Phys 412-1 Quantum Mechanics

Kaan Simsek

ksimsek@u.northwestern.edu | kagsimsek.github.io

(Dated: 2022 Fall)

(Last updated: November 11, 2022)

Abstract

This document contains the TA notes for the course Phys 412-1 Quantum Mechanics in the fall quarter of the academic year 2022-2023.

CONTENTS

I. DISCUSSION 1 (SEP 29)

What is an observable? It is an *abstract operator* which can be *represented* in a matrix in terms of some entries using a suitable *basis*. Suppose we have an observable A. The matrix representation of A is hermitian. We need to know what we are talking about when we say a matrix is hermitian.

Hermitian matrices have the three important properties:

- They have *real eigenvalues*.
- Their *eigenvectors* are *orthogonal* [1].
- Their eigenvectors form a *complete basis*.

The last item is purely mathematical but it is physically useful.

Now suppose I want to write down the eigenvalue equation for my operator A:

$$A|n\rangle = a_n|n\rangle \tag{1}$$

I'll be very simplistic in my notation and just put the quantum number into my kets and bras. A more proper notation is for the eigenkets is $|A;n\rangle$; namely, you'd indicate the operator, as well. Now, here we know that $a_n \in \mathbb{R}$ for all n [2]. As for the eigenkets, we know that they are orthogonal. But let's not worry about the magnitudes, so suppose everything is normalized. Then, we have the orthonormality relation

$$\langle n|m\rangle = \delta_{nm} \tag{2}$$

Let's talk about the completeness now. But to do that, I'll introduce the *projection operator*, Λ_n. Suppose I write down an operator like $|n\rangle\langle m|$. What does this guy do for a living? It is clearly not a ket vector due to the bra on the right. It is a tensor, which can be represented by some matrix in a proper basis. Consider the operator $|n\rangle\langle n|$. What does this do? If I give you a generic state ket, say $|\psi\rangle$ and tell you to act this operator on this state, you'd get $|n\rangle\langle n|\psi\rangle$. The second factor, i.e. $\langle n|\psi\rangle$, is certainly a scalar. We have a ket, $|n\rangle$, next to it, so the operator $|n\rangle\langle n|$ most certainly produces this ket, times some complex number. We call this number, the *probability amplitude* (or the *transition amplitude*). It measures *how much* $|n\rangle$ *is contained in* $|\psi\rangle$, for lack of a better explanation.

This is no different than the usual projection of vectors. Suppose I have this vector $\boldsymbol{v}$ and I want to project it onto an axis in the direction of some other vector $\boldsymbol{n}$. Then, we express this new, projected vector as $(\boldsymbol{v} \cdot \hat{\boldsymbol{n}})\hat{\boldsymbol{n}}$, where I normalized my direction vector. This object, $\langle n|\psi\rangle$ has 1:1 correspondence to the object $\hat{\boldsymbol{n}} \cdot \boldsymbol{v}$.

The moral of the story is, $|n\rangle\langle n|$ projects any state onto the state $|n\rangle$. We define our projection operator then as

$$\Lambda_n := |n\rangle\langle n| \tag{3}$$

Let's study some properties of this projection operator. What happens if I act $\Lambda_n\Lambda_m$ ($m \neq n$) on some general state $|\psi\rangle$? Without doing explicit mathematics, we immediately see that we get zero. The reason is, we project our initial state onto $|m\rangle$ first, but then we try to project this ket onto $|n\rangle$. We cannot do that since our eigenkets are orthogonal. Let's prove this mathematically:

$$\Lambda_n\Lambda_n|\psi\rangle = |n\rangle\langle n|m\rangle\langle m|\psi\rangle \tag{4}$$

This is equal to zero because $\langle n|m\rangle = 0$. Suppose now I act Λ_n on $|\psi\rangle$ twice. What do I get? I get $|n\rangle$ again. Once I'm in this state, I don't go anywhere. So we may write

$$\Lambda_n\Lambda_m = \Lambda_n\delta_{nm} \tag{5}$$

Now let's go back to the completeness. Suppose I have this vector $\boldsymbol{v}$ in $3D$. How do I get the components along the Cartesian axes? I take the dot product of the vector and tack in the unit vector in that direction. So I'm actually projecting my vector onto all these axes. I also know that the three-dimensional real space is spanned by $\{\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}}\}$ *completely*. There is no vector left out. You can express any vector using these unit vectors. An identical story takes place here. If I have a general state $|\psi\rangle$, then, in order to be able express it using the eigenkets of this operator A, I need to project it onto each and every one of these eigenkets:

$$|\psi\rangle = |1\rangle\langle 1|\psi\rangle + |2\rangle\langle 2|\psi\rangle + \cdots = \sum_n |n\rangle\langle n|\psi\rangle = \sum_n \Lambda_n|\psi\rangle \tag{6}$$

We have discovered something now. Notice that we have $|\psi\rangle$ on both sides, namely a *vector* equal to itself with some operator in front. This can mean only one thing:

$$\sum_n \Lambda_n = \mathbb{1} \tag{7}$$

This is called the *completeness* relation. Inserting a sum of eigenkets and bras between two operators, which we'll do every now and then, is called the *resolution of identity*.

* * *

I like my projection operators because that's how I (used to) compute my transitions (and hopefully we'll do it together in the third quarter). If I want to measure the probability of an event taking place between some initial and final states, I immediately write down the transition amplitude as

$$T_{f\leftarrow i} = \langle f|\Lambda|i\rangle \tag{8}$$

where Λ is effectively the projection operator that causes this transition. By squaring this element, we'll get the probability. This may seem a bit vague but it will become much clearer when we start talking about time-dependent phenomena.

Let's do a little exercise now in which I'll stress the importance of these projections operators again. Suppose I have this function f and, whatever it is, I want to evaluate it at my observable A represented in a suitable basis. What can I do?

Whenever in doubt, expand in a power series. That's one of my top quotes by anyone on this planet. Fermi said that, and that's what we are going to do here. Suppose I pretend to expand my function f in a Taylor series but I'll do it in a vague manner, i.e. I'll absorb all the factors into a constant, c_j, and just focus on the expansion parameter, namely our operator, A:

$$f(A) = \sum_{j\geq 0} c_j A^j \tag{9}$$

Now, I want to express my operator in terms of *something*. That something will turn out to be its eigenvalues and eigenkets. Suppose I take my operator A and multiply it by 1:

$$A = A1 = A\left[\sum_n \Lambda_n\right] = A\sum_n |n\rangle\langle n| \tag{10}$$

Then I know how to act A on the ket:

$$A = \sum_n a_n |n\rangle\langle n| \tag{11}$$

This is called the *spectral decomposition* of an operator in terms of its eigenvalues and eigenkets. Let's take the powers of this operator:

$$A = \sum_n a_n |n\rangle\langle n| = \sum_n a_n \Lambda_n \tag{12}$$

$$A^2 = \left[\sum_n a_n |n\rangle\langle n|\right]\left[\sum_m a_m |m\rangle\langle m|\right] = \sum_n \sum_m a_n a_m |n\rangle \underbrace{\langle n|m\rangle}_{\delta_{nm}} \langle m| = \sum_n a_n^2 |n\rangle\langle n| = \sum_n a_n^2 \Lambda_n \tag{13}$$

or, let's use the properties of the projection operator:

$$A^3 = \left[\sum_n a_n \Lambda_n\right]\left[\sum_m a_m \Lambda_m\right]\left[\sum_k a_k \Lambda_k\right] = \sum_n \sum_m \sum_k a_n a_m a_k \underbrace{\underbrace{\Lambda_n \Lambda_m}_{\Lambda_n \delta_{nm}} \Lambda_k}_{\Lambda_n \delta_{nm} \delta_{nk}} = \sum_n a_n^3 \Lambda_n = \sum_n a_n^3 \Lambda_n \tag{14}$$

So we get the idea:

$$A^j = \sum_n a_n^j \Lambda_n \tag{15}$$

With this, our series expansion becomes

$$f(A) = \sum_{j \geq 0} c_j \left[\sum_n a_n^j \Lambda_n\right] = \sum_n \left[\sum_{j \geq 0} c_j a_n^j\right] \Lambda_n = \sum_n f(a_n) \Lambda_n \tag{16}$$

This way, we generalize the spectral decomposition into any functional form. Now, we don't need to worry about the operator itself. We just deal with the projections, which are easier to work with.

* * *

Now let's do an example on representations and measurements. Suppose I give you an observable A with the following properties:

$$\langle 1|A|1\rangle = a \tag{17}$$

$$\langle 1|A|2\rangle = b \tag{18}$$

$$\langle 2|A|1\rangle = b \tag{19}$$

$$\langle 2|A|2\rangle = a \tag{20}$$

Then, I give you a general state:

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|1\rangle + i|2\rangle\right) \tag{21}$$

The question is, what are the probabilities that if I measure A on this state (say, the energy or the spin or any other observable if you want to leave this abstract realm), I get a, b, $a+b$, $a-b$ (in separate measurements, not within the same measurement)?

$$* * *$$

Measurement is nothing but taking projections. Using the projection operator and just the very basic assumptions about a measurement, Schwinger created a *measurement formalism* of quantum mechanics, which turned out to be identical to Schrödinger's wavefunction approach and Heisenberg's matrix mechanics.

We start with the *spectrum* of this operator A. Here, spectrum is just a fancy way of saying its eigenvalues and eigenstates. At this point, we are going to resort to matrices because it is highly convenient.

Now, we are given the A *sandwiches* above. What do we do with them? They appear to be the matrix elements of the operator A in the $\{|1\rangle, |2\rangle\}$ basis. So we may write

$$A \doteq \begin{pmatrix} a & b \\ b & a \end{pmatrix} \tag{22}$$

where $\doteq$ is read *can be represented by*. I just put a in the $(1,1)$ and $(2,2)$ entries and b in the $(1,2)$ and $(2,1)$ entries. How did I do that? This is a step that's usually taken without saying but it's crucial to be aware of that. We have tacitly assumed the representations

$$|1\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |2\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{23}$$

Why? Because it's convenient. Now, I want to obtain the eigenvalues and the eigenvectors of this matrix. Let's use Mathematica [3].

```
A = {
    {a, b},
    {b, a}
};
```

Mathematica is row-major, so we enter the components row by row as we write them on paper. The newlines and the spacing that I'm using here are totally optional. Note the semicolon at the end. It just suppresses the output if you are working with a Mathematica notebook. Let's now play with this.

```
EigenValues[A]
EigenVectors[A]
```

When you evaluate this, the first output should be like

```
{a - b, a + b}
```

and the second output is

```
{{-1, 1}, {1, 1}}
```

So the eigenvalues are

$$\lambda_{\pm} = a \pm b \tag{24}$$

The unnormalized eigenvector corresponding to the *plus* eigenvalue is

$$|\widetilde{+}\rangle \doteq \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{25}$$

and the unnormalized eigenvector corresponding to the *minus* eigenvalue is

$$|\widetilde{-}\rangle \doteq \begin{pmatrix} -1 \\ 1 \end{pmatrix} \tag{26}$$

I'm going to normalize the *plus* eigenket by $1/\sqrt{2}$ and the *minus* eigenket by $-1/\sqrt{2}$. There is nothing that prevents me from choosing some other factor—even a complex one—but this is what I like. So, the normalized eigenkets are

$$|\pm\rangle \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix} \tag{27}$$

But remember these are just representations. We don't want to work with representations because they are just tools. Let's express the eigenkets in the basis $\{|1\rangle, |2\rangle\}$—even though they look abstract, it's these states that are our reality:

$$|\pm\rangle \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \pm \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}}|1\rangle \pm \frac{1}{\sqrt{2}}|2\rangle = \frac{|1\rangle \pm |2\rangle}{\sqrt{2}} \tag{28}$$

With this, let's perform a spectral decomposition of the observable A:

$$\begin{aligned} A &= \sum_n a_n |n\rangle\langle n| \\ &= (a+b)|+\rangle\langle +| + (a-b)|-\rangle\langle -| \end{aligned} \tag{29}$$

where the projection operator to the *plus* eigenstate is given by

$$\Lambda_+ = |+\rangle\langle +| = \frac{|1\rangle + |2\rangle}{\sqrt{2}} \frac{\langle 1| + \langle 2|}{\sqrt{2}} \tag{30}$$

and the projection operator to the *minus* eigenstate is given by

$$\Lambda_- = |-\rangle\langle -| = \frac{|1\rangle - |2\rangle}{\sqrt{2}} \frac{\langle 1| - \langle 2|}{\sqrt{2}} \tag{31}$$

Finally, let's do some measurement. The spectral decomposition tells you what you can get if you try to measure the observable A on any state. You can either get $a+b$ or $a-b$. There is no other possibility. In a way, you can only measure what is in your spectrum [4]. Therefore, without doing any other calculation, we can directly conclude that the probability of measuring A to be a or b on any state is 0.

We can get $a+b$ and $a-b$, so let compute the probability amplitudes. We have

$$\Lambda_+ |\psi\rangle = \frac{|1\rangle + |2\rangle}{\sqrt{2}} \frac{\langle 1| + \langle 2|}{\sqrt{2}} \frac{|1\rangle + i|2\rangle}{\sqrt{2}} \tag{32}$$

Noting that $\langle 1|2\rangle = 0$ and $\langle 1|1\rangle = \langle 2|2\rangle = 1$ (because that's how we chose these basis vectors), we get

$$\Lambda_+ |\psi\rangle = |+\rangle \frac{1+i}{2} \tag{33}$$

so the probability of measuring A to be $a+b$ is

$$P(a+b) = \left| \frac{1+i}{2} \right|^2 = \frac{1}{2} \tag{34}$$

It is just a cute exercise of taking the inner product of 1's and 2's to show that measuring A to be $a-b$ has the same probability (so that the total probability is 1).

* * *

At some point in this discussion session, I mentioned that you can do a lot of amazing things on Mathematica other than just symbolic calculation. I claimed that you can even

compose music with it. Just for the sake of fun, here is a Mathematica notebook where you can create a random composition, which you then may extract as a .wav or .mp3 file using the `Export` command.

https://github.com/kagsimsek/random_compose (35)

If you run it a couple of times, a good melody will come out eventually.

II. DISCUSSION 2 (OCT 6)

Let's do some index manipulation. We start with the well-known mathematical objects that carry some indices:

- Position vector, momentum vector, electric-field vector, etc.: v_i
- The moment of inertia tensor, the quadrupole moment tensor, the field-strength tensor, etc.: T_{ij}
- The unit symmetric tensor, i.e. the Kronecker delta: δ_{ij} (given by `KroneckerDelta[i, j]` on Mathematica)
- The *totally antisymmetric structure constants*: f_{ijk}. Note that the Levi-Civita tensor falls into this class. It's just the structure constant of $SU(2)$, if you want some formalism. Note that *totally symmetric* means here that if you flip any pair of indices, your tensor will receive a minus sign. With this in mind, we can safely say that if any of the two indices are the same, the value of such tensors is zero for that component; i.e. $\epsilon_{112} = -\epsilon_{112}$, where I flipped the 1's, so we have $\epsilon_{112} = 0$. (The Levi-Civita tensor is a built-in tensor on Mathematica, as well. We write ϵ_{ijk} as `LeviCivitaTensor[3][[i, j, k]]`, where 3 here means that we are in the three-dimensional space.)

Now, we say that quantities with only one index are rank-1 tensor, those with two are rank-2, etc. if you want even more formalism. But that's just a language. The important point is the ranges of these indices, which define the size of the matrices that we can form to represent these objects. For an elementary vector, we have i running from 1 to 3 in the good ol' Cartesian space. But why not generalize this? We can have an infinitely large vector space if we let i run from 1 to infinity. That's important to keep in mind but for what follows, we won't really dig into that.

Let's do some index manipulation. We start with the dot product of two vectors. Suppose I have $\boldsymbol{a}$ and $\boldsymbol{b}$ with the same size, i.e. suppose they are both in $\mathbb{R}^n$. Then,

$$\boldsymbol{a} \cdot \boldsymbol{b} = a_1 b_1 + \cdots a_n b_n = \sum_{i=1}^{n} a_i b_i \tag{36}$$

Here, we have expressed the same quantity using an explicit summation over the indices. i here—or any index that's being summed over—is called a *dummy* index. With this, we are ready to drop the summation symbol:

$$\boldsymbol{a} \cdot \boldsymbol{b} = a_i b_i \tag{37}$$

How comfortable is that! Einstein once that his greatest contribution to physics was to drop this summation symbol over repeated indices. By convention, we assume any two repeated indices are meant to be summed over. If you see something like $a_i b_i c_i$, be careful. We need to be aware of the context here, so I won't go into that.

Now, let's do a cross product. Suppose I have $\boldsymbol{a}$ and $\boldsymbol{b}$ in the familiar three-dimensional space. Then, their cross product is given by the following determinant:

$$\boldsymbol{a} \wedge \boldsymbol{b} = \det \begin{pmatrix} \hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix} \tag{38}$$

Note the use of the wedge product here. In most cases, it is used interchangeably by the traditional $\times$ symbol but the latter looks like x, so I prefer the former. Let's expand this cross product:

$$\boldsymbol{a} \wedge \boldsymbol{b} = \hat{\boldsymbol{x}}(a_2 b_3 - a_3 b_2) + \hat{\boldsymbol{y}}(a_3 b_1 - a_1 b_3) + \hat{\boldsymbol{z}}(a_1 b_2 - a_2 b_1) \tag{39}$$

Suppose I want the first component of this cross product.

$$(\boldsymbol{a} \wedge \boldsymbol{b})_1 = a_2 b_3 - a_3 b_2 \tag{40}$$

We have this beautiful cyclic pattern here. Whenever we see something like this, it's best to use resort to the Levi-Civita tensor. Now, I just want to test this out. Let's try the following:

$$(\boldsymbol{a} \wedge \boldsymbol{b})_1 \stackrel{?}{=} \epsilon_{1jk} a_j b_k \quad \text{(sum implied)} \tag{41}$$

For this summation, we say that j and k runs from 1 to 3. Let's expand this:

$$(\boldsymbol{a} \wedge \boldsymbol{b})_1 = \epsilon_{111} a_1 b_1 + \epsilon_{112} a_1 b_2 + \epsilon_{113} a_1 b_3 + \ldots + \epsilon_{133} a_3 b_3 \tag{42}$$

We have nine terms here. But most of them are zero. We want to remember the convention that $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1$ and that if you flip any indices, you'll get -1, and that

if any of the indices are repeated, you get zero. So, we have only two nontrivial terms in this summation, for which all the indices are different:

$$(\boldsymbol{a} \wedge \boldsymbol{b})_1 = \epsilon_{123} a_2 b_3 + \epsilon_{132} a_3 b_2 \tag{43}$$

The first Levi-Civita is just $+1$ and the second one is -1. So we have

$$(\boldsymbol{a} \wedge \boldsymbol{b})_1 = a_2 b_3 - a_3 b_2 \tag{44}$$

It seems to be working. That's what we want, then. Notice how the fact that the cross product is anticommutative is consistent here with the definition of the Levi-Civita tensor:

$$(\boldsymbol{b} \wedge \boldsymbol{a})_1 = b_2 a_3 - b_3 a_2 = -(a_2 b_3 - a_3 b_2) = -\epsilon_{1jk} a_j b_k = -(\boldsymbol{a} \wedge \boldsymbol{b})_1 \tag{45}$$

Next, let's *contract* $\boldsymbol{a} \wedge \boldsymbol{b}$ with another vector. Here—and most of the time—*contraction* means taking the dot product to produce a scalar. Let's investigate the *triple product*. The question is, how would you indexify $\boldsymbol{a} \wedge \boldsymbol{b} \cdot \boldsymbol{c}$? Notice that I don't have to put some parentheses to emphasize the multiplication order, it's just what it is; namely, $\boldsymbol{a} \wedge (\boldsymbol{b} \cdot \boldsymbol{c})$ is a meaningless quantity because $\boldsymbol{b} \cdot \boldsymbol{c}$ is a scalar and we can't take the cross of of a scalar with a vector.

We start with the dot product because it's easier to expand, compared to the cross product. Suppose I define $\boldsymbol{u} := \boldsymbol{a} \wedge \boldsymbol{b}$ for a moment to write

$$\boldsymbol{u} \cdot \boldsymbol{c} = u_i c_i \tag{46}$$

I'm going to insert a Kronecker delta here:

$$\boldsymbol{u} \cdot \boldsymbol{c} = \delta_{ij} u_i c_j \tag{47}$$

Why does it work? Well, let's expand it to confirm:

$$\boldsymbol{u} \cdot \boldsymbol{c} = \delta 11 u_1 c_1 + \delta_{12} u_1 c_2 + \delta_{13} u_1 c_3 + \cdots + \delta_{33} u_3 c_3 \tag{48}$$

But all the terms expect for the ones that go like δ_{11}, δ_{22}, and δ_{33} drop out because of the properties of this tensor:

$$\boldsymbol{u} \cdot \boldsymbol{c} = \delta_{11} u_1 c_1 + \delta_{22} u_2 c_2 + \delta_{33} u_3 c_3 \tag{49}$$

But now all these delta terms are just 1:

$$\boldsymbol{u} \cdot \boldsymbol{c} = u_1 c_1 + \boldsymbol{u}_2 c_2 + \boldsymbol{u}_3 c_3 \tag{50}$$

This is like inserting an identity matrix in between when we express these vectors in arrays:

$$\boldsymbol{u} \cdot \boldsymbol{c} = \begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \tag{51}$$

So, we are cool on this. Let's get back to the triple product:

$$\boldsymbol{u} \cdot \boldsymbol{c} = \delta_{ij} u_i c_j \tag{52}$$

Now let's switch back to $\boldsymbol{a}$ and $\boldsymbol{b}$:

$$\boldsymbol{a} \wedge \boldsymbol{b} \cdot \boldsymbol{c} = \delta_{ij} (\boldsymbol{a} \wedge \boldsymbol{b})_i c_j \tag{53}$$

We will now express the i^{th} component of this cross product using the Levi-Civita but we have to be extra careful with all the other indices lurking around. We can't use j anywhere else. Let's use the next couple of letters:

$$\boldsymbol{a} \wedge \boldsymbol{b} \cdot \boldsymbol{c} = \delta_{ij} (\epsilon_{ik\ell} a_k b_\ell) c_j = \delta_{ij} \epsilon_{ik\ell} a_k b_\ell c_j \tag{54}$$

How many summations do we have here? Well, there is no *free* index, meaning there is no index on the left-hand side of this equality, so all the indices are *dummy*. We have these indices i, j, k, and ℓ, so all four of them are meant to be summed over from 1 to 3:

$$\boldsymbol{a} \wedge \boldsymbol{b} \cdot \boldsymbol{c} = \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=1}^{3} \sum_{\ell=1}^{3} \delta_{ij} \epsilon_{ik\ell} a_k b_\ell c_j \tag{55}$$

Now, what does this δ_{ij} do for a living? It just replaces indices. You have the option to let $i \to j$ or $j \to i$ in all the terms that are multiplied by this Kronecker delta. Then you drop the delta term. I think I'll just let $j \to i$:

$$\boldsymbol{a} \wedge \boldsymbol{b} \cdot \boldsymbol{c} = \epsilon_{ik\ell} a_k b_\ell c_i \tag{56}$$

Now, this looks hideous. We have all these indices but they are successive letters. For some unknown reason, I find this disturbing. We note that, since these are all dummy

indices, meaning they will not be used anymore once we carry out the summation, I can relabel them as I wish. N.B. this relabeling is important only when you have some free indices around. Otherwise, you can just go crazy on them however you want. Here, I want to have i under a, j under b, and k under c. That is, I want to let $k \to i$, $\ell \to j$, and $i \to k$ (or if you want, you can always define some intermediate dummy variables by putting some primes on them before doing these replacements). Let's see what we get:

$$\boldsymbol{a} \wedge \boldsymbol{b} \cdot \boldsymbol{c} = \epsilon_{kij} a_i b_j c_k \tag{57}$$

Well, now it looks better. But we have this epsilon messed up. I want to bring it to the alphabetical order, as well—just for the sake of visual pleasure, nothing more. I want to have ϵ_{ijk}, so first, I flip k and i to have $-\epsilon_{ikj}$, where the minus sign comes from this flipping, and then I flip k and j to get $-(-\epsilon_{ijk})$. Here, the second minus sign comes from the second flipping. Then we get

$$\boldsymbol{a} \wedge \boldsymbol{b} \cdot \boldsymbol{c} = \epsilon_{ijk} a_i b_j c_k \tag{58}$$

Notice how everything is in order here.

Now, I want to play around with these indices to prove a useful identity. Suppose I shift my indices in a cyclic manner, i.e. I let $i \to j$ and $j \to k$ and $k \to i$. Again, I can do that because these are all dummy indices:

$$\boldsymbol{a} \wedge \boldsymbol{b} \cdot \boldsymbol{c} = \epsilon_{jki} a_j b_k c_i \tag{59}$$

Let's reorder these vector components here so that the indices are ordered alphabetically, namely

$$\boldsymbol{a} \wedge \boldsymbol{b} \cdot \boldsymbol{c} = \epsilon_{jki} c_i a_j b_k \tag{60}$$

Good one. Now I want to bring the indices of the epsilon to the alphabetical order, as well: $\epsilon_{jki} = -\epsilon_{jik} = +\epsilon_{ijk}$. Now we have

$$\boldsymbol{a} \wedge \boldsymbol{b} \cdot \boldsymbol{c} = \epsilon_{ijk} c_i a_j b_k \tag{61}$$

If we compare this to $\epsilon_{ijk} a_i b_j c_k = \boldsymbol{a} \wedge \boldsymbol{b} \cdot c$, we see that the right-hand side can be written compactly as

$$\boldsymbol{a} \wedge \boldsymbol{b} \cdot \boldsymbol{c} = \boldsymbol{c} \wedge \boldsymbol{a} \cdot b \tag{62}$$

If you want, try showing that this is also equal to $\boldsymbol{b} \wedge \boldsymbol{c} \cdot \boldsymbol{a}$.

These types of manipulations are essential when you have some objects with indices in an expanded form (as opposed to a compact one) because you'll want to recognize the patterns of contractions.

Let's do some exercises before we connect these to the Pauli matrices and all the other good stuff. I'll just to some basic examples and see where we get with them.

- $x_i x_i = \boldsymbol{x} \cdot \boldsymbol{x} = |\boldsymbol{x}|^2$
- $x_i y_j \delta_{ij} = \boldsymbol{x} \cdot \boldsymbol{y}$
- $(a_i)_r (a_j)_s \epsilon_{rst} = (\boldsymbol{a}_i \wedge \boldsymbol{a}_j)_t$: This will turn out to be important if you want to solve the last problem of the first homework with index manipulation. Note that I pulled a trick on you by not saying that what a_i is. It's not the i^{th} component of some vector $\boldsymbol{a}$. In reality, we have three $\boldsymbol{a}_i$ vectors and it's just one of them. Otherwise, the cross product would be meaningless.
- Let's make it spicier by including the derivative. There is this Laplace identity in electromagnetic theory. Suppose you have some magnetic field, $\boldsymbol{B}$, and you are interested in its curl, i.e. $\boldsymbol{\nabla} \wedge \boldsymbol{B}$. We want to express the magnetic field in terms of the magnetic potential, $\boldsymbol{A}$, which satisfies $\boldsymbol{B} = \boldsymbol{\nabla} \wedge \boldsymbol{A}$. Thus, we want to actually simplify $\boldsymbol{\nabla} \wedge (\boldsymbol{\nabla} \wedge \boldsymbol{A})$. Notice the importance of parentheses here. I can't just omit them because the order is important here.

  Now, the question is, what is this object? In terms its rank maybe? It's a certainly a vector. So, we can investigate its i^{th} component:

  $$[\boldsymbol{\nabla} \wedge (\boldsymbol{\nabla} \wedge \boldsymbol{A})]_i \tag{63}$$

  We start with the outermost objects and expand our way inward:

  $$[\boldsymbol{\nabla} \wedge (\boldsymbol{\nabla} \wedge \boldsymbol{A})]_i = \epsilon_{ijk} \nabla_j (\boldsymbol{\nabla} \wedge \boldsymbol{A})_k \tag{64}$$

  (N.B. some people write $\nabla_i = \partial_i$ and that's totally fine.) I've taken my Levi-Civita with i as the first index, and then my first vector here, which is the derivative, with an index that matches the second index of the Levi-Civita, and then I have my second vector—whatever it is, it's a vector—with an index that matches the

third index of the Levi-Civita. So far, so good. Let's now expand this second cross product. I have to be careful to choose indices different than i (which is a free index here because it appears on both sides of the equality) and j and k. But we are interested in the k^{th} component of this vector, so our Levi-Civita starts with a k. And I'm going to pick r and s as my new indices, why not:

$$[\boldsymbol{\nabla} \wedge (\boldsymbol{\nabla} \wedge \boldsymbol{A})]_i = \epsilon_{ijk} \nabla_j (\epsilon_{krs} \nabla_r A_s) \tag{65}$$

Remember that ϵ_{ijk} or ϵ_{krs} are just 1's and 0's—they are just numbers. They come out of derivative:

$$[\boldsymbol{\nabla} \wedge (\boldsymbol{\nabla} \wedge \boldsymbol{A})]_i = \epsilon_{ijk} \epsilon_{krs} \nabla_j \nabla_r A_s \tag{66}$$

We have something interesting here: two Levi-Civitas but with one common index. Notice that I can write $\epsilon_{krs} = -\epsilon_{rks} = +\epsilon_{rsk}$, so I have

$$[\boldsymbol{\nabla} \wedge (\boldsymbol{\nabla} \wedge \boldsymbol{A})]_i = \epsilon_{ijk} \epsilon_{rsk} \nabla_j \nabla_r A_s \tag{67}$$

(Confirm that you can write $\epsilon_{ijk}\epsilon_{rsk} = \epsilon_{kij}\epsilon_{krs}$ because some people like to see the common index as the first index. I don't know why.) We have this beautiful theorem for this contraction (N.B. the contraction is over k only. It's still a dot product but for only some parts of these tensors or matrices.):

$$\epsilon_{ijk}\epsilon_{imn} = \det \begin{pmatrix} \delta_{jm} & \delta_{km} \\ \delta_{jn} & \delta_{kn} \end{pmatrix} = \delta_{jm}\delta_{kn} - \delta_{km}\delta_{jn} \tag{68}$$

Notice how we've just managed to get rid of these antisymmetric tensors carrying three indices and now we just have the Kronecker deltas [5]. The Levi-Civita takes the cross product but the Kronecker just replaces indices. The latter has the easiest job in the world. Let's finish our calculation now:

$$\begin{aligned} [\boldsymbol{\nabla} \wedge (\boldsymbol{\nabla} \wedge \boldsymbol{A})]_i &= (\delta_{ir}\delta_{js} - \delta_{jr}\delta_{is}) \nabla_j \nabla_r A_s \\ &= \delta_{ir}\delta_{js} \nabla_j \nabla_r A_s - \delta_{jr}\delta_{is} \nabla_j \nabla_r A_s \end{aligned} \tag{69}$$

In the first term, using our Kroneckers, we'll replace r by i and s by j, and in the second term, we'll do the other way around:

$$[\boldsymbol{\nabla} \wedge (\boldsymbol{\nabla} \wedge \boldsymbol{A})]_i = \nabla_j \nabla_i A_j - \nabla_j \nabla_j A_i$$

$$= \nabla_i \nabla_j A_j - \nabla_j \nabla_j A_i \tag{70}$$

where, in the second line, I've just flipped the order of the derivatives as they commute. Now, what's the contraction pattern here, meaning who talks to whom? In the first term, the second derivative talks to the A vector because they have the same index, so we have a dot product in between. In the second term, the derivatives will talk to each other, i.e. we'll get the Laplacian, $\nabla^2 = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla}$:

$$[\boldsymbol{\nabla} \wedge (\boldsymbol{\nabla} \wedge \boldsymbol{A})]_i = \nabla_i (\boldsymbol{\nabla} \cdot \boldsymbol{A}) - \nabla^2 \boldsymbol{A}_i \tag{71}$$

so, in the full vector notation now, without just the i^{th} index, we have

$$\boldsymbol{\nabla} \wedge (\boldsymbol{\nabla} \wedge \boldsymbol{A}) = \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{A}) - \nabla^2 \boldsymbol{A} \tag{72}$$

- Here are some little exercises:
  - Show that $v_i v_j$ is a tensor symmetric with respect to i and j.
  - Show that the contraction of a symmetric tensor with an antisymmetric one over their symmetric and antisymmetric indices is zero; to be more precise, i.e. show that $\epsilon_{ijk} v_i v_j = 0$. (This explains why $\boldsymbol{v} \wedge \boldsymbol{v} = 0$.)
  - Prove that divergence of curl of a vector is zero.
  - Prove that curl of a gradient of a scalar is zero.

  Just to visualize, I mean, show that $\boldsymbol{\nabla} \cdot \boldsymbol{\nabla} \wedge V = 0$ and $\boldsymbol{\nabla} \wedge \boldsymbol{\nabla} f = 0$ using the index manipulation.

III. HW1 SUPPLEMENTS

A. Problem 1

1. *Part (a)*

I think we can all agree that $\langle z|\boldsymbol{\sigma}|z\rangle = \hat{\boldsymbol{n}}$ is an object with three scalar, and in general complex, components, so it's a vector in the usual sense. But here we are looking at the diagonal matrix elements of some hermitian operators, so all the components of $\hat{\boldsymbol{n}}$ *have to* be real:

$$\hat{n}_i \in \mathbb{R} \quad \forall i = 1, 2, 3 \tag{73}$$

Let's show that it has unit norm.

- Method 1: Let's use the sigma matrices explicitly.

```
In[1]:= Z = Subscript[z, #] & /@ Range[2];
        n = Conjugate[Z].PauliMatrix[#].Z & /@ Range[3];

In[3]:= n // MatrixForm // TraditionalForm
```

Out[3]//TraditionalForm=

$$\begin{pmatrix} z_1 (z_2)^* + z_2 (z_1)^* \\ i\, z_1 (z_2)^* - i\, z_2 (z_1)^* \\ z_1 (z_1)^* - z_2 (z_2)^* \end{pmatrix}$$

```
In[4]:= n.n // Simplify[#, Assumptions → {Conjugate[Z].Z == 1}] &

Out[4]= 1
```

- Method 2: Let's use the sigma matrices with indices without employing the Einstein summation convention for the sake of clarity—if the explicit summation symbols offend you, well, it's easier to ignore than to figure out. We can write the i^{th} component of this $\hat{\boldsymbol{n}}$ vector as

$$\begin{aligned} \hat{n}_i &= \langle z|\sigma_i|z\rangle \\ &= \sum_{\alpha=1}^{2} \sum_{\beta=1}^{2} z_\alpha^* (\sigma_i)_{\alpha\beta} z_\beta \end{aligned} \tag{74}$$

Let's take two copies of it to compute its norm:

$$\begin{aligned}\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{n}} &= \sum_{i=1}^{3}\hat{n}_i\hat{n}_i \\ &= \sum_{i=1}^{3}\left[\sum_{\alpha=1}^{2}\sum_{\beta=1}^{2}z_\alpha^*(\sigma_i)_{\alpha\beta}z_\beta\right]\left[\sum_{\gamma=1}^{2}\sum_{\delta=1}^{2}z_\gamma^*(\sigma_i)_{\gamma\delta}z_\delta\right] \\ &= \sum_{i=1}^{3}\sum_{\alpha=1}^{2}\sum_{\beta=1}^{2}\sum_{\gamma=1}^{2}\sum_{\delta=1}^{2}z_\alpha^*z_\beta z_\gamma^*z_\delta(\sigma_i)_{\alpha\beta}(\sigma_i)_{\gamma\delta} \end{aligned} \tag{75}$$

where all the indexed quantities are just numbers now so I have the liberty to move them around as I wish. Now I want to claim and prove an identity, namely the *completeness relation of the Pauli matrices*:

Claim.

$$\sum_{i=1}^{3}(\sigma_i)_{\alpha\beta}(\sigma_i)_{\gamma\delta} = 2\delta_{\alpha\delta}\delta_{\beta\gamma} - \delta_{\alpha\beta}\delta_{\gamma\delta} \tag{76}$$

Proof. We have an object that carries four indices, but we select these indices in groups of two from two identical quantities. Thus, the resultant tensor should be totally symmetric in any pair of indices, i.e. we take the following as our ansätz:

$$\sum_{i=1}^{3}(\sigma_i)_{\alpha\beta}(\sigma_i)_{\gamma\delta} = A\delta_{\alpha\beta}\delta_{\gamma\delta} + B\delta_{\alpha\gamma}\delta_{\beta\delta} + C\delta_{\alpha\delta}\delta_{\beta\gamma} \tag{77}$$

[If, at this point, the index δ of the Kronecker delta bothers you, well, I don't know what to say, *c'est la vie*.] [Convince yourselves that there is no other way than this ansätz.] Now, the rest is just the determination of these unknown coefficients A, B, and C by proper contractions:

– Take $\gamma = \beta$ and $\delta = \alpha$ and then sum over α and β (from 1 to 2):

$$\begin{aligned}\sum_{i=1}^{3}\sum_{\alpha=1}^{2}\sum_{\beta=1}^{2}(\sigma_i)_{\alpha\beta}(\sigma_i)_{\beta\alpha} &= \sum_{i=1}^{3}\sum_{\alpha=1}^{2}(\sigma_i\sigma_i)_{\alpha\alpha} \\ &= \sum_{\alpha=1}^{2}\left(\sum_{i=1}^{3}\sigma_i\sigma_i\right)_{\alpha\alpha} \\ &= \text{tr}\left(\sum_{i=1}^{3}\sigma_i\sigma_i\right) \\ &= \text{tr}\left(\sigma_1^2+\sigma_2^2+\sigma_3^2\right)\end{aligned} \tag{78}$$

Noting that $\sigma_i^2 = \mathbb{1}_{2\times 2}$ for all $i = 1, 2, 3$, we have

$$\begin{aligned}\sum_{i=1}^{3}\sum_{\alpha=1}^{2}\sum_{\beta=1}^{2}(\sigma_i)_{\alpha\beta}(\sigma_i)_{\beta\alpha} &= \operatorname{tr}(3\mathbb{1}_{2\times 2}) \\ &= 3\operatorname{tr}(\mathbb{1}_{2\times 2}) \\ &= 3\times 2 \\ &= 6 \end{aligned} \tag{79}$$

Let's check the right-hand side:

$$\begin{aligned}\sum_{\alpha=1}^{2}\sum_{\beta=1}^{2}\left(A\delta_{\alpha\beta}\delta_{\beta\alpha} + B\delta_{\alpha\beta}\delta_{\beta\alpha} + C\delta_{\alpha\alpha}\delta_{\beta\beta}\right) &= A\sum_{\alpha=1}^{2}\delta_{\alpha\alpha} + B\sum_{\alpha=1}^{2}\delta_{\alpha\alpha} + C\left[\sum_{\alpha=1}^{2}\delta_{\alpha\alpha}\right]^2 \\ &= 2A + 2B + 4C \end{aligned} \tag{80}$$

Then, we have our first equation:

$$6 = 2A + 2B + 4C \tag{81}$$

– Take $\gamma = \delta$ and sum over γ:

$$\sum_{i=1}^{3}\sum_{\gamma=1}^{2}(\sigma_i)_{\alpha\beta}(\sigma_i)_{\gamma\gamma} = \sum_{i=1}^{3}(\sigma_i)_{\alpha\beta}\operatorname{tr}(\sigma_i) \tag{82}$$

Noting that all the Pauli matrices are traceless, we get 0 here. Let's check the right-hand side of the ansätz:

$$\begin{aligned}\sum_{\gamma=1}^{2}\left(A\delta_{\alpha\beta}\delta_{\gamma\gamma} + B\delta_{\alpha\gamma}\delta_{\beta\gamma} + C\delta_{\alpha\gamma}\delta_{\beta\gamma}\right) &= 2A\delta_{\alpha\beta} + B\delta_{\alpha\beta} + C\delta_{\alpha\beta} \\ &= \delta_{\alpha\beta}(2A + B + C) \end{aligned} \tag{83}$$

The Kronecker delta is not zero in general, so the coefficient must vanish:

$$0 = 2A + B + C \tag{84}$$

– Finally, take $\gamma = \alpha$ and $\delta = \beta$ and then sum over α and β:

$$\sum_{i=1}^{3}\sum_{\alpha=1}^{2}\sum_{\beta=1}^{2}(\sigma_i)_{\alpha\beta}(\sigma_i)_{\alpha\beta} = \sum_{i=1}^{3}\sum_{\alpha=1}^{2}\sum_{\beta=1}^{2}(\sigma_i)_{\alpha\beta}(\sigma_i^\dagger)_{\alpha\beta}$$

$$
\begin{aligned}
&= \sum_{i=1}^{3}\sum_{\alpha=1}^{2}\sum_{\beta=1}^{2}(\sigma_i)_{\alpha\beta}(\sigma_i^*)_{\beta\alpha} \\
&= \sum_{i=1}^{3}\sum_{\alpha=1}^{2}(\sigma_i\sigma_i^*)_{\alpha\alpha} \\
&= \sum_{i=1}^{3}\mathrm{tr}(\sigma_i\sigma_i^*) \\
&= \mathrm{tr}\left(\sum_{i=1}^{3}\sigma_i\sigma_i^*\right) \\
&= \mathrm{tr}(\sigma_1 \underbrace{\sigma_1^*}_{\sigma_1} + \sigma_2 \underbrace{\sigma_2^*}_{-\sigma_2} + \sigma_3 \underbrace{\sigma_3^*}_{\sigma_3}) \\
&= \mathrm{tr}(\mathbb{1}_{2\times 2} - \mathbb{1}_{2\times 2} + \mathbb{1}_{2\times 2}) \\
&= \mathrm{tr}(\mathbb{1}_{2\times 2}) \\
&= 2
\end{aligned} \tag{85}
$$

Now let's check the right-hand side:

$$
\sum_{\alpha=1}^{2}\sum_{\beta=1}^{2}\left(A\delta_{\alpha\beta}\delta_{\beta\alpha} + B\delta_{\alpha\alpha}\delta_{\beta\beta} + C\delta_{\alpha\beta}\delta_{\beta\alpha}\right) = 2A + 4B + 2C \tag{86}
$$

Then we get our third equation:

$$
2 = 2A + 4B + 2C \tag{87}
$$

We now have three equations in three unknowns, so let's solve them:

```
In[5]:= Solve[
         {
          6 == 2 A + 2 B + 4 C,
          0 == 2 A + B + C,
          2 == 2 A + 4 B + 2 C
         },
         {A, B, C}
        ]

Out[5]= {{A → -1, B → 0, C → 2}}
```

Thus, we obtain

$$
\sum_{i=1}^{3}(\sigma_i)_{\alpha\beta}(\sigma_i)_{\gamma\delta} = -\delta_{\alpha\beta}\delta_{\gamma\delta} + 2\delta_{\alpha\delta}\delta_{\beta\gamma} \tag{88}
$$

qed.

With this identity, let's go back to our calculation:

$$\begin{aligned}
\hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{n}} &= \sum_{i=1}^{3} \sum_{\alpha=1}^{2} \sum_{\beta=1}^{2} \sum_{\gamma=1}^{2} \sum_{\delta=1}^{2} z_\alpha^* z_\beta z_\gamma^* z_\delta (\sigma_i)_{\alpha\beta} (\sigma_i)_{\gamma\delta} \\
&= \sum_{\alpha=1}^{2} \sum_{\beta=1}^{2} \sum_{\gamma=1}^{2} \sum_{\delta=1}^{2} z_\alpha^* z_\beta z_\gamma^* z_\delta \left(2\delta_{\alpha\delta}\delta_{\beta\gamma} - \delta_{\alpha\beta}\delta_{\gamma\delta} \right) \\
&= \sum_{\alpha=1}^{2} \sum_{\beta=1}^{2} \sum_{\gamma=1}^{2} \sum_{\delta=1}^{2} (2\delta_{\alpha\delta}\delta_{\beta\gamma} z_\alpha^* z_\beta z_\gamma^* z_\delta - \delta_{\alpha\beta}\delta_{\gamma\delta} z_\alpha^* z_\beta z_\gamma^* z_\delta) \\
&= \sum_{\alpha=1}^{2} \sum_{\beta=1}^{2} (2 z_\alpha^* z_\beta z_\beta^* z_\alpha - z_\alpha^* z_\alpha z_\gamma^* z_\gamma) \\
&= 2 \left[\sum_{\alpha=1}^{2} z_\alpha^* z_\alpha \right]^2 - \left[\sum_{\alpha=1}^{2} z_\alpha^* z_\alpha \right]^2 \\
&= 2\langle z|z\rangle^2 - \langle z|z\rangle^2 \\
&= 2 - 1 \\
&= 1
\end{aligned} \tag{89}$$

Thus, we have shown that $\hat{\boldsymbol{n}}$ is a three-dimensional unit vector with real scalar components.

2. Part (b)

We are given the projection operator $|z\rangle\langle z|$. Let's expand it in the 2-by-2 basis, i.e. in the basis of $\{1, \boldsymbol{\sigma}\}$ [Note that I'm switching from $\mathbb{1}_{2\times 2}$ to just 1, hoping that it will be clear whether it's a matrix or a scalar from the context.]. First, let's observe that this operator is hermitian:

$$\begin{aligned}
[|z\rangle\langle z|]^\dagger &= (\langle z|)^\dagger (|z\rangle)^\dagger \\
&= |z\rangle\langle z|
\end{aligned} \tag{90}$$

Thus, we know that the expansion coefficients will all be real. Now, let's do the expansion:

$$|z\rangle\langle z| = \zeta_0 + \sum_{i=1}^{3} \zeta_i \sigma_i \tag{91}$$

with $\zeta_\mu \in \mathbb{R}$ for all $\mu = 0, 1, 2, 3$. Here, the coefficients are given by some traces:

$$\zeta_0 = \frac{1}{2}\,\mathrm{tr}(|z\rangle\langle z|) \tag{92}$$

$$\zeta_i = \frac{1}{2}\,\mathrm{tr}(\sigma_i|z\rangle\langle z|), \quad i = 1, 2, 3 \tag{93}$$

The trace of ket-bra may look strange. Just for the sake of completeness, let's do the following:

Claim.

$$\mathrm{tr}(\Gamma|z\rangle\langle z|) = \langle z|\Gamma|z\rangle \tag{94}$$

where Γ is any operator.

Proof. Take a complete basis and write the trace as a sum over the diagonal matrix elements:

$$\mathrm{tr}(\Gamma|z\rangle\langle z|) = \sum_n \langle n|\left[\Gamma|z\rangle\langle z|\right]|n\rangle \tag{95}$$

where $\{|n\rangle\}$ may be the canonical basis, i.e. in its representation as a column vector, it has a 1 in the n^{th} entry and 0 everywhere else. The dimension does not matter for this proof, but for concreteness, we can just focus on $\{|n\rangle\} = \{|1\rangle, |2\rangle\}$ because we are working with 2-by-2 matrices. Now let's play around with this.

$$\mathrm{tr}(\Gamma|z\rangle\langle z|) = \sum_n \langle n|\Gamma|z\rangle\langle z|n\rangle \tag{96}$$

Each of these two factors is just a number, so I can reorder them anyway I want:

$$\begin{aligned}\mathrm{tr}(\Gamma|z\rangle\langle z|) &= \sum_n \langle z|n\rangle\langle n|\Gamma|z\rangle \\ &= \langle z|\left[\sum_n |n\rangle\langle n|\right]\Gamma|z\rangle \\ &= \langle z|\Gamma|z\rangle\end{aligned} \tag{97}$$

where I've used the completeness of this basis, $\sum_n |n\rangle\langle n| = 1$. qed.

With this, let's go back to the expansion coefficients:

$$\zeta_0 = \frac{1}{2}\,\mathrm{tr}(|z\rangle\langle z|) = \frac{1}{2}\langle z|z\rangle = \frac{1}{2} \tag{98}$$

$$\zeta_i = \frac{1}{2}\,\mathrm{tr}(\sigma_i|z\rangle\langle z|) = \frac{1}{2}\langle z|\sigma_i|z\rangle = \frac{1}{2}\hat{n}_i \tag{99}$$

Therefore,

$$|z\rangle\langle z| = \frac{1}{2} + \sum_{i=1}^{3} \hat{n}_i \sigma_i$$
$$= \frac{1}{2}\left(1 + \hat{\boldsymbol{n}} \cdot \boldsymbol{\sigma}\right) \tag{100}$$

3. Part (c)

$|z\rangle\langle z|$ is a projection operator, so without doing any calculations, I can directly say that its eigenvalues are 0 and 1. Why? Because I know that if I square this operator, I get the same thing.

Let's do this mathematically now. We have the eigenvalue equation:

$$|z\rangle\langle z|\lambda\rangle = \lambda|\lambda\rangle \tag{101}$$

Multiply from left by $|z\rangle\langle z|$:

$$|z\rangle \underbrace{\langle z|z\rangle}_{1}\langle z|\lambda\rangle = \lambda|z\rangle\langle z|\lambda \tag{102}$$

$$|z\rangle\langle z|\lambda\rangle = \lambda\left[\lambda|\lambda\rangle\right] \tag{103}$$

$$\lambda|\lambda\rangle = \lambda^2|\lambda\rangle \tag{104}$$

$$\lambda = \lambda^2 \tag{105}$$

so we must have $\lambda = 0, 1$. Let's do the eigenkets now one by one. First, I want to assume that my eigenkets have the form

$$|0,1\rangle = c^{+}_{0,1}|+\rangle + c^{-}_{0,1}|-\rangle \tag{106}$$

where I'm using the basis spanned by the eigenkets of the third Pauli matrix, $\sigma_3|\pm\rangle = \pm|\pm\rangle$, i.e. the canonical basis in two dimensions. Note that in this basis, we can also write

$$|z\rangle = z_1|+\rangle + z_2|-\rangle \tag{107}$$

- $\lambda = 0$:

$$|z\rangle\langle z|0\rangle = 0 \tag{108}$$

$$(z_1|+\rangle + z_2|-\rangle)(z_1^*\langle+| + z_2^*\langle-|)(c_0^+|+\rangle + c_0^-|-\rangle) = 0 \tag{109}$$

$$\underbrace{(z_1|+\rangle + z_2|-\rangle)}_{\neq 0}\underbrace{(z_1^* c_0^+ + z_2^* c_0^-)}_{=0} = 0 \tag{110}$$

so we have

$$c_0^- = -\frac{z_1^*}{z_2^*} c_0^+ \tag{111}$$

With this, we can write

$$|0\rangle = c_0^+ \left(|+\rangle - \frac{z_1^*}{z_2^*}|-\rangle\right) \tag{112}$$

Normalize it:

$$\begin{aligned}\langle 0|0\rangle &= 1 \\ &= \left|c_0^+\right|^2 \left(1 + \frac{|z_1|^2}{|z_2|^2}\right) \\ &= \left|c_0^+\right|^2 \frac{1}{|z_2|^2}\end{aligned} \tag{113}$$

so

$$\left|c_0^+\right| = |z_2| \tag{114}$$

or

$$c_0^+ = |z_2|\, e^{i\varphi} \tag{115}$$

Note that for a suitable phase φ, we can have $c_0^+ = z_2$. That's your liberty. At the end of the day, we obtain

$$|0\rangle = |z_2|\, e^{i\varphi} \left(|+\rangle - \frac{z_1^*}{z_2^*}|-\rangle\right) \tag{116}$$

- $\lambda = 1$:

$$|z\rangle\langle z| = |1\rangle \tag{117}$$

$$(z_1|+\rangle + z_2|-\rangle)(z_1^*\langle+| + z_2^*\langle-|)(c_1^+|+\rangle + c_1^-|-\rangle) = (c_1^+|+\rangle + c_1^-|-\rangle) \tag{118}$$

$$(z_1|+\rangle + z_2|-\rangle)(z_1^* c_1^+ + z_2^* c_1^-) = (c_1^+|+\rangle + c_1^-|-\rangle) \tag{119}$$

Since the two basis kets are linearly independent, we can match the coefficients of the ket $|+\rangle$ on both sides:

$$z_1 z_1^* c_1^+ + z_1 z_2^* c_1^- = c_1^+ \tag{120}$$

so

$$\begin{aligned} c_1^- &= \frac{1}{z_1 z_2^*}\left(1 - z_1 z_1^*\right) c_1^+ \\ &= \frac{z_2 z_2^*}{z_1 z_2^*} c_1^+ \\ &= \frac{z_2}{z_1} c_1^+ \end{aligned} \tag{121}$$

Then, we can write

$$|1\rangle = c_1^+ \left(|+\rangle + \frac{z_2}{z_1}|-\rangle\right) \tag{122}$$

Normalize it:

$$\begin{aligned} \langle 1|1\rangle &= \left|c_1^+\right|^2 \left(1 + \frac{|z_2|^2}{|z_1|^2}\right) \\ &= \left|c_1^+\right|^2 \frac{1}{|z_1|^2} \end{aligned} \tag{123}$$

This gives us

$$\left|c_1^+\right| = |z_1| \tag{124}$$

or

$$c_1^+ = |z_1|\, e^{i\varphi'} \tag{125}$$

Again, for a suitable phase φ', we can have $c_1^+ = z_1$. In general, we have the other eigenket given by

$$|1\rangle = |z_1|\, e^{i\varphi'} \left(|+\rangle + \frac{z_2}{z_1}|-\rangle\right) \tag{126}$$

$|z\rangle\langle z|$ is hermitian and its eigenvalues are nondegenerate. Therefore, $|0\rangle$ and $|1\rangle$ must be orthogonal [Check it and show that it's true whatever your phases are.].

Let's perform a spectral decomposition of $|z\rangle\langle z|$ [see Eq. (11)]:

$$|z\rangle\langle z| = \sum_\lambda \lambda |\lambda\rangle\langle\lambda|$$

$$
\begin{aligned}
&= 0|0\rangle\langle 0| + 1|1\rangle\langle 1| \\
&= |1\rangle\langle 1| \\
&= \left[|z_1|\, e^{i\varphi'}\left(|+\rangle + \frac{z_2}{z_1}|-\rangle\right)\right]\left[|z_1|\, e^{i\varphi'}\left(|+\rangle + \frac{z_2}{z_1}|-\rangle\right)\right]^\dagger \\
&= \left[|z_1|\, e^{i\varphi'}\left(|+\rangle + \frac{z_2}{z_1}|-\rangle\right)\right]\left[|z_1|\, e^{-i\varphi'}\left(\langle +| + \frac{z_2^*}{z_1^*}\langle -|\right)\right] \\
&= |z_1|^2\,|+\rangle\langle +| + |z_2|^2\,|-\rangle\langle -| + z_1^* z_2|-\rangle\langle +| + z_1 z_2^*|+\rangle\langle -|
\end{aligned} \tag{127}
$$

which is exactly what you'd get if you simply take the outer product of the given z vector.

4. Part (d)

Now we act this projection operator on the *plus/minus* (or canonical) basis, though the notation in the problem statement is a bit different:

$$
\begin{aligned}
|z\rangle\langle z|+\rangle &= |z_1|^2\,|+\rangle + z_1^* z_2|-\rangle \\
&= z_1^*(z_1|+\rangle + z_2|-\rangle) \\
&= z_1^*|z\rangle
\end{aligned} \tag{128}
$$

and

$$
\begin{aligned}
|z\rangle\langle z|-\rangle &= |z_2|^2\,|-\rangle + z_1 z_2^*|+\rangle \\
&= z_2^*(z_1|+\rangle + z_2|-\rangle) \\
&= z_2^*|z\rangle
\end{aligned} \tag{129}
$$

We have just projected these kets in the direction of the initial z ket. The coefficients measure *how much z is contained in these plus and minus kets*, most roughly.

5. Part (e)

Let's remember the definition of the vector $\hat{\boldsymbol{n}}$ and let $|z\rangle \to e^{i\alpha}|z\rangle$ and simultaneously $\langle z| \to \langle z|e^{-i\alpha}$:

$$
\begin{aligned}
\hat{\boldsymbol{n}} &= \langle z|\boldsymbol{\sigma}|z\rangle \\
&\to \langle z|e^{-i\alpha}\boldsymbol{\sigma}e^{i\alpha}|z\rangle
\end{aligned} \tag{130}
$$

But these exponentials are just constant phases, so they commute with the Pauli matrices:

$$\begin{aligned}\hat{\boldsymbol{n}} &\to \langle z|\boldsymbol{\sigma}e^{-i\alpha}e^{i\alpha}|z\rangle \\ &= \langle z|\boldsymbol{\sigma}|z\rangle \\ &= \hat{\boldsymbol{n}}\end{aligned} \tag{131}$$

B. Problem 2

We are given the following:

$$\sigma_i' = \hat{\boldsymbol{n}}_i \cdot \boldsymbol{\sigma} = \sum_{j=1}^{3} (\hat{\boldsymbol{n}}_i)_j \sigma_j \tag{132}$$

for $i = 1, 2, 3$. Notice that we have three $\hat{\boldsymbol{n}}_i$ vectors.

1. *Part (a)*

Let's demonstrate the power of index notation, where I'll keep my summation symbols for clarity:

$$[\sigma_i', \sigma_j'] = \left[\sum_{r=1}^{3} (\hat{\boldsymbol{n}}_i)_r \sigma_r, \sum_{s=1}^{3} (\hat{\boldsymbol{n}}_j)_s \sigma_s\right] \tag{133}$$

Here, $(\hat{\boldsymbol{n}}_i)_j$ are really some numbers—it's the j^{th} component of the i^{th} $\hat{\boldsymbol{n}}$ vector. Thus, they can safely leave the commutator:

$$[\sigma_i', \sigma_j'] = \sum_{r=1}^{3}\sum_{s=1}^{3} (\hat{\boldsymbol{n}}_i)_r (\hat{\boldsymbol{n}}_j)_s [\sigma_r, \sigma_s] \tag{134}$$

Now we are left with the usual commutator of two Pauli matrices:

$$\begin{aligned}[\sigma_i', \sigma_j'] &= \sum_{r=1}^{3}\sum_{s=1}^{3} (\hat{\boldsymbol{n}}_i)_r (\hat{\boldsymbol{n}}_j)_s \left[\sum_{t=1}^{3} 2i\epsilon_{rst}\sigma_t\right] \\ &= 2i \sum_{r=1}^{3}\sum_{s=1}^{3}\sum_{t=1}^{3} \epsilon_{rst} (\hat{\boldsymbol{n}}_i)_r (\hat{\boldsymbol{n}}_j)_s \sigma_t\end{aligned} \tag{135}$$

Compare this to Eq. (58). We have the same structure but with different indices—but that's totally okay because we are summing over them anyways. Now, the important point is to be able to make the connection that $\boldsymbol{a} = \hat{\boldsymbol{n}}_i$, $\boldsymbol{b} = \hat{\boldsymbol{n}}_j$, and $\boldsymbol{c} = \boldsymbol{\sigma}$. Even though we have these i and j indices, they are not our enemies, we can still live with them:

$$[\sigma_i', \sigma_j'] = 2i\hat{\boldsymbol{n}}_i \wedge \hat{\boldsymbol{n}}_j \cdot \boldsymbol{\sigma} \tag{136}$$

2. Part (b)

Now we do the anticommutator:

$$\begin{aligned}\left\{\sigma_i', \sigma_j'\right\} &= \left\{\sum_{r=1}^{3}(\hat{\boldsymbol{n}}_i)_r\sigma_r, \sum_{s=1}^{3}(\hat{\boldsymbol{n}}_j)_s\sigma_s\right\} \\ &= \sum_{r=1}^{3}\sum_{s=1}^{3}(\hat{\boldsymbol{n}}_i)_r(\hat{\boldsymbol{n}}_j)_s\left\{\sigma_r, \sigma_s\right\} \\ &= \sum_{r=1}^{3}\sum_{s=1}^{3}(\hat{\boldsymbol{n}}_i)_r(\hat{\boldsymbol{n}}_j)_s\left[2\delta_{rs}\right] \end{aligned} \tag{137}$$

Now we remember that the Kronecker delta has the easiest job on the planet, which is just replacing indices:

$$\begin{aligned}\left\{\sigma_i', \sigma_j'\right\} &= 2\sum_{r=1}^{3}(\hat{\boldsymbol{n}}_i)_r(\hat{\boldsymbol{n}}_j)_r \\ &= 2\hat{\boldsymbol{n}}_i \cdot \hat{\boldsymbol{n}}_j \end{aligned} \tag{138}$$

3. Part (c)

We are changing the basis from $\{\boldsymbol{\sigma}\}$ to $\{\boldsymbol{\sigma}'\}$:

$$\begin{pmatrix}\sigma_1' \\ \sigma_2' \\ \sigma_3'\end{pmatrix} = \begin{pmatrix}(\hat{\boldsymbol{n}}_1)_1 & (\hat{\boldsymbol{n}}_1)_2 & (\hat{\boldsymbol{n}}_1)_3 \\ (\hat{\boldsymbol{n}}_2)_1 & (\hat{\boldsymbol{n}}_2)_2 & (\hat{\boldsymbol{n}}_2)_3 \\ (\hat{\boldsymbol{n}}_3)_1 & (\hat{\boldsymbol{n}}_3)_2 & (\hat{\boldsymbol{n}}_3)_3\end{pmatrix}\begin{pmatrix}\sigma_1 \\ \sigma_2 \\ \sigma_3\end{pmatrix} \tag{139}$$

It's like going from the usual Cartesian basis vectors $\boldsymbol{x}$ to some rotated basis vectors $\boldsymbol{x}' = R\boldsymbol{x}$. We don't really care about the structure of the primed sigma matrices here. The entire story is about these unit vectors.

Now, for these primed sigma matrices to form a basis, the transformation should be given by a rotation matrix. If we have real unit vectors, then we have an orthogonal transformation, or if we have complex unit vectors, we have a unitary transformation. Whatever it is, we know that it will be a matrix with all the columns (or the row) orthogonal to each other. The reason is, only an orthogonal (or unitary) matrix preserves lengths. If the inner product gives different results after a transformation, then these new basis vectors (or matrices) are not properly defining a basis. We don't want that.

Therefore, for $\{\boldsymbol{\sigma}'\}$ to be a proper basis, we must have $\hat{\boldsymbol{n}}_i \cdot \hat{\boldsymbol{n}}_j = \delta_{ij}$.

C. Problem 3

We are given this unitary matrix U:

$$U = e^{i\frac{\theta}{2}\hat{\varphi}\cdot\boldsymbol{\sigma}} \tag{140}$$

which we know how to free from the exponential function—by a series expansion:

$$\begin{aligned}
U &= e^{i\frac{\theta}{2}\hat{\varphi}\cdot\boldsymbol{\sigma}} \\
&= \sum_{k\geq 0}\frac{1}{k!}\left(i\frac{\theta}{2}\hat{\varphi}\cdot\boldsymbol{\sigma}\right)^k \\
&= \sum_{k\geq 0}\frac{1}{(2k)!}\left(\frac{i\theta}{2}\right)^{2k}(\hat{\varphi}\cdot\boldsymbol{\sigma})^{2k} + \sum_{k\geq 0}\frac{1}{(2k+1)!}\left(\frac{i\theta}{2}\right)^{2k+1}(\hat{\varphi}\cdot\boldsymbol{\sigma})^{2k+1} \\
&= \sum_{k\geq 0}\frac{1}{(2k)!}\left(\frac{i\theta}{2}\right)^{2k} + \sum_{k\geq 0}\frac{1}{(2k+1)!}\left(\frac{i\theta}{2}\right)^{2k+1}\hat{\varphi}\cdot\boldsymbol{\sigma} \\
&= \cos\left(\frac{\theta}{2}\right) + i\sin\left(\frac{\theta}{2}\right)\hat{\varphi}\cdot\boldsymbol{\sigma}
\end{aligned} \tag{141}$$

With that, we also have

$$U^\dagger = \cos\left(\frac{\theta}{2}\right) - i\sin\left(\frac{\theta}{2}\right)\hat{\varphi}\cdot\boldsymbol{\sigma} \tag{142}$$

1. *Part (a)*

Let me employ the Einstein summation convention now because we'll have lots of terms in each line. In what follows, assume all the repeated indices are summed from 1 to 3:

$$\begin{aligned}
\Omega_j &= U\sigma_j U^\dagger \\
&= \left[\cos\left(\frac{\theta}{2}\right) + i\sin\left(\frac{\theta}{2}\right)\hat{\varphi}_r\sigma_r\right]\left[\cos\left(\frac{\theta}{2}\right) - i\sin\left(\frac{\theta}{2}\right)\hat{\varphi}_s\sigma_s\right] \\
&= \cos\left(\frac{\theta}{2}\right)^2\sigma_j - i\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)\hat{\varphi}_r\sigma_j\sigma_r + i\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)\hat{\varphi}_r\sigma_r\sigma_j + \sin\left(\frac{\theta}{2}\right)^2\hat{\varphi}_r\hat{\varphi}_s\sigma_r\sigma_j\sigma_s \\
&= \cos\left(\frac{\theta}{2}\right)^2\sigma_j + i\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)\hat{\varphi}_r\underbrace{[\sigma_r,\sigma_j]}_{2i\epsilon_{rjt}\sigma_t} + \sin\left(\frac{\theta}{2}\right)^2\hat{\varphi}_r\hat{\varphi}_s\sigma_r\underbrace{\sigma_j\sigma_s}_{\delta_{js}+i\epsilon_{jst}\sigma_t} \\
&= \cos\left(\frac{\theta}{2}\right)^2\sigma_j + \sin(\theta)(\hat{\varphi}\wedge\boldsymbol{\sigma})_j + \sin\left(\frac{\theta}{2}\right)^2(\hat{\varphi}_j\hat{\varphi}_r\sigma_r + i\hat{\varphi}_r\hat{\varphi}_s\epsilon_{jst}\underbrace{\sigma_r\sigma_t}_{\delta_{rt}+i\epsilon_{rtu}\sigma_u})
\end{aligned}$$

$$= \cos\left(\frac{\theta}{2}\right)^2 \sigma_j + \sin(\theta)(\hat{\boldsymbol{\varphi}} \wedge \boldsymbol{\sigma})_j + \sin\left(\frac{\theta}{2}\right)^2 (\hat{\varphi}_j \hat{\boldsymbol{\varphi}} \cdot \boldsymbol{\sigma} + i \underbrace{\hat{\varphi}_s \hat{\varphi}_t \epsilon_{jst}}_{0} - \underbrace{\epsilon_{jst} \underbrace{\epsilon_{rtu}}_{\epsilon_{urt}}}_{\delta_{ju}\delta_{sr} - \delta_{jr}\delta_{su}} \hat{\varphi}_r \hat{\varphi}_s \sigma_u)$$

$$= \cos\left(\frac{\theta}{2}\right)^2 \sigma_j + \sin(\theta)(\hat{\boldsymbol{\varphi}} \wedge \boldsymbol{\sigma})_j + \sin\left(\frac{\theta}{2}\right) (\hat{\varphi}_j \hat{\boldsymbol{\varphi}} \cdot \boldsymbol{\sigma} - \boldsymbol{\sigma}_j + \hat{\varphi}_j \hat{\boldsymbol{\varphi}} \cdot \boldsymbol{\sigma}) \tag{143}$$

or

$$\Omega_j = \cos(\theta)\sigma_j + \sin(\theta)(\hat{\boldsymbol{\varphi}} \wedge \boldsymbol{\sigma})_j + 2\sin\left(\frac{\theta}{2}\right)^2 \hat{\varphi}_j \hat{\boldsymbol{\varphi}} \cdot \boldsymbol{\sigma} \tag{144}$$

A lot of things have been done without saying. Questions…Just ask.

2. *Part (b)*

Expand Ω_j in the usual 2-by-2 basis:

$$\Omega_j = (m_j)_0 + (m_j)_r \sigma_r \tag{145}$$

where

$$(m_j)_0 = \frac{1}{2}\,\mathrm{tr}(\Omega_j) \tag{146}$$

$$(m_j)_r = \frac{1}{2}\,\mathrm{tr}(\sigma_r \Omega_j) \tag{147}$$

Note that $(m_j)_0 = 0$ for all $j = 1, 2, 3$ because each Ω_j has one and only one power of sigma:

$$\Omega_j = \left[\cos(\theta)\delta_{sj} + \sin(\theta)\epsilon_{jps}\hat{\varphi}_p + 2\sin\left(\frac{\theta}{2}\right)^2 \hat{\varphi}_j \hat{\varphi}_s\right] \sigma_s \tag{148}$$

Now, let's see what happens when we multiply this by σ_r and take the trace:

$$\sigma_r \Omega_j = [\cdots]\,\sigma_r \sigma_s \tag{149}$$

$$\mathrm{tr}(\sigma_r \Omega_j) = [\cdots]\,\mathrm{tr}(\sigma_r \sigma_s) \tag{150}$$

where the terms in the square brackets are given in (148). Let's do the trace. Recall the anticommutation relation:

$$\{\sigma_r, \sigma_s\} = 2\delta_{rs}$$

$$= \sigma_r\sigma_s + \sigma_s\sigma_r \tag{151}$$

so

$$\begin{aligned}
\operatorname{tr}(2\delta_{rs}) &= 2\delta_{rs}\operatorname{tr}(\mathbb{1}_{2\times 2}) \\
&= 4\delta_{rs} \\
&= \operatorname{tr}(\sigma_r\sigma_s + \sigma_s\sigma_r) \\
&= \operatorname{tr}(\sigma_r\sigma_s) + \operatorname{tr}(\sigma_s\sigma_r) \\
&= 2\operatorname{tr}(\sigma_r\sigma_s)
\end{aligned} \tag{152}$$

and hence

$$\operatorname{tr}(\sigma_r\sigma_s) = 2\delta_{rs} \tag{153}$$

This gives us

$$\operatorname{tr}(\sigma_r\Omega_j) = 2\cos(\theta)\delta_{rj} + 2\sin(\theta)\epsilon_{rjp}\hat{\varphi}_p + 4\sin\left(\frac{\theta}{2}\right)^2\hat{\varphi}_r\hat{\varphi}_j \tag{154}$$

and therefore

$$(m_j)_r = \cos(\theta)\delta_{rj} + \sin(\theta)\epsilon_{rjp}\hat{\varphi}_p + 2\sin\left(\frac{\theta}{2}\right)^2\hat{\varphi}_r\hat{\varphi}_j \tag{155}$$

Now let's construct this matrix:

In[6]:=
```
φ = {-Sin[ϕ], Cos[ϕ], 0};
m = Array[ (* r = #1, j = #2 *)
    Cos[θ] KroneckerDelta[#1, #2]
      + Sin[θ] Sum[LeviCivitaTensor[3][[#1, #2, p]] φ[[p]], {p, Range[3]}]
      + 2 Sin[θ/2]^2 φ[[#1]] φ[[#2]] &,
    {3, 3}
  ];
```

In[8]:=
```
m // MatrixForm // TraditionalForm
```

Out[8]//TraditionalForm=

$$\begin{pmatrix} 2\sin^2\left(\frac{\theta}{2}\right)\sin^2(\phi) + \cos(\theta) & -2\sin^2\left(\frac{\theta}{2}\right)\sin(\phi)\cos(\phi) & \sin(\theta)\,(-\cos(\phi)) \\ -2\sin^2\left(\frac{\theta}{2}\right)\sin(\phi)\cos(\phi) & 2\sin^2\left(\frac{\theta}{2}\right)\cos^2(\phi) + \cos(\theta) & -\sin(\theta)\sin(\phi) \\ \sin(\theta)\cos(\phi) & \sin(\theta)\sin(\phi) & \cos(\theta) \end{pmatrix}$$

Here, the three $\hat{\boldsymbol{n}}$ vectors are given by the rows due to our construction of the m matrix. Let's show them they are orthonormal:

```
In[9]:= n1 = m[[1]];
        n2 = m[[2]];
        n3 = m[[3]];

In[12]:= n1.n1 // Simplify
         n1.n2 // Simplify
         n1.n3 // Simplify
         n2.n2 // Simplify
         n2.n3 // Simplify
         n3.n3 // Simplify
Out[12]=
         1
Out[13]=
         0
Out[14]=
         0
Out[15]=
         1
Out[16]=
         0
Out[17]=
         1
```

Or we could just show that $m^{-1} = m^{\mathrm{T}}$:

```
In[18]:= Inverse[m] == Transpose[m] // Simplify
Out[18]=
         True
```

3. Part (c)

The R matrix here is the transpose of the m matrix from the previous part due to our construction. Now, since $m^{-1} = m^{\mathrm{T}}$, we also have $R^{-1} = R^{\mathrm{T}}$. This makes R most definitely a rotation matrix because it's orthogonal. If it's hard to see that, just do an example. Set, e.g., $\phi = \pi/2$, why not:

```
In[20]:= R /. ϕ → π / 2 // Simplify // MatrixForm // TraditionalForm
```

Out[20]//TraditionalForm=

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

This is a rotation about the x axis in the clockwise direction when viewed from the positive x direction. Nice and easy.

IV. DISCUSSION 3 (OCT 13)

Why do we have this Dirac Hamiltonian? Where does it come from? Since we have mentioned this Dirac Hamiltonian in today's lecture, let's very briefly mention it here.

Suppose you want to solve the Schrödinger equation in the good ol' position basis. What is the most general form of the Hamiltonian that you can write? We have, most naively, a kinetic term plus a potential:

$$H = K + U = \frac{\boldsymbol{p}^2}{2m} + V(\boldsymbol{x}) \tag{156}$$

where the *representation* of the momentum operator in the position basis is just the derivative:

$$\boldsymbol{p} \doteq \frac{\hbar}{i}\boldsymbol{\nabla} \tag{157}$$

I'm a lazy guy, so I'll just set $\hbar = c = 1$, though we don't deal with the speed of light, yet. Now, with this Hamiltonian, you write down the Schrödinger equation in the position basis:

$$i\hbar\frac{\partial\psi(t,\boldsymbol{x})}{\partial t} = H\psi(t,\boldsymbol{x}) \tag{158}$$

where I've indicated the $\hbar$ explicitly for this last time. Schrödinger essentially took the classical expression for the Hamiltonian an promoted it to be an *operator*. That's how we do quantum mechanics.

Now, consider the case of a free particle:

$$i\frac{\partial\psi(t,\boldsymbol{x})}{\partial t} = -\frac{1}{2m}\boldsymbol{\nabla}^2\psi(t,\boldsymbol{x}) \tag{159}$$

Now, this equation is disturbing, if you know where to look at. The complex unit may bother you, that's okay, but that's how quantum mechanics works. We have a diffusion equation and we work essentially with imaginary time (if you let $t \to -it$).

The keyword is relativity. What are the most common objects in relativity? Well, that's a really broad question but I have 4-vectors in my mind. We promote the usual position 3-vector to a 4-vector by marrying it to time:

$$x^\mu = (ct, \boldsymbol{x}) \tag{160}$$

and this will be the last time where I explicit write c. Similarly, for the momentum, we have

$$p^\mu = (E, \boldsymbol{p}) \tag{161}$$

You can also promote the force 3-vector to $4d$. What's the quantity that should go into the 0^{th} or the *time* component? It's power:

$$F^\mu = (P, \boldsymbol{F}) \tag{162}$$

In the most beautiful theory ever written down on this planet, namely in Einstein's theory of relativity, we treat time and position on an equal footing. They become just some parameters using which you write down your *fields*—okay, that's a new concept, which I'll not pursue here.

Now, in light of these promotions, let's ask again. What's that thing in the Schrödinger equation that *should* bother you? It's the fact that we have only one time derivative but two position derivatives. You can't do relativity with this (Schrödinger tried that). Why? Because we don't treat time and position on an equal footing.

So, what can we do? First, let's ask this: What's the energy that we should promote to be an operator in relativity? It's the relativistic energy:

$$E^2 = \boldsymbol{p}^2 + m^2 \tag{163}$$

We want to be able to promote this to an operator so as to write a Hamiltonian like

$$H \stackrel{?}{=} \sqrt{\boldsymbol{p}^2 + m^2} \tag{164}$$

(Again, I'm just considering the case of a free particle for the sake of simplicity). Now, in the position basis, the Schrödinger equation becomes

$$i\frac{\partial \psi}{\partial t} \stackrel{?}{=} \sqrt{-\boldsymbol{\nabla}^2 + m^2}\psi \tag{165}$$

Having a Laplacian under a square root causes lots of troubles. So what can we do about this?

At this point, Dirac came in and said *hey, I'm just going to take the square root of this expression*. He chose the following as ansätz:

$$H = \sum_{j=1}^{3} \alpha_j p_j + \beta m \tag{166}$$

Then, he took two copies of this to get the usual relativistic energy formula:

$$\begin{aligned} H^2 &= (\alpha_j p_j + \beta m)(\alpha_k p_k + \beta m) \\ &= \alpha_j \alpha_k p_j p_k + \beta^2 m^2 + \alpha_j \beta p_j m + \beta \alpha_k p_k m \\ &= \alpha_j \alpha_k p_j p_k + \beta^2 m^2 + (\alpha_j \beta + \beta \alpha_j) p_j m \end{aligned} \tag{167}$$

Now we want to have

$$H^2 = \delta_{jk} p_j p_k + m^2 \tag{168}$$

So we see that $\beta^2 = 1$ and that the last term above should be zero:

$$\alpha_j \beta + \beta \alpha_j = 0 \quad \forall j = 1, 2, 3 \quad \text{(no sum!)} \tag{169}$$

And we also the alpha factors in the first term to give δ_{jk}—but we need to symmetrize it first: you can't just write $\alpha_j \alpha_k = \delta_{jk}$ because we know that Kronecker delta is explicitly symmetric.

$$\alpha_j \alpha_k = \frac{1}{2} \left\{ \alpha_j, \alpha_k \right\} + \frac{1}{2} \left[\alpha_j, \alpha_k \right] \tag{170}$$

So we see that this product of alphas contain an antisymmetric part—the commutator—as well. Now we have:

$$\frac{1}{2} \left\{ \alpha_j, \alpha_k \right\} = \delta_{jk} \tag{171}$$

Okay, let's write the two relations we found in a simplified way:

$$\left\{ \alpha_j, \alpha_k \right\} = 2\delta_{jk}, \quad \left\{ \alpha_j, \beta \right\} = 0 \tag{172}$$

[For my particle-theorist colleagues, yes, these are the gamma matrices—not directly but in a way.] Dirac realized that he can't just take the alphas and the beta to be scalars because he can't satisfy these relations. Therefore, he said, they must be matrices. So, that's why we said earlier in today's lecture that we express the Dirac Hamiltonian using anticommuting matrices!

So, the next question any sane person would ask is, what's the size of these matrices? Let's try 2-by-2 because we have the Pauli matrices that satisfy very similar anticommutation relations. Can we take $\alpha_j = \sigma_j$ and $\beta = 1$? Well, the alpha part works but the

beta part is problematic. With beta being the identity, you can't just satisfy the above anticommutations. So, what do we do? We have four 2-by-2 matrices and we can't express them using the basis of the 2-by-2 hermitian matrices (Show that $\boldsymbol{\alpha}$ and β should be hermitian!). The basis is simply *not sufficient* to cover these matrices. Then, we go for the next option, namely we assume these matrices to be 3-by-3. We have the 8 Gell-Mann matrices and the identity matrix, so we have a lot more freedom to express these alpha and beta matrices than with the Pauli matrices. But... There is a problem.

We have

$$\alpha_j\beta + \beta\alpha_j = 0 \tag{173}$$

Recalling that $\beta^2 = 1$, multiply both sides from left by beta and take the trace. You'll find that

$$\mathrm{tr}(\alpha_j) = 0 \tag{174}$$

Multiply the same equation by α_j, sum over j, and take the trace. You'll get

$$\mathrm{tr}(\beta) = 0 \tag{175}$$

so we conclude that our Hamiltonian must be traceless. Now that's important.

The eigenvalues of the Dirac Hamiltonian are, as we have also shown in today's lecture, $\pm\sqrt{\boldsymbol{p}^2 + m^2}$. There is this relation from linear algebra:

$$\mathrm{tr}(A) = \sum_\lambda \lambda \tag{176}$$

In words, we say that the trace of any nice enough matrix is given by the sum of its eigenvalues. Our Hamiltonian is traceless and we have the eigenvalues $\pm\lambda$. If these alpha and beta matrices, hence our Hamiltonian, were 3-by-3, then there is no way to make our Hamiltonian traceless. You can't get zero by adding $+\lambda$ and $-\lambda$ even in the presence of degeneracies. (Suppose $+\lambda$ is 2-fold generate. Then we get $\mathrm{tr}(H) = \lambda$. Suppose $-\lambda$ is 2-fold degenerate. Then we get $\mathrm{tr}(H) = -\lambda$. There is always a left-over).

This means that we should have equal numbers of these energy eigenvalues, i.e. both $+\lambda$ and $-\lambda$ should be N-fold degenerate. If we have $N + N = 2N$ eigenvalues, then our matrix should be $2N$-by-$2N$. This shows that we can just use 3-by-3 (or any odd-dimensional) matrices. The smallest number is 4. This is what we have discussed in today's lecture—in a different language.

* * *

I'm going to introduce the machinery of ladder operators. I'll rely on your earlier exposure to quantum mechanics in writing down some parts, but if you see these for the first time, it should still be followable.

If we have a spin-1/2 particle, if we measure the z component of its spin, we can have only $+\hbar/2$ and $-\hbar/2$. This has been given to you since modern physics courses. So let's either accept this or wait until we discuss angular momentum in the class. Take your pick.

From linear algebra, we know that if we know the eigenvalues of a matrix, then we can express it as a diagonal matrix with its eigenvalues on the diagonal entries in the canonical basis. Here, by the canonical basis, I mean the unit vectors where we have a 1 somewhere and the rest of the entries are 0. Then we can represent the S_z operator as

$$S_z \doteq \begin{pmatrix} +1/2 & 0 \\ 0 & -1/2 \end{pmatrix} \tag{177}$$

where I've now set $\hbar = 1$. Notice that this is just the third Pauli matrix multiplied by $1/2$. With that, we can write the spin operators as the spin times the Pauli matrices:

$$\boldsymbol{S} = \frac{1}{2}\boldsymbol{\sigma} \tag{178}$$

Now, we have said that the eigenkets are the kets that are reprenseted by the vectors in the canonical basis. Let's introduce the notation $|s, m\rangle$, where s is the spin, which is fixed here, and m is the z component of the spin. We have

$$\left|\frac{1}{2}, +\frac{1}{2}\right\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{179}$$

$$\left|\frac{1}{2}, -\frac{1}{2}\right\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{180}$$

The eigenvalue equation for the S_z operator is

$$S_z \left|\frac{1}{2}, \pm\frac{1}{2}\right\rangle = \pm\frac{1}{2}\left|\frac{1}{2}, \pm\frac{1}{2}\right\rangle \tag{181}$$

So when I say *spin*, I essentially mean the Pauli matrices—at least for this case.

Now, I'll introduce two nonhermitian operators, S_+ and S_-, which can represented in the canonical basis by

$$S_+ \doteq \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{182}$$

$$S_- \doteq \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{183}$$

$$\tag{184}$$

But this is just for visual pleasure. Let's express them in terms of the eigenkets of the S_z operator:

$$S_\pm = \left|\frac{1}{2}, \pm\frac{1}{2}\right\rangle \left\langle\frac{1}{2}, \mp\frac{1}{2}\right| \tag{185}$$

So they are basically projection operators. Let's see what these guys do for a living:

$$S_+ \left|\frac{1}{2}, +\frac{1}{2}\right\rangle = 0 \tag{186}$$

$$S_+ \left|\frac{1}{2}, -\frac{1}{2}\right\rangle = \left|\frac{1}{2}, -\frac{1}{2}\right\rangle \tag{187}$$

$$S_- \left|\frac{1}{2}, +\frac{1}{2}\right\rangle = \left|\frac{1}{2}, +\frac{1}{2}\right\rangle \tag{188}$$

$$S_- \left|\frac{1}{2}, -\frac{1}{2}\right\rangle = 0 \tag{189}$$

where I've used the orthonormality of the eigenkets. So essentially, S_+ increases the z component of the spin by one or annihilates the state if it's already in the state with the highest possible m value, and S_- works in the opposite way.

Now let's see what we get if we take the combos $S_+ \pm S_-$:

$$S_+ + S_- \doteq \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \sigma_1 = 2S_x \tag{190}$$

$$S_+ - S_- \doteq \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = i\sigma_2 = 2iS_y \tag{191}$$

so we have

$$S_x = \frac{S_+ + S_-}{2} \tag{192}$$

$$S_y = \frac{S_+ - S_-}{2i} \tag{193}$$

$$\tag{194}$$

Since we are taking the half of the sigma matrices to define these spin matrices, we have this simplified relation for the commutator:

$$[S_x, S_y] = iS_z \quad \text{(cyclic)} \tag{195}$$

By simplified, I mean there is no factor of 2.

Now I'm going to generalize these to any spin value. We always start with the z component of the spin. The eigenvalue equation for the S_z operator is given by

$$S_z|m\rangle = m|m\rangle \quad (\text{cf. } \sigma_3|\pm\rangle = \pm|\pm\rangle) \tag{196}$$

I don't really care about what these m values are. I know that they form a complete basis:

$$\sum_{m=-s}^{s} |m\rangle\langle m| = \mathbb{1}_{(2s+1)\times(2s+1)} \quad \left(\text{or } \sum_m = |m\rangle\langle m| = 1 \text{ for short}\right) \tag{197}$$

Multiply the eigenvalue equation by $\langle m|$ from right and sum over mL

$$S_z = \sum_m m|m\rangle\langle m| \tag{198}$$

which is in fact nothing but the usual spectral decomposition of this operator. Now, we have the ladder operators:

$$S_\pm|m\rangle = c_m^\pm|m \pm 1\rangle \tag{199}$$

where $c_m^\pm = \sqrt{s(s+1) - m(m \pm 1)}$. Do the same trick of multiplying by the bra:

$$S_\pm = \sum_m c_m^\pm|m \pm 1\rangle\langle m| \tag{200}$$

so we have

$$S_x = \frac{1}{2}\sum_m \left[c_m^+|m+1\rangle\langle m| + c_m^-|m-1\rangle\langle m|\right] \tag{201}$$

$$S_y = \frac{1}{2i}\sum_m \left[c_m^+|m+1\rangle\langle m| - c_m^-|m-1\rangle\langle m|\right] \tag{202}$$

Let's now try to visualize these operators using matrices. I'll consider $s = 1$ and start with S_z:

$$S_z = \sum_{m=-1}^{1} m|m\rangle\langle m| \tag{203}$$

$$= -1|-1\rangle\langle -1| + 0|0\rangle\langle 0| + 1|1\rangle\langle 1| \tag{204}$$

Now, we match these eigenkets $\{|m\rangle\}$ to the canonical basis starting with the ket with the highest m value:

$$|1\rangle \doteq \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \tag{205}$$

$$|0\rangle \doteq \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \tag{206}$$

$$|-1\rangle \doteq \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{207}$$

so we have

$$S_z \doteq \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{208}$$

Let's do S_+ and S_-:

$$\begin{aligned} S_+ &= \sum_{m=-1}^{1} c_m^+ |m+1\rangle\langle m| \\ &= c_1^+ \underbrace{|2\rangle}_{\text{no such state}} \langle 1| + c_0^+ |1\rangle\langle 0| + c_{-1}^+ |0\rangle\langle -1| \\ &\doteq \begin{pmatrix} 0 & c_0^+ & 0 \\ 0 & 0 & c_{-1}^+ \\ 0 & 0 & 0 \end{pmatrix} \end{aligned} \tag{209}$$

Note that $S_- = S_+^\dagger$:

$$S_- \doteq \begin{pmatrix} 0 & 0 & 0 \\ c_0^+ & 0 & 0 \\ 0 & c_{-1}^+ & 0 \end{pmatrix} \tag{210}$$

so

$$S_x = \frac{1}{2} \begin{pmatrix} 0 & c_0^+ & 0 \\ c_0^+ & 0 & c_{-1}^+ \\ 0 & c_{-1}^+ & 0 \end{pmatrix} \tag{211}$$

$$S_y = \frac{1}{2i} \begin{pmatrix} 0 & c_0^+ & 0 \\ -c_0^+ & 0 & c_{-1}^+ \\ 0 & -c_{-1}^+ & 0 \end{pmatrix} \tag{212}$$

$$\tag{213}$$

If you compute these numbers, you'll get

$$S_x = \begin{pmatrix} 0 & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 0 \end{pmatrix} \tag{214}$$

$$S_y = \begin{pmatrix} 0 & -i/\sqrt{2} & 0 \\ i/\sqrt{2} & 0 & -i/\sqrt{2} \\ 0 & i/\sqrt{2} & 0 \end{pmatrix} \tag{215}$$

Note that we still have $[S_x, S_y] = iS_z$. Try this using the kets and bras—there is nothing boring than multiplying matrices. This commutation relation doesn't really care about the particular spin value.

So now, anything can be accomplished with the ket-bra expansions given in Eqs. (198) and (200). Study the case $s = 3/2$. Or not, I'm not a cop.

* * *

For an earlier course, I wrote a mini Mathematica code that gives these spin matrices for any spin value [6]. I'll drop it here and you can use it to check your calculations.

```
(* play with the s value only *)
s = 1;
(* black box *)
dim =
    2*s + 1;
Sp =
    ConstantArray[0, {dim, dim}];
Sm =
    ConstantArray[0, {dim, dim}];
Sz =
    ConstantArray[0, {dim, dim}];

i = 1;

For[mp = s, mp >= -s, mp--, j = 1;
    For[m = s, m >= -s, m--,
        Sp[[i]][[j]] =
            \[HBar]*Sqrt[s*(s + 1) - m*(m + 1)]*KroneckerDelta[mp, m + 1];
        Sz[[i]][[j]] =
            \[HBar]*m*KroneckerDelta[mp, m];
        Sm[[i]][[j]] =
            \[HBar]*Sqrt[s*(s + 1) - m*(m - 1)]*KroneckerDelta[mp, m - 1];
        j++;
    ];
    i++;
];

Sx = (Sp + Sm)/2;
Sy = (Sp - Sm)/(2*I);

Print["\!\(\*SubscriptBox[\(S\), \(x\)]\) = ", \[HBar], Sx/\[HBar] // MatrixForm]
Print["\!\(\*SubscriptBox[\(S\), \(y\)]\) = ", \[HBar], Sy/\[HBar] // MatrixForm]
```

```
Print["\!\(\*SubscriptBox[\(S\), \(z\)]\) = ", \[HBar], Sz/\[HBar] // MatrixForm]
```

* * *

I'm not so fond of visualization but there was a comment/question about how to visualize a $3D$ vector that depends on 4 parameters, t and $\boldsymbol{x}$. Here is how I would do it on Mathematica.

```
n[1, ct_, x_, y_, z_] := x (* or any crazy function you can think of *)
n[2, ct_, x_, y_, z_] := y (* or any crazy function you can think of *)
n[3, ct_, x_, y_, z_] := Sin[ct] z (* or any crazy function you can think of *)
n[ct_, x_, y_, z_] := n[#, ct, x, y, z] & /@ Range[3]
```

So we have now a $3D$ vector that depends on time and position. I'll use the command `VectorPlot3D`:

```
VectorPlot3D[
    n[1, x, y, z],
    {x, -5, 5},
    {y, -5, 5},
    {z, -5, 5}
]
```

Just evaluate it and see what you get. I've just defined some arbitrary ranges for the position. Notice that I've taken `ct` equal to 1 here. So, how do we visualize the evolution of this vector in time? Use `Animate`:

```
Animate[
    VectorPlot3D[
        n[ct, x, y, z],
        {x, -5, 5},
        {y, -5, 5},
        {z, -5, 5}
    ],
    {ct, 0, 10}
]
```

It's also possible to export this to a gif.

V. DISCUSSION 4 (OCT 20)

What are the eigenkets of the position operator? Well, we should have a relation like

$$X|x\rangle = x|x\rangle \tag{216}$$

Note that x here is a continuous parameter. If you take the inner product of two different eigenkets of the position operator, they will give you the Dirac delta:

$$\langle x|x'\rangle = \delta(x - x') \tag{217}$$

The completeness relation is then given by an integral instead of a sum:

$$\int \mathrm{d}x \ |x\rangle\langle x| = \mathbb{1} \tag{218}$$

Now, what is the quantity $\langle x(t)|x(0)\rangle$? I'm aware of that we haven't done any time evolution yet. I'll just throw this at you.

In a conference long ago, Dirac said that this object represents a Lagrangian. WTF, right? But guess what happened next. One of the participants was Feynman and he thought long about this mysterious remark. This eventually led him to the path-integral formalism of quantum mechanics.

There is a very similar-looking object that goes like this:

$$\langle 0|x(t)x(0)|0\rangle \tag{219}$$

What does this guy do for a living? It looks nasty, first of all. Remember your Green's functions from elementary electrodynamics? It looks like some $i\boldsymbol{k}\cdot\boldsymbol{x}$ in the exponent, and the entire thing is divided by the length of this position vector. That's okay if you don't recall this right now. When we compute this object within the context of simple harmonic oscillator, we'll see that it has the same exact form but the position replaced by the energy. So, it's the propagator in the energy space. Now, for my particle-theorist colleagues, if you promote the position operator to fields, you will get the good ol' propagator.

This was interesting so I thought I should just throw this at you and embrace the consequences.

* * *

I can't describe how much I like my measurements. As I stated in the first discussion session and as we have seen in today's lecture, measurements are all about projections. You take your initial state in the ket, then your final state in the bra, and make a projection sandwich in between. This gives you the transition amplitude. If you take the square of this guy, you'll get the probability. Note that transition amplitudes are not physical observables. What we—well, my experimentalist colleagues—measure in their labs are these probabilities. That's important to remember.

Now, I want to fry your brains and give you something to think about. Let's start with the two-slit interference in the classical optics. How do you write the interference pattern? Well, you have an electric-field component coming from one slit, say $|\boldsymbol{E}|$, and then you take the one coming from the other slit with a phase difference, $|\boldsymbol{E}|e^{i\delta}$—assuming that we have a single, coherent source. Then you sum them up, take the absolute square, and obtain the intensity profile. This is very much like the measurement. The electric field here corresponds to your wave function and the intensity is the probability. You don't measure (or see) your electric-field strength, but rather the intensity.

Now suppose I put a small piece of circuit around one of the slits. It can be just a simple wire with an ammeter. What happens to the interference pattern on the screen? We see only a single bright spot now. Why? Because we are doing a measurement on the slit. We are, in a way, selecting which *channel* [Yeah, I'm going to use this word a lot] the photons go—or the electrons due to the wave-particle duality in nature. This is equivalent to say that we block out the other channel.

This is the same phenomenon that takes place with the Stern-Gerlach apparatus—well, almost. Using the magnetic field (like the slits), we essentially split our beam into two. And we have the liberty to pick one of the beams that come out of the SG apparatus. It's not exactly like the double-slit experiment because we see only two *bright spots* (for a spin-1/2 particle), where as we have a couple of them in a double-slit experiment.

Now, any sane person would ask at this point, why the heck am I talking about this?

Let's continue with a funny story. When young Feynman was a PhD student, in an EMT course, I think, he started asking his professor *what happens if there is another slit in the double slit experiment?* His professor said, just take another copy of the electric field with twice the phase difference and add it to the other two. He then asked what happens if we have a fourth one. Then the professor replied in a similar fashion. I think after

the tenth slit, the professor got mad. But do you see where this is going? What if we have infinitely many slits on the cardboard? Well, then we wouldn't have the cardboard anymore?! The heck is going on? Some time later, Feynman would conclude that the particles you send at your cardboard go through all the slits. Then they interfere. When you are working in a lab and send an electron beam to a measuring device, there is a finite probability that your electrons visit the nearest galaxy—whatever it is, I'm not good at cosmology.

This is interesting. If you try to enforce your beam to pass through a single slit, you don't see the interference. What you are doing is that you collapse the wavefunction to that particular state. And by the way, it goes usually without saying but it was Dirac who postulated that when you perform a measurement, the wavefunction collapses one of the eigenkets of the operator with which you are doing the measurement. Anyways, let's consider the following experiment now.

Suppose you have an oven, or any other form of source for some beam of particles. Suppose I have an SG-like device but I'm measuring some arbitrary observable, say A. Since it's a hermitian operator, we know that it has real eigenvalues:

$$A|a\rangle = a|a\rangle, \quad a \in \mathbb{R} \tag{220}$$

I don't really care about the dimensionality now. Suppose we have 10 nondegenerate outcomes in the spectrum, but that's not really important. Now, I'm sending my beam in the initial state $|\psi\rangle$ and after it comes of this first SG-like device, out of all possible components, I'm going to just select one of them by blocking the rest. Let's call this selected state $|a\rangle$. Now, suppose I take another SG-like device but this time, for another observable, say B. Assume for concreteness that A and B are incompatible observables, so that $|a\rangle$ just doesn't go straight through this second device. In this case, we'll have lots of beams coming out of this second device, assuming that we have again an observable with an extended spectrum than just having two possible outcomes. Let's select one of them. I mean, like we did for the double-slit experiment, let's emphasize (or enforce) that we want to have just a certain state coming out of this second device. Call it $|b\rangle$—one of the eigenstates of the B operator. Block out the rest and direct this beam to a third device, which measure some other observable C, which we assume doesn't commute with A or B. Okay, things are getting spicier now, but I promise you that this is the last device I'm

going to consider. Now, after this third measurement, I want to select just one eigenstate again, which I may call $|c\rangle$. The question is, what is the probability for this process?

Let's take a step back and focus on the *transition amplitude* of this process. We have something like this, for which I'm going to consider what happens just after my beam leaves the A machine:

$$T^{b}_{c\leftarrow a} = \langle c|b\rangle\langle b|a\rangle = \langle c|P_b|a\rangle \tag{221}$$

Just as I promised, just evaluate the appropriate projection operator between the initial and final states. (That's why we write $c \leftarrow a$, actually.) Note that the superscript indicates the channel I'm choosing. Now, take its square to get the probability.

Suppose now I repeat this single experiment N_B times, where N_B is the size of the spectrum of the B observable. But each time, I select only one channel in this B machine and I record it down in my notebook—so as to keep track of what I'm doing. What is that I'm doing here? Each time, in each experiment, I'm collapsing my intermediate state to a very specific eigenstate of the B operator. At the end of the end, I want to compute the probability of getting $|c\rangle$ after each experiment. How do I do that? Since each time I'm recording (*selecting*) the intermediate channel (just as in the double-slit experiment), I'm collapsing my state to a single intermediate ket. But I'm repeating this a couple of times. And these experiments are statistically independent (or *uncorrelated*, if you wish). So I just add up my probabilities that I get at each of the experiments:

$$\begin{aligned} P_{c\leftarrow a} &= |T^{b_1}_{c\leftarrow a}|^2 + |T^{b_2}_{c\leftarrow a}|^2 + \cdots + |T^{b_{N_B}}_{c\leftarrow a}|^2 \\ &= \sum_b |T^{b}_{c\leftarrow a}|^2 \end{aligned} \tag{222}$$

where I've suppressed the summation limits for the sake of convenience. Now let's expand this expression:

$$\begin{aligned} P_{c\leftarrow a} &= \sum_b \langle c|P_b|a\rangle^* \langle c|P_b|a\rangle \\ &= \sum_b \langle a|b\rangle\langle b|c\rangle\langle c|b\rangle\langle b|a\rangle \end{aligned} \tag{223}$$

I can't simply get rid of this summation using the completeness of the eigenstates of the B operator because there are two factors of ket-bra combinations. I can do this, instead:

$$P_{c\leftarrow a} = \sum_b |\langle b|a\rangle|^2 |\langle b|c\rangle|^2 \tag{224}$$

I think this is as far as I can go.

Now, let's repeat the experiment (i.e. sending a beam from one side of the A machine, letting it pass through B, and measuring a very specific eigenstate of the C machine) in a slightly adjusted manner. Now, I will not record anything on the B device. It will be there, but I'll choose to be voluntarily blind to its actions. What happens now? What's the probability of getting this $|c\rangle$ state from $|a\rangle$?

This is the point where we remember the funny story involving Feynman. We just focus on the kets $|a\rangle$ and $|c\rangle$ and write the transition amplitde as

$$T_{c\leftarrow a} = \langle c|a\rangle \tag{225}$$

But here, if you think hard about this, we see that in fact our beam goes through all the B channels!

$$T_{c\leftarrow a} = \sum_b \langle c|P_b|a\rangle \tag{226}$$

What is the probability now?

$$\begin{aligned} P_{c\leftarrow a} &= |T_{c\leftarrow a}|^2 \\ &= \left[\sum_b \langle c|P_b|a\rangle\right]^* \left[\sum_{b'} \langle c|P_{b'}|a\rangle\right] \\ &= \sum_{b,b'} \langle a|b\rangle\langle b|c\rangle\langle c|b'\rangle\langle b'|a\rangle \end{aligned} \tag{227}$$

I think I'm bored with this, so I'll stop messing around with these expressions. The moral of the story is this. If you measure intermediate steps, the individual probabilities add up. Otherwise, you'll be totally blind to what this intermediate device does.

* * *

It's Mathematica time. Let me drop you a couple of hints that you may find useful. More on them will be in my supplementary notes for the key to the second homework.

- There are two types of outer product. Perhaps they have distinct names but I'll just pass on them. We have this first one that we use to construct projections operators from eigenkets, $P_\lambda = |\lambda\rangle\langle\lambda|$, which gives you an N-by-N matrix if your dimensionality is N, and there is this other one that we use to construct, e.g., the Gamma

matrices from the Pauli matrices, $\Gamma = \sigma \otimes \sigma'$, where σ and σ' are some Pauli matrices, If you have the second one, you'll get a large matrix at the end, e.g. the product $A_{m\times n} \otimes B_{p\times q}$ will return a $(m+p)$-by-$n+q$ matrix. Let's show how they work on Mathematica.

The first one is given by the function `Outer`. Suppose I denote my ket by the $\boldsymbol{x}$ vector and my dimensionality is 3:

```
dim = 3;
x = x /@ Range[dim];
Px = Outer[Times, x, Conjugate[x]]
```

It's important to take the conjugate here because we have a bra as the second factor.

The other outer product is given by the function `KroneckerProduct`. Suppose you want to compute $\Gamma = \sigma_2 \otimes \sigma_0$. Please don't define the Pauli matrices by hand!

```
s[i_] := PauliMatrix[i]
G = KroneckerProduct[s[2], s[0]]
```

I use the `MatrixForm` (when applicable) and `TraditionalForm` on my outputs:

```
Px // MatrixForm // TraditionalForm
```

Now this is very important. If you have anything to simplify, do it before you change the form of the output:

```
Px //
    Simplify[#, Assumptions -> {(* some assumptions here *)}] //
    MatrixForm //
    TraditionalForm
```

otherwise, you'll encounter bugs.

- Now we have this `Px` matrix but it looks ugly. And I like pretty things—I mean, who doesn't?! So, let's write a command that will prettify the output.

```
Prettify[expr_] :=
    expr /. {
        x[i_] :> Subscript[x, i]
    } //
        Style[
            #,
            {
                FontSize -> 20,
                FontColor -> Black,
                FontFamily -> "Latin Modern Roman"
            }
            ]
```

  Try running `Prettify[Px] // MatrixForm // TraditionalForm` (Try also doing it without the TraditionalForm—you'll see the function heads `Conjugate` lurking around. This `TraditionalForm` function converts things to a more human-friendly form).

- Now suppose you have this $\boldsymbol{x}$ vector with real components. How do you explain this Mathematica?

```
$Assumptions =
    {
        x[1] \[Element] Reals,
        x[2] \[Element] Reals,
        x[3] \[Element] Reals
    };
```

Boring, right? Suppose you have a 10-component vector like this. Nobody has time for that. Use the pattern-recognition of Mathematica:

```
$Assumptions =
    {
        x[_] \[Element] Reals
    };
```

Doing so, whatever your dimension is, all the x components will be assumed real. Furthermore, suppose you want to have a normalized $\boldsymbol{x}$.

```
$Assumptions =
    {
        x[1]^2 + x[2]^2 + x[3]^2 == 1
    };
```

Again, long and boring. Do this instead:

```
$Assumptions =
    {
        Total[(x /@ Range[dim])^2] == 1
    };
```

Let's break it down. That funny symbol, i.e. `/@`, means mapping over a list by taking whatever it is on the left as a function. First, Mathematica will expand `Range[3]` into `{1, 2, 3}`. Then, it will take `x` and distribute it over this list, bringing each entry into the argument of `x`:

```
x /@ Range[3] (* = {x[1], x[2], x[3]}*)
```

Notice that we square this array. In some programming languages, this may be the square of the array, e.g. like a dot product with itself. But in Mathematica, the power of an array is carried out entry by entry. So, `(x /@ Range[3])^2` will give you $\{x_1^2, x_2^2, x_3^2\}$. Now, finally, we have this `Total` function. It just adds all the entries of an array together. That's it.

- Note that `$Assumptions` works at a global level, so as long as you don't quit your kernel, and if you evaluate your assumptions at the beginning before anything else, you should be just fine.

- In the item before the previous one, I played a very dangerous game. I defined the `$Assumptions` variable twice. If you do that, your most recent `$Assumptions` will overwrite the ones that precede it. So, it's a good strategy to either put all the assumptions in a single variable, or join them as you go down the line. I mean, either do

```
$Assumptions =
    {
        x[_] \[Element] Reals,
        Total[(x /@ Range[dim])^2] == 1
    };
```

or

```
$Assumptions =
    {
        x[_] \[Element] Reals
    };
(* some other code pieces here *)
$Assumptions =
    Join[
        $Assumptions,
        {Total[(x /@ Range[dim])^2] == 1}
```

```
];
```

Well, the third option would be a mix usage of them, which is perfectly fine. The problem with Mathematica is that, this `$Assumptions` variable is not a dynamic one.

- Now, you have this `Px` matrix. And if you introduce your assumptions to Mathematica in a solid way and then evaluate the `Px` matrix again, you'll see that you don't get rid of the complex conjugation right away. You need to simplify it. Unless there is good reason (e.g. all the variables being numerical and in the float form), Mathematica will just give you what you write back (Do you remember the Casio scientific calculators? The ones that'd give you, for instance when you enter 13/6, 13/6? Yeah, like that.) So, run a `Simplify` on this projection matrix.

- There is a really smart trick if you don't want to deal with the assumptions. It is the `ComplexExpand` function. Essentially, it assumes that everything other than the complex unit is real. Cool, huh? But use it with caution. Sometimes you won't be able to simplify the argument function, `Arg`. In this case, just try to break down your code into smaller pieces and use a `Simplify` on the variables which you know are real definitely.

- But still, this is just a machine doing some symbolic manipulations based on the most general assumptions, right? There will be points at which an outside, intelligent human-being should intervene. For instance, if you get the `Sign` function somewhere in your eigenvectors, then try using the gauge freedom (the phase-redefinition, basically) to see if you can extract this sign in a consistent manner so that the diagonalizing matrix that you'd construct using your eigenvectors still diagonalizes your initial matrix properly.

That's all, folks!

VI. DISCUSSION 5 (OCT 27)

Problem 1

$$|\psi\rangle = \begin{pmatrix} \cos\left(\frac{\alpha}{2}\right) \\ \sin\left(\frac{\alpha}{2}\right) e^{i\beta} \end{pmatrix}$$

$$\alpha = \arccos\left(\frac{1}{\sqrt{3}}\right), \quad \beta = \frac{\pi}{4}$$

(a) $O_1 = \hat{n}_i \cdot \sigma_i$

$O_1^2 = 1 \Rightarrow \lambda_1 = \pm 1$

$$P_1^+ = \frac{O_1 - (-1)}{(1) - (-1)} = \frac{O_1 + 1}{2}$$

$$P_1^- = \frac{O_1 - (1)}{(-1) - (1)} = \frac{O_1 - 1}{-2}$$

$$\Big\} \; P_1^{\pm} = \frac{1 \pm O_1}{2}$$

$T_{1_1 \leftarrow \psi} = \langle 1_1 | P_1^+ | \psi \rangle$

$T_{-1_1 \leftarrow \psi} = \langle -1_1 | P_1^- | \psi \rangle$

$$P_{1_1 \leftarrow \psi} = |T_{1_1 \leftarrow \psi}|^2$$

$$= |\langle 1_1 | P_1^+ | \psi \rangle|^2$$

$$= \langle \psi | P_1^+ | 1_1 \rangle \langle 1_1 | P_1^+ | \psi \rangle$$

$$= \langle \psi | P_1^+ | \psi \rangle$$

$$P_{-1_1 \leftarrow \psi} = |T_{-1_1 \leftarrow \psi}|^2$$

$$= \langle \psi | P_1^- | \psi \rangle$$

(b)

$$O_2 = \hat{n}_i \cdot \sigma_i$$

$$O_2^2 = 1 \Rightarrow \lambda_2 = \pm 1$$

$$P_2^+ = \frac{O_2 - (-1)}{(1) - (-1)} = \frac{O_2 + 1}{2}$$

$$P_2^- = \frac{O_2 - (1)}{(-1) - (1)} = \frac{O_2 - 1}{-2}$$

$$\Bigg\} \; P_2^{\pm} = \frac{1 \pm O_2}{2}$$

$$T_{1_2 \leftarrow -1_1 \leftarrow \psi} = \langle 1_2 | P_2^+ P_1^- | \psi \rangle$$

$$T_{-1_2 \leftarrow -1_1 \leftarrow \psi} = \langle -1_2 | P_2^- P_1^- | \psi \rangle$$

$$P_{1_2 \leftarrow -1_1 \leftarrow \psi} = |T_{1_2 \leftarrow -1_1 \leftarrow \psi}|^2$$

$$= \langle \psi | P_1^- P_2^+ | -1_2 \rangle \langle -1_2 | P_2^+ P_1^- | \psi \rangle$$

$$= \langle \psi | P_1^- P_2^+ P_1^- | \psi \rangle$$

$$P_{-1_2 \leftarrow -1_1 \leftarrow \psi} = |T_{-1_2 \leftarrow -1_1 \leftarrow \psi}|^2$$

$$= \langle \psi | P_1^- P_2^- | -1_2 \rangle \langle -1_2 | P_2^- P_1^- | \psi \rangle$$

$$= \langle \psi | P_1^- P_2^- P_1^- | \psi \rangle$$

problem 1

part (a)

In[1]:=
```
alpha = ArcCos[1 / Sqrt[3]];
beta = Pi / 4;
psi = {Cos[alpha / 2], Sin[alpha / 2] Exp[I beta]};
n = {Cos[alpha] Cos[beta], Cos[alpha] Sin[beta], -Sin[alpha]};
s = PauliMatrix /@ Range[3];
O1 = Total[n s];
I2 = IdentityMatrix[2];
P1[1] = (I2 + O1) / 2;
P1[-1] = (I2 - O1) / 2;
p[psiToO1plus] = Conjugate[psi].P1[1].psi;
p[psiToO1minus] = Conjugate[psi].P1[-1].psi;
```

In[12]:=
```
p[psiToO1plus] // N
p[psiToO1minus] // N
```

Out[12]= $0.5 - 1.54311 \times 10^{-17}\, i$

Out[13]= $0.5 - 5.38561 \times 10^{-18}\, i$

part (b)

In[14]:=
```
n1 = {-Sin[beta], Cos[beta], 0};
O2 = Total[n1 s];
P2[1] = (I2 + O2) / 2;
P2[-1] = (I2 - O2) / 2;
p[psiToO1minusToO2plus] = Conjugate[psi].P1[-1].P2[1].P1[-1].psi;
p[psiToO1minusToO2minus] = Conjugate[psi].P1[-1].P2[1].P1[-1].psi;
```

In[20]:=
```
p[psiToO1minusToO2plus] // N
p[psiToO1minusToO2minus] // N
```

Out[20]= $0.25 + 2.27328 \times 10^{-17}\, i$

Out[21]= $0.25 + 2.27328 \times 10^{-17}\, i$

Problem 2

$$M = \hat{n}_i M_i$$

$$M^{2k} = 1 - P_n, \quad k \in \mathbb{Z}^+$$

$$M^{2k+1} = M$$

$$H = \alpha M + \beta M^2$$

$$M \Rightarrow \mu = 1, 0, -1$$

$$[H, M] = 0 \Rightarrow \exists |\mu\eta\rangle : H|\mu\eta\rangle = \eta|\mu\eta\rangle$$

$$M|\mu\eta\rangle = \mu|\mu\eta\rangle$$

(a) $H|\mu\eta\rangle = \eta|\mu\eta\rangle$

$$= (\alpha M + \beta M^2)|\mu\eta\rangle$$

$$= (\alpha\mu + \beta\mu^2)|\mu\eta\rangle$$

μ	η
1	$\alpha+\beta$
0	0
-1	$-\alpha+\beta$

spectrum is nondegenerate.

$$P_H^{\alpha+\beta} = \frac{H-(0)}{(\alpha+\beta)-(0)} \frac{H-(-\alpha+\beta)}{(\alpha+\beta)-(-\alpha+\beta)}$$

$$P_H^{0} = \frac{H-(\alpha+\beta)}{(0)-(\alpha+\beta)} \frac{H-(-\alpha+\beta)}{(0)-(-\alpha+\beta)}$$

$$P_H^{-\alpha+\beta} = \frac{H-(\alpha+\beta)}{(-\alpha+\beta)-(\alpha+\beta)} \quad \frac{H-(0)}{(-\alpha+\beta)-(0)}$$

(b) $\alpha = \beta$: $\eta = 2\alpha,\ 0,\ 0$

$$P_H^{2\alpha} = \frac{H-(0)}{(2\alpha)-(0)}$$

$$P_H^{0} = \frac{H-(2\alpha)}{(0)-(2\alpha)}$$

$\alpha = -\beta$: $\eta = 0,\ 0,\ -2\alpha$

$$P_H^{0} = \frac{H-(-2\alpha)}{(0)-(-2\alpha)}$$

$$P_H^{-2\alpha} = \frac{H-(0)}{(-2\alpha)-(0)}$$

(c) $\alpha = \beta$

$$T_{0_H \leftarrow \psi} = \langle 0_H | P_H^0 | \psi \rangle$$

$$P_{0_H \leftarrow \psi} = |T_{0_H \leftarrow \psi}|^2$$

$$= \langle \psi | P_H^0 | 0_H \rangle \langle 0_H | P_H^0 | \psi \rangle$$

$$= \langle \psi | P_H^0 | \psi \rangle$$

problem 2

part (a)

In[1]:=
```
M /: M^k_?(# > 2 && OddQ[#] &) = M;
M /: M^k_?(# > 1 && EvenQ[#] &) = M^2;
H = alpha M + beta M^2;
PH[alphaPlusbeta] = (H - (0))/((alpha + beta) - 0) (H - (-alpha + beta))/((alpha + beta) - (-alpha + beta));
PH[0] = (H - (alpha + beta))/((0) - (alpha + beta)) (H - (-alpha + beta))/((0) - (-alpha + beta));
PH[malphaPlusbeta] = (H - (alpha + beta))/((-alpha + beta) - (alpha + beta)) (H - (0))/((-alpha + beta) - (0));
```

In[7]:=
```
PH[alphaPlusbeta] // Expand // Simplify
```

Out[7]=
$$\frac{1}{2} M (1 + M)$$

In[8]:=
```
PH[0] // Expand // Simplify
```

Out[8]=
$$1 - M^2$$

In[9]:=
```
PH[malphaPlusbeta] // Expand // Simplify
```

Out[9]=
$$\frac{1}{2} (-1 + M) M$$

part (b)

In[10]:=
```
PH[alphaeqbeta, twoalpha] = (H - (0))/((2 alpha) - (0));
PH[alphaeqbeta, 0] = (H - (2 alpha))/((0) - (2 alpha));
PH[alphaeqmbeta, 0] = (H - (-2 alpha))/((0) - (-2 alpha));
PH[alphaeqmbeta, mtwoalpha] = (H - (0))/((-2 alpha) - (0));
```

2 |

In[14]:=
```
PH[alphaeqbeta, twoalpha] // Simplify
PH[alphaeqbeta, 0] // Simplify
PH[alphaeqmbeta, 0] // Simplify
PH[alphaeqmbeta, mtwoalpha] // Simplify
```

Out[14]= $\frac{M\,(\text{alpha} + \text{beta}\,M)}{2\,\text{alpha}}$

Out[15]= $1 - \frac{M}{2} - \frac{\text{beta}\,M^2}{2\,\text{alpha}}$

Out[16]= $\frac{\text{beta}\,M^2 + \text{alpha}\,(2 + M)}{2\,\text{alpha}}$

Out[17]= $-\frac{M\,(\text{alpha} + \text{beta}\,M)}{2\,\text{alpha}}$

part (c)

In[18]:=
```
nj = {Sin[theta] Cos[phi], Sin[theta] Sin[phi], Cos[theta]};
Mj = {{{0, 0, 0}, {0, 0, -I}, {0, I, 0}},
   {{0, 0, I}, {0, 0, 0}, {-I, 0, 0}}, {{0, -I, 0}, {I, 0, 0}, {0, 0, 0}}};
PHzero = (PH[alphaeqbeta, 0] // Expand) /. M^2 → M.M /. M → Total[nj Mj];
```

In[21]:=
```
psi = {Sin[theta] Cos[phi], Sin[theta] Sin[phi], Cos[theta]};
P[psiToHzero] = Conjugate[psi].PHzero.psi;
```

In[23]:=
```
P[psiToHzero] // ComplexExpand // Simplify
```

Out[23]= $(\text{Cos}[\text{theta}] + (\text{Cos}[\text{phi}] + \text{Sin}[\text{phi}])\,\text{Sin}[\text{theta}])^2$

Problem 3

$$H = d_\alpha \Gamma_\alpha$$

Γ_α : anticommuting and traceless 4-by-4 matrices

$$\{\Gamma_\alpha, \Gamma_\beta\} = 2\delta_{\alpha\beta}$$

(a) $$H^2 = d_\alpha d_\beta \underbrace{\Gamma_\alpha \Gamma_\beta}_{\frac{1}{2}\{\Gamma_\alpha, \Gamma_\beta\}} = \vec{d}^{\,2}$$

$\eta = \pm\sqrt{\vec{d}^{\,2}}$: $tr(H) = 0$ ∴ each EV is 2-fold degen.

$$P_H^+ = \frac{H-(-E)}{(E)-(-E)} = \frac{H+E}{2E}$$

$$P_H^- = \frac{H-(E)}{(-E)-(E)} = \frac{H-E}{-2E}$$

$$P_H^\pm = \frac{E \pm H}{2E}$$

$$E = \sqrt{\vec{d}^{\,2}}$$

(b) $$M = d_i\, \sigma_0 \otimes \sigma_i, \quad i \neq 0$$

$$H = d_\alpha \Gamma_\alpha = d_i\, \sigma_1 \otimes \sigma_i + d_4\, \sigma_2 \otimes \sigma_0 + d_5\, \sigma_3 \otimes \sigma_0$$

$$[d_i \sigma_0 \otimes \sigma_i,\ d_j \sigma_1 \otimes \sigma_j]$$

$$= \underbrace{d_i d_j}_{\text{sym.}}\ \sigma_1 \otimes \underbrace{[\sigma_i, \sigma_j]}_{\text{antisym.}}$$

$$= 0$$

$$[\sigma_0 \otimes \sigma_i,\ \sigma_j \otimes \sigma_0] = \sigma_j \otimes \sigma_i - \sigma_j \otimes \sigma_i = 0$$

$\therefore [M, H] = 0$

(c) $\exists\, |\mu\eta\rangle : H|\mu\eta\rangle = \eta|\mu\eta\rangle$

$M|\mu\eta\rangle = \mu|\mu\eta\rangle$

$$|++\rangle = P_M^+ P_H^+ |1\rangle$$
$$|+-\rangle = P_M^+ P_H^- |2\rangle$$
$$|-+\rangle = P_M^- P_H^+ |3\rangle$$
$$|--\rangle = P_M^- P_H^- |4\rangle$$

} up to normalization

$$\mathcal{U} = \left(|++\rangle \;\; |+-\rangle \;\; |-+\rangle \;\; |--\rangle\right)$$

$$\mathcal{U}^\dagger H \mathcal{U} = \mathrm{diag}(E, -E, E, -E)$$

$$\mathcal{U}^\dagger M \mathcal{U} = \mathrm{diag}(m, m, -m, -m)$$

$|1\rangle, |2\rangle, |3\rangle, |4\rangle$: canonical basis

In[1]:=
```
s[i_] := PauliMatrix[i]
ss[i_, j_] := KroneckerProduct[s[i], s[j]]
G[1] = ss[1, 1];
G[2] = ss[1, 2];
G[3] = ss[1, 3];
G[4] = ss[2, 0];
G[5] = ss[3, 0];
G1[1] = ss[0, 1];
G1[2] = ss[0, 2];
G1[3] = ss[0, 3];
I4 = IdentityMatrix[4];
H = Total[(d /@ Range[5]) (G /@ Range[5])];
M = Total[(d /@ Range[3]) (G1 /@ Range[3])];
e = Sqrt[Total[(d /@ Range[5])^2]];
m = Sqrt[Total[(d /@ Range[3])^2]];
$Assumptions = {d[_] ∈ Reals, e^2 > 0, e > 0, m^2 > 0, m > 0};
P[H, sgn_] := (e I4 + sgn H)/(2 e)
P[M, sgn_] := (m I4 + sgn M)/(2 m)
CB[i_] := Table[If[i == k, 1, 0], {k, Range[4]}];
```

In[20]:=
```
EVec[h_, m_, i_] := Normalize[P[H, h].Normalize[P[M, m].CB[i]]]
U = {
    EVec[1, 1, 1],
    EVec[1, -1, 2],
    EVec[-1, 1, 3],
    EVec[-1, -1, 4]
   } // Transpose;
```

In[22]:=
```
DH = ConjugateTranspose[U].H.U;
DM = ConjugateTranspose[U].M.U;
```

In[24]:=
```
Prettify[expr_] := expr /.
   d[i_] :> Subscript[d, i] //
  Style[#, {FontSize → 18, FontFamily → "Latin Modern Roman"}] &
```

In[25]:=
```
DH // ComplexExpand // Simplify // FullSimplify // MatrixForm // Prettify
```

Out[25]=

$$\begin{pmatrix} \sqrt{d_1^2+d_2^2+d_3^2+d_4^2+d_5^2} & 0 & 0 & 0 \\ 0 & \sqrt{d_1^2+d_2^2+d_3^2+d_4^2+d_5^2} & 0 & 0 \\ 0 & 0 & -\sqrt{d_1^2+d_2^2+d_3^2+d_4^2+d_5^2} & 0 \\ 0 & 0 & 0 & -\sqrt{d_1^2} \end{pmatrix}$$

In[26]:=
```
DM // ComplexExpand // Simplify // FullSimplify // MatrixForm // Prettify
```

Out[26]=

$$\begin{pmatrix} \sqrt{d_1^2+d_2^2+d_3^2} & 0 & 0 & 0 \\ 0 & -\sqrt{d_1^2+d_2^2+d_3^2} & 0 & 0 \\ 0 & 0 & \sqrt{d_1^2+d_2^2+d_3^2} & 0 \\ 0 & 0 & 0 & -\sqrt{d_1^2+d_2^2+d_3^2} \end{pmatrix}$$

VII. DISCUSSION 6 (NOV 3)

Did you know it was Dirac who first derived the Heisenberg equation? This is usually not stated. But yes, he was such a cool and humble guy, he just wanted to call it the Heisenberg equation, the equation he derived when he was one of his famous, all-day Sunday walks on the countryside. But he couldn't recall the expression for the Poisson brackets that day, so, he had to wait until the morning for the library to open...

* * *

Let's prove the following Baker-Campbell-Hausdorff (BCH) formula—as a matter of fact, there is a family of BCH formulae and this is just one member.

Claim.

$$e^{i\lambda G}Ae^{-i\lambda G} = A + i\lambda[G,A] + \frac{(i\lambda)^2}{2!}[G,[G,A]] + \frac{(i\lambda)^3}{3!}[G,[G,[G,A]]] + \cdots \tag{228}$$

Proof. Define the left-hand side to be a function of λ and expand it in a Taylor series around $\lambda = 0$.

$$f(\lambda) = e^{i\lambda G}Ae^{-i\lambda G} \tag{229}$$

$$f(0) = A \tag{230}$$

$$f'(\lambda) = iGe^{i\lambda G}Ae^{-i\lambda G} + e^{i\lambda G}Ae^{-i\lambda G}(-iG) \tag{231}$$

$$f'(0) = iGA - iAG = i[G,A] \tag{232}$$

$$f''(\lambda) = iGiGe^{i\lambda G}Ae^{-i\lambda G} + iGe^{i\lambda G}Ae^{-i\lambda G} + iGe^{i\lambda G}Ae^{-i\lambda G}(-iG) + e^{i\lambda G}Ae^{-i\lambda G}(-iG)(-iG) \tag{233}$$

$$f''(0) = i^2GGA - i^2GAG - i^2GAG + i^2AGG = i^2[G,[G,A]] \tag{234}$$

Continue in this fashion:

$$f^{(k)}(0) = i^k[\underbrace{G,\ldots[G}_{k\text{ times}},A]] \tag{235}$$

Taylor expansion of f:

$$f(\lambda) = \sum_{k\geq 0}\frac{(i\lambda)^k}{k!}[\underbrace{G,\ldots[G}_{k\text{ times}},A]] \tag{236}$$

qed.

Let's make an application of this formula. Suppose I take $G = S_3$ and $A = S_1$, where S_1 and S_3 are the x and z components of the spin operator—but mind that I'm not telling you what the spin value is. You'll see it doesn't matter.

$$e^{i\lambda S_3} S_1 e^{-i\lambda S_1} = S_1 + i\lambda[S_3, S_1] + \frac{(i\lambda)^2}{2!}[S_3, [S_3, S_1]] + \cdots \tag{237}$$

Now this expression may look like a beast but in reality, no, it's the other one, the beauty. For very specific cases, like the spin operators or the position and momentum, we see some alternating terms here so that you can express the result in a compact manner. Let's start evaluating the commutators. I'll do the first four:

$$[S_3, S_1] = i\hbar S_2 \tag{238}$$

$$[S_3, [S_3, S_1]] = -(i\hbar)^2 S_1 = \hbar^2 S_1 \tag{239}$$

$$[S_3, [S_3, [S_3, S_1]]] = i\hbar^3 S_2 \tag{240}$$

$$[S_3, [S_3, [S_3, [S_3, S_1]]]] = \hbar^2 S_1 \tag{241}$$

where I've made successive use of $[S_3, S_1] = i\hbar S_2$ and $[S_3, S_2] = -i\hbar S_1$. Now we see that

$$[\underbrace{S_3, \ldots [S_3}_{2k \text{ times}}, S_1]] = \hbar^{2k} S_1 \tag{242}$$

$$[\underbrace{S_3, \ldots [S_3}_{2k+1 \text{ times}}, S_1]] = i\hbar^{2k+1} S_1 \tag{243}$$

Then we see that

$$e^{i\lambda S_3} S_1 e^{-i\lambda S_1} = \sum_{k\geq 0} \frac{(i\lambda)^k}{k!} [\underbrace{S_3, \ldots [S_3}_{k \text{ times}}, S_1]] \tag{244}$$

$$= \sum_{k\geq 0} \frac{(i\lambda)^{2k}}{(2k)!} \hbar^{2k} S_1 + \sum_{k\geq 0} \frac{(i\lambda)^{2k+1}}{(2k+1)!} i\hbar^{2k+1} S_2 \tag{245}$$

$$= S_1 \cos(\lambda\hbar) - S_2 \sin(\lambda\hbar) \tag{246}$$

Let $\lambda = -\omega t/\hbar$:

$$e^{-i\omega t S_3/\hbar} S_1 e^{i\omega t S_3/\hbar} = S_1 \cos(\omega t) + S_2 \sin(\omega t) \tag{247}$$

Looks familiar? As an exercise, show that

$$e^{-i\omega t S_3/\hbar} S_2 e^{i\omega t S_3/\hbar} = S_2 \cos(\omega t) - S_1 \sin(\omega t) \tag{248}$$

Let's keep going:

$$e^{-i\omega t S_3/\hbar} S_3 e^{i\omega t S_3} = S_3 \tag{249}$$

Then, we see something worth looking twice:

$$e^{-i\omega t S_z/\hbar} \boldsymbol{S} e^{i\omega t S_z/\hbar} = R\boldsymbol{S} \tag{250}$$

where

$$R = \begin{pmatrix} \cos(\omega t) & \sin(\omega t) & 0 \\ -\sin(\omega t) & \cos(\omega t) & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{251}$$

$$* * *$$

In quantum mechanics, in addition to the Schrödinger and Heisenberg pictures, there is another one, called the Dirac (or *interaction*) picture. In this framework, both states and operators are evolving in time.

We define an *intermediate* state ket, say α_I, via

$$|\alpha(t)\rangle = U_0(t)|\alpha_I(t)\rangle \tag{252}$$

where $U_0(t) = e^{-iH_0 t/\hbar}$ is some time-evolution operator. Now, what is this Hamiltonian, H_0?

We write down the *total* Hamiltonian for some problem as

$$H = H_0 + H_1 \tag{253}$$

Here, H_0 is the part we love so much that we'd , and we know everything that exists to know about it—the spectrum, eigenkets, all the good stuff. For the time being, we'll assume that this part is time-independent. Consider, for example, a spin-1 particle under some magnetic field aligned specifically along the z axis, i.e. $H_0 = -\mu B_0 S_z$. H_1, on the other hand, can be any wild-looking Hamiltonian for our purposes. As an example, consider a rotating magnetic field on the xy plane on top of this existing one, i.e. $H_1 = -\mu B_1[\cos(\omega t)S_x + \sin(\omega t)S_y]$. The main motivation of employing this Dirac picture is that we treat this second piece of the Hamiltonian as a "perturbation". I'm using

quotation marks here because there are problems that can be solved exactly, so it's just a name, you know.

Now, let's write down the Schrödinger equation and play around with it:

$$i\hbar\frac{\partial}{\partial t}|\alpha(t)\rangle = H|\alpha(t)\rangle \tag{254}$$

The first order of business is to write down the Hamiltonian explicitly as the two parts given above and next, we just use the definition of the intermediate state ket:

$$i\hbar\frac{\partial}{\partial t}\left[e^{-iH_0t/\hbar}|\alpha_I(t)\rangle\right] = H_0e^{-iH_0t/\hbar}|\alpha_I(t)\rangle + H_1e^{-iH_0t/\hbar}|\alpha_I(t)\rangle \tag{255}$$

Use the chain rule on the left-hand side:

$$i\hbar\frac{-iH_0}{\hbar}e^{-iH_0t/\hbar}|\alpha_I(t)\rangle + i\hbar e^{-iH_0t/\hbar}\frac{\partial}{\partial t}|\alpha_I(t)\rangle = H_0e^{-iH_0t/\hbar}|\alpha_I(t)\rangle + H_1e^{-iH_0t/\hbar}|\alpha_I(t)\rangle \tag{256}$$

The first term on the left cancels the first term on the right:

$$i\hbar e^{-iH_0t/\hbar}\frac{\partial}{\partial t}|\alpha_I(t)\rangle = H_1e^{-iH_0t/\hbar}|\alpha_I(t)\rangle \tag{257}$$

Multiply both sides from left by $e^{iH_0t/\hbar}$:

$$i\hbar\frac{\partial}{\partial t}|\alpha_I(t)\rangle = e^{iH_0t/\hbar}H_1e^{-iH_0t/\hbar}|\alpha_I(t)\rangle \tag{258}$$

Let

$$H_I(t) = e^{iH_0t/\hbar}H_1e^{-iH_0t/\hbar} \tag{259}$$

so we have

$$i\hbar\frac{\partial}{\partial t}|\alpha_I(t)\rangle = H_I(t)|\alpha_I(t)\rangle \tag{260}$$

This is the Schrödinger equation in the Dirac picture.

Most naively, we can assert the existence of a time-evolution operator $U_I(t)$ such that

$$U_I(t)|\alpha_I(0)\rangle = |\alpha_I(t)\rangle \tag{261}$$

With this, the Schrödinger equation is written in terms of the operators now:

$$i\hbar\frac{\partial}{\partial t}U_I(t) = H_I(t)U_I(t) \tag{262}$$

The story doesn't end here. We still have to answer the question *how do we go from* $|\alpha(0)\rangle$ *to* $|\alpha(t)\rangle$*?* That's the million-dollar problem. Let's observe something. If we have

$$|\alpha(t)\rangle = e^{-iH_0 t/\hbar}|\alpha_I(t)\rangle \tag{263}$$

then we see that

$$|\alpha(0)\rangle = |\alpha_I(0)\rangle \tag{264}$$

Now let's bring everything together:

$$\begin{aligned} |\alpha(t)\rangle &= U_0(t)|\alpha_I(t)\rangle \\ &= U_0(t)U_I(t)|\alpha_I(0)\rangle \\ &= U_0(t)U_I(t)|\alpha(0)\rangle \end{aligned} \tag{265}$$

This is simply marvelous. It is just so beautiful that it makes you cry, you know, like the sight of Grand Canyon. We managed to separate the time-evolution into two parts. First, we carry our given, initial state to an intermediate state by performing a time translation that's generated by the Hamiltonian H_I, which is given in terms of the perturbing part of the total Hamiltonian. Then, we let it evolve further in time but this time with the fully known part of the Hamiltonian.

As it also happens in life, the beautiful are mostly unachievable. If we get an intermediate Hamiltonian, H_I, that looks easy (in the sense that it may depend on time but nevertheless commutes with itself at different times), then that's great. If that's not the case, we employ some other machinery to get there. It's the *particle way* of doing calculations, which I myself especially enjoy. We'll see it in action below. I want to cover just one more point before that.

Even though we mentioned only the *intermediate Hamiltonian,* H_I, you can essentially define an intermediate operator for any of your operators. Suppose I have some operator V in the Schrödinger picture. Now I'm making the explicit assumption that I'm working with operators in the Schrödinger picture. Hamiltonian can depend on time and in fact, it generates the time evolution, so in a way, it has a special place in our hearts; nevertheless, we should be really careful with any other operator. Note that operators in the Schrödinger picture *can* have time-dependence. The essential difference is this: What causes this time-dependence? Is it like $\cos(\omega t)$ tacked in to a spin operator? Yes!

It's something that you can change physically—or something that changes physically. Heisenberg operators have time dependence via the unitary time-evolution operator. If you have an operator like $S_z \cos(\omega t)$ in the Schrödinger picture, which you *can*, then your operator in the Heisenberg picture would look like $U(t)^\dagger S_z \cos(\omega t) U(t)$. If this is confusing, well, it should be. Heisenberg operators evolve this way so as to prevent states from evolving, if you like. Therefore, it doesn't really matter if your operator has some explicit, crazy-looking time-dependence tacked in to it—it will always be your Schrödinger operator... until you let it evolve with a unitary time-evolution operator relevant to your problem.

Now, I can define the corresponding *intermediate operator* or the *interaction operator* by

$$V_I(t) = e^{iH_0 t/\hbar} V_S e^{-iH_0 t/\hbar} \tag{266}$$

where I think this is the only place I'll use the subscript S to denote an operator in the Schrödinger picture. Then look what happens if I take the time derivative of both sides:

$$\begin{aligned}
\frac{\mathrm{d}V_I(t)}{\mathrm{d}t} &= e^{iH_0 t/\hbar} \left(\frac{iH_0}{\hbar} V - V \frac{iH_0}{\hbar} \right) e^{-iH_0 t/\hbar} \\
&= \frac{i}{\hbar} e^{iH_0 t/\hbar} [H_0, V] e^{-iH_0 t/\hbar} \\
&= \frac{i}{\hbar} [H_0, e^{iH_0 t/\hbar} V e^{-iH_0 t/\hbar}] \\
&= \frac{i}{\hbar} [H_0, V_I] \\
&= \frac{[V_I, H_0]}{i\hbar}
\end{aligned} \tag{267}$$

[If your Schrödinger operator has explicit time dependence, we'd have to include a term like $\partial V_I / \partial t$ on the right-hand side. But let's avoid that to keep things simple.] This looks nice: the interaction operators evolve with the known part of the Hamiltonian. It just keeps getting better, doesn't it?

Let's solve one problem (Sakurai Problem 5.30) to make things concrete. Suppose I have a spin-1/2 particle under a magnetic field that has a stationary component along the z and that has other components that rotate on the xy plane. The Hamiltonian is then given by something like this:

$$H = -\mu B_0 S_3 - \mu B_1 [\cos(\omega t) S_1 + \sin(\omega t) S_2] \tag{268}$$

Now we make the idenfitication

$$H_0 = -\mu B_0 S_3 \tag{269}$$

Why? Because we know the eigenvalues and the eigenkets of the S_3 operator. This leaves us with

$$H_1 = -\mu B_1[\cos(\omega t)S_1 + \sin(\omega t)S_2] \tag{270}$$

Now if you try computing the H_I operator, you see it won't be any different from H_1 because you'll still get S_1 times cosine of something plus S_2 times sine of something, and this interaction Hamiltonian will not commute with itself at different times.

Now we try something else. We'll exploit the known spectrum of the known Hamiltonian. Suppose I know the eigenkets of H_0 and I can write the following:

$$H_0|n\rangle = E_n|n\rangle \tag{271}$$

I'll exploit the fact that these eigenkets form a complete basis, so that I can expand my intermediate state ket as

$$|\alpha_I(t)\rangle = \sum_n c_n(t)|n\rangle \tag{272}$$

Note that for our actual state ket, we have

$$|\alpha(t)\rangle = e^{-iH_0t/\hbar}|\alpha_I(t)\rangle = \sum_n e^{-iE_nt/\hbar}c_n(t)|n\rangle \tag{273}$$

Let's use the expansion of the intermediate state ket in the Schrödinger equation in the Dirac picture:

$$i\hbar\frac{\partial}{\partial t}|\alpha_I(t)\rangle = H_I(t)|\alpha_I(t)\rangle \tag{274}$$

$$i\hbar\frac{\partial}{\partial t}\sum_n c_n(t)|n\rangle = e^{iH_0t/\hbar}H_1e^{-iH_0t/\hbar}\sum_n c_n(t)|n\rangle \tag{275}$$

$$\sum_n \dot{c}_n(t)|n\rangle = \sum_n e^{iH_0t/\hbar}\frac{H_1}{i\hbar}e^{-iH_0t/\hbar}c_n(t)|n\rangle \tag{276}$$

$$\sum_n \dot{c}_n(t)|n\rangle = \sum_n e^{iH_0t/\hbar}\frac{H_1}{i\hbar}e^{-iE_nt/\hbar}c_n(t)|n\rangle \tag{277}$$

Multiply by $\langle m|$ from left:

$$\sum_n \dot{c}_n(t)\underbrace{\langle m|n\rangle}_{\delta_{mn}} = \sum_n \langle m|e^{iH_0t}\frac{H_1}{i\hbar}e^{-iE_nt/\hbar}c_n(t)|n\rangle \tag{278}$$

Then,

$$\dot{c}_m(t) = \sum_n e^{iE_m t/\hbar} \langle m| \frac{H_1}{i\hbar} e^{-iE_n t/\hbar} c_n(t) |n\rangle \tag{279}$$

$$= \sum_n e^{i(E_m - E_n)t/\hbar} \langle m| \frac{H_1}{i\hbar} |n\rangle c_n(t) \tag{280}$$

$$= \sum_n e^{i\omega_{mn} t} \frac{H_1^{mn}}{i\hbar} c_n(t) \tag{281}$$

where I've defined $\omega_{mn} = \frac{E_m - E_n}{\hbar}$, and the H_1^{mn} are the matrix elements of the perturbing Hamiltonian in the basis spanned by the eigenkets of the known Hamiltonian, H_0. We'll now have N first-order coupled ODEs. We know how to solve this, right?

Let's go back to our spin-1/2 particle in a rotating magnetic field. The known Hamiltonian is

$$H_0 = -\mu B_0 S_3 \tag{282}$$

so it has eigenvalues $E_1 = -\mu B_0 \hbar/2$ and $E_2 = +\mu B_0 \hbar/2$. Its eigenkets are the eigenkets of the third component of the spin operator, i.e. the eigenkets of the third Pauli matrix, which are just the $2D$ canonical basis vectors when represented by matrices:

$$H_0|1\rangle = E_1|1\rangle, \quad H_0|2\rangle = E_2|2\rangle \tag{283}$$

$$|1\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |2\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{284}$$

The perturbing Hamiltonian can be represented by the matrix

$$H_1 = -\mu B_1 [\cos(\omega t) S_1 + \sin(\omega t) S_2] \tag{285}$$

$$\doteq -\frac{\mu B_1 \hbar}{2} \begin{pmatrix} 0 & e^{-i\omega t} \\ e^{i\omega t} & 0 \end{pmatrix} \tag{286}$$

so we see that

$$\frac{H_1^{11}}{i\hbar} = 0 \tag{287}$$

$$\frac{H_1^{12}}{i\hbar} = \frac{i\mu B_1}{2} e^{-i\omega t} \tag{288}$$

$$\frac{H_1^{21}}{i\hbar} = \frac{i\mu B_1}{2} e^{i\omega t} \tag{289}$$

$$\frac{H_1^{22}}{i\hbar} = 0 \tag{290}$$

Now, our two coupled first-order ODEs are given by

$$\dot{c}_1 = e^{i\omega_{11}t}\langle 1|\frac{H_1^{11}}{i\hbar}|1\rangle c_1 + e^{i\omega_{12}t}\langle 1|\frac{H_1^{12}}{i\hbar}|2\rangle c_2 \tag{291}$$

$$\dot{c}_2 = e^{i\omega_{21}t}\langle 2|\frac{H_1^{21}}{i\hbar}|1\rangle c_1 + e^{i\omega_{22}t}\langle 2|\frac{H_1^{22}}{i\hbar}|2\rangle c_2 \tag{292}$$

After simplification, we obtain

$$\dot{c}_1 = \frac{i\mu B_1}{2}e^{i(\omega_{12}-\omega)t}c_2 \tag{293}$$

$$\dot{c}_2 = \frac{i\mu B_1}{2}e^{-i(\omega_{12}-\omega)t}c_1 \tag{294}$$

where I've used the fact that $\omega_{21} = -\omega_{12}$. Let $\beta = \mu B_1/2$ and $\delta = \omega_{12} - \omega$. Then let's rewrite these two equations using arrays:

$$\dot{\boldsymbol{c}} = M\boldsymbol{c} \tag{295}$$

where

$$\boldsymbol{c} = \begin{pmatrix} c_1 & c_2 \end{pmatrix} \tag{296}$$

and

$$M := \begin{pmatrix} 0 & i\beta e^{i\delta t} \\ i\beta e^{-i\delta t} & 0 \end{pmatrix} \tag{297}$$

is the mixing matrix. Suppose there exists a matrix Q that diagonalizes M:

$$Q^\dagger M Q = M_* \tag{298}$$

You know that I'm using the subscript star for diagonalized stuff. With this, we can write

$$\dot{\boldsymbol{c}} = M\boldsymbol{c} \tag{299}$$

$$\dot{\boldsymbol{c}} = MQQ^\dagger\boldsymbol{c} \tag{300}$$

$$Q^\dagger\dot{\boldsymbol{c}} = Q^\dagger MQQ^\dagger\boldsymbol{c} \tag{301}$$

Let

$$\boldsymbol{c}_* = Q^\dagger\boldsymbol{c} \tag{302}$$

so

$$\dot{\boldsymbol{c}}_* = M_* \boldsymbol{c}_* \tag{303}$$

Even though this is a matrix equation, since we have diagonalized the mixing matrix, I can integrate it as if it were a usual scalar first-order ODE:

$$\boldsymbol{c}_*(t) = e^{M_* t} \boldsymbol{c}_*(0) \tag{304}$$

Let's switch back to $\boldsymbol{c}(t)$:

$$\boldsymbol{c}(t) = Q\boldsymbol{c}_*(t) = Qe^{M_* t}\boldsymbol{c}_*(0) = Qe^{M_* t}Q^\dagger \boldsymbol{c}(0) \tag{305}$$

Consider the n^{th} component. I'll do it with turtle steps:

$$\begin{aligned} c_n(t) &= \sum_m (Qe^{M_* t}Q^\dagger)_{nm} c_m(0) \\ &= \sum_m \langle n|Qe^{M_* t}Q^\dagger|m\rangle c_m(0) \\ &= \sum_m \langle n|Qe^{M_* t}Q^\dagger|m\rangle \langle m|\alpha_I(0)\rangle \\ &= \sum_m \langle n|Qe^{M_* t}Q^\dagger|m\rangle \langle m|\alpha(0)\rangle \\ &= \langle n|Qe^{M_* t}Q^\dagger|\alpha(0)\rangle \end{aligned} \tag{306}$$

Then,

$$\begin{aligned} |\alpha_I(t)\rangle &= \sum_n c_n(t)|n\rangle \\ &= \sum_n |n\rangle c_n(t) \\ &= \sum_n |n\rangle\langle n|Qe^{M_* t}Q^\dagger|\alpha(0)\rangle \\ &= Qe^{M_* t}Q^\dagger|\alpha(0)\rangle \end{aligned} \tag{307}$$

Finally,

$$\begin{aligned} |\alpha(t)\rangle &= e^{-iH_0 t/\hbar}|\alpha_I(t)\rangle \\ &= e^{-iH_0 t/\hbar}Qe^{M_* t}Q^\dagger|\alpha(0)\rangle \end{aligned} \tag{308}$$

Therefore, we conclude that

$$U(t) = e^{-iH_0 t/\hbar}Qe^{M^* t}Q^\dagger \tag{309}$$

As a humble suggestion, do not expand anything in your calculations if you want to do it this way. Do it by hand. Don't even actually diagonalize the M matrix to obtain the Q matrix. When you have the final result in a simple form as this, then proceed with all the computational power you have. There are some huge simplifications here. For instance, H_0 goes like S_z, so we know how to expand this—we eliminate this exponent easily, so that we have a nice matrix to deal with. Then, we have the Q matrices, which will be again just matrices—no exponentiation, no funny business. And finally, $e^{M_* t}$ is just $\mathrm{diag}(e^{\lambda_1 t}, e^{\lambda_2 t})$, where λ_1 and λ_2 are the eigenvalues of the mixing matrix M.

As you see, this is the ugly way of doing things. That's why Schwinger was a total genius.

I think I just got bored here, now that the rest is just fourth-grade algebra (or whenever you were introduced the matrices). If you've been reading up until the end, well, all these should be sufficient for this week before we all burn out. Go take a walk under the rain. Isn't this city even more beautiful when it rains…

VIII. HW3 SUPPLEMENTS

Below is the link to a Mathematica-based supplementary solutions to the third homework:

https://www.dropbox.com/s/dkzwj52xlencz03/2022_fall_phys412-1_hw3.nb?dl=1

Below is the print version of the same notebook.

2022 Fall Phys 412-1 Quantum Mechanics HW3

Mathematica-based supplementary materials

The sole purpose of these notes is to show the minimum amount of work that can be done using Mathematica to get full credits from assignments. It's not my intention to show new or different ways of solving the given problems compared to the professor's solutions; after all, we are doing the same old physics. But the problem is, sometimes having a great source of power as that of Mathematica can be so tempted that you'd want to do it all on Mathematica. There are *proper* ways to do that, and then there are *other* ways to do that. It's important to see when it'd be a good time for you to switch to Mathematica (or any other language for that matter) during your calculations to employ your full potential. That being said, it's my intention to present the shortest ways to do good stuff on Mathematica.

Problem 1

Out[•]=

$$|\psi(0)\rangle \doteq \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) e^{i\varphi} \end{pmatrix}$$

Out[•]=

$$H = -\frac{\mu\hbar}{2}\boldsymbol{B}\cdot\boldsymbol{\sigma}$$

Out[•]=

$$\boldsymbol{B} = B\begin{pmatrix} \cos(\theta)\cos(\varphi) \\ \cos(\theta)\sin(\varphi) \\ -\sin(\theta) \end{pmatrix} =: B\hat{\boldsymbol{n}}$$

Out[•]=

$$H = -\frac{\mu\hbar B}{2}\hat{\boldsymbol{n}}\cdot\boldsymbol{\sigma} =: -\frac{\mu\hbar B}{2}\sigma_n$$

Part (a)

Out[•]=

$$\frac{\partial H}{\partial t} = 0$$

Out[•]=

$$\begin{aligned}
U(t) &= e^{-iHt/\hbar} \\
&= e^{-i(-\mu\hbar B\sigma_n/2)t/\hbar} \\
&= e^{i\frac{\mu Bt}{2}\sigma_n} \\
&= e^{i\alpha\sigma_n}, \quad \alpha := \frac{\mu Bt}{2} \\
&= \sum_{k\geq 0}\frac{1}{k!}(i\alpha)^k\sigma_n^k \\
&= \sum_{k\geq 0}\frac{1}{(2k)!}(i\alpha)^{2k}\underbrace{\sigma_n^{2k}}_{1} + \sum_{k\geq 0}\frac{1}{(2k+1)!}(i\alpha)^{2k+1}\underbrace{\sigma_n^{2k+1}}_{\sigma_n} \\
&= \cos(\alpha) + i\sin(\alpha)\sigma_n \\
&= \cos\left(\frac{\mu Bt}{2}\right) + i\sin\left(\frac{\mu Bt}{2}\right)\sigma_n
\end{aligned}$$

Out[•]=

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle$$

In[ ]:=
```
psi[0] =
  {
   Cos[theta/2],
   Sin[theta/2] Exp[I phi]
  };
I2 =
  IdentityMatrix[2];
n =
  {
   Cos[theta] Cos[phi],
   Cos[theta] Sin[phi],
   -Sin[theta]
  };
s =
  PauliMatrix /@ Range[3];
U[t_] := (* order matters! not s.n *)
 Cos[mu B t / 2] I2 + I Sin[mu B t / 2] n.s
psi[t_] :=
 U[t].psi[0] //
  FullSimplify
```

In[•]:=
```
Prettify[expr_] :=
 expr /. {
     mu → μ,
     theta → ϴ,
     phi → φ,
     hbar → ħ
    } //
   TraditionalForm //
  Style[
    #,
    {
     FontSize → 24,
     FontFamily → "Palatino"
    }
   ] &
```

In[•]:=
```
Row[{
  tex["\\ket{\\psi(t)} \\repr "],
  psi[t] // MatrixForm // Prettify
 }]
```

Out[•]=

$$|\psi(t)\rangle \doteq \begin{pmatrix} \cos\left(\frac{\theta}{2}\right)\cos\left(\frac{B\mu t}{2}\right) - i\sin\left(\frac{\theta}{2}\right)\sin\left(\frac{B\mu t}{2}\right) \\ e^{i\varphi}\left(\sin\left(\frac{\theta}{2}\right)\cos\left(\frac{B\mu t}{2}\right) + i\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{B\mu t}{2}\right)\right) \end{pmatrix}$$

Part (b)

Out[•]=

$$P(t,\psi') = \left|\langle\psi'|\psi(t)\rangle\right|^2$$

Out[ ]=

$$|\psi_1\rangle \doteq \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) e^{i\varphi} \end{pmatrix}$$

Out[ ]=

$$|\psi_2\rangle \doteq \begin{pmatrix} -\sin\left(\frac{\theta}{2}\right) e^{-i\varphi} \\ \cos\left(\frac{\theta}{2}\right) \end{pmatrix}$$

In[ ]:=

```
P[t_, psi1_] :=
 Conjugate[Conjugate[psi1].psi[t]]
    Conjugate[psi1].psi[t] //
   ComplexExpand //
  FullSimplify
psi1 =
  {
   Cos[theta/2],
   Sin[theta/2] Exp[I phi]
  };
psi2 =
  {
   -Sin[theta/2] Exp[-I phi],
   Cos[theta/2]
  };
```

In[•]:=
```
Table[
  Row[{
    tex["P(t, \\psi_" <> ToString[i] <> ") = "],
    P[t, ToExpression["psi" <> ToString[i]]] // Prettify
   }],
  {i, Range[2]}
 ] //
 TableForm
```

Out[•]//TableForm=

$$P(t,\psi_1) = \cos^2\left(\frac{B\,\mu\,t}{2}\right)$$
$$P(t,\psi_2) = \sin^2\left(\frac{B\,\mu\,t}{2}\right)$$

Part (c)

Out[•]=

$$\langle \boldsymbol{S} \rangle_t := \frac{\hbar}{2} \langle \psi(t) | \boldsymbol{\sigma} | \psi(t) \rangle$$

In[•]:=
```
Sav[t_] :=
 hbar / 2 Conjugate[psi[t]].#.psi[t] & /@ s //
   ComplexExpand //
  FullSimplify
```

In order to see the spin precession, we need the averages at $t = 0$, as well.

In[•]:=
```
Table[
  Grid[{
    {tex["\\braket{S_" <> ToString[{x, y, z}[[i]]] <>
       "}_0 = "], Sav[0][[i]] // Expand // Prettify},
    {tex["\\braket{S_" <> ToString[{x, y, z}[[i]]] <>
       "}_t = "], Sav[t][[i]] // Expand // Prettify}
   },
   Alignment → Left
  ],
  {i, Range[3]}
 ] //
 TableForm
```

Out[•]//TableForm=

$$\langle S_x\rangle_0 = \frac{1}{2}\,\hbar\sin(\theta)\cos(\varphi)$$
$$\langle S_x\rangle_t = \frac{1}{2}\,\hbar\sin(\theta)\cos(\varphi)\cos(B\,\mu\,t) - \frac{1}{2}\,\hbar\sin(\varphi)\sin(B\,\mu\,t)$$
$$\langle S_y\rangle_0 = \frac{1}{2}\,\hbar\sin(\theta)\sin(\varphi)$$
$$\langle S_y\rangle_t = \frac{1}{2}\,\hbar\sin(\theta)\sin(\varphi)\cos(B\,\mu\,t) + \frac{1}{2}\,\hbar\cos(\varphi)\sin(B\,\mu\,t)$$
$$\langle S_z\rangle_0 = \frac{1}{2}\,\hbar\cos(\theta)$$
$$\langle S_z\rangle_t = \frac{1}{2}\,\hbar\cos(\theta)\cos(B\,\mu\,t)$$

Now we can write the following:

Out[•]=

$$\langle S_x\rangle_t = \langle S_x\rangle_0\cos(B\mu t) - \langle S_y\rangle_0\sin(\theta)^{-1}\sin(B\mu t)$$
$$\langle S_y\rangle_t = \langle S_x\rangle_0\sin(\theta)^{-1}\sin(B\mu t) + \langle S_y\rangle_0\cos(B\mu t)$$
$$\langle S_z\rangle_t = \langle S_z\rangle_0\cos(B\mu t)$$

Out[ ]=

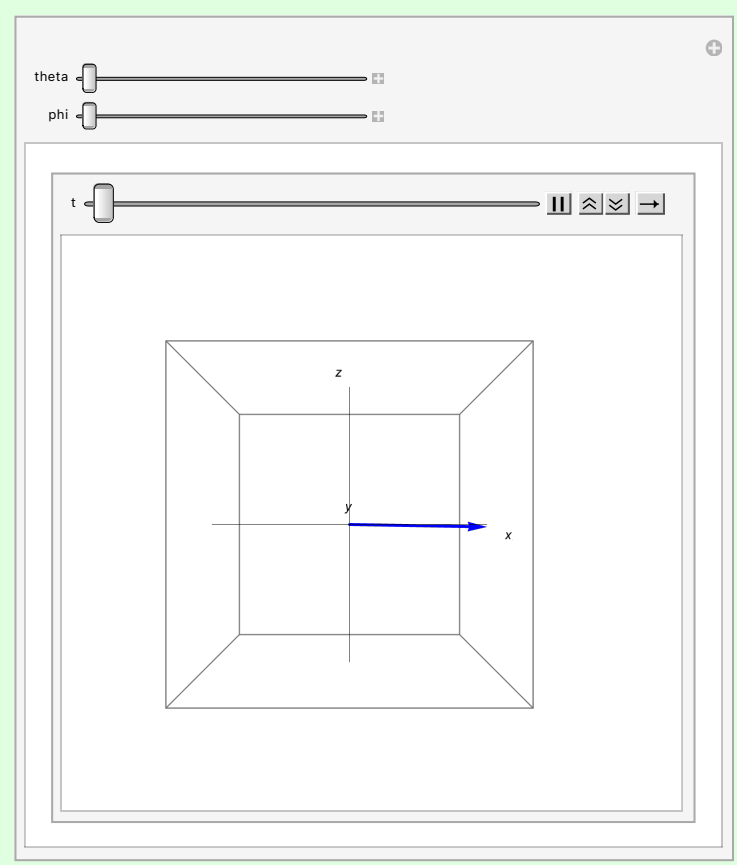
$$\langle \boldsymbol{S} \rangle_t = \begin{pmatrix} \cos(B\mu t) & -\sin(\theta)^{-1}\sin(B\mu t) & 0 \\ \sin(\theta)^{-1}\sin(B\mu t) & \cos(B\mu t) & 0 \\ 0 & 0 & \cos(B\mu t) \end{pmatrix}$$

In[ ]:=

```
R[B_, mu_, t_, theta_] :=
 {
  {Cos[B mu t], -Sin[B mu t] / Sin[theta], 0},
  {Sin[B mu t], Cos[B mu t], 0},
  {0, 0, Cos[B mu t]}
 }
S0[hbar_, theta_, phi_] :=
 Sav[0] //
  Evaluate
St[hbar_, B_, mu_, t_, theta_, phi_] :=
 R[B, mu, t, theta].S0[hbar, theta, phi]
```

In[ ]:=
```
Manipulate[
 Animate[
  ParametricPlot3D[
    {
      {Cos[phi Degree] Cos[theta Degree], Cos[theta Degree]
         Sin[phi Degree], -Sin[theta Degree]} r,
      St[1, 1, 1, t, theta Degree, phi Degree] r
     } // Evaluate,
    {r, 0, 1},
    PlotStyle → {Magenta, Blue},
    AxesOrigin → {0, 0, 0},
    PlotRange → {{-1, 1}, {-1, 1}, {-1, 1}},
    ViewPoint → (*{1.3, -2.4, 2.}*)Front,
    Ticks → None,
    AxesLabel → {x, y, z}
   ] /. Line → Arrow,
  {t, 0, 10 Pi, Pi/50}
 ],
 {theta, 1, 180, 1},
 {phi, 0, 360, 1}
]
```

Out[ ]=



Part (d)

Out[ ]=

$$\dot{\boldsymbol{S}} = \frac{[\boldsymbol{S}, H]}{i\hbar}$$

Out[•]=

$$\begin{aligned}[S_i, H] &= [S_i, -\mu B_j S_j] \\ &= -\mu B_j [S_i S_j] \\ &= -\mu B_j (i\hbar \varepsilon_{ijk} S_k) \\ &= -i\hbar\mu (\boldsymbol{B} \wedge \boldsymbol{S})_i \\ &= i\hbar\mu (\boldsymbol{S} \wedge \boldsymbol{B})_i\end{aligned}$$

Out[•]=

$$\dot{\boldsymbol{S}} = \mu \boldsymbol{S} \wedge \boldsymbol{B}$$

Any equation of this form is called a precession equation. Show that the (time-dependent) averages we've found earlier satisfy this relation:

In[•]:=
```
D[Sav[t], t] == mu Cross[Sav[t], B n] // Simplify
```

Out[•]=
```
True
```

Problem 2

Part (a)

Suppose you have your laboratory oriented in a certain direction and you choose the longitudinal axis of your machine to be the $+z$ axis. You measure the spin of some particle and find the spectrum of the z component of the spin to be $\hbar$, 0, and $-\hbar$. Suppose you change the orientation of your machine and you call $\hat{\boldsymbol{n}}$ the new direction pointing along the longitudinal axis. If you measure the spectrum of the z component of the spin, what do you get? The very same spectrum. The moral of the story is, due to the isotropy of the space, it doesn't matter if you measure S_z or $S_n := \hat{\boldsymbol{n}} \cdot \boldsymbol{S}$. You will get the same spectrum.

Now, write down the secular equation for this operator with said eigenvalues.

Out[•]=

$$(\lambda - \hbar)(\lambda - 0)(\lambda + \hbar) = 0$$

Out[•]=

$$\lambda^3 = \hbar^2 \lambda$$

The Cayley-Hamilton theorem tells us that this is the equation satisfied by the matrix itself:

Out[•]=

$$S_n^3 = \hbar^2 S_n$$

Then we observe that, while S_n^2 is independent of S_n, S_n^3 is given in terms of S_n:

Out[•]=

$$\begin{aligned} S_n &= S_n \\ S_n^2 &= S_n^2 \\ S_n^3 &= \hbar^2 S_n \\ S_n^4 &= \hbar^2 S_n^2 \\ &\cdots \\ S_n^{2k} &= \hbar^{2k-2} S_n^2, \quad k \in \mathbb{Z}^+ \\ S_n^{2k+1} &= \hbar^{2k} S_n \end{aligned}$$

(Try this with spin-3/2 matrices. Again, S_n will have the same eigenvalues as S_z, i.e. $3/2$, $1/2$, $-1/2$, and $-3/2$. You'll see that the second and third powers of S_n are independent of S_n, the fourth power depends on all the three.)

Out[•]=

$$H = -\mu \boldsymbol{B} \cdot \boldsymbol{S}$$

Out[•]=

$$\frac{\partial H}{\partial t} = 0$$

Out[•]=

$$
\begin{aligned}
U(t) &= e^{-iHt/\hbar} \\
&= e^{-i(-\mu \boldsymbol{B}\cdot\boldsymbol{S})t/\hbar} \\
&= e^{i\frac{\mu Bt}{\hbar}S_n}, \quad S_n := \hat{\boldsymbol{n}}\cdot\boldsymbol{S} \\
&= e^{i\alpha S_n}, \quad \alpha := \frac{\mu Bt}{\hbar} \\
&= \sum_{k\geq 0}\frac{1}{k!}(i\alpha)^k S_n^k \\
&= \sum_{k\geq 0}\frac{1}{(2k)!}(i\alpha)^{2k}S_n^{2k} + \sum_{k\geq 0}\frac{1}{(2k+1)!}(i\alpha)^{2k+1}S_n^{2k+1} \\
&= 1 + \sum_{k\geq 1}\frac{1}{(2k)!}(i\alpha)^{2k}\underbrace{S_n^{2k}}_{\hbar^{2k-2}S_n^2} + \sum_{k\geq 0}\frac{1}{(2k+1)!}(i\alpha)^{2k+1}\underbrace{S_n^{2k+1}}_{\hbar^{2k}S_n} \\
&= 1 - [1-\cos(\hbar\alpha)]\frac{S_n^2}{\hbar^2} + i\sin(\alpha\hbar)\frac{S_n}{\hbar} \\
&= 1 - [1-\cos(\mu Bt)]\frac{S_n^2}{\hbar^2} + i\sin(\mu Bt)\frac{S_n}{\hbar}
\end{aligned}
$$

and I don't even really care what the $\hat{\boldsymbol{n}}$ vector is, as long as it's a unit vector.

Part (b)

In[•]:=

```
psi[0] =
  {
   1,
   0,
   0
  };
S =
  {
```

```
    hbar / Sqrt[2] {
      {0, 1, 0},
      {1, 0, 1},
      {0, 1, 0}
     },
    hbar / Sqrt[2] {
      {0, -I, 0},
      {I, 0, -I},
      {0, I, 0}
     },
    hbar DiagonalMatrix[{1, 0, -1}]
   };
n =
   {
    Sin[theta] Cos[phi],
    Sin[theta] Sin[phi],
    Cos[theta]
   };
Sn = (* order matters! not S.n *)
   n.S;
I3 =
   IdentityMatrix[3];
U[t_] :=
 I3 - (1 - Cos[mu B t]) Sn.Sn / hbar ^ 2 + I Sin[mu B t] Sn / hbar
psi[t_] :=
 U[t].psi[0] //
   FullSimplify
```

In[•]:=
```
Prettify[expr_] :=
 expr /. {
     mu → μ,
     theta → Θ,
     phi → φ,
     hbar → ħ
    } //
   TraditionalForm //
  Style[
    #,
    {
     FontSize → 24,
     FontFamily → "Palatino"
    }
   ] &
```

As an aside, confirm that $S_n^3 = \hbar^2 S_n$:

In[•]:=
```
Sn.Sn.Sn == hbar^2 Sn // Simplify
```

Out[•]=
```
True
```

Back to the problem:

Out[•]=

$$P(t,i) = \left|\langle i|\psi(t)\rangle\right|^2, \quad i = 1,2,3$$

In[•]:=
```
CB[i_] :=
 Table[If[i == k, 1, 0], {k, Range[3]}]
P[t_, i_] :=
 Conjugate[Conjugate[CB[i]].psi[t]]
    Conjugate[CB[i]].psi[t] //
   ComplexExpand //
  FullSimplify
```

In[•]:=
```
Table[
  Row[{
    tex["P(t, " <> ToString[i] <> ") = "],
    P[t, i] // Prettify
   }],
  {i, Range[3]}
 ] //
 TableForm
```

Out[•]//TableForm=

$$P(t,1) = \left(\cos^2(\theta)\sin^2\left(\frac{B\mu t}{2}\right) + \cos^2\left(\frac{B\mu t}{2}\right)\right)^2$$

$$P(t,2) = \frac{1}{2}\sin^2(\theta)\sin^2\left(\frac{B\mu t}{2}\right)\left(2\cos(2\theta)\sin^2\left(\frac{B\mu t}{2}\right) + \cos(B\mu t) + 1\right)$$

$$P(t,3) = \sin^4(\theta)\sin^4\left(\frac{B\mu t}{2}\right)$$

Confirm that $\sum_{i=1}^{3} P(t, i) = 1$ (Why?):

In[•]:=
```
Sum[P[t, i], {i, Range[3]}] // Simplify
```

Out[•]=
```
1
```

Part (c)

In[•]:=
```
Sav[t_] :=
 Conjugate[psi[t]].#.psi[t] & /@ S //
   ComplexExpand //
  FullSimplify
```

In order to confirm that we see precession, show that $\langle \boldsymbol{S} \rangle_t$ satisfies the precession equation, $\dot{\boldsymbol{S}} = \mu\, \boldsymbol{S} \wedge \boldsymbol{B}$:

In[•]:=
```
D[Sav[t], t] == mu Cross[Sav[t], B n] // Simplify
```

Out[•]=
```
True
```

Problem 3

Part (a)

Out[•]=
$$\begin{aligned} H(t) &= -\mu \boldsymbol{B}(t) \cdot \boldsymbol{S} \\ &= -\mu B_j(t) S_j \end{aligned}$$

Out[•]=
$$\begin{aligned} [H(t), H(t')] &= [-\mu B_i(t) S_i, -\mu B_j(t') S_j] \\ &= \mu^2 B_i(t) B_j(t') [S_i, S_j] \\ &= \mu^2 B_i(t) B_j(t') (i\hbar \varepsilon_{ijk} S_k) \\ &= i\hbar \mu^2 \boldsymbol{B}(t) \wedge \boldsymbol{B}(t') \cdot \boldsymbol{S} \end{aligned}$$

t is like an index here that destroys the symmetry of the otherwise symmetric tensor $B_i\, B_j$, so the commutator is not zero. This is the most general case for a time-dependent Hamiltonian. There are a couple of ways to solve this problem. Let's do one of them.

Method I

Let's start with the Baker-Campbell-Hausdorff formula. Actually, there is a family of Baker-Campbell-Hausdorff formulae, and here we'll do only one member.

Claim.

Out[•]=

$$e^{i\lambda G} A e^{-i\lambda G} = A + i\lambda [G, A] + \frac{(i\lambda)^2}{2!}[G, [G, A]] + \frac{(i\lambda)^3}{3!}[G, [G, [$$

Proof.

Out[•]=

$$\begin{aligned} f(\lambda) &= e^{i\lambda G} A e^{-i\lambda G} \\ f(0) &= A \\ f'(\lambda) &= iGe^{i\lambda G} A e^{-i\lambda G} + e^{i\lambda G} A e^{-i\lambda G}(-iG) \\ f'(0) &= iGA - iAG = i[G, A] \\ f''(\lambda) &= iGiGe^{i\lambda G} A e^{-i\lambda G} + iGe^{i\lambda G} A e^{-i\lambda G}(-iG) + iGe^{i\lambda G} A e^{-i} \\ f''(0) &= i^2 GGA - i^2 GAG - i^2 GAG + i^2 AGG = i^2 [G, [G, A]] \end{aligned}$$

Continue in this fashion:

Out[•]=

$$f^{(k)}(0) = i^k [\underbrace{G, \ldots [G}_{k \text{ times}}, A]]$$

Taylor expansion of f:

Out[•]=

$$f(\lambda) = \sum_{k \geq 1} \frac{(i\lambda)^k}{k!} [\underbrace{G, \ldots [G}_{k \text{ times}}, A]]$$

qed.

Now consider $G = S_3$ and $A = S_1$, where S_1 and S_3 are the x and z components of the spin operator---but mind that I'm not telling you what the spin value is. You'll see that it doesn't matter.

Out[•]=

$$e^{i\lambda S_3} S_1 e^{-i\lambda S_1} = S_1 + i\lambda [S_3, S_1] + \frac{(i\lambda)^2}{2!}[S_3, [S_3, S_1]] + \cdots$$

Out[•]=

$$\begin{aligned}
[S_3, S_1] &= i\hbar S_2 \\
[S_3, [S_3, S_1]] &= -(i\hbar)^2 S_1 = \hbar^2 S_1 \\
[S_3, [S_3, [S_3, S_1]]] &= i\hbar^3 S_2 \\
[S_3, [S_3, [S_3, [S_3, S_1]]]] &= \hbar^4 S_1 \\
&\cdots \\
[\underbrace{S_3, \ldots [S_3}_{2k \text{ times}, k\in\mathbb{Z}^+}, S_1]] &= \hbar^{2k} S_1 \\
[\underbrace{S_3, \ldots [S_3}_{2k+1 \text{ times}, k\in\mathbb{N}}, S_1]] &= i\hbar^{2k+1} S_1
\end{aligned}$$

Out[•]=

$$\begin{aligned}
e^{i\lambda S_3} S_1 e^{-i\lambda S_1} &= S_1 + \sum_{k\geq 1} \frac{(i\lambda)^k}{k!} [\underbrace{S_3, \ldots [S_3}_{k}, S_1]] \\
&= S_1 + \sum_{k\geq 1} \frac{(i\lambda)^{2k}}{(2k)!} \hbar^{2k} S_1 + \sum_{k\geq 0} \frac{(i\lambda)^{2k+1}}{(2k+1)!} i\hbar^{2k+1} S_2 \\
&= S_1 \cos(\lambda\hbar) - S_2 \sin(\lambda\hbar)
\end{aligned}$$

Let $\lambda = -\omega\, t/\hbar$:

Out[•]=

$$e^{-i\omega t S_3/\hbar} S_1 e^{i\omega t S_3/\hbar} = S_1 \cos(\omega t) + S_2 \sin(\omega t)$$

We are given the Hamiltonian

Out[•]=

$$H(t) = -\mu B_0 S_3 - \mu B_1 [S_1 \cos(\omega t) + S_2 \sin(\omega t)]$$

Let $Q := e^{i\,\omega\, t\, S_3/\hbar}$ and let's play around:

Out[•]=

$$\begin{aligned} H(t) &= -\mu B_0 S_3 - \mu B_1 [S_1 \cos(\omega t) + S_2 \sin(\omega t)] \\ &= -\mu B_0 S_3 - \mu B_1 Q^\dagger S_1 Q \\ &= -\mu B_0 S_3 Q^\dagger Q - \mu B_1 Q^\dagger S_1 Q \\ &= -\mu B_0 Q^\dagger S_3 Q - \mu B_1 Q^\dagger S_1 Q \\ &= Q^\dagger (-\mu B_0 S_3 - \mu B_1 S_1) Q \\ &= Q^\dagger H(0) Q \end{aligned}$$

Claim. For any matrix M and any invertible matrix Q, we have

Out[•]=

$$f(Q^{-1} M Q) = Q^{-1} f(M) Q$$

Proof. "Whenever in doubt, expand in a power series." -E. Fermi

Out[•]=

$$\begin{aligned} f(Q^{-1} M Q) &= \sum_{j \geq 0} c_j (Q^{-1} M Q)^j \\ &= \sum_{j \geq 0} c_j \underbrace{(Q^{-1} M Q) \ldots (Q^{-1} M Q)}_{j \text{ times}} \\ &= \sum_{j \geq 0} c_j Q^{-1} M^j Q \\ &= Q^{-1} \left[\sum_{j \geq 0} c_j M^j \right] Q \\ &= Q^{-1} f(M) Q \end{aligned}$$

qed.

Our Q operator above is unitary ($Q^\dagger = Q^{-1}$), so we can somehow make use of this identity.

Out[ ]=

$$H(t) = Q^\dagger(t) H(0) Q(t)$$
$$e^{-iH(T)\delta/\hbar} = Q^\dagger(T) e^{-iH(0)\delta/\hbar} Q(T)$$

where I've chosen the f function to be $f(x) = e^{-i\,x\,\delta/\hbar}$. T and δ have no special meaning right now, but there will be soon.

Now the time-evolution operator. Suppose we look at a time interval from t_i to t_f and we slice it into N intervals of interval size δ in such a way that when we take the limit as $N \to \infty$ and $\delta \to 0$, the product $N\,\delta$ will give the total time elapsed, i.e. $t_f - t_i$.

Out[ ]=

$$U(t_f, t_i) = \lim_{N\to\infty, \delta\to 0} U(t_i + N\delta, t_i + (N-1)\delta) \cdots U(t_i + \delta, t_i)$$

Note that only for infinitesimal time separation, we have $U(T + \delta,\ T) = e^{-i\,H(T)\,\delta/\hbar}$:

Out[ ]=

$$\begin{aligned} U(t_f, t_i) &= \lim_{N\to\infty, \delta\to 0} e^{-iH(t_i+(N-1)\delta)\delta/\hbar} \cdots e^{-iH(t_i)\delta/\hbar} \\ &= \lim_{N\to\infty, \delta\to 0} Q^\dagger(t_i + (N-1)\delta) e^{-iH(0)\delta/\hbar} Q(t_i + (N-1 \\ &\qquad Q^\dagger(t_i + \delta) e^{-iH(0)\delta/\hbar} Q(t_i + \delta) Q^\dagger(t_i) e^{-iH(0} \end{aligned}$$

Notice that $Q(t)\ Q^\dagger(t') = Q(t - t')$:

Out[ ]=

$$\begin{aligned} U(t_f, t_i) &= \lim_{N\to\infty, \delta\to 0} Q^\dagger(t_i + (N-1)\delta) \left[e^{-iH(0)\delta/\hbar} Q(\delta)\right]^{N-1} e^{-} \\ &= \lim_{N\to\infty, \delta\to 0} Q^\dagger(t_i + (N-1)\delta) \left[e^{-iH(0)\delta/\hbar} e^{i\omega\delta S_3/\hbar}\right]^{N-1} \end{aligned}$$

Note that $e^{\delta\,A}\ e^{\delta\,B} = e^{\delta(A+B)} + O(\delta^2)$:

Out[•]=

$$\begin{aligned} U(t_f, t_i) &= \lim_{N\to\infty,\delta\to 0} Q^\dagger(t_i + (N-1)\delta) e^{i(N-1)\delta(\omega S_3 - H(0))} e^{-iH(0} \\ &= \lim_{N\to\infty,\delta\to 0} Q^\dagger(t_i + N\delta - \delta) e^{i(N\delta\omega S_3 - \delta\omega S_3 - N\delta H(0))/\hbar} Q \\ &= Q^\dagger(t_f) e^{i(\omega S_3 - H(0))(t_f - t_i)/\hbar} Q(t_i) \end{aligned}$$

or

Out[•]=

$$U(t_f, t_i) = e^{-i\omega t_f S_3/\hbar} e^{i[(\omega+\mu B_0)S_3 + \mu B_1 S_1](t_f - t_i)/\hbar} e^{i\omega t_i S_3/\hbar}$$

or if we set $t_i = 0$ and $t_f = t$, we get

Out[•]=

$$U(t) = e^{-i\omega t S_3/\hbar} e^{i[(\omega+\mu B_0)S_3 + \mu B_1 S_1]t/\hbar}$$

Now we need to simplify this; otherwise, we can barely make any progress. Consider the first factor:

Out[•]=

$$-\frac{\omega t}{\hbar} S_3 = \boldsymbol{p} \cdot \boldsymbol{S}$$

Out[•]=

$$\boldsymbol{p} = \begin{pmatrix} 0 \\ 0 \\ -\frac{\omega t}{\hbar} \end{pmatrix}$$

Out[•]=

$$\alpha = -\frac{\omega t}{\hbar}$$

Out[•]=

$$\hat{\boldsymbol{p}} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Consider the second factor:

Out[•]=

$$[\mu B_1 S_1 + (\omega + \mu B_0) S_3] \frac{t}{\hbar} = \boldsymbol{q} \cdot \boldsymbol{S}$$

Out[•]=

$$\boldsymbol{q} = \begin{pmatrix} \frac{\mu B_1 t}{\hbar} \\ 0 \\ \frac{(\omega+\mu B_0) t}{\hbar} \end{pmatrix}$$

Out[•]=

$$\beta = \sqrt{(\mu B_1)^2 + (\omega + \mu B_0)^2} \frac{t}{\hbar}$$

Out[•]=

$$\hat{\boldsymbol{q}} = \begin{pmatrix} \frac{\mu B_1}{\sqrt{(\mu B_1)^2+(\omega+\mu B_0)^2}} \\ 0 \\ \frac{\omega+\mu B_0}{\sqrt{(\mu B_1)^2+(\omega+\mu B_0)^2}} \end{pmatrix}$$

Then we can write

Out[•]=

$$U(t) = e^{i\alpha(t) S_p} e^{i\beta(t) S_q}$$
$$= \left\{ 1 - [1 - \cos(\alpha(t))] \frac{S_p^2}{\hbar^2} + i \sin(\alpha(t)) \frac{S_p}{\hbar} \right\} \left\{ 1 - [1 - \text{co} \right.$$

Don't bother expanding this expression.

Part (b)

Out[•]=

$$P(t, i) = |\langle i | \psi(t) \rangle|^2$$

In[•]:=

```
alpha[t_] :=
 -omega t / hbar
beta[t_] :=
 Sqrt[(mu B1) ^ 2 + (omega + mu B0) ^ 2] t / hbar
```

```
p =
  {
   0,
   0,
   1
  };
q =
  {
   mu B1 / Sqrt[(mu B1) ^ 2 + (omega + mu B0) ^ 2],
   0,
   (omega + mu B0) / Sqrt[(mu B1) ^ 2 + (omega + mu B0) ^ 2]
  };
S = {
   hbar / Sqrt[2] {
     {0, 1, 0},
     {1, 0, 1},
     {0, 1, 0}
    },
   hbar / Sqrt[2] {
     {0, -I, 0},
     {I, 0, -I},
     {0, I, 0}
    },
   hbar DiagonalMatrix[{1, 0, -1}]
  };
Sn[n_] :=
 n.S
I3 =
  IdentityMatrix[3];
U[t_] :=
 (I3 - (1 - Cos[alpha[t]]) Sn[p].Sn[p] / hbar ^ 2 +
```

```
    I Sin[alpha[t]] Sn[p] / hbar) .
   (I3 - (1 - Cos[beta[t]]) Sn[q].Sn[q] / hbar ^ 2 +
    I Sin[beta[t]] Sn[q] / hbar)
CB[i_] :=
 Table[If[i == k, 1, 0], {k, Range[3]}]
psi[0] =
  CB[1];
psi[t_] :=
 U[t].psi[0]
$Assumptions =
  {
   B0 > 0,
   B1 > 0,
   mu > 0,
   omega > 0,
   t > 0
  };
P[t_, i_] :=
 Conjugate[Conjugate[CB[i]].psi[t]]
    Conjugate[CB[i]].psi[t] //
   ComplexExpand //
  FullSimplify
```

In[ ]:=
```
Prettify[expr_] :=
 expr /. {
     B0 → Subscript[B, 0],
     B1 → Subscript[B, 1],
     hbar → ℏ,
     omega → ω,
     mu → μ
    } //
   TraditionalForm //
  Style[
    #,
    {
     FontSize → 24,
     FontFamily → "Palatino"
    }
   ] &
```

In[•]:=
```
Table[
  Row[{
    tex["P(t," <> ToString[i] <> ") = "],
    P[t, i] // Prettify
   }],
  {i, Range[3]}
 ] // TableForm
```

Out[•]//TableForm=

$$P(t,1) = \frac{\left(\left(2\,B_0^2+B_1^2\right)\mu^2+4\,B_0\,\mu\,\omega+B_1^2\,\mu^2\,\cos\left(\frac{t\,\sqrt{B_1^2\,\mu^2+(B_0\,\mu+\omega)^2}}{\hbar}\right)+2\,\omega^2\right)^2}{4\left(B_1^2\,\mu^2+(B_0\,\mu+\omega)^2\right)^2}$$

$$P(t,2) = \frac{B_1^2\,\mu^2\,\sin^2\left(\frac{t\,\sqrt{B_1^2\,\mu^2+(B_0\,\mu+\omega)^2}}{2\,\hbar}\right)\left(\left(2\,B_0^2+B_1^2\right)\mu^2+4\,B_0\,\mu\,\omega+B_1^2\,\mu^2\,\cos\left(\frac{t\,\sqrt{B_1^2\,\mu^2+(B_0\,\mu+\omega}\right.\right.}{\left(B_1^2\,\mu^2+(B_0\,\mu+\omega)^2\right)^2}$$

$$P(t,3) = \frac{B_1^4\,\mu^4\,\sin^4\left(\frac{t\,\sqrt{B_1^2\,\mu^2+(B_0\,\mu+\omega)^2}}{2\,\hbar}\right)}{\left(B_1^2\,\mu^2+(B_0\,\mu+\omega)^2\right)^2}$$

Confirm that $\sum_{i=1}^{3} P(t,\,i) = 1$ (Why?):

In[•]:=
```
Sum[P[t, i], {i, Range[3]}] // Simplify
```

Out[•]=
```
1
```

IX. DISCUSSION 7 (NOV 10)

I think today we mostly focused on spin couplings, particularly the coupling of two spin-1/2 particles (as if in a bound state). My intention was to emphasize the usage and/or importance of Kronecker product. Let's see if I can prepare today's notes without getting bored, lol.

Suppose I take two electron and by some magical means, I let them form a bound state. (For the sake of simplicity, suppose the temperature is so low that the rotational modes are not activated yet, i.e. ignore the orbital angular momentum.) Then, I can consider a vector operator of the form

$$\boldsymbol{S} = \boldsymbol{S}_1 + \boldsymbol{S}_2 \tag{310}$$

where $\boldsymbol{S}_i$ is the i^{th} electron spin and $\boldsymbol{S}$ is the total spin. Now, we know that electrons are spin-1/2 particles. This means, the z component of their spin can be $\pm 1/2$ (I'm working with natural units). (This z axis in fact is determined by the longitudinal axis of my experimental setup. It's just easy to call it the z axis and embrace the consequences; otherwise, it has no particular implications whatsoever.) We usually call these states spin up or down, right? Now, suppose my electrons are both in the up state, so that I can describe this bound state with the ket $|\uparrow\uparrow\rangle$.

Let's have a deeper look at this state ket. It is actually a short-hand notation for the combined state $|\frac{1}{2}\frac{1}{2}\rangle_1 \otimes |\frac{1}{2}\frac{1}{2}\rangle_2$, where I'm using the notation $|s_i m_i\rangle_i$ for the i^{th} electron. Since for any given problem the electron's spin is a constant number, we usually ignore the first quantum number in the kets that go into this product state and even usually write $|\uparrow\rangle_1 \otimes |\uparrow\rangle_2$. There is a really special meaning attached to this Kronecker product here.

We call this product state the *decoupled* state just because we treat the two electrons as if they were free—I mean, they live in their respective kets, that's all. There are three other decoupled states, i.e. $|\uparrow\rangle_1 \otimes |\downarrow\rangle_2$, $|\downarrow\rangle_1 \otimes |\uparrow\rangle_2$, and $|\downarrow\rangle_1 \otimes |\downarrow\rangle_2$. These kets span the $(2+2)\otimes(2+2)$ space in a *partial* way, meaning not every state that lives in the $4 \otimes 4$ space can be decomposed into the product of these $2d$ kets, which has already been discussed earlier today in the lecture. Even though these decoupled kets form a $4d$ space, we need decoupled operators acting on them. Let me explain what I mean.

There are some spin addition rules—which I'm aware we haven't covered in the class yet, and I'm not going to go into details, no worries—and one of them says that if you add two spin vectors, the z components just add up, i.e. we can safely focus on the z component of the total spin given by (310):

$$S_z = S_{1z} + S_{2z} \tag{311}$$

This notation has always bugged me, and it should bug you, as well. On the left, we have an observable that is supposed to act on the *coupled* state (which I'll define shortly) and on the right, we have two *decoupled* observables, meaning the S_{1z} term doesn't do anything on the second electron and the S_{2z} term doesn't do anything on the first one. Let's make it more concrete. What does this equation tell you—or, how would you evaluate it:

$$(S_{1z} + S_{2z})|\uparrow\uparrow\rangle = (S_{1z} + S_{2z})|\uparrow\rangle_1 \otimes |\uparrow\rangle_2 = ? \tag{312}$$

The operators S_{1z} and S_{2z} has their separate rooms in this $4 \otimes 4$ apartment, each having dimensionality 2. And we have a product state that is being acted on...

After staring at this expression long enough, we see that the proper way of expressing it should involve a Kronecker product:

$$(S_{1z} \otimes \mathbb{1} + \mathbb{1} \otimes S_{2z})|\uparrow\rangle_1 \otimes |\uparrow\rangle_2 = ? \tag{313}$$

Now this is just beautiful. The reason is, now we see who talks to whom. In the first term in the parentheses, S_{1z} will act on $|\uparrow\rangle_1$ and the identity will hit on $|\uparrow\rangle_2$ and in the second term, the identity will act on $|\uparrow\rangle_1$ and S_{2z} will go crazy on $|\uparrow\rangle_2$. Recalling the eigenvalue equation $S_z|sm\rangle = m|sm\rangle$ from your earlier exposure to quantum mechanics, we see that

$$\begin{aligned}
(S_{1z} \otimes \mathbb{1} + \mathbb{1} \otimes S_{2z})|\uparrow\rangle_1 \otimes |\uparrow\rangle_2 &= (S_{1z}|\uparrow\rangle_1) \otimes (\mathbb{1}|\uparrow\rangle_2) + (\mathbb{1}|\uparrow\rangle_1) \otimes (S_{2z}|\uparrow\rangle_2) \\
&= (\frac{1}{2}|\uparrow\rangle_1) \otimes (1|\uparrow\rangle_2) + (1|\uparrow\rangle_1) \otimes (\frac{1}{2}|\uparrow\rangle_2) \\
&= \frac{1}{2}|\uparrow\rangle_1 \otimes |\uparrow\rangle_2 + \frac{1}{2}|\uparrow\rangle_1 \otimes |\uparrow\rangle_2 \\
&= |\uparrow\rangle_1 \otimes |\uparrow\rangle_2 \\
&= |\uparrow\uparrow\rangle
\end{aligned} \tag{314}$$

indicating that this state is in fact an eigenstate of the S_z operator with eigenvalue 1, as expected (Why? Because the m_i values ($i = 1, 2$) add up to 1).

Now let's do something else and I'll get back to this point shortly. As I mentioned, there is this *decoupled basis*, given in terms of the product kets of the two electrons, namely

$$\text{decoupled basis} = \{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\} \tag{315}$$

and there is this other basis, called the *coupled basis*. Let's construct this. But before that, we need to discuss something important.

The question is, if you manage to create a bound state of two electrons, and if you try to measure the total spin (not just the z component), what'd you get?

In one extreme case, the individual spin vectors may be parallel, so the spins add up, i.e. s can be equal to $s_1 + s_2 = \frac{1}{2} + \frac{1}{2} = 1$. In the other extreme, they may be antiparallel, totally cancelling each other, i.e. $s = |s_1 - s_2| = |\frac{1}{2} - \frac{1}{2}| = 0$. Since the spin is quantized in integer steps, we see that these two are the only possibilities. We write this *composition* in a group-theoretical language as

$$\frac{1}{2} \otimes \frac{1}{2} = 0 \oplus 1 \tag{316}$$

But we are not doing group theory here; this is just for notational purposes. In words, it just says *when you add two spin-1/2 particles together, the combined state can have a total spin of 0 or 1*. That's it. (For my particle-theorist colleagues, note that some textbooks like Srednicki's use a very slightly different version of these composition identities. He like to express the composition not as $s_1 \otimes s_2$ but rather as $(2s_1 + 1) \otimes (2s_2 + 1)$, so Srednicki would write this as $2 \otimes 2 = 1 \oplus 3$—I don't know if you'd find this useful, but it's important to be aware of this.) Now, there is an important point here. The dimensionality of this composition is 4 (up and down for two particles) but we see only a total spin of 0 and 1. There must be some multiplicities attached to these total spins. Again, without doing any group theory, we can directly resolve these multiplicities (They are like the degeneracies of the Dirac Hamiltonian, which we resolved using the trace technology). Since for any spin value s, the m values go from $-s$ to $+s$ in integer steps, we see that there is only one state with $s = 0$, which is $|00\rangle$, and there are three states with $s = 1$, i.e. $|11\rangle$, $|10\rangle$, $|1-1\rangle$. Now I'm using the notation $|sm\rangle$, where s is the total spin and m is its z component. These four states form the *coupled basis*:

$$\text{coupled basis} = \{|11\rangle, |10\rangle, |1-1\rangle, |00\rangle\} \tag{317}$$

The spin-1 states constitute a *triplet* and the spin-0 state is just a *singlet*. This is just semantics. What's more crucial is, the triplet states are symmetric and the singlet state is antisymmetric... with respect to what? With respect to the *exchange* operator, i.e. if you flip your electrons, your state will get a minus sign. We'll prove this shortly.

The next question is, can we make a connection between these two bases? Yes. How? Let's start with an observation. If both of my electrons are in the up state, then my $m = m_1 + m_2$ value will be one, and note that there is no other way to achieve this. Therefore, the coupled state $|11\rangle$ can be given only by the decoupled state $|\uparrow\uparrow\rangle$:

$$|11\rangle = |\uparrow\uparrow\rangle \tag{318}$$

(I'm not sure if you are still reading or have stopped giving an f long ago, but note that the coupled-basis state are different from what we have seen in the lecture earlier today. Just as a clarification. Let's continue.) I want to generate the $|10\rangle$ using this relation and only this relation. I like this exercise because it makes fine illustration of the Kronecker product.

Long ago, I've mentioned some ladder operators in these discussion sessions. Let's briefly remember them. Suppose I give you this operator:

$$S_+ \doteq \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{319}$$

What does this guy do for a living? Well, you can act it on a known state and see what you get. Noting that I've represented this operator in the basis spanned by the eigenkets of the operator S_3, I can write

$$|\uparrow\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\downarrow\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{320}$$

Then, I see that

$$S_+|\uparrow\rangle \doteq \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0 \tag{321}$$

Similarly,

$$S_+|\downarrow\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |\uparrow\rangle \tag{322}$$

So it increases the m value by 1. There is this cousin of this operator, given by its hermitian conjugate, $S_- = S_+^\dagger$:

$$S_- \doteq \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{323}$$

You can show that

$$S_- |\uparrow\rangle = |\downarrow\rangle \tag{324}$$

and

$$S_- |\downarrow\rangle = 0 \tag{325}$$

These are the ladder, or raising and lowering, operators. Notice that, in the very same basis, I have

$$S_x = \frac{S_+ + S_-}{2} \tag{326}$$

and

$$S_y = \frac{S_+ - S_-}{2i} \tag{327}$$

We see that for the case of spin $1/2$, these ladder operators are *normalized* in a way, meaning they just flip the operators without bringing in any numerical factor in front of the state. In the general case, you may want to remember the following:

$$S_\pm |sm\rangle = \sqrt{s(s+1) - m(m \pm 1)}|sm \pm 1\rangle \tag{328}$$

(Note that I'm not saying whether this is a coupled state of a certain number of particles or just the state of a single particle—it doesn't matter!) Let's go back and derive $|10\rangle$ now.

We have

$$|11\rangle = |\uparrow\uparrow\rangle \tag{329}$$

From the spin addition, we have

$$S_\pm = S_{1\pm} + S_{2\pm} \tag{330}$$

But again, this is a subtle notation! We should write this as

$$S_\pm = S_{1\pm} \otimes \mathbb{1} + \mathbb{1} \otimes S_{2\pm} \tag{331}$$

Now act this on $|11\rangle$. Notice that $S_\pm$ likes to talk to the coupled state but the right-hand side will talk to the decoupled state:

$$\begin{aligned} S_-|11\rangle &= (S_{1-} \otimes \mathbb{1} + \mathbb{1} \otimes S_{2-})|\uparrow\uparrow\rangle \\ &= (S_{1-} \otimes \mathbb{1} + \mathbb{1} \otimes S_{2-})|\uparrow\rangle_1 \otimes |\uparrow\rangle_2 \end{aligned} \tag{332}$$

On the left, we simply have $\sqrt{1(1+1)-1(1-1)}|10\rangle = \sqrt{2}|10\rangle$ and on the right,

$$\begin{aligned} \sqrt{2}|10\rangle &= (S_{1-}|\uparrow\rangle_1) \otimes (\mathbb{1}|\uparrow\rangle_2) + (\mathbb{1}|\uparrow\rangle_1) \otimes (S_{2-}|\uparrow\rangle_2) \\ &= (|\downarrow\rangle_1) \otimes (|\uparrow\rangle_2) + (|\uparrow\rangle_1) \otimes (|\downarrow\rangle_2) \\ &= |\downarrow\rangle_1 \otimes |\uparrow\rangle_2 + |\uparrow\rangle_1 \otimes |\downarrow\rangle_2 \\ &= |\downarrow\uparrow\rangle + |\uparrow\downarrow\rangle \end{aligned} \tag{333}$$

so

$$|10\rangle = \frac{|\downarrow\uparrow\rangle + |\uparrow\downarrow\rangle}{\sqrt{2}} \tag{334}$$

For the coupled state $|1-1\rangle$, you don't need to perform any muscle work; just consider the fact that the total m value is -1 and there is only one way to achieve this: put both electrons in the down state:

$$|1-1\rangle = |\downarrow\downarrow\rangle \tag{335}$$

Notice in these three coupled states that if you flip the arrows, you'll get the same state. That's why they are called the symmetric combinations.

Now we need to take care of the singlet. Note the following. If you have a Hilbert space spanned by $\{|n\rangle\}$—it can be infinite-dimensional or just $2d$, doesn't matter—then if your basis is orthonormal and complete, you can write

$$\langle n|n'\rangle = \delta_{nn'} \tag{336}$$

We can generalize to Hilbert space spanned by kets that carry multiple quantum numbers—after all, they should be complete and orthonormal to form a proper basis (the language

here is a bit vague, I must admit; a mathematician would certainly do a better job here). I mean, we have the following:

$$\langle nm\ell \ldots | n'm'\ell' \ldots \rangle = \delta_{nn'}\delta_{mm'}\delta_{\ell\ell'} \ldots \tag{337}$$

Noting that both the coupled and decoupled basis kets obey this construction, we can directly read off $|00\rangle$ from $|10\rangle$—it must be the orthogonal combination:

$$|00\rangle = \frac{|\downarrow\uparrow\rangle - |\uparrow\downarrow\rangle}{\sqrt{2}} \tag{338}$$

up to an overall phase, which is usually taken to be $e^{i\pi}$. Notice that this singlet is odd under the exchange of the two electrons.

I'm getting bored here so let's finish it up by doing two exercises. Maybe three, I don't know. We'll see. I don't really plan out these notes, you know.

Let's act the $\boldsymbol{S}^2$ operator on the coupled state $|11\rangle$. This is just an exercise, you know, nothing profound will emerge here.

$$\boldsymbol{S}^2|11\rangle = (\boldsymbol{S}_1 + \boldsymbol{S}_2)^2|\uparrow\uparrow\rangle \tag{339}$$

The left-hand side is easy. There is only one operator and it really loves to talk to this coupled state. Just to remind you (to refresh your memories of modern physics courses at worst), we have $\boldsymbol{S}^2|sm\rangle = s(s+1)|sm\rangle$—again, it doesn't matter if the $|sm\rangle$ ket here denotes a coupled state or a single-particle state. Now, the left-hand side then just gives $2|11\rangle$. Let's work out the right-hand side. We have

$$(\boldsymbol{S}_1 + \boldsymbol{S}_2)^2 = \boldsymbol{S}_1^2 + \boldsymbol{S}_2^2 + 2\boldsymbol{S}_1 \cdot \boldsymbol{S}_2 \tag{340}$$

This notation is highly annoying in terms of who is acting on whom—so we need to *Kroneckerize* this. The dot product is especially disturbing. Let's first expand it:

$$\begin{aligned}(\boldsymbol{S}_1 + \boldsymbol{S}_2)^2 &= \boldsymbol{S}_1^2 + \boldsymbol{S}_2^2 + 2\boldsymbol{S}_1 \cdot \boldsymbol{S}_2 \\ &= \boldsymbol{S}_1^2 + \boldsymbol{S}_2^2 + 2(S_{1x}S_{2x} + S_{1y}S_{2y} + S_{1z}S_{2z})\end{aligned} \tag{341}$$

Eliminate S_{ix} and S_{iy} using the ladder operators:

$$(\boldsymbol{S}_1 + \boldsymbol{S}_2)^2 = \boldsymbol{S}_1^2 + \boldsymbol{S}_2^2 + S_{1z}S_{2z} + \frac{1}{2}(S_{1+}S_{2-} + S_{1-}S_{2+}) \tag{342}$$

Okay, still disturbing. Let's work on the proper notation now. We know that $\boldsymbol{S}_1^2$ will act on the first electron's ket and won't do anything on the second electron's state. $\boldsymbol{S}_2^2$ will do it the other way around. Now we have these products, $S_{1z}S_{2z}$, $S_{1+}S_{2-}$, and $S_{1-}S_{2+}$. Noting that the first factor will always act on the first electron here, we should write these as

$$S_{1z}S_{2z} = (S_{1z} \otimes \mathbb{1})(\mathbb{1} \otimes S_{2z}) = S_{1z} \otimes S_{2z} \tag{343}$$

and similarly for the other two. Thus, we have

$$(\boldsymbol{S}_1 + \boldsymbol{S}_2)^2 = \boldsymbol{S}_1^2 \otimes \mathbb{1} + \mathbb{1} \otimes \boldsymbol{S}_2^2 + 2S_{1z} \otimes S_{2z} + \frac{1}{2}(S_{1+} \otimes S_{2-} + S_{1-} \otimes S_{2+}) \tag{344}$$

Everything is crystal clear here:

$$\begin{aligned}
(\boldsymbol{S}_1 + \boldsymbol{S}_2)^2|\uparrow\uparrow\rangle &= \left[\boldsymbol{S}_1^2 \otimes \mathbb{1} + \mathbb{1} \otimes \boldsymbol{S}_2^2 + 2S_{1z} \otimes S_{2z} + \frac{1}{2}(S_{1+} \otimes S_{2-} + S_{1-} \otimes S_{2+})\right] |\uparrow\uparrow\rangle \\
&= \left\{\begin{array}{l}
(\boldsymbol{S}_1^2|\uparrow\rangle_1) \otimes (\mathbb{1}|\uparrow\rangle_2) \\
+(\mathbb{1}|\uparrow\rangle_1) \otimes (\boldsymbol{S}_2^2|\uparrow\rangle_2) \\
+2(S_{1z}|\uparrow\rangle_1) \otimes (S_{2z}|\uparrow\rangle_2) \\
+\frac{1}{2}(S_{1+}|\uparrow\rangle_1) \otimes (S_{2-}|\uparrow\rangle_2) \\
+\frac{1}{2}(S_{1-}|\uparrow\rangle_1) \otimes (S_{2+}|\uparrow\rangle_2)
\end{array}\right\} \\
&= \frac{3}{4}|\uparrow\uparrow\rangle + \frac{3}{4}|\uparrow\uparrow\rangle + 2 \times \frac{1}{4}|\uparrow\uparrow\rangle + 0 \\
&= 2|\uparrow\uparrow\rangle
\end{aligned} \tag{345}$$

As our second exercise, let's consider a Hamiltonian of the form

$$H = -g\boldsymbol{S}_1 \cdot \boldsymbol{S}_2 \tag{346}$$

which may represent the spin-spin interaction of two electrons in a bound state or of the electron and the proton inside the hydrogen atom. Let's express this Hamiltonian in the basis spanned by decoupled kets, i.e. ups and downs. For concreteness, I mean, if you don't like this much abstractness, let's take $S_i = \frac{\hbar}{2}\boldsymbol{\sigma}$ for $i = 1, 2$. But I'll do this later.

We have

$$\begin{aligned}
\boldsymbol{S}_1 \cdot \boldsymbol{S}_2 &= S_{1x}S_{2x} + S_{1y}S_{2y} + S_{1z}S_{2z} \\
&= \frac{1}{2}S_{1+}S_{2-} + \frac{1}{2}S_{1-}S_{2+} + S_{1z}S_{2z}
\end{aligned} \tag{347}$$

and after *Kroneckerizing*, we get

$$\boldsymbol{S}_1 \cdot \boldsymbol{S}_2 = \frac{1}{2} S_{1+} \otimes S_{2-} + \frac{1}{2} S_{1-} \otimes S_{2+} + S_{1z} \otimes S_{2z} \tag{348}$$

Now, we choose the decoupled kets as our basis:

$$\text{decoupled basis} = \{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\} \tag{349}$$

and we know how to act each operator in (348) on each of these kets:

$$\langle\uparrow\uparrow|\boldsymbol{S}_1 \cdot \boldsymbol{S}_2|\uparrow\uparrow\rangle = \left\{ \begin{array}{l} \frac{1}{2}\langle\uparrow\uparrow|[(S_{1+} \otimes S_{2-})(|\uparrow\rangle_1 \otimes |\uparrow\rangle_2)] \\ +\frac{1}{2}\langle\uparrow\uparrow|[(S_{1+} \otimes S_{2-})(|\uparrow\rangle_1 \otimes |\uparrow\rangle_2)] \\ +\frac{1}{2}\langle\uparrow\uparrow|[(S_{1z} \otimes S_{2z})(|\uparrow\rangle_1 \otimes |\uparrow\rangle_2)] \end{array} \right\} \tag{350}$$

Now act each operator on their respective kets and if you get some flipped states facing each other at the end, e.g. $\langle\uparrow\uparrow | \downarrow\uparrow\rangle$, then realize that it gives you zero and move on. If all the arrows match on each side of the inner product, then you get 1. I really got bored here and I don't have the energy to type all these. I'm not sure if you managed to come to this point, but let me know if you have any questions. If not, have a lovely rest of the week.

---

[1] This is true in the absence of degeneracies. When we have degeneracies, we say that the eigenvectors corresponding to the nondegenerate subset are orthogonal.

[2] Suppose we are working in an infinite-dimensional basis.

[3] To get a copy of Mathematica, go to https://www.it.northwestern.edu/software/mathematica-fac/index.html and click Students Wolfram Activation Key Request Form site under Activation and Download.

[4] But when you perform lots of measurement, your *average* may be some other value, depending on your state.

[5] The extension to the most general case is trivial, i.e. when we have no common indices:

$$\epsilon_{ijk}\epsilon_{rst} = \det \begin{pmatrix} \delta_{ir} & \delta_{jr} & \delta_{kr} \\ \delta_{is} & \delta_{js} & \delta_{ks} \\ \delta_{it} & \delta_{jt} & \delta_{kt} \end{pmatrix} \tag{351}$$

Yes, you will end up more terms than you could imagine, but all of them are deltas. Just expand your terms and start replacing indices. That's what this tells you.

[6] Now it's important to be cautious: these are the components of the spin operator in general, not the basis matrices for the matrix space of the relevant dimensionality.