: $\mathcal{H}_{\frac{1}{2}} = \text{span}\{|\frac{1}{2}, \frac{1}{2}\rangle \equiv |\uparrow\rangle, |\frac{1}{2}, -\frac{1}{2}\rangle \equiv |\downarrow\rangle\}$.

$$\mathbf{J}^2 \left| \frac{1}{2}, \pm \frac{1}{2} \right\rangle = \frac{3}{4} \hbar^2 \left| \frac{1}{2}, \pm \frac{1}{2} \right\rangle$$

$$\mathbf{J}_z \left| \frac{1}{2}, \pm \frac{1}{2} \right\rangle = \pm \frac{1}{2} \hbar \left| \frac{1}{2}, \pm \frac{1}{2} \right\rangle$$

$$\mathbf{J}_+ \left| \frac{1}{2}, \frac{1}{2} \right\rangle = 0, \quad \mathbf{J}_+ \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \stackrel{(4.11)}{=} \hbar \sqrt{\frac{1}{2} \cdot \frac{3}{2} - \left(-\frac{1}{2}\right) \left(\frac{1}{2}\right)} \left| \frac{1}{2}, \frac{1}{2} \right\rangle = \hbar \left| \frac{1}{2}, \frac{1}{2} \right\rangle$$

Similarly,

$$\mathbf{J}_- \left| \frac{1}{2}, -\frac{1}{2} \right\rangle = 0, \quad \mathbf{J}_- \left| \frac{1}{2}, +\frac{1}{2} \right\rangle = \hbar \left| \frac{1}{2}, \frac{1}{2} \right\rangle.$$

Their matrix elements in this basis are therefore

$$\mathbf{J}_z \rightarrow \frac{\hbar}{2} \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} = \frac{\hbar}{2} \sigma_3, \quad \mathbf{J}_+ \rightarrow \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \hbar \sigma_+ \mathbf{J}_- \rightarrow \frac{\hbar}{2} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \hbar \sigma_-$$

and

$$\mathbf{J}_x = \frac{1}{2} (\mathbf{J}_+ + \mathbf{J}_-) = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \frac{\hbar}{2} \sigma_x, \quad \mathbf{J}_y = \frac{1}{2i} (\mathbf{J}_+ - \mathbf{J}_-) = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \frac{\hbar}{2} \sigma_y,$$

so that $\vec{\mathbf{J}} = \frac{\hbar}{2}\vec{\sigma}$. Notice that these satisfy the $\mathfrak{so}(3)$ algebra $[\mathbf{J}_i, \mathbf{J}_j] = i\hbar\epsilon_{ijk}\mathbf{J}_k$.

If we keep going to spin-1, $j = 1$, we'll reproduce the 3×3 infinitesimal rotation matrices we started with. Using the steps for spin- $\frac{1}{2}$, we could continue further to $j = \frac{3}{2}$ and construct 4×4 matrices, and so on.

Rotation matrices: There is a useful standard notation for the finite rotations, which is different from the axis-angle representation, namely *Euler angles*. The definition is

$$R(\alpha, \beta, \gamma) = R_z(\alpha)R_y(\beta)R_z(\gamma).$$

Any rotation can be written this way. To relate $R(\alpha, \beta, \gamma)$ and $R(\tilde{n}, \theta)$ is an exercise in geometry which you can do at home. This representation plays nicely with the $|j, m\rangle$ basis because of the $\mathbf{J}_z$ s on the outside as follows: Its matrix elements are

$$\langle j'm' | R(\alpha, \beta, \gamma) | jm \rangle \stackrel{[\mathbf{J}^2, \vec{\mathbf{J}}]=0}{=} \delta_{j,j'} \mathcal{D}_{mm'}^{(j)}(\alpha, \beta, \gamma)$$

where

$$\begin{aligned} \mathcal{D}_{mm'}^{(j)}(\alpha, \beta, \gamma) &= \langle j, m' | e^{-i\alpha\mathbf{J}_z/\hbar} e^{-i\beta\mathbf{J}_y/\hbar} e^{-i\gamma\mathbf{J}_z/\hbar} | j, m \rangle \\ &= e^{-im'\alpha} \underbrace{\langle j, m' | e^{-i\beta\mathbf{J}_y/\hbar} | j, m \rangle}_{\equiv d_{mm'}^{(j)}(\beta)} e^{-im\gamma}. \end{aligned} \quad (4.12)$$

Here $d_{mm}^{(j)}(\beta) \equiv \langle j, m' | e^{-i\beta\mathbf{J}_y/\hbar} | j, m \rangle$ is real since the matrix elements of $\mathbf{J}_y$ are pure imaginary in the j, m basis.

The matrix $\mathcal{D}$ is what we need to know to figure out how R acts on $|jm\rangle$:

$$R(\alpha, \beta, \gamma) |jm\rangle = \sum_{m'} |j, m'\rangle \langle j, m' | R(\alpha, \beta, \gamma) | j, m \rangle = \sum_{m'} |j, m'\rangle \mathcal{D}_{mm'}^{(j)}(\alpha, \beta, \gamma).$$

For $j = 0$: we have $d_{m'm}^{(0)}(\beta) = 1, \mathcal{D}^{(0)} = 1$.

For $j = \frac{1}{2}$:

$$e^{-i\beta\mathbf{J}_y/\hbar} = e^{-i\beta\sigma^y/2} = \cos \frac{\beta}{2} \mathbb{1} - i\sigma^y \sin \frac{\beta}{2} \rightarrow \begin{pmatrix} \cos \frac{\beta}{2} & -\sin \frac{\beta}{2} \\ \sin \frac{\beta}{2} & \cos \frac{\beta}{2} \end{pmatrix}$$

and hence

$$\mathcal{D}^{(\frac{1}{2})} = \begin{pmatrix} e^{-i\alpha/2} & 0 \\ 0 & e^{+i\alpha/2} \end{pmatrix} \begin{pmatrix} \cos \frac{\beta}{2} & -\sin \frac{\beta}{2} \\ \sin \frac{\beta}{2} & \cos \frac{\beta}{2} \end{pmatrix} \begin{pmatrix} e^{-i\gamma/2} & 0 \\ 0 & e^{+i\gamma/2} \end{pmatrix} \quad (4.13)$$

For $j = 1$:

$$\mathcal{D}^{(1)} = \begin{pmatrix} e^{-i\alpha} & & \\ & 0 & \\ & & e^{+i\alpha} \end{pmatrix} \begin{pmatrix} \frac{1}{2}(1 + \cos \beta) & -\frac{1}{\sqrt{2}} \sin \beta & \frac{1}{2}(1 - \cos \beta) \\ \frac{1}{\sqrt{2}} \sin \beta & \cos \beta & -\frac{1}{\sqrt{2}} \sin \beta \\ \frac{1}{2}(1 - \cos \beta) & \frac{1}{\sqrt{2}} \sin \beta & \frac{1}{2}(1 + \cos \beta) \end{pmatrix} \begin{pmatrix} e^{-i\gamma} & & \\ & 0 & \\ & & e^{+i\gamma} \end{pmatrix}$$

SO(3) and SU(2): I have to comment on the relation between these two things.

SU(2) can be defined as the group of 2×2 unitary matrices with determinant one. What does this have to do with the proper rotation group, SO(3)? If we follow the same strategy we took for rotations, and write the general such matrix as $U = e^{-iA}$ we discover that A is a hermitian 2×2 matrix. But such things are spanned by Pauli matrices: $A = \vec{a} \cdot \vec{\sigma}$, as we saw on the first homework. Therefore the generators of SU(2) transformations satisfy the same algebra as proper rotations (in fact, if we normalize them using the Hilbert-Schmidt norm, even the normalization factor of $\frac{1}{2}$ agrees): As Lie algebras (hence lowercase letters), $\mathfrak{su}(2) = \mathfrak{so}(3)$.

But this is not the end of the story. Although the Lie *algebras* are the same, the Lie *groups* are not.

Notice from (4.13) that a rotation by 2π (for example about y) gives $\mathcal{D}^{(\frac{1}{2})}(0, 2\pi, 0) = -\mathbb{1}$ (same for any 2π rotation, *e.g.* about z : $\mathcal{D}^{(\frac{1}{2})}(0, 0, 2\pi) = -\mathbb{1}$). But a 2π rotation does nothing to a vector! This extra sign means that spin- $\frac{1}{2}$ is in fact a *projective* representation of SO(3). It is an actual representation of SU(2).

The same pattern persists for all j : integer j give actual unitary representations of SO(3), while odd-half-integer j give projective representations of SO(3). There is a closed-form expression for the rotation matrices for general j which can be found in Commins' book (equation (7.63)).

4.5 Orbital angular momentum

The d -dimensional angular momentum algebra is realized on the Hilbert space of a particle in d dimensions, with $[\mathbf{x}_i, \mathbf{p}_j] = i\hbar\delta_{ij}$. The essential case arises for $d \geq 2$, where we can focus on two dimensions:

$$\begin{pmatrix} \mathbf{x}' \\ \mathbf{y}' \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}.$$

For small θ , this is

$$\begin{pmatrix} \mathbf{x}' \\ \mathbf{y}' \end{pmatrix} \simeq \begin{pmatrix} 1 & -\theta \\ \theta & 1 \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} - \frac{i}{\hbar} \left[\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}, \mathbf{G} \right]$$

where $\mathbf{G}$ is the generator. Since the momentum is also a d -vector of operators, a similar expression must hold, with the *same* generator $\mathbf{G}$:

$$\begin{pmatrix} \mathbf{p}'_x \\ \mathbf{p}'_y \end{pmatrix} \simeq \begin{pmatrix} 1 & -\theta \\ \theta & 1 \end{pmatrix} \begin{pmatrix} \mathbf{p}_x \\ \mathbf{p}_y \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} \mathbf{p}_x \\ \mathbf{p}_y \end{pmatrix} - \frac{i}{\hbar} \left[\begin{pmatrix} \mathbf{p}_x \\ \mathbf{p}_y \end{pmatrix}, \mathbf{G} \right]$$

It is perhaps not a surprise that these two equations are satisfied by the operator

$$\mathbf{G} = \mathbf{x}\mathbf{p}_y - \mathbf{y}\mathbf{p}_x = \mathbf{L}_z$$

– in the last equality we used the familiar relation in $d = 3$. Angular momentum generates rotations. More generally, the other components are

$$\mathbf{L} = \vec{\mathbf{r}} \times \vec{\mathbf{p}}, \quad \mathbf{L}_i \equiv \epsilon_{ijk} \mathbf{x}_j \mathbf{p}_k$$

(where I use $\vec{\mathbf{r}} \equiv (\mathbf{x}, \mathbf{y}, \mathbf{z})$). Note that there is no ordering ambiguity in this expression, since only the $\mathbf{x}$ and $\mathbf{p}$ associated with the same direction fail to commute, but the ϵ prevents them from appearing together. [\[End of Lecture 18\]](#)

You can check explicitly by commutator algebra that the $\mathbf{L}_i$ satisfy the $\mathfrak{so}(3)$ algebra:

$$[\mathbf{L}_i, \mathbf{L}_j] = i\hbar \epsilon_{ijk} \mathbf{L}_k.$$

It is worth noting that the expression

$$[\mathbf{L}_i, \mathbf{v}_j] = i\hbar \epsilon_{ijk} \mathbf{v}_k$$

says that the operator $\mathbf{v}_j$ transforms like a vector under rotations. It is satisfied by $\mathbf{v}_j \in \{\mathbf{L}_j, \mathbf{x}_j, \mathbf{p}_j\}$.

A consequence of the $\mathfrak{so}(3)$ algebra is that

$$\mathbf{L}^2 = \sum_{i=x,y,z} \mathbf{L}_i^2$$

is a Casimir: $[\mathbf{L}^2, \mathbf{L}_i] = 0$. For the same reasons as the general case, the eigenvalues of $\mathbf{L}^2$ are of the form $\ell(\ell+1)\hbar^2$. However, ℓ is restricted further to be integer (rather than integer or half-integer). One way to see this is to notice that we can write $\mathbf{L}_z$ in terms of harmonic oscillator algebras:

$$\mathbf{L}_z = \mathbf{a}_+^\dagger \mathbf{a}_+ - \mathbf{a}_-^\dagger \mathbf{a}_-,$$

where these are the number operators for two independent SHO algebras:

$$[\mathbf{a}_\pm, \mathbf{a}_\pm^\dagger] = 1, \quad [\mathbf{a}_+, \mathbf{a}_-] = 0, \quad [\mathbf{a}_+, \mathbf{a}_-^\dagger] = 0.$$

They are made of

$$\mathbf{a}_\pm \equiv \frac{1}{2} (\mathbf{x} + i\mathbf{p}_x \pm i(\mathbf{y} + i\mathbf{p}_y)).$$

This has the consequence that the eigenvalues of $\mathbf{a}_\pm^\dagger \mathbf{a}_\pm$ are non-negative integers, and hence the m_z are integers (not half-integers), so ℓ is integer. (This oscillator algebra may seem unmotivated; we'll see another argument later.)

Spherical harmonics

Let us factor out the radial component of the wavefunction of a particle in 3d, $|x, y, z\rangle = |r\rangle \otimes |\theta, \varphi\rangle$. Even better, let's imagine that the radial potential localizes the particle strongly to a particular sphere, so that the radial degree of freedom is stuck in its groundstate.

I am using coordinates where

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = r \begin{pmatrix} \sin \theta \cos \varphi \\ \sin \theta \sin \varphi \\ \cos \theta \end{pmatrix}.$$

In the position-basis for these coordinates, the angular momentum operators take the form

$$L_z = -i\hbar(x\partial_y - y\partial_x) = -i\hbar\partial_\varphi, \quad L_\pm = \hbar e^{\pm i\varphi}(\pm\partial_\theta + i\cot\theta\partial_\varphi).$$

A check on my algebra is that these satisfy the raising and lower algebra: $[L_\pm, L_z] = \pm\hbar L_\pm$. Finally, the total orbital angular momentum takes the form

$$L^2 = -\left(\frac{1}{\sin\theta}\partial_\theta(\sin\theta\partial_\theta) + \frac{1}{\sin^2\theta}\partial_\varphi^2\right).$$

A reason to care about this form of L^2 is that it can be used to rewrite the kinetic term in the Hamiltonian in polar coordinates:

$$\begin{aligned} \mathbf{L}^2 &= (\vec{\mathbf{r}} \times \vec{\mathbf{p}}) \cdot (\vec{\mathbf{r}} \times \vec{\mathbf{p}}) \\ &= \mathbf{r}^2 \mathbf{p}^2 - (\vec{\mathbf{r}} \cdot \vec{\mathbf{p}})^2 + i\hbar \underbrace{\vec{\mathbf{r}} \cdot \vec{\mathbf{p}}}_{=-i\hbar r\partial_r} \\ &= \mathbf{r}^2 \mathbf{p}^2 + \hbar^2 \underbrace{r\partial_r(r\partial_r)}_{=r^2\partial_r^2 + r\partial_r} + \hbar^2 r\partial_r \end{aligned} \tag{4.14}$$

which we can solve for $\mathbf{p}^2$:

$$\mathbf{p}^2 = -\hbar^2 \left(\partial_r^2 + \frac{2}{r}\partial_r \right) + \frac{\mathbf{L}^2}{r^2}.$$

Let's define the wavefunctions of the angular momentum eigenstates to be

$$\langle\theta, \varphi|\ell, m\rangle \equiv Y_{\ell m}(\theta, \varphi) = e^{im\varphi} f_{\ell, m}(\theta). \tag{4.15}$$

These are defined by the eigenvalue equations

$$\mathbf{L}_z |\ell, m\rangle = \hbar m |\ell, m\rangle, \quad \mathbf{L}^2 |\ell, m\rangle = \hbar^2 \ell(\ell+1) |\ell, m\rangle$$

In the second step of (4.15) we used $-\mathbf{i}\partial_\varphi Y_{\ell,m} = mY_{\ell,m}$.

Here's an easy way to see that $\ell \in \mathbb{Z}$: The wavefunction in φ must be periodic under $\varphi \rightarrow \varphi + 2\pi$. The condition $e^{\mathbf{i}m(\varphi+2\pi)} = e^{\mathbf{i}m\varphi}$ says that we need $m \in \mathbb{Z}$. Since $m \in \{-\ell, -\ell+1, \dots, \ell-1, \ell\}$, we must then have $\ell \in \mathbb{Z}$ as well.

The eigenvalue equation for L^2 says:

$$\hbar^2 \ell(\ell+1)Y_{\ell,m} = L^2 Y_{\ell,m} = - \left(\frac{1}{\sin \theta} \partial_\theta (\sin \theta \partial_\theta) + \frac{1}{\sin^2 \theta} \partial_\varphi^2 \right) Y_{\ell,m}$$

which looks like a horrible second-order differential equation. Here are two ways to find the solutions. A bad way is to somehow know that by changing variables to $\mathfrak{x} \equiv \cos \theta$, it takes the form

$$\partial_{\mathfrak{x}} \left((1 - \mathfrak{x}^2) \partial_{\mathfrak{x}} f \right) + \left(\ell(\ell+1) - \frac{m^2}{1 - \mathfrak{x}^2} \right) f = 0.$$

This is the equation satisfied by the associated Legendre polynomials.

A better way to find the solution is to use the $\mathfrak{su}(2)$ algebra: The highest weight state, with $m = \ell$, satisfies $0 = \mathbf{L}_+ |\ell, \ell\rangle$ and hence

$$0 = \langle \theta, \varphi | \mathbf{L}_+ |\ell, \ell\rangle = -\mathbf{i}\hbar e^{\mathbf{i}\varphi} (\mathbf{i}\partial_\theta - \cot \theta \partial_\varphi) e^{\mathbf{i}\ell\varphi} f_{\ell\ell}(\theta) \propto \mathbf{i}\partial_\theta f_{\ell\ell} - \cot \theta \mathbf{i}\ell f_{\ell\ell}$$

which says

$$\partial_\theta f_{\ell\ell} = \ell \cot \theta f_{\ell\ell}(\theta) = \ell \frac{\cos \theta}{\sin \theta} f_{\ell\ell}(\theta)$$

which first-order ODE is solved by

$$f_{\ell\ell}(\theta) = c_{\ell\ell} \sin^\ell \theta.$$

Hence

$$Y_{\ell\ell}(\theta, \varphi) = c_{\ell\ell} e^{\mathbf{i}\ell\varphi} \sin^\ell \theta.$$

Normalization requires

$$1 = |c_{\ell\ell}|^2 \int d\Omega \sin^{2\ell}(\theta) \implies c_{\ell\ell} = \frac{(-1)^\ell}{2^\ell \ell!} \sqrt{\frac{(2\ell+1)!}{4\pi}}$$

where the $(-1)^\ell$ is a useful convention.

To find the general $Y_{\ell,m}$, then we just use the lowering operator:

$$\mathbf{L}_- |\ell, m\rangle = \sqrt{\ell(\ell+1) - m(m-1)} |\ell, m-1\rangle$$

and hence

$$Y_{\ell,m}(\theta, \varphi) = \sqrt{\frac{2\ell+1}{4\pi} \cdot \frac{(\ell-m)!}{(\ell+m)!}} (-1)^m e^{\mathbf{i}m\varphi} P_\ell^m(\cos \theta)$$

where $P_\ell^m(\cos \theta)$ is the associated Legendre polynomial mentioned above. ($Y_{\ell,0}(\theta, \varphi) = \sqrt{\frac{2\ell+1}{4\pi}} P_\ell(\cos \theta)$ is the Legendre polynomial.)

Since the $|\ell, m\rangle$ states are orthonormal, we learn an integral identity:

$$\delta_{\ell\ell'}\delta_{mm'} = \langle \ell', m' | \ell, m \rangle = \int \underbrace{d\Omega}_{\equiv \sin \theta d\theta d\varphi} \langle \ell', m' | \theta, \varphi \rangle \langle \theta, \varphi | \ell, m \rangle = \int d\Omega (Y_{\ell'm'}(\theta, \varphi))^* Y_{\ell m}(\theta, \varphi).$$

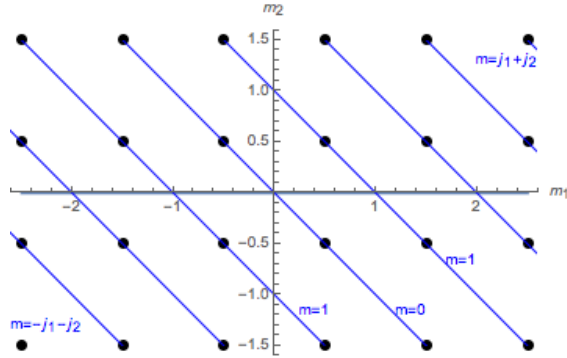
4.6 Addition of angular momentum

Consider what happens if we tensor together two representations of the rotation group, $\vec{J}_1$ on $\mathcal{H}_1$ and $\vec{J}_2$ on $\mathcal{H}_2$. Then on $\mathcal{H} \equiv \mathcal{H}_1 \times \mathcal{H}_2$, the operators

$$\vec{J} \equiv \vec{J}_1 \otimes \mathbb{1}_2 + \mathbb{1}_1 \otimes \vec{J}_2 \equiv \vec{J}_1 + \vec{J}_2$$

also satisfy the angular momentum algebra (4.6). Suppose further that $\mathcal{H}_{1,2}$ are irreps with spin $j_{1,2}$ respectively. Then (the definition of tensor product) $\mathcal{H} = \text{span}\{|j_1, m_1\rangle \otimes |j_2, m_2\rangle\}$. The dimension of $\mathcal{H}$ is $(2j_1 + 1) \times (2j_2 + 1)$. How do rotations act on it?

The states are plotted by (m_1, m_2) in the figure at right for $j_1 = \frac{5}{2}, j_2 = \frac{3}{2}$. The basis states are already eigenstates of $\mathbf{J}_z \equiv (\mathbf{J}_1)_z + (\mathbf{J}_2)_z$ with eigenvalues $m = m_1 + m_2$. So the diagonal blue lines have fixed m . What j do they have? These values range up to $j_1 + j_2$, so this must be the largest value of j of the resulting reducible representation. But the



action on $\mathcal{H}$ is definitely reducible since this spin- $(j_1 + j_2)$ irrep only accounts for $2(j_1 + j_2) + 1$ of the states.

To find the rest, we use induction: remove the spin- $(j_1 + j_2)$ irrep from $\mathcal{H}$. This leaves a single state with $m = j_1 + j_2 - 1$, which must be the highest-weight state for a spin- $(j_1 + j_2 - 1)$ irrep. We can keep doing this, moving to the left in the figure and removing one state from each blue line, until we reach the point where we run out. This happens when $j = |j_1 - j_2|$ (in the figure, $j_1 > j_2$ so $j = |j_1 - j_2| = j_1 - j_2$ (which happens to be $5 - 3 = 2$)). Therefore, if we denote the irrep by $\mathcal{H}_j \equiv \underline{j}$

$$\underline{j}_1 \times \underline{j}_2 = (\underline{j}_1 + \underline{j}_2) \oplus (\underline{j}_1 + \underline{j}_2 - 1) \oplus \cdots \oplus \underline{|j_1 - j_2|}.$$

In the figure, we have $\frac{3}{2} \otimes \frac{5}{2} = \underline{4} \oplus \underline{3} \oplus \underline{2} \oplus \underline{1}$. Notice that the dimensions add up:

(assuming WLOG that $j_1 \geq j_2$),

$$\dim \mathcal{H} = \sum_{j=j_1-j_2}^{j_1+j_2} (2j+1) = (1+j_1+j_2)^2 - (1+j_1-j_2)^2 = (2j_1+1)(2j_2+1)$$

Notice that if j_1 and j_2 have different parity, then we get only half-integer m , while if j_1 and j_2 have the same parity (as in the figure), we only get integer m .

An important special case is when $j_2 = j_1$. Then an interesting operator which commutes with $\vec{\mathbf{J}}$ is the *swap* operator

$$\mathbf{S} \equiv \sum_{m_1, m_2 = -j_1}^{j_1} |j_1, m_1\rangle \otimes |j_1, m_2\rangle \langle j_1, m_2| \otimes \langle j_1, m_1|$$

which interchanges the two factors. Since it commutes with the angular momenta, it can be simultaneously diagonalized. What are the swap eigenvalues of the various j ? The state (at the top right corner)

$$|j = 2j_1, m = 2j_1\rangle = |j_1, m_1 = j_1\rangle_1 \otimes |j_1, m_2 = j_1\rangle_2$$

is the highest-weight state of $\mathcal{H}_{j_1+j_2}$ and is clearly symmetric under interchange:

$$\mathbf{S} |j_1 + j_2, j_1 + j_2\rangle = |j_1 + j_2, j_1 + j_2\rangle.$$

The rest of this multiplet is obtained by acting repeatedly on this state with the lowering operator $\mathbf{J}_- \equiv (\mathbf{J}_1)_- + (\mathbf{J}_2)_-$ which is symmetric $[\mathbf{J}_-, \mathbf{S}] = 0$, so that whole multiplet is symmetric. In the line with $m = j_1 + j_2 - 1$ there is a second state, which must be orthogonal to

$$|j = j_1 + j_2, m = j_1 + j_2 - 1\rangle \propto \mathbf{J}_- |j_1 + j_2, j_1 + j_2\rangle = |j_1 - 1\rangle_1 \otimes |j_1\rangle_2 + |j_1\rangle_1 \otimes |j_1 - 1\rangle_2$$

and hence must have a minus sign - *i.e.* it must be antisymmetric. This pattern continues at each step, and we have

$$\underline{\mathbf{j}}_1 \times \underline{\mathbf{j}}_1 = (\underline{2\mathbf{j}}_1)_+ \oplus \cdots \oplus (\underline{2\mathbf{j}}_1 - 1)_- \oplus (\underline{0})_{(-1)^{2j}}$$

where the $\mathbf{S}$ eigenvalue is indicated in the subscript.

For example:

$$\frac{1}{2} \times \frac{1}{2} :$$

$$|1, 1\rangle = |\uparrow\uparrow\rangle \quad \text{is symmetric.}$$

$$|1, 0\rangle = \frac{1}{\sqrt{2}} \mathbf{J}_- |1, 1\rangle = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \quad \text{is symmetric.}$$

Lowering again (and being careful about the coefficient) we get

$$|1, -1\rangle = |\downarrow\downarrow\rangle \quad \text{is symmetric.}$$

$$|0, 0\rangle = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \quad \text{is antisymmetric.}$$

The singlet $|0, 0\rangle$ has to have a minus sign to make it orthogonal to the $|1, 0\rangle$ state.

The two bases $|j, m\rangle$ and $|m_1\rangle \otimes |m_2\rangle$ are related by

$$(|1, 1\rangle, |1, 0\rangle, |1, -1\rangle, |0, 0\rangle) = (|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

The matrix elements on the right are called Clebsch-Gordan (C-G) coefficients.

More generally:

$$|j, m\rangle = \sum_{m_1, m_2} |j_1, m_1\rangle \otimes |j_2, m_2\rangle \begin{pmatrix} j_1 & j_2 & j \\ m_1 & m_2 & m \end{pmatrix}$$

Here

$$\langle j_1, m_1; j_2, m_2 | j, m \rangle \equiv \begin{pmatrix} j_1 & j_2 & j \\ m_1 & m_2 & m \end{pmatrix}$$

is a unitary matrix since it implements a change of basis. In fact, it is actually real (up to an arbitrary overall phase which we may set to 1): Therefore the inverse formula is

$$|j_1, m_1; j_2, m_2\rangle = \sum_j |jm\rangle \langle j_1 m_1; j_2 m_2 | jm \rangle.$$

An important application of this formula is to determine the rotation matrices on the tensor product of two representations: since the generators on the two factors $\vec{\mathbf{J}}_1$ and $\vec{\mathbf{J}}_2$ commute, a rotation on the product space is $R = R_1 \otimes R_2$ and its matrix elements are

$$\mathcal{D}_{m'm}^{(j)}(\alpha\beta\gamma) = \sum_{m'_1, m'_2} \langle j_1 m_1; j_2 m_2 | jm \rangle \langle j_1 m'_1; j_2 m'_2 | jm' \rangle \mathcal{D}_{m'_1 m_1}^{(j_1)}(\alpha\beta\gamma) \mathcal{D}_{m'_2 m_2}^{(j_2)}(\alpha\beta\gamma).$$

This equation can be used to determine the matrices for a given j in terms of those for lower j .

The C-G coefficients may be determined recursively by starting with the largest- j highest weight state, repeatedly applying the lowering operator using

$$\mathbf{J}_\pm |\ell, m\rangle = C_\pm^{\ell, m} |\ell, m \pm 1\rangle, \quad C_\pm^{\ell, m} \equiv \sqrt{\ell(\ell+1) - m(m \pm 1)}$$

and then using Gram-Schmidt to make an ON basis for the remaining states, which will have a new largest- j highest weight state. The key step is the recursion relation

$$C_{\mp}^{j'm'} \langle j', m' \mp 1 | j'' m''; j, m \rangle - C_{\pm}^{j''m''} \langle j', m' | j'' m'' \pm 1; j, m \rangle = C_{\pm}^{jm} \langle j' m'; j, m \pm 1 | j' m' \rangle. \quad (4.16)$$

(which we will need in the next subsection). This relation comes by comparing the state you get by raising and then lowering to the one you get by lowering and then raising. They are related by the algebra $[\mathbf{J}_+, \mathbf{J}_-] = 2\hbar \mathbf{J}_3$. A (not so useful) general formula (the solution of this recursion) for the C-G coefficients can be found in equation (7.119) of Commins' book.

An important application of the addition of angular momentum is to the situation where a particle with spin has a coupling between its spin and its orbital degrees of freedom (spin-orbit coupling) which means that only the combination $\vec{\mathbf{S}} + \vec{\mathbf{L}}$ is a symmetry.

4.7 Wigner-Eckhart Theorem

Recall that I said that the transformation rule

$$[\mathbf{J}_i, \mathbf{v}_j] = i\epsilon_{ijk} \mathbf{v}_k$$

says that the operator $\vec{\mathbf{v}}$ transforms as a vector under rotations. We can decompose this multiplet into definite ℓ, m by writing

$$\mathbf{v}_{\ell=1}^{m=0} = \mathbf{v}_z, \quad \mathbf{v}_{\ell=1}^{m=\pm 1} = \mathbf{v}_x \pm i\mathbf{v}_y \equiv \mathbf{v}_{\pm}.$$

They satisfy

$$[\mathbf{J}_z, \mathbf{v}_{\ell}^m] = m\mathbf{v}_{\ell}^m, \quad [\mathbf{J}_{\pm}, \mathbf{v}_{\ell}^m] = C_{\pm}^{\ell, m} \mathbf{v}_{\ell}^{m\pm 1} \quad (4.17)$$

(for $\ell = 1$ so far).

This means that under a finite rotation, it does

$$\mathbf{v}_{\ell}^m \rightarrow R \mathbf{v}_{\ell}^m R^{-1} = \sum_{m'} \mathcal{D}_{m', m}^{\ell} \mathbf{v}_{\ell}^{m'}. \quad (4.18)$$

You can see that we can generalize this to other values of ℓ . Such things are called (*irreducible*) *spherical tensors*. They are defined by their transformation property, either the infinitesimal one (4.17) or the finite one (4.18). The proof that these are equivalent follows just by taking an infinitesimal rotation in (4.18) and can be found in Commins §7.10.

The Wigner-Eckhart theorem follows by sandwiching (4.17) between $\langle j', m' |$ and $| j'', m'' \rangle$. The left equation gives:

$$(m' - m'') \langle j', m' | \mathbf{v}_\ell^m | j'', m'' \rangle = m \langle j', m' | \mathbf{v}_\ell^m | j, m'' \rangle$$

which says that the matrix element $\langle j', m' | \mathbf{v}_\ell^m | j'', m'' \rangle$ is only nonzero if the J^z s add up: $m' = m + m''$.

The right equations give recursion relations:

$$C_{\mp}^{j'm'} \langle j', m' \mp 1 | \mathbf{v}_\ell^m | j'', m'' \rangle - C_{\pm}^{j''m''} \langle j', m' | \mathbf{v}_\ell^m | j'', m'' \pm 1 \rangle = C_{\pm}^{jm} \langle j', m' | \mathbf{v}_\ell^{m \pm 1} | j'', m'' \rangle .$$

On the other hand, the C-G coefficients themselves satisfy the same recursion relation (4.16)! Therefore, the matrix element in question is proportional to the C-G coefficient: $\langle j', m' | \mathbf{v}_\ell^m | j'', m'' \rangle \propto \langle j' m' | j'' m''; \ell, m \rangle$. The key piece of information in this statement (the Wigner-Eckhart theorem)

$$\langle j', m' | \mathbf{v}_\ell^m | j'', m'' \rangle = \underbrace{\langle j' || \mathbf{v}_\ell || j'' \rangle}_{\text{reduced matrix elt}} \langle j' m' | j'' m''; \ell, m \rangle$$

is that the proportionality constant (called the *reduced matrix element*) does not depend on m, m', m'' , but only on ℓ, j', j'' : This contains all the dynamical information, while the m -dependent bit is purely geometric, independent of the operator in question $\mathbf{v}$, and can be looked up in a table.

This equation leads to *selection rules*: not only must $m + m'' = m'$ in order for the matrix element to be nonzero, but j' must lie in the tensor product of j'' and ℓ , that is

$$j' \in \{j'' + \ell, j'' + \ell - 1, \dots, |j'' - \ell|\}.$$

So for example, a scalar operator only connects states of the same j , a vector operator $\ell = 1$ changes j only by $-1, 0$ or 1 , and so on.

These selection rules are extremely important, for example in the study of emission and absorption of radiation by atoms.

5 Entanglement

It is unfortunate that we didn't get to discuss this important subject (which I claimed was the essence of quantum mechanics at some point) in more detail this quarter. As a consolation, here is a brief discussion extracted from a previous set of notes. Come to my office hours and ask me questions about it.

Consider the following state of two spin- $\frac{1}{2}$ particles

$$|\text{Bohm}\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle \otimes |\downarrow\rangle - |\downarrow\rangle \otimes |\uparrow\rangle) \equiv \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) . \quad (5.1)$$

(Sometimes called an EPR pair.) Notice that this state is a *singlet* of the total spin:

$$\vec{\sigma}_T \equiv \vec{\sigma}_A \otimes \mathbb{1}_B + \mathbb{1}_A \otimes \vec{\sigma}_B , \quad (5.2)$$

(We will sometimes write things like this as $\sigma_T = \sigma_A + \sigma_B$.) that is (this is three equations):

$$\vec{\sigma}_T |\text{Bohm}\rangle = 0 . \quad (5.3)$$

This fact (that this is the state of two qbits with zero total angular momentum) can guarantee that when a spinless object decays into this pair of spin-1/2 particles, this is the spin state they end up in, and lets us imagine preparing this 2-qbit state.

In the state $|\text{Bohm}\rangle$, the two spins are *maximally entangled*. The meaning of this term is: if we compute the reduced density matrix for A , tracing over the states of B , we get an answer proportional to the identity:

$$\begin{aligned} \rho_A &\equiv \text{tr}_B |\text{Bohm}\rangle \langle \text{Bohm}| \\ &= \sum_{\alpha_B = \uparrow, \downarrow} \langle \alpha_B| \frac{1}{\sqrt{2}} (|\uparrow\rangle_A \otimes |\downarrow\rangle_B - |\downarrow\rangle_A \otimes |\uparrow\rangle_B) \frac{1}{\sqrt{2}} (\langle \uparrow|_A \otimes \langle \downarrow|_B - \langle \downarrow|_A \otimes \langle \uparrow|_B) |\alpha_B\rangle_B \\ &= \frac{1}{2} (-|\downarrow\rangle_A \langle \downarrow|_A + |\uparrow\rangle_A \langle \uparrow|_A) = \frac{1}{2} \mathbb{1} \end{aligned}$$

(the proportionality constant is fixed by $\text{tr} \rho = 1$ to be $1/\dim \mathcal{H}_A$). Recall that this reduced density matrix for A encodes all the results of measurements that we can do if we only have access to A . This means that if we can't play with B also, we have absolutely no information about A !

This is very dramatic: we cannot speak about 'the wavefunction of A ' independently of B . A consequence of this is that the outcomes of experiments done on A and B are correlated in a way which is impossible to produce classically with local rules. *Bell's inequalities* are a quantification of this statement.

A more dramatic and easier to explain example is the following.

5.0.1 GHZM states

[Le Bellac 6.3.4, [this lecture by Sidney Coleman](#)] Something even more dramatic can happen with three qbits: a single measurement outcome with a perfect anti-correlation between the correct answer from QM and the expectation from local realism.

Consider the following state of 3 qbits [GHZM stands for [Greenberger](#), [Horne](#), [Zeilinger](#), Mermin]:

$$|GHZM\rangle = \frac{1}{\sqrt{2}} (|\uparrow\uparrow\uparrow\rangle - |\downarrow\downarrow\downarrow\rangle) \in \mathcal{H} \equiv \mathcal{H}_a \otimes \mathcal{H}_b \otimes \mathcal{H}_c .$$

I am writing $|\uparrow\uparrow\uparrow\rangle \equiv |\uparrow_z\rangle_a \otimes |\uparrow_z\rangle_b \otimes |\uparrow_z\rangle_c$ to avoid clutter. We can imagine preparing three qbits in this state and sending them to distant laboratories, labelled a, b, c . The observers at these laboratories measure the spins along various axes. Let A_x denote the result of measuring the x -spin of particle a by the observer who catches it, B_y denote the result of measuring the y -spin of particle b , etc...; each of these satisfies $A_x^2 = 1$ since $A_x = \pm 1$.

The classical situation to which we'd like to compare is if these measurements are all completely independent and don't care about each others' results. If all these measurements could be done independently (as in a classical world with local hidden variables), the outcomes of these results would obviously satisfy

$$A_x B_x C_x = (A_x B_y C_y)(A_y B_x C_y)(A_y B_y C_x) \quad (5.4)$$

(using $A_y^2 = B_y^2 = C_y^2 = 1$).

Now we do quantum mechanics instead. Consider the following operators acting on $\mathcal{H}_a \otimes \mathcal{H}_b \otimes \mathcal{H}_c$:

$$\Sigma_a \equiv \sigma_x \otimes \sigma_y \otimes \sigma_y, \quad \Sigma_b \equiv \sigma_y \otimes \sigma_x \otimes \sigma_y, \quad \Sigma_c \equiv \sigma_y \otimes \sigma_y \otimes \sigma_x .$$

Now I am relying on the order to keep track of which qbit each sigma is acting on – the fully explicit expression is:

$$\Sigma_a \equiv \sigma_x \otimes \sigma_y \otimes \sigma_y \equiv \sigma_{ax} \otimes \sigma_{by} \otimes \sigma_{cy} .$$

Notice that the label on the Σ_i indicates which of the three factors has the σ_x . The operator

$$\Sigma \equiv \sigma_x \otimes \sigma_x \otimes \sigma_x$$

will also be useful.

Claims:

- (a) These operators Σ_i all commute with each other. This can be checked using the properties of the Paulis – the important point is that Σ_i and Σ_j differ by an even number of anti commutations. They also commute with Σ .
- (b) It is therefore possible for them to be simultaneously measured. Notice that this will require some coordinated effort between the observers in those distant laboratories.
- (c) The Σ s square to 1 and hence have eigenvalues ± 1 .
- (d) $\Sigma_i |GHZM\rangle = |GHZM\rangle, i = a, b, c$. The state $|GHZM\rangle$ above is an eigenstate of all three Σ s with eigenvalue +1. Act with each Σ_i on it and see what happens! (Recall that $\sigma_x |\uparrow\rangle = |\downarrow\rangle \dots$)
- (e) $\Sigma |GHZM\rangle = -|GHZM\rangle$ – it's also an eigenstate of Σ , with eigenvalue -1 .

Now comes the icepick to the forehead: If the observers measure one x component and two y components of the qbits in the GHZM state, the outcome is an eigenvalue of one of the operators Σ_i , with eigenvalue +1:

$$A_x B_y C_y = +1, \quad A_y B_x C_y = +1, \quad A_y B_y C_x = +1 \quad . \quad (5.5)$$

On the other hand, if two of the observers decide to rotate their polarizers so that they all measure in the x direction, they get the eigenvalue of Σ :

$$A_x B_x C_x = -1 \quad . \quad (5.6)$$

Now compare this with the classical expectation in (5.4)!

They disagree completely. The quantum answer is the correct one in our world. There are no numbers $\pm$ that we can assign to A_x, A_y etc which will produce both results (5.5) and (5.6) – they simply don't exist.

Notice that although the Σ s all commute with each other, they are made from operators which do not individually commute (*e.g.* σ_x and σ_y acting on any one of the qbits). We ran into trouble in the classical calculation because we assigned outcomes to these measurements of σ_x that we didn't do!

[\[End of Lecture 19\]](#)

We can use this discovery here to do something that you can't do without QM [I learned this from Boccio]:

Consider a game for three players A, B, C . They are told that they will be separated from each other and each one will be asked one of two questions, say X or Y , whose allowed answers are +1 or -1 . And *either*

(a) all players will be asked the same question X

or

(b) one of the three players will be asked X and the other two will be asked Y .

After having been asked X or Y no player can communicate with the others until all three players have given all their answers. To win, the players must give answers such that, in case (a) the product of the three answers is -1 , and in case (b) the product of the answers is $+1$.

What we've shown above is:

(1) There is no classical strategy that gives a certain win for the team.

(2) There is a quantum strategy which gives a certain win, as long as each player can take one of three qbits along with them, and the three qbits are in the GHZ state.
