This packet contains active learning lectures provided by a Stanford faculty members for a variety of advanced courses.

It is meant to give you a sense of the breadth of topics which are being taught in this method as well as the breadth of implementation strategies faculty are employing.

At least 10 courses are being taught in this manner at Stanford, by over 10 different faculty members (I believe something like 15 and counting!).

There was nearly universal adoption by junior faculty, as well as a few key senior faculty. Many senior faculty are now moving to adopt these methods.

All of the faculty volunteered to share materials if desired, so please contact me or them if you are interested!

Lauren Tompkins — laurenat@stanford.edu  
April 25th, 2019
Course: Physics 130 — Introduction to Quantum Mechanics I

Instructor: Patricia Burchat

Target Audience: Sophomores/Juniors

Brief summary of examples:

A ~30 minute “summary activity” on notation and concepts related to linear algebra in quantum mechanics; a lecture with exercises on the postulates of quantum mechanics. These were in ~week 4 of this first quarter in our quantum sequence.

Notes are in black/red
Student exercises are in blue
Solutions (which the students do not see) are in cyan
<table>
<thead>
<tr>
<th>What’s its name?</th>
<th>Bra-ket</th>
<th>Matrix</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. a ket/state vector</td>
<td>$</td>
<td>\alpha\rangle$</td>
<td>$\bar{a}$</td>
</tr>
<tr>
<td>2. a different state vector</td>
<td>$</td>
<td>\beta\rangle$</td>
<td>$\bar{b}$</td>
</tr>
<tr>
<td>3. inner product</td>
<td>$\langle \beta</td>
<td>\alpha \rangle$</td>
<td>$\bar{b}^\dagger \bar{a}$</td>
</tr>
<tr>
<td>4. bra/dual vector/functional</td>
<td>$\langle \beta</td>
<td>$</td>
<td>$\bar{b}^\dagger$</td>
</tr>
<tr>
<td>5. transformation/matrix/operator</td>
<td>$\hat{Q}$</td>
<td>$Q$ (a matrix!)</td>
<td>$\hat{Q}$ (an operator!)</td>
</tr>
<tr>
<td>6. inner product of $</td>
<td>\beta\rangle$ with $\hat{Q}</td>
<td>\alpha\rangle$</td>
<td>$\langle \beta</td>
</tr>
<tr>
<td>7. norm</td>
<td>$\sqrt{\langle \alpha</td>
<td>\alpha \rangle}$</td>
<td>$\sqrt{\bar{a}^\dagger \bar{a}}$</td>
</tr>
<tr>
<td>8. TI Schrödinger equation</td>
<td>$\hat{H}</td>
<td>\psi\rangle = E</td>
<td>\psi\rangle$</td>
</tr>
<tr>
<td>9. expectation value of $Q$</td>
<td>$\langle Q \rangle = \langle \psi</td>
<td>\hat{Q}\psi \rangle$</td>
<td>$\langle \tilde{\psi}</td>
</tr>
<tr>
<td>10. eigenvalue equation</td>
<td>$\hat{Q}</td>
<td>\alpha\rangle = \lambda</td>
<td>\alpha\rangle$</td>
</tr>
<tr>
<td>11. Hermitian conjugation</td>
<td>$\langle \hat{Q}^\dagger</td>
<td>\alpha\rangle = (Q^\dagger \bar{b})^\dagger \bar{a}$</td>
<td>$\int (\hat{Q}^\dagger g(x))^* f(x)dx = \int g(x)^* \hat{Q}f(x)dx$</td>
</tr>
<tr>
<td>12. Hermitian operator</td>
<td>$\hat{Q}^\dagger = \hat{Q}$</td>
<td>$Q^\dagger = Q$</td>
<td>$\hat{Q}^\dagger = \hat{Q}$</td>
</tr>
</tbody>
</table>

A functional is a linear map from a vector space to its field of scalars. The functional notation $[...]$ indicates where the operand is plugged into the operator expression for the functional. This is also called the “covector” or “one-form.”

Note that the time-independent Schrödinger equation is just a special case of an eigenvalue problem, with operator $\hat{Q} = \hat{H}$ and eigenvalue $\lambda = E$. A stationary state wave function is therefore an eigenvector of the Hamiltonian $H$.
Recap of inner products & Hermitian transformations:

In the previous lecture, we introduced an “inner product space” as a vector space over the field of (complex) scalars, together with an inner product that satisfies three defining properties:

(i) $\langle \beta | \alpha \rangle = \langle \alpha | \beta \rangle^*$,

(ii) $\langle \alpha | \alpha \rangle \geq 0$,

(iii) $\langle \alpha | (b \beta + c \gamma) \rangle = b \langle \alpha | \beta \rangle + c \langle \alpha | \gamma \rangle$.

We also introduced the notation $|a\alpha\rangle \equiv a |\alpha\rangle$ and $|\hat{T}\rangle \equiv \hat{T} |\alpha\rangle$.

You used the defining properties of the inner product to show that $\langle a\alpha | = a^* \langle \alpha |$.

We defined the Hermitian conjugate of $\hat{T}$ to be the transformation $\hat{T}^\dagger$ that satisfies

$$\langle \hat{T}^\dagger \alpha | \beta \rangle = \langle \alpha | \hat{T} \beta \rangle.$$  \hspace{1cm} (145)

And we defined a Hermitian transformation to be a transformation $\hat{T}$ that satisfies

$$\langle \hat{T} \alpha | \beta \rangle = \langle \alpha | \hat{T} \beta \rangle.$$  \hspace{1cm} (146)

You then proved the first two of these three important properties of the eigenvalues and eigenvectors of Hermitian transformations:

1. The eigenvalues of a Hermitian transformation are real.

2. The eigenvectors of a Hermitian transformation belonging to distinct eigenvalues are orthogonal.

3. The eigenvectors of a Hermitian transformation span the space – i.e., they form a complete set of basis vectors.

We are now ready to move on to the deep connection between Hermitian transformations and operators in quantum mechanics!
Earlier this quarter, we defined the position and momentum operators:

\[ \hat{x} = x; \quad \hat{p} = -i\hbar \frac{\partial}{\partial x}. \]  

(147)

Exercise 1: Are the position and momentum operators Hermitian?

(a) Working together on your white board, use the definition of the inner product in terms of an integral to determine whether the position operator \( \hat{x} \) is Hermitian.

Need to show that \( \langle f|xg \rangle = \langle xf|g \rangle \). Start with right-hand side:

\[
\langle xf|g \rangle = \langle g|xf \rangle^* = \left( \int_{-\infty}^{+\infty} g(x)^* xf(x) dx \right)^* = \int_{-\infty}^{+\infty} g(x)xf(x)^* dx = \int_{-\infty}^{+\infty} f(x)^* xg(x) dx \]  

since \( f, g \) and \( x \) all commute.

Therefore the position operator is Hermitian.

(b) Now determine whether the momentum operator \( \hat{p} \) is Hermitian.

Need to show that \( \langle f|pg \rangle = \langle pf|g \rangle \). Start with right-hand side:

\[
\langle pf|g \rangle = \langle g|pf \rangle^* = \left( \int_{-\infty}^{+\infty} g(x)^* (-i\hbar \frac{\partial}{\partial x}) f(x) dx \right)^* = \left( \int_{-\infty}^{+\infty} g(x)(+i\hbar \frac{\partial}{\partial x}) f(x)^* dx \right) \]

Take complex conjugate.

Integrate by parts.

Surface term is zero.

Therefore the momentum operator is Hermitian.
The Hamiltonian $H$ as a Hermitian Operator

Recall that the Hamiltonian $\hat{H}$ in the Schrödinger equation can be expressed in terms of $\hat{x}$ and $\hat{p}$:

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(x).$$  \hfill (148)

Since the operators $\hat{x}$ and $\hat{p}$ are Hermitian, $\hat{H}$ is also Hermitian. (Think about it...)

Now, the statements on the previous page about the eigenvalues and eigenvectors of a Hermitian transformation are general for any vector space in which there exists an inner product that satisfies the defining properties – called an inner product space or a Hilbert space$^5$.

$\implies$ We can apply all the statements regarding Hermitian transformations to the Hamiltonian $\hat{H}$, as well as to $\hat{x}$ and $\hat{p}$.

The Postulates of Quantum Mechanics

We are now ready to state the postulates of quantum mechanics in the language of linear algebra (physics $\iff$ linear algebra).

(1) The state of a particle is represented by a normalized vector $|\Psi\rangle$ in an inner product space.

(2) Observable quantities $Q$ are represented by Hermitian transformations $\hat{Q}$.

(3) The expectation value of $Q$, for a particle in the state $|\Psi\rangle$, is given by the inner product of $|\Psi\rangle$ with $\hat{Q}|\Psi\rangle$:

$$\langle Q \rangle_\Psi = \langle \Psi | \hat{Q} | \Psi \rangle = \langle \hat{Q} | \Psi \rangle \langle \Psi | \rangle.$$

(4) If we measure an observable $Q$ for a particle in the state $|\Psi\rangle$, we are certain to measure a value that is one of the eigenvalues of $\hat{Q}$.

(5) The probability of measuring the particular eigenvalue $\lambda$ is equal to the absolute square of the $\lambda$ component of $|\Psi\rangle$ when $|\Psi\rangle$ is expressed in the (complete) orthonormal basis of eigenvectors of $\hat{Q}$.

(6) Immediately after the measurement, the particle is in the state represented by the eigenvector corresponding to the measured eigenvalue.

To more fully interpret postulate (5), we’ll introduce some useful notation. In today’s Lecture, we’ll deal with the case of a discrete set of eigenvalues (possibly countably infinite in size). In a later Lecture, we’ll deal with the case of a continuous spectrum of eigenvalues (uncountably infinite).

$^5$See footnote 3 on p. 94 in Griffiths for a brief discussion of Hilbert space.
Discrete (non-degenerate) eigenvalues of a Hermitian operator $Q$:

- Label the eigenvalues and eigenvectors with an integer $n$:
  \[ \hat{Q} |e_n \rangle = \lambda_n |e_n \rangle . \]
  The $|e_n \rangle$ could be represented as column vectors or as functions.
- The eigenvectors can always be chosen to be orthonormal:
  \[ \langle e_n | e_m \rangle = \delta_{nm} . \]
  The inner product could be matrix multiplication or an integral.
- The completeness relation for the eigenvectors takes the form of a sum:
  \[ |\Psi \rangle = \sum_n c_n |e_n \rangle . \]
  $|\Psi \rangle$ could be a column vector or a function.
- Postulate (5) above can then be written
  \[ \text{Prob}(\lambda_n) = |c_n|^2 . \]
  \[ \sum_n \text{Prob}(\lambda_n) = \sum_n |c_n|^2 = 1. \]

**Exercise 2:** Show that $c_n = \langle e_n | \Psi \rangle$. *Hint:* Use “Fourier’s trick.”

Take the inner product of $|e_n \rangle$ with $|\Psi \rangle = \sum_m c_m |e_m \rangle$:

\[ \langle e_n | \Psi \rangle = \langle e_n | \sum_m c_m |e_m \rangle = \sum_m \langle e_n | c_m |e_m \rangle = \sum_m c_m \langle e_n | e_m \rangle = \sum_m c_m \delta_{nm} = c_n . \]

where we used $\langle e_n | e_m \rangle = \delta_{nm}$ to pick off only the $c_n$ term. So we get the result $c_n = \langle e_n | \psi \rangle$.

**Exercise 3:** Postulates (3) and (5) give two different ways of computing $\langle Q \rangle$: (3) dictates a formula, and (5) provides the probabilities necessary to compute it directly. Show that computing the expectation value using postulate (5) gives the same result as beginning with postulate (3) and expanding $|\Psi \rangle$ and $\langle \Psi |$ in terms of $\{|e_n\}\}$.

Postulate (5) $\implies$ $\langle Q \rangle = \sum_n \lambda_n |c_n|^2$.

Postulate (3) $\implies$ $\langle Q \rangle = \sum_n \lambda_n |c_n|^2$.

\[
\langle Q \rangle = \langle \Psi | \hat{Q} \Psi \rangle \\
= \left( \sum_n c_n^* \langle e_n | \right) \left( \sum_m \hat{Q} c_m |e_m \rangle \right) \\
= \sum_{n,m} c_n^* c_m \langle e_n | \hat{Q} e_m \rangle \\
= \sum_{n,m} c_n^* c_m \langle e_n | \lambda_m e_m \rangle \\
= \sum_{n,m} c_n^* c_m \lambda_m \langle e_n | e_m \rangle \\
= \sum_n \lambda_n |c_n|^2 \\
\]

Use orthonormality of $\{|e_n\}$.
Summary: For discrete eigenvalues \( \lambda_n \), the probability of measuring the value \( \lambda_n \) when we make a measurement on the state \( |\Psi\rangle \) is given by

\[
\text{Prob}(\lambda_n) = |\langle e_n | \Psi \rangle|^2 = |c_n|^2 \quad \text{and} \quad \langle Q \rangle = \sum_n \lambda_n |c_n|^2.
\]

(149)

Since \( |\Psi\rangle \) can depend on time, \( \text{Prob}(\lambda_n) \) can also depend on time through \( c_n(t) \).

Now let’s see how we can express “completeness”.

The Projection Operator

Exercise 4: Define an operator \( \hat{P}_n = |e_n\rangle \langle e_n| \). Find \( \hat{P}_n |\Psi\rangle \).

\[
\hat{P}_n |\Psi\rangle = |e_n\rangle \langle e_n| |\Psi\rangle \\
= c_n |e_n\rangle
\]

Use \( c_n = \langle e_n | \Psi \rangle \).

Exercise 5: Given your result, discuss with your classmates why \( P_n \) is called the projection operator.

\( \hat{P}_n \) projects \( |\Psi\rangle \) onto the one-dimensional subspace spanned by \( |e_n\rangle \).

Exercise 6: Show that \( \sum_n |e_n\rangle \langle e_n| \) is the identity operator — i.e., show that \( \sum_n |e_n\rangle \langle e_n| = 1 \). \text{Hint:} \text{Apply the operator } \sum_n |e_n\rangle \langle e_n| \text{ to } |\Psi\rangle.

\[
\sum_n |e_n\rangle \langle e_n| |\Psi\rangle = \sum_n |e_n\rangle c_n = \sum_n c_n |e_n\rangle = |\Psi\rangle.
\]

Note that the operator \( \sum_n |e_n\rangle \langle e_n| \) acts on a vector (e.g., a column vector or a function, depending on the nature of the vector space) and returns a vector.

Closure relation — a statement of completeness

The identity

\[
\sum_n |e_n\rangle \langle e_n| = 1
\]

is called the closure relation and is equivalent to saying that the set of orthonormal basis vectors \( |e_n\rangle \) is complete. The sum on the LHS is over a complete set of basis vectors; the identity allows you to insert this complete set of states on the left side of a vector \( |\Psi\rangle \), or on the right side of \( \langle \Psi| \). As you will see below, inserting a complete set of states allows for some very elegant proofs.
Matrix elements of operators in a particular basis

• Now consider the quantity $\langle e_n | \hat{Q} e_m \rangle$ where the $\{|e_n\}\} form a complete set of orthonormal basis vectors.

• Note that $\langle e_n | \hat{Q} e_m \rangle$ is an inner product and therefore is just a complex number.

• Just as a vector can be represented with respect to a particular basis $\{|e_n\}\}$ by its components $c_n$—i.e., $|\psi\rangle = \sum_n c_n |e_n\rangle$ with $c_n = \langle e_n | \psi \rangle$—an operator $\hat{Q}$ can be represented with respect to a particular basis by its matrix elements $Q_{nm} = \langle e_n | \hat{Q} e_m \rangle$.

• If the operator $\hat{Q}$ is Hermitian, then

$$Q_{nm} = \langle e_n | \hat{Q} e_m \rangle = \langle \hat{Q} e_n | e_m \rangle = \langle e_m | \hat{Q} e_n \rangle^* = Q_{mn}^*, \quad (150)$$

where we used the first defining property of an inner product. Therefore, the matrix $Q$ representing a Hermitian operator satisfies $Q^\dagger = Q$—i.e., $Q$ is equal to its transpose conjugate. Quick question: What does this imply about the diagonal matrix elements $Q_{nn}$ for a Hermitian operator? They are real since they must be equal to their complex conjugate.

Exercise 7: Use the closure relation to show that we can write $\langle \phi|\}$ in the $\{|e_n\}$ basis as $\langle \phi| = \sum_n b_n^* \langle e_n|\}$ where $b_n = \langle e_n|\phi\rangle$. Hint: Insert a complete set of states on the right of $\langle \phi|\}$.

$$\langle \phi| = \sum_n \langle \phi|e_n\rangle \langle e_n| = \sum_n \langle e_n|\phi\rangle^* \langle e_n| = \sum_n b_n^* \langle e_n|.$$ 

Exercise 8: Now insert a complete set of states between $\langle \phi|\}$ and $|\psi\rangle$ to show that

$$\langle \phi|\psi\rangle = \sum_n b_n^* c_n,$$

where $b_n$ and $c_n$ are the components of $|\phi\rangle$ and $|\psi\rangle$ in the $\{|e_n\}$ basis, as already defined: $b_n = \langle e_n|\phi\rangle$ and $c_n = \langle e_n|\psi\rangle$.

$$\langle \phi|\psi\rangle = \sum_n \langle \phi|e_n\rangle \langle e_n|\psi\rangle = \sum_n b_n^* c_n .$$

$\implies$ This is just the rule for matrix multiplication of a row vector with a column vector $(b^\dagger c)$.
Exercise 9: Now insert a complete set of states twice to show that
\[ \langle \phi | Q \psi \rangle = \sum_{n,m} b_n^* Q_{nm} c_m. \]

\[ \langle \phi | Q \psi \rangle = \sum_{n,m} \langle \phi | e_n \rangle \langle e_n | Q | e_m \rangle \langle e_m | \psi \rangle \]
\[ = \sum_{n,m} b_n^* Q_{nm} c_m \]

We can use the usual rules for matrix multiplication to evaluate inner products, expectation values, etc., once we evaluate the components of the vectors and the matrix elements of the operators!

Exercise 10: Suppose that the basis vectors \( \{ | e_n \rangle \} \) correspond to the eigenvectors of the Hermitian operator \( \hat{Q} \) with corresponding eigenvalues \( \{ \lambda_n \} \). Formally write out the matrix elements \( Q_{nm} = \langle e_n | \hat{Q} e_m \rangle. \)

\[ Q_{nm} = \langle e_n | \hat{Q} e_m \rangle \]
\[ = \langle e_n | \lambda_m e_m \rangle \]
\[ = \lambda_m \langle e_n | e_m \rangle \]
\[ = \lambda_m \delta_{nm} \]

On your white board, write out the matrix \( Q \) whose elements are \( Q_{nm} \).

We can see from the above calculation that the matrix will be diagonal:
\[
Q = \begin{pmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \lambda_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
\end{pmatrix}
\]

Activity 12: Matrix Representation of Operators – Applications to the Harmonic Oscillator Potential
**Course:** Physics 130 — Introduction to Quantum Mechanics II

**Instructor:** Sean Hartnoll

**Target Audience:** Sophomores/Juniors

**Brief summary of examples:**

Sean writes the notes as mini lectures on the board, and has the students fill out the worksheets during class.
The variational principle

Perturbation theory is useful if there's a small number in the Hamiltonian, what if not?

Variational principle -> fairly simple to get upper bounds on ground state energy. [and sometimes excited states too]

Theorem: Pick any normalized state \( \psi_1 \). Then
\[
E_\psi \leq \langle \psi | H | \psi_1 \rangle = \langle \psi \mid H \rangle \psi
\]

Proof: GOTO WORKSHEET 22

Name of the game: pick as good as possible a wavefunt to get a good bound! Do something of an art.

Very useful! Much of physical chemistry, BCS theory, PECHE...

Eg, if we can find a state st. \( \langle \psi | H | \psi_1 \rangle < 0 \), we know the potential admits a bound state.

Example: Ground state of helium revisited.
\[
H = -\frac{\epsilon^2}{2m} \left( \nabla_1^2 + \nabla_2^2 \right) - \frac{\epsilon^2}{4\pi\alpha} \left( \frac{1}{r_1} + \frac{1}{r_2} - \frac{1}{r_1 - r_2} \right)
\]

Experiment: \( E_\psi = -78.975 \text{ eV} \)
saw before that ignoring the interaction: $E = -10.9 \text{eV}$, off.

But now suppose we use the non-interacting ground state as a trial wavefunction:

$$\psi_0(r_1, r_2) = \psi_{1s}(r_1) \psi_{1s}(r_2)$$

$$\psi_0 = \frac{8}{\pi a^3} e^{-2(r_1 + r_2)/a}$$

Then:

$$\langle H \rangle_{\psi_0} = -10.9 \text{eV} + \left[ \frac{\langle (r_1 - r_2)^2 \rangle}{\langle r_1, r_2 \rangle} \right] + \frac{e^2}{4 \pi \hbar^2} \frac{1}{a}$$

$$= -7.5 \text{eV}$$

So now we know that this is an upper bound, not a good one!

Note, this is the same computation as in first order perturbation theory, always gives an upper bound.

We can do better!

**GOTO WORKSHEET 23**

$$\langle H \rangle = \frac{1}{2} (\frac{3}{2})^2 E_1 = -7.75 \text{eV}, \text{ within 2\% of experimental values}$$
Hydrogen molecule ion

Does H\textsuperscript{+} exist?

I.e. is the energy of an electron bound to two protons less than the energy of Hydrogen + a free proton.

I.e. can one electron lead to a covalent bond of two protons?

GOTO WORKSHEET 24

Let \( \psi_0(r) = \frac{1}{\sqrt{4\pi a^3}} e^{-r/a} \), ground state of hydrogen.

Try \( \psi = A \left[ \psi_0(r_1) \pm \psi_0(r_2) \right] \) this type of wavefunction is called LCAO, linear combination of above orbitals.

The integrals needed to obtain \( \langle H \rangle \) are a mess.

- See section 7.3 of Griffiths.

\[
\langle H \rangle + \frac{2a}{R} E_1
\]

Equilibrium: \( R = 2.5a \approx 1.3 \text{Å} \)

\( E = 1.8 \text{eV} \)

Extrapolation? \( R = 1.06a \) ?

\( E = 2.8 \text{eV} \) ?
Given a Hamiltonian $H$, write a wavefunction $\psi$ as an expansion in the orthonormal basis of energy eigenstates $\psi_n$.

Obtain or just write down the formula for $\langle H \rangle$ in the state $\psi$ in terms of the coefficients in the expansion above.

From this expansion, why can you conclude that $E_{\text{gs}} \leq \langle H \rangle$?

Suppose that you are given a complicated potential that is negative everywhere and goes to zero as $|x| \to \infty$. Does such a potential necessarily admit a bound state? For a given potential, how might you try to show that it admits a bound state?
We have just used the ground state of the non-interacting Hamiltonian as a trial wavefunction for the ground state of Helium. I.e. we took

$$\psi = \frac{8}{\pi a^3} e^{-2(r_1+r_2)/a}. \quad (1)$$

You can do better by trying a different wavefunction that incorporates the effects of screening. You might think that screening would reduce the effective nuclear charge seen by the electrons to some number $Z < 2$. What new trial wavefunction might you try that incorporates this effect? This wavefunction should depend on an unspecified number $Z$.

$\langle H \rangle_\psi$ can be computed for this wavefunction with integrals that you have already done (see Griffiths section 7.2). The answer is

$$\langle H \rangle_\psi = \left(-2Z^2 + \frac{27}{4}Z\right) \times (-13.6 \text{ eV}). \quad (2)$$

What is the best possible bound on the ground state energy that you can get from this expression? How does it compare to the experimental value?
Let the protons be a distance $R$ apart. Write down the Hamiltonian for the electron when it is a distance $r_1$ from the first proton and $r_2$ from the second proton. Draw a picture showing the electron and the two protons.

How could we use the variational principle for the electron with above Hamiltonian to determine whether or not a bound state of the two protons exists?

What might be a good variational wavefunction?
Course: Physics 120/121 — Advance Electricity and Magnetism I & II

Instructor: Jason Hogan

Target Audience: Sophomores — typically first advanced physics course they take

Brief summary of examples:

“The example from Physics 120 is from week 6 (lecture 16) and is intended to help the students develop a physical picture of polarization and bound charge. I do a similar activity for magnetization and bound current later in the quarter.

The first example from Physics 121 (lecture 3) shows how I like to handle longer derivations in class (in this case Poynting’s theorem). The students are first guided through the steps, and then we do a practice problem with the result. The second example is from late in the quarter (lecture 20) and shows how I try to use tables/lists as scaffolding -- here the students construct Maxwell’s equations in tensor form. Usually I partially fill in these kind of tables (as here) to give them a few hints. Most students pick up on the pattern.”

Students are given the worksheets without solutions and Jason works through it with them in class.
Physics 120: Worksheet for Lecture 16

Learning Goals: By the end of this class you should be able to:

• Explain polarizability and to be able to calculate its magnitude for a simple atom.
• Define the polarization $\vec{P}$.
• Explain the physical origin of bound charge.
• Calculate both volume and surface charge given $\vec{P}$.
• Explain the relationship of the electric displacement, $\vec{D}$, to $\vec{E}$ and $\vec{P}$.

Work on these questions as a team, without using your notes or the textbook.

Useful Results from Prior Lectures:

• Boundary Conditions

\[
(E_1 - E_2) \cdot \hat{n} = \frac{\sigma}{\epsilon_0} \\
(E_1 - E_2) \times \hat{n} = 0 \\
V_1 = V_2 \\
\frac{\partial V_1}{\partial n} - \frac{\partial V_2}{\partial n} = -\frac{\sigma}{\epsilon_0}
\]

• Dipole moment of a charge distribution and associated potential:

\[
\vec{p} = \int \vec{r}'\rho(\vec{r}') d\tau' \\
V(\vec{r}) = \frac{\vec{p} \cdot \hat{r}}{4\pi\epsilon_0 r^2}
\]
1 Polarizability of an Atom

In this section you are going to derive the polarizability of an atom, using a simple classical model of a point nucleus $+q$ surrounded by a uniform electron cloud of radius $a$, and total charge $-q$. When the atom is placed in an electric field $\vec{E}$, the cloud is displaced by a distance $d$ from the nucleus, as shown below. It is assumed that the cloud does not deform.

\begin{equation*}
q = Ze
\end{equation*}

1. Write down the charge density $\rho(\vec{r}')$ and find the dipole moment using the center of the cloud as the origin of coordinates.

Say the displacement is in the $\hat{y}$ direction: $\rho(\vec{r}') = \rho_{\text{cloud}}(\vec{r}') + Ze \delta(\vec{r}' - \hat{y}d)$

where $\rho_{\text{cloud}}(\vec{r}') = \begin{cases} 
\rho_o & r' < a \\
0 & r' > a
\end{cases}$ and $\rho_o$ is a constant. (Note origin of $\vec{r}'$ is at sphere center.)

\begin{equation*}
\vec{p} = \int \vec{r}' \rho(\vec{r}') d\tau' = \int \rho_{\text{cloud}}(\vec{r}') \vec{r}' d\tau' + Ze \int \vec{r}' \delta(\vec{r}' - \hat{y}d) d\tau'
\end{equation*}

\begin{equation*}
= \rho_o \int \vec{r}' d\tau' + Ze d \hat{y}
\end{equation*}

\begin{equation*}
= 0 \text{ by symmetry}
\end{equation*}

\begin{equation*}
\vec{p} = Ze d \hat{y}
\end{equation*}

2. Find an expression for the electric field at the location of the nucleus (remember, the cloud is assumed to be uniformly charged).

At the nucleus, the electric field consists of the external applied field $\vec{E}$ plus the field due to the cloud $\vec{E}_{\text{cloud}}$.

Due to spherical symmetry we can use Gauss's law to find $\vec{E}_{\text{cloud}}$:

\begin{equation*}
\rho_o \int \vec{E}_{\text{cloud}} \cdot d\vec{a} = \frac{1}{\varepsilon_0} Q_{\text{enc}} \Rightarrow 4\pi r^2 |\vec{E}_{\text{cloud}}| = \frac{1}{\varepsilon_0} \frac{Ze r}{3} \rho_o
\end{equation*}

where $\rho_o = \frac{-q}{4\pi a^3} \Rightarrow \vec{E}_{\text{cloud}} = \frac{1}{4\pi \varepsilon_0} \frac{Ze}{a^3} (-\hat{r})$

and $q = Ze$

At the nucleus, $\vec{r} = \hat{y}d$ \Rightarrow $\vec{E}_{\text{cloud}}(\hat{y}d) = -\frac{1}{4\pi \varepsilon_0} \frac{Ze d}{a^3} \hat{y}$
3. Use this expression to find the atomic polarizability $\alpha$ (the constant of proportionality between the induced dipole $p$ and applied electric field $E$, such that $p = \alpha E$.)

In equilibrium, the total electric field at the nucleus is zero:

$$E_{\text{tot}} = E + E_{\text{cloud}}(\hat{y}d) = 0 \implies E = -E_{\text{cloud}}$$

\[\begin{cases}
\vec{E} = \frac{1}{4\pi\varepsilon_0} \frac{Zed}{a^3} \hat{y} \\
\vec{p} = Zed \hat{y}
\end{cases}\]

\[\vec{E} = \frac{1}{4\pi\varepsilon_0} \frac{\vec{p}}{a^3}\]

\[\vec{p} = 4\pi\varepsilon_0 a^3 \vec{E} = \alpha \vec{E} \quad \alpha = 4\pi\varepsilon_0 a^3\]

4. Optional if you finish the other parts early. Show that the dipole moment you found in part (a) doesn’t actually depend on the choice of origin.

Consider a new origin displaced by $\vec{r}_0$ from the original:

W.r.t. the new origin $\tilde{O}$,

\[\vec{p} = \int \vec{r} \rho(\vec{r}') d\tau' \quad \text{and} \quad \vec{r} = \vec{r}' - \vec{r}_0\]

\[\vec{p} = \int (\vec{r}' - \vec{r}_0) \rho(\vec{r}') d\tau' = \int \vec{r}' \rho(\vec{r}') d\tau' - \vec{r}_0 \int \rho(\vec{r}') d\tau' = \vec{p} - \vec{r}_0 \rho = Q\]

\[\vec{p} = \vec{p} - \vec{r}_0 Q \quad \text{but} \quad Q = 0 \quad \text{for this problem, so}\]

\[\vec{p} = \vec{p} \quad \text{(unchanged)}\]
2 Polarization; Free and Bound Charge

1. A slab of plastic is placed within a charged capacitor. In the absence of the slab, there is a uniform electric field \( E_{\text{ext}} \) inside the capacitor. Assume that each molecule in the plastic can be modeled as two charges, \( +q \) and \( -q \), with a separation \( d \).

(a) Sketch, on the figure, a cartoon of the first few rows of atoms inside the plastic.
(b) Show on the sketch where the net charge is positive, negative, and zero.
(c) If there are \( N \) atoms per unit volume in the plastic write simple expressions for:
   i. The volume charge density \( \rho \) within the slab.
   ii. The surface charge density, \( \sigma \) on all surfaces.
   iii. The polarization \( \vec{P} \), equal to the dipole moment per unit volume. Which way does \( \vec{P} \) point?

\( \text{(i)} \) Inside the slab, each + charge is very near each - and they effectively cancel. So \( \rho = 0 \)

\( \text{(ii)} \) On the sides of the slab there are no charges because the dipoles are aligned along \( \vec{E} \). On the top (and bottom) surface, there is a set of + charges that extend a distance \( d \) above the last row of - charges, and these are not cancelled.

\( \text{(iii)} \) \( \vec{P} = N \vec{p} = \left( \frac{\text{Dipole moment}}{\text{volume}} \right) \) dipole moment of one atom

\[ |\vec{P}| = N\vec{p} = \sigma \]

\[ \vec{P} \text{ points along } \vec{E}_{\text{ext}} \]
(d) Now suppose that the top of the slab is sliced off at an angle \( \theta \) as shown below. Calculate the new surface charge on the top surface.

Since the dipoles are still aligned vertically along \( \vec{E}_{\text{ext}} \), the total number of charges on the surface does not change:

\[ Q = q(NA\delta) \quad \text{(from part (c))} \]

However, the area of the top surface is now larger:

\[ A_n = \frac{A}{\cos \theta} \]

where unit vector \( \hat{n} \) is \( \perp \) to surface \( A_n \)

\[ \Rightarrow \quad \sigma = \frac{Q}{A_n} = \frac{qN\delta}{(A/\cos \theta)} = \rho N \cos \theta \quad \sigma = \left| \vec{P} \right| \cos \theta \]

\[ \sigma = \vec{P} \cdot \hat{n} \]

(e) By comparing to the equations for calculating bound charge, show that the answers to (c)-(d) are what you expect for the polarization you calculated.

We found a uniform polarization \( \vec{P} \).

The expected bound charge densities are

\[
\begin{align*}
\sigma_b &= \vec{P} \cdot \hat{n} & \checkmark \\
\rho_b &= -\nabla \cdot \vec{P} = 0 & \checkmark \\
\vec{P} &= \text{constant}
\end{align*}
\]

As expected.
3 The Electric Displacement $\mathbf{D}$

(a) Starting from Gauss’s law in differential form, construct a vector field $\mathbf{D}$ which satisfies a “Gauss-like” law but that only depends on the free charge density $\rho_f$.

$$\nabla \cdot \mathbf{E} = \frac{1}{\varepsilon_0} \rho = \frac{1}{\varepsilon_0} (\rho_f + \rho_b) = \frac{1}{\varepsilon_0} (\rho_f - \nabla \cdot \mathbf{P})$$

Free \quad \text{(Bound)}

$$\nabla \cdot (\varepsilon_0 \mathbf{E} + \mathbf{P}) = \rho_f \quad \Rightarrow \quad \nabla \cdot \mathbf{D} = \rho_f \quad \mathbf{D} \equiv \varepsilon_0 \mathbf{E} + \mathbf{P}$$

(b) Express this law for $\mathbf{D}$ in integral form.

Integrating both sides over volume $V$:

$$\int_V (\nabla \cdot \mathbf{D}) \, dV = \int_V \rho_f \, dV$$

$$\oint_{\partial V} \mathbf{D} \cdot d\mathbf{a} = Q_{f,\text{enc}}$$

(c) Use your answer to part 2 to derive boundary conditions for the normal component of $\mathbf{D}$. Draw a picture to help define all terms.

$$\oint_{\partial \mathbf{D}_1} \mathbf{D} \cdot d\mathbf{a} = \mathbf{D}_1 \cdot \hat{n} A - \mathbf{D}_2 \cdot \hat{n} A = \sigma_f A$$

$$\mathbf{D}_1 \cdot \hat{n} - \mathbf{D}_2 \cdot \hat{n} = \sigma_f$$

$$\mathbf{D}_1 - \mathbf{D}_2 = \sigma_f$$

(d) What about the parallel component of $\mathbf{D}$?

From $\nabla \times \mathbf{E} = 0$ we know $\mathbf{E}_1^\parallel - \mathbf{E}_2^\parallel = 0$. Using $\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P}$

$$\mathbf{D}_1^\parallel - \mathbf{D}_2^\parallel = \mathbf{P}_1^\parallel - \mathbf{P}_2^\parallel$$
Physics 121: Worksheet for Lecture 3

Learning Goals: By the end of this class you should be able to:

- Explain why conservation laws for energy and momentum need to be expanded in electrodynamics.
- Derive conservation of energy for electromagnetic fields.
- Use Poynting's theorem to calculate the rate of change of electromagnetic energy in a region.

Work on these questions as a team, without using your notes or the textbook.

Useful Results from Prior Lectures:

\[ \nabla \cdot \vec{E} = \frac{\rho}{\varepsilon_0} \]
\[ \nabla \cdot \vec{B} = 0 \]
\[ \nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \]
\[ \nabla \times \vec{B} = \mu_0 \vec{J} + \mu_0 \varepsilon_0 \frac{\partial \vec{E}}{\partial t} \]
1 Poynting’s Theorem

You will find the following useful in this section:

- The energy density of electromagnetic fields is: \( u = \frac{1}{2} \varepsilon_0 E^2 + \frac{1}{2\mu_0} B^2 \)

- Current density: \( \vec{J} = \rho \vec{V} \)

1. Write down the work \( dW \) done on a charge \( q \) in time \( dt \) due to the Lorentz force from arbitrary \( \vec{E} \) and \( \vec{B} \) fields.

\[
\vec{B} \rightarrow \quad d\vec{l} = \vec{V} dt
\]

\[
dW = \vec{F} \cdot d\vec{l} = \vec{E} \cdot d\vec{l}
\]

\[
dW = q \left( \vec{E} + \vec{V} \times \vec{B} \right) \cdot \vec{V} dt
\]

2. What do you notice about the \( \vec{B} \) term?

\[
\vec{V} \cdot (\vec{V} \times \vec{B}) = 0 \quad \vec{B} - \text{fields do no work.}
\]

3. We now know the power \( \frac{dW}{dt} \). Generalize this result for an arbitrary charge density \( \rho \).

Express in terms of \( \vec{J} \).

\[
\frac{dW}{dt} = \sum_{\text{charges}} q \vec{V} \cdot \vec{E} \quad \quad \frac{dW}{dt} = \int_{V} \rho \vec{V} \cdot \vec{E} dt
\]

Using \( \vec{J} = \rho \vec{V} \)

\[
\frac{dW}{dt} = \int_{V} (\vec{E} \cdot \vec{J}) dt
\]

4. In the above, \( \vec{E} \cdot \vec{J} = \frac{\text{power}}{\text{volume}} \). For now, just work with \( \vec{E} \cdot \vec{J} \) by itself. We want to express this entirely in terms of the fields. Eliminate \( \vec{J} \) using Ampere’s law.

\[
\vec{E} \cdot \vec{J} = \vec{E} \cdot \left( \frac{1}{\mu_0} \nabla \times \vec{B} - \varepsilon_0 \frac{\partial \vec{E}}{\partial t} \right)
\]

\[
\vec{E} \cdot \vec{J} = \frac{1}{\mu_0} \vec{E} \cdot (\nabla \times \vec{B}) - \varepsilon_0 \vec{E} \cdot \frac{\partial \vec{E}}{\partial t}
\]
5. Note the following vector identity (product rule 6 in Griffiths):

\[ \nabla \cdot (\vec{E} \times \vec{B}) = \vec{B} \cdot (\nabla \times \vec{E}) - \vec{E} \cdot (\nabla \times \vec{B}) \]

Using this identity (and Maxwell’s equations), get rid of all the curls in \( \vec{E} \cdot \vec{J} \).

\[
\vec{E} \cdot \vec{J} = \frac{1}{\mu_0} \left( \vec{B} \cdot (\nabla \times \vec{E}) - \nabla \cdot (\vec{E} \times \vec{B}) \right) - \varepsilon_0 \vec{E} \cdot \frac{\partial \vec{E}}{\partial t} = -\frac{\partial \vec{B}}{\partial t} \quad \text{(Faraday's law)}
\]

\[
\vec{E} \cdot \vec{J} = -\varepsilon_0 \vec{E} \cdot \frac{\partial \vec{E}}{\partial t} - \frac{1}{\mu_0} \vec{B} \cdot \frac{\partial \vec{B}}{\partial t} - \frac{1}{\mu_0} \nabla \cdot (\vec{E} \times \vec{B})
\]

6. Final result: Now write down \( \frac{dW}{dt} \) using the electromagnetic energy density \( u \).

We can simplify \( \vec{E} \cdot \vec{J} \) by noticing a total time derivative:

\[
\vec{E} \cdot \vec{J} = -\frac{\partial}{\partial t} \left( \frac{1}{2} \varepsilon_0 \vec{E}^2 + \frac{1}{2\mu_0} \vec{B}^2 \right) - \frac{1}{\mu_0} \nabla \cdot (\vec{E} \times \vec{B})
\]

Plugging this into part (3) and using the divergence theorem:

\[
\frac{dW}{dt} = -\frac{d}{dt} \int_V \left( \frac{1}{2} \varepsilon_0 \vec{E}^2 + \frac{1}{2\mu_0} \vec{B}^2 \right) \, dV - \oint_S \left( \frac{1}{\mu_0} \vec{E} \times \vec{B} \right) \cdot d\vec{a}
\]

Power delivered to particles in \( V \)

Field energy content

Flow of energy through boundary

Poynting's Theorem:

\[
\frac{dW}{dt} = -\frac{d}{dt} \int_V u \, dV - \oint_S \vec{S} \cdot d\vec{a}
\]

In empty space: \( \frac{dW}{dt} = 0 \) (no particles around) so:

\[
\frac{d}{dt} \int_V u \, dV = -\oint_S \vec{S} \cdot d\vec{a} \quad \iff \quad \frac{d}{dt} u = -\nabla \cdot \vec{S} \quad \text{(Integral)}
\]

\( \text{(Differential)} \)
2 Increasing Current Solenoid, Revisited

Consider a long solenoid of length $L$ and radius $R$ with $n$ turns per unit length that is aligned along the $z$-axis. The current $I$ in the solenoid is increasing at a rate $\dot{I}$.

1. Calculate the electric field inside $(s < R)$ and outside $(s > R)$ the solenoid.

   \[
   \overrightarrow{\mathbf{B}} = \mu_0 n I \hat{z} \quad \overrightarrow{\mathbf{E}} \text{ forms closed loops.} 
   \]

   \[
   \text{Faraday: } \oint_C \overrightarrow{\mathbf{E}} \cdot d\overrightarrow{\ell} = -\frac{d}{dt} \int_S \overrightarrow{\mathbf{B}} \cdot d\overrightarrow{\mathbf{a}} 
   \quad \overrightarrow{\mathbf{E}} = E(s) \hat{\phi} 
   \]

   \[
   2\pi s |\overrightarrow{\mathbf{E}}| = -\frac{d}{dt} \left\{ \begin{array}{ll}
   \mu_0 n I \pi s^2 & (s < R) \\
   \mu_0 n I \pi R^2 & (s > R) 
   \end{array} \right. 
   \]

   \[
   \overrightarrow{\mathbf{E}} = \left\{ \begin{array}{ll}
   -\frac{\mu_0 n I}{2} \pi s^2 \hat{\phi} & (s < R) \\
   -\frac{\mu_0 n I}{2} \frac{R^2}{s} \hat{\phi} & (s > R) 
   \end{array} \right. 
   \]

2. Calculate the Poynting vector $\vec{S} = \frac{1}{\mu_0} \overrightarrow{\mathbf{E}} \times \overrightarrow{\mathbf{B}}$ at the surface of the solenoid $(s = R)$.

   \[
   \vec{S} = R : \quad \vec{S} = \frac{1}{\mu_0} \overrightarrow{\mathbf{E}} \times \overrightarrow{\mathbf{B}} = \frac{1}{\mu_0} \left( \frac{\mu_0 n I}{2} \pi s^2 \right) (\mu_0 n I) (\hat{\phi} \times \hat{z}) 
   \quad \vec{S} = -\frac{\mu_0 n^2 I^2 \pi R}{2} \hat{\phi} 
   \quad \phi \times \hat{z} = \hat{S} 
   \]

3. Use a surface integral of $\vec{S}$ (and Poynting’s theorem) to calculate the rate of change of electromagnetic energy inside the solenoid.

   \[
   \text{Poynting’s theorem: } \frac{d}{dt} \int_S \mathbf{u} \cdot d\mathbf{a} = -\oint_S \mathbf{S} \cdot d\overrightarrow{\mathbf{a}} 
   \]

   \[
   = \frac{\mu_0 n^2 I^2 \pi R}{2} \int_S d\mathbf{a} 
   \]

   \[
   = 2\pi R L 
   \]

   \[
   \frac{d}{dt} \int_S \mathbf{u} \cdot d\mathbf{a} = \mu_0 n^2 I^2 (\pi R^2 L) 
   \]
4. Now explicitly calculate the total magnetic energy $U_B = \int \frac{1}{2\mu_0} \mathbf{B}^2 dv$ inside the solenoid. Is this consistent with part (3) and Poynting’s theorem?

\[ \mathbf{B} = \mu_0 n \mathbf{I} \hat{z} \quad \quad U_B = \int \frac{1}{2\mu_0} (\mu_0 n \mathbf{I})^2 d\tau = \frac{\mu_0 n^2 I^2}{2} \int d\tau \]

\[ U_B = \frac{\mu_0 n^2 I^2}{2} (\pi R^2 L) \]

\[ \frac{dU_B}{dt} = \mu_0 n^2 \mathbf{I} \cdot \mathbf{I} (\pi R^2 L) \]

This agrees with part (3) and Poynting’s Theorem.

5. What about the total electric energy $U_E = \int \frac{\varepsilon_0}{2} \mathbf{E}^2 dv$ inside the solenoid? Is this consistent with part (5) and Poynting’s theorem?

\[ U_E = \int_{S<R} \frac{\varepsilon_0}{2} \left( -\frac{\mu_0 n \mathbf{I} \cdot s}{2} \right)^2 d\tau = \frac{\varepsilon_0 \mu_0}{8} n^2 I^2 \int_{0}^{R} s^2 (2\pi s L) ds = \int_{0}^{R} \frac{\varepsilon_0 \mu_0 n^2 I^2 R^2 (2\pi R^2)}{4} L \]

\[ U_E = \frac{\pi}{16} \varepsilon_0 \mu_0 n^2 I^2 R^4 L \]

So \[ \frac{dU_E}{dt} \neq 0 \], unless \[ I = 0 \]

Generally this seems inconsistent with parts (3) and (4) since Poynting’s theorem says \[ \frac{d}{dt} \int \mathbf{u} d\tau = \frac{d}{dt} (U_B + U_E) = -\oint \mathbf{S} \cdot d\mathbf{a} \]

But part (4) showed that \[ \frac{dU_B}{dt} = -\oint \mathbf{S} \cdot d\mathbf{a} \]

So our answer seems to violate Poynting’s theorem. What’s wrong?
6. What approximations (if any) did we make when originally solving Maxwell’s equations for $\vec{E}$ and $\vec{B}$?

The quasistatic approximation in our original solution for $\vec{B}$.

The Biot-Savart law ignores the displacement current $\frac{d\vec{B}}{dt}$.

Since we expect $\frac{d\vec{E}}{dt} = 0$, our $\vec{B}$-field is slightly wrong.

Poynting's theorem is exact, so it must be that our fields are only approximate. How bad is our error?

7. Examine the ratio $U_E/U_B$. Given Poynting’s theorem, what must be true about the ratio $I/I$ for our solution to be consistent? Interpret this physically.

To calculate the error that we made, we can compare $U_E$ to $U_B$. As long as $U_E << U_B$ our solution will be approximately consistent with Poynting’s theorem.

$$U_E = \frac{\frac{1}{2} E_0 I^2 R^2 L}{\frac{1}{2} \mu_0 n^2 I^2 R^2 L} = \frac{1}{8} \left( \frac{R}{C} \right)^2 \left( \frac{I}{I} \right)^2$$

where $C = \frac{1}{\sqrt{\mu_0 \epsilon_0}}$

$speed$ $of$ $light$

So we will have $U_E << U_B$ when $\left( \frac{R}{C} \right)^2 \left( \frac{I}{I} \right)^2 \ll 1$

$$\frac{I}{I} \ll \frac{C}{R}$$

Note that $(I/I)$ is a measure of how quickly the current changes.

And $\frac{R}{C}$ is the time it takes light to cross the radius $R$.

So our solution is approximately consistent as long as $I$ changes slowly compared to the light travel time across the solenoid.

(That is, as long as the current changes are non-relativistic.)
Physics 121: Worksheet for Lecture 20

Learning Goals: By the end of this class you should be able to:
- Express Maxwell’s equations using tensor notation.
- Explain what it means for Maxwell’s equations in tensor form to be covariant.
- Write Maxwell’s equations using the relativistic 4-potential.
- Perform gauge transformation using tensor notation.

Work on these questions as a team, without using your notes or the textbook.

Useful Results from Prior Lectures:

Lorentz Transformation:

\[
\Lambda^\mu_\nu = \begin{pmatrix}
\gamma & -\gamma \beta & 0 & 0 \\
-\gamma \beta & \gamma & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix} \quad \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \beta = \frac{v}{c}
\]

Field Transformation. Frame \(\vec{S}\) moves along +x at speed \(v\) with respect to frame \(S\):

\[
\begin{align*}
\overline{E}_x &= E_x \\
\overline{E}_y &= \gamma(E_y - vB_z) \\
\overline{E}_z &= \gamma(E_z + vB_y) \\
\overline{B}_x &= B_x \\
\overline{B}_y &= \gamma(B_y + \frac{v}{c^2}E_z) \\
\overline{B}_z &= \gamma(B_z - \frac{v}{c^2}E_y)
\end{align*}
\]

Some 4-vectors:

\[
\eta^\mu = \begin{pmatrix} \gamma c \\ \gamma \vec{v} \end{pmatrix} \quad p^\mu = \begin{pmatrix} E/c \\ \vec{p} \end{pmatrix} \quad J^\mu = \begin{pmatrix} c\rho \\ \vec{J} \end{pmatrix}
\]

\[
\partial^\mu = \frac{\partial}{\partial x^\mu} \quad \partial_\mu = \frac{\partial}{\partial x^\mu}
\]

Fields in terms of the potentials:

\[
\vec{E} = -\nabla V - \frac{\partial \vec{A}}{\partial t} \quad \vec{B} = \nabla \times \vec{A}
\]

In the Lorenz Gauge:

\[
\nabla \cdot \vec{A} + \frac{1}{c^2} \frac{\partial V}{\partial t} = 0
\]

Transformation rules:

Vector: \( \vec{a}^\mu = \Lambda^\mu_\nu a^\nu \)

Covector: \( a_\mu = \Lambda^\nu_\mu \vec{a}_\nu \)

Tensor: \( \overline{T}^{\mu\nu} = \Lambda^\mu_\alpha \Lambda^\nu_\beta T^{\alpha\beta} \)
1. **Tensor Maxwell Equations**

1. Fill in the table below by writing Maxwell’s equations using tensor notation.

<table>
<thead>
<tr>
<th></th>
<th>( \nabla \cdot \mathbf{E} = \frac{1}{\varepsilon_0} \rho )</th>
<th>( \frac{1}{c} \partial_k E^k = \mu_0 c \rho = \mu_0 \mathbf{J}^0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>( \nabla \times \mathbf{B} - \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} = \mu_0 \mathbf{J} )</td>
<td>( \epsilon^{ijk} \partial_j B_k - \frac{1}{c} \partial_0 E^i = \mu_0 J^i ) ( i = 1, 2, 3 )</td>
</tr>
<tr>
<td>3</td>
<td>( \nabla \cdot \mathbf{B} = 0 )</td>
<td>( \partial_k B^k = 0 )</td>
</tr>
<tr>
<td>4</td>
<td>( \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 )</td>
<td>( \frac{1}{c} \epsilon^{ijk} \partial_j E_k + \partial_0 B^i = 0 )</td>
</tr>
</tbody>
</table>

2. Let’s try to simplify further by making the following definitions:

\[
F^{0i} = \frac{1}{c} E^i \quad F^{ij} = \epsilon^{ijk} B_k \\
G^{0i} = B^i \quad G^{ij} = -\frac{1}{c} \epsilon^{ijk} E_k
\]

Fill in the table below for each of Maxwell’s equations using \( F \) and \( G \).

<table>
<thead>
<tr>
<th></th>
<th>( \partial_k F^{0k} = \mu_0 \mathbf{J}^0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>( \partial_j F^{ij} - \partial_0 F^{0i} = \mu_0 \mathbf{J}^i )</td>
</tr>
<tr>
<td>3</td>
<td>( \partial_k G^{0k} = 0 ) ( k = 1, 2, 3 )</td>
</tr>
<tr>
<td>4</td>
<td>( -\partial_j G^{ij} + \partial_0 G^{0j} = 0 )</td>
</tr>
</tbody>
</table>
3. We want to convert all our sums to 4-vector notation, but there are still some annoying minus signs in equations (2) and (4)! However, notice that \( F^{ij} \) and \( G^{ij} \) are anti-symmetric because of the \( e^{ijk} \). Let’s make the whole tensor anti-symmetric by choosing:

\[
F^{i0} = -F^{0i} \quad G^{i0} = -G^{0i}
\]

Use these definitions to write Maxwell’s equations one more time, now using 4-vector summation notation:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \partial_k F^{0k} = \mu_0 J^0 )</td>
</tr>
<tr>
<td>2</td>
<td>( \partial_j F^{ij} + \partial_0 F^{i0} = \mu_0 J^i \Rightarrow \partial_j F^{i\nu} = \mu_0 J^i )</td>
</tr>
<tr>
<td>3</td>
<td>( \partial_\nu G^{0\nu} = 0 ) Note, by definition: ( G^{00} = -G^{00} ), so ( G^{00} = 0 ).</td>
</tr>
<tr>
<td>4</td>
<td>( \partial_j G^{ij} + \partial_0 G^{i0} = 0 \Rightarrow \partial_\nu G^{i\nu} = 0 )</td>
</tr>
</tbody>
</table>

4. Combine (1) with (2) and (3) with (4) to form two tensor equations. Success!

\[
\partial_\nu F^{\mu\nu} = \mu_0 J^\mu \\
\partial_\nu G^{\mu\nu} = 0
\]

\( F^{\mu\nu} \): "Field Tensor"

\( G^{\mu\nu} \): "Dual Tensor"

5. Based on our definitions above, write out the components of \( F^{\mu\nu} \) and \( G^{\mu\nu} \) in matrix form.

\[
F^{\mu\nu} = \begin{pmatrix}
0 & E_x/c & E_y/c & E_z/c \\
-E_x/c & 0 & B_z & -B_y \\
-E_y/c & -B_z & 0 & B_x \\
-E_z/c & B_y & -B_x & 0
\end{pmatrix}
\]

\( G^{\mu\nu} = \begin{pmatrix}
0 & B_x & B_y & B_z \\
-B_x & 0 & -E_z/c & E_y/c \\
-B_y & E_x/c & 0 & -E_x/c \\
-B_z & -E_y/c & E_x/c & 0
\end{pmatrix}
\]
2 Covariance of the Field Equations

Maxwell’s equations written in tensor form should be invariant under Lorentz transformations. Let’s check.

We found: \( \partial_\gamma = \Lambda_\gamma^\mu \partial_\mu \) and \( \bar{J}^\alpha = \Lambda_\alpha^\mu \bar{J}^\mu \)

\( \partial_\gamma F^{\mu\nu} = M_0 \bar{J}^\mu \)

\( \Lambda^\alpha_\mu (\Lambda_\nu^\beta \partial_\beta) F^{\mu\nu} = \Lambda^\alpha_\mu (M_0 \bar{J}^\mu) \)

\( \bar{\partial}_\beta (\Lambda_\alpha^\mu \Lambda_\nu^\beta F^{\mu\nu}) = M_0 \bar{J}^\alpha \)

\( \bar{\partial}_\beta \bar{F}^{\alpha\beta} = M_0 \bar{J}^\alpha \)

where \( \bar{F}^{\alpha\beta} = \Lambda_\alpha^\mu \Lambda_\beta^\nu F^{\mu\nu} \)

1. Does the tensor transformation law reproduce the transformation laws we found for \( E \) and \( B \) (see cover sheet)? How many unique terms in \( \bar{F}^{\mu\nu} \) would you have to consider to check this?

There are 6 unique terms in \( \bar{F}^{\mu\nu} \) that correspond to the six components of \( \bar{E} \) and \( \bar{B} \).

2. Check component \( \bar{F}^{02} \) of the tensor transformation law for \( F^{\mu\nu} \) by explicitly expanding out all the sums over the Lorentz transformations.

\[
\begin{align*}
\bar{F}^{02} &= \Lambda^0_\mu \Lambda^2_\nu F^{\mu\nu} = \Lambda^0_0 \Lambda^2_2 F^{02} + \Lambda^0_1 \Lambda^2_2 F^{12} \\
&= \gamma \cdot 1 \cdot F^{02} + (-\gamma \frac{V}{C}) \cdot 1 \cdot F^{12} \\
&= \gamma (F^{02} - \frac{V}{C} F^{12})
\end{align*}
\]

\( \frac{E_x}{C} = \gamma \left( \frac{E_y}{C} - \frac{V}{C} B_z \right) \)

Agrees with field transformation we found.
Course: Physics 170 — Statistical Mechanics

Instructor: Monika Schleier-Smith

Target Audience: Junior/Seniors

Brief summary of examples:

“1) Random Walk: lecture 1 of Physics 170
2) Statistical Derivation of the Ideal Gas Law: lecture 8 (week 4) of Physics 170

For each of these, I include one file that is just the empty worksheet and another file that includes the filled out worksheet and any additional notes/slides from that day. For anybody who wants to browse through more examples, my Canvas sites from both of these courses are public and can be accessed via the following link."

Students are given the worksheets without solutions and Monika works through it with them in class.
* Introductions

Chris Overstreet
Quinn MacPherson
Ryan Hazelton — Ph.D. in physics education research

We are interested understanding what concepts are hardest to grasp in learning statistical mechanics. To finding ways to assessing what works in teaching them.

* What is statistical mechanics?

* Course Information

* The Random Walk

N steps, each left or right ... Probabilities $p, q = 1-p$ right, left

Probability distribution $P_n(x)$ for where he ends up?
EXERCISE 1A: RANDOM WALK

Objectives:
- Derive the binomial distribution describing the statistics of a random walk
- Calculate and visualize the mean and variance of the binomial distribution

Reference: Kittel & Kroemer, Ch. 1

1. A drunkard, initially standing under a lamppost, sets out on a random walk along a road, taking either a step to the right with probability \( p \) or a step to the left with probability \( q = 1 - p \).

   a. After a total of \( N \) steps, what is the probability \( P_N(X) \) to find the drunkard a distance of \( X \) steps to the right of the lamppost \( (-N \leq X \leq N) \)? Derive your answer based on the following considerations:
      i. How many different sequences of \( N \) steps are possible?

         \[
         \begin{array}{c}

         2^N \text{ sequences}

         \[ \]

      ii. How many of these sequences include \( R \) steps to the right (and \( N - R \) steps to the left)? Explain your reasoning. (Extra: prove!)

         As suggested by the diagram above, the problem is equivalent to asking for the # of different ways of putting \( R \) right arrows into \( N \) boxes:

         \[
         \frac{N(N-1)(N-2)\ldots(N-R+1)}{R!} = \frac{N!}{(N-R)!R!} \equiv \binom{N}{R}
         \]

         we all \( \$ \)s are identical.

         \[
         \]

         * We can prove this rigorously by induction (see attached notes...)

         or algebraically (on homework).

         * For the algebraic approach: where does \( \binom{N}{R} \) show up?

         \[
         (a+b)^N = \sum_{R=0}^{N} \binom{N}{R} a^R b^{N-R}
         \]
Inductive proof of the # of combinations \((\binom{N}{R})\):

* Suppose our formula is valid for some \(N \land R \leq N\):

Is it also valid for \(N+1\)?

\[
\binom{N+1}{R} = \binom{N}{R} + \binom{N}{R-1} = \frac{N!}{(N-R)! \cdot R!} + \frac{N!}{(N-R+1)! \cdot (R-1)!} = \frac{N!\left[(N-R+1) + R\right]}{(N-R+1)! \cdot R!}
\]

\[
= \frac{(N+1)!}{[(N+1)-R]! \cdot R!}
\]

Also: \(\binom{N}{N} = \frac{N!}{0! \cdot N!} = 1 \quad \forall N\) \(\checkmark\)

Base case: \(\binom{1}{1} = 1 \quad \checkmark\)

\(\therefore\) valid \(\forall N, R\)
iii. What is the probability of executing a specific one of the sequences in ii.?

\[ P^R \binom{N-R}{R} = p^R (1-p)^{N-R} \]

iv. What is the probability \( P_N(R) \) for the drunkard to take \( R \) steps to the right?

\[ P_N(R) = \binom{N}{R} p^R (1-p)^{N-R} \]  \( \xrightarrow{p=1/2} \)  \( \frac{1}{2^N} \binom{N}{R} \) \( \checkmark \)

v. Verify that the distribution \( P_N(R) \) is properly normalized.

\[ \sum_{R=0}^{N} P_N(R) = \sum_{R=0}^{N} \binom{N}{R} p^R (1-p)^{N-R} = \left[ p + (1-p) \right]^N = 1 \]

\( \xrightarrow{\text{binomial theorem}} \)

vi. What is the probability \( P_N(X) \) for the drunkard to end up a distance \( X \) from the lamppost? Write out explicitly including factorials.

\[ X = R-L = R-(N-R) \]
\[ = 2R-N \quad \Rightarrow \quad R = \frac{N+X}{2} \]

\[ P_N(X) = \binom{N}{\frac{N+X}{2}} \frac{N+X}{2} \left( 1-p \right)^{\frac{N-X}{2}} = \frac{N!}{\left( \frac{N+X}{2} \right)! \left( \frac{N-X}{2} \right)!} \left[ p^\frac{N+X}{2} \left( 1-p \right)^{\frac{N-X}{2}} \right] \]

This is a bit messy ... to get an intuition for how \( P_N(X) \) looks, let's calculate the mean and variance.

* Looks reasonable, e.g., symmetric under \( X \rightarrow -X \) and \( p \rightarrow 1-p \).
b. Find the mean $\langle X \rangle$ and standard deviation $\Delta X$ of the drunkard’s final position:

i. Write down generic expressions for the mean $\langle Y \rangle$ and second moment $\langle Y^2 \rangle$ of a random variable $Y$ described by a discrete probability distribution $P(Y)$.

\[
\langle Y \rangle = \sum_Y P(Y) \cdot Y
\]
\[
\langle Y^2 \rangle = \sum_Y P(Y) \cdot Y^2
\]

Above assumes $\sum_Y P(Y) = 1$ properly normalized.

(What would you do otherwise?)

ii. How would you obtain the standard deviation $\Delta Y$ from $\langle Y \rangle$ and $\langle Y^2 \rangle$?

\[
\Delta Y = \sqrt{\langle (Y - \langle Y \rangle)^2 \rangle} \quad \text{root-mean-square deviation from mean}
\]
\[
= \sqrt{\langle Y^2 - 2\langle Y \rangle \langle Y \rangle + \langle Y \rangle^2 \rangle}
\]
\[
= \sqrt{\langle Y^2 \rangle - 2 \langle Y \rangle^2 + \langle Y \rangle^2}
\]
\[
= \sqrt{\langle Y^2 \rangle - \langle Y \rangle^2}
\]

iii. For the case of the random walk, it will be convenient first to calculate the mean $\langle R \rangle$ and variance $\Delta R$ of the number of steps to the right. How are these quantities related to $\langle X \rangle$ and $\Delta X$?

\[
X = 2R - N \quad \Rightarrow \quad \langle X \rangle = 2\langle R \rangle - N
\]
\[
\Delta X = 2\Delta R \quad \text{(at fixed } N, \text{i.e., } \Delta N = 0\text{)}
\]
iv. Calculate the average number of steps \( \langle R \rangle \) taken to the right using \( P_N(R) \).

\[
\langle R \rangle = \sum_{R=0}^{N} P_N(R) \cdot R = \sum_{R=0}^{N} \binom{N}{R} p^R q^{N-R} \cdot R
\]

Where \( q = 1 - p \)

**Setup:**

**Algebraic trick:** We know that generally, \( \sum_{R=0}^{N} \binom{N}{R} p^R q^{N-R} = (p+q)^N \) \( \forall p, q \)

Let's pull down a factor of \( R \) by taking a derivative:

\[
\frac{\partial}{\partial p} \sum_{R=0}^{N} \binom{N}{R} p^R q^{N-R} = \sum_{R=0}^{N} \binom{N}{R} R p^{R-1} q^{N-R}
\]

\[
\Rightarrow \langle R \rangle = p \frac{\partial}{\partial p} \sum_{R=0}^{N} \binom{N}{R} p^R q^{N-R}
\]

\[
= p \frac{\partial}{\partial p} [(p+q)^N] = p N (p+q)^{N-1} = p N \quad \left( q = 1 - p \right)
\]

\[
\therefore \langle R \rangle = p N \quad \checkmark
\]

In retrospect perhaps this was obvious, but the same trick will come in handy again and again...
v. Calculate the standard deviation $\Delta R = \sqrt{\langle R^2 \rangle - \langle R \rangle^2}$ of the number of steps taken to the right.

Now we also need $\langle E^2 \rangle$

\[
\langle E^2 \rangle = \sum_{R=0}^{N} \binom{N}{R} p^R q^{N-R} E_{-R}^2 = \left\{ \frac{\partial}{\partial p} \left[ p \sum_{R=0}^{N} \binom{N}{R} p^R q^{N-R} \right] \right\}_{q=1-p} = p \frac{\partial}{\partial p} \left[ p N(p+q)^{N-1} \right] = p \left[ N(p+q)^{N-1} + p N(N-1)(p+q)^{N-2} \right] = p \left[ N + p N(N-1) \right]
\]

\[
\Rightarrow \langle E^2 \rangle - \langle E \rangle^2 = pN + p^2 N(N-1) - p^2 N = (p-p^2)N = Np(1-p) = (\Delta R)^2
\]

\[
\therefore \Delta R = \sqrt{Np(1-p)}
\]

vi. Based on iv.-v., determine $\langle X \rangle$ and $\Delta X$. Evaluate your results explicitly for the special cases $p = 0$, $p = 1/2$, $p = 1$. Do they make sense?

\[
\langle X \rangle = 2\langle E \rangle - N = (2p-1)N
\]

\[
\Delta X = 2\Delta R = 2\sqrt{Np(1-p)}
\]

$p = 0$: $\langle X \rangle = -N$ (all steps left), $\Delta X = 0$ ✓

$p = 1/2$: $\langle X \rangle = 0$, $\Delta X = \sqrt{N}$ — characteristic scaling for random walk variance $\propto N$

$p = 1$: $\langle X \rangle = N$ (all steps right), $\Delta X = 0$ ✓
c. Compare the random walk of the human drunkard with a random walk undertaken by an ant, both tending towards the right with probability \( p = \frac{3}{4} \). The drunkard takes steps of length \( \ell = 40 \text{ cm} \), whereas the ant takes steps of length \( \ell = 1 \text{ mm} \). Each walks the same total distance \( L = N\ell = 10 \text{ m} \), but each with a different number of steps \( N \). Sketch, on a single set of axes, probability distributions for the final positions \( x = X\ell \) of the drunkard and the ant.

\[
\langle x \rangle = N(2p-1) = \frac{N}{2}
\]

\[
\langle X \rangle = L\langle x \rangle = \frac{L}{2}
\]

\[
\Delta x = 2\sqrt{Np(1-p)} = \frac{\sqrt{3}N}{2}
\]

\[
\Delta X = \frac{L}{2}\sqrt{3/N}
\]

Comment: in large-\( N \) limit, the probability distribution becomes sharply peaked. This will be important for understanding systems with many particles: the behavior is very well described by the mean in the "thermodynamic limit" \( (N \to \infty) \).

Two other commonly encountered probability distributions arise as limiting cases of the binomial distribution. The limit \( p \ll 1 \) yields the Poisson distribution (the "law of rare events"), which you will derive in the homework. We will next consider the large-\( N \) limit, where the size of each individual step is small (i.e., ant-like) compared to the total length of the walk. This limit will lead us to the Gaussian distribution, and to a profound theorem that explains why we encounter the Gaussian in so many different contexts.

For next time: think of other phenomena that can be modeled as random walks...
* Reminder to submit #1 from Wed's worksheet (4A) on gradescope

* How many have started problem set?
  - I fixed one missing part yesterday... make sure you have latest version

---

Last time

Fund. thermo. relation: \[ dE = -pdV + \tau d\sigma \]

\[ \Rightarrow \quad df = -pdV - \sigma dt \]

* Let's us determine \( p, \sigma \) from derivatives of free energy

* \( F = -\ln Z \) is easy to calculate if we can write down partition function

Today: derive ideal gas law & understand what we mean by "ideal"
EXERCISE 4B: STATISTICAL DERIVATION OF THE IDEAL GAS LAW

Objectives:

- Derive the ideal gas law from first principles (quantum and statistical mechanics)
- Calculate the energy and entropy of an ideal (non-interacting) gas

Reading: Kittel & Kroemer, Ch. 3

Useful past results:

- \( dF = -pdV - \sigma d\tau \)
- \( F = -\tau \ln Z \)

1. **Partition function of an ideal gas.** Consider a gas of \( N \) non-interacting particles in a cubic box of volume \( V = L^3 \). Recall that the energy eigenstates of a single particle in a three-dimensional box are given by

   \[
   \varepsilon(n_x, n_y, n_z) = \frac{\pi^2 \hbar^2}{2mL^2} \left( n_x^2 + n_y^2 + n_z^2 \right). \tag{1}
   \]

   a. Let \( z_1 \) denote the partition function of a single particle in a one-dimensional box of length \( L \) (which we will calculate below). Find an expression for the partition function of \( N \) particles in a box of dimensions \( L \times L \times L \) in terms of \( z_1 \):
   
   i. ...assuming that the particles are *distinguishable.*

   \[
   Z = z_1^{3N}.
   \]
   
   Check: \( Z = \sum_{n=0}^{\infty} \sum_{n=0}^{\infty} \sum_{n=0}^{\infty} e^{-\alpha \varepsilon(n_x, n_y, n_z)} = \sum_{n=0}^{\infty} e^{-\alpha \varepsilon(n_x, n_y, n_z)} \)

   \[
   = \left( \sum_{n=0}^{\infty} e^{-\alpha \varepsilon(n_x, n_y, n_z)} \right)^3 \]

   \[
   = z_1^{3N} \quad \text{where} \quad z_1 = \sum_{n=0}^{\infty} e^{-\alpha \varepsilon(n_x, n_y, n_z)}
   \]

   ii. ...assuming that the particles are *indistinguishable.*

   We have overcounted by a factor \( N! \), i.e., each configuration of indistinguishable particles appeared \( N! \) times in the sum in (1).

   \[
   \Rightarrow Z = z_1^{3N} / N!
   \]

   Assumption: terms with 2 particles in same orbital (\( n_x, n_y, n_z \)) are negligible.

   Should be valid except at very low temperature...
b. Write out the partition function \( z_1 \) for a single particle in a 1D box and simplify it by approximating the sum as an integral.

\[
\begin{align*}
z_1 &= \sum_{n=1}^{\infty} e^{-\beta n^2} \\ &\approx \frac{1}{2} \int_{-\infty}^{\infty} e^{-\beta x^2} \, dx \\ &= \frac{1}{2} \sqrt{\frac{\pi}{\alpha \beta}} = \frac{1}{2} \sqrt{\frac{\pi \tau^2 \cdot 2m}{\hbar^2}}
\end{align*}
\]

\[= L \cdot \frac{\sqrt{2\pi m \tau}}{\hbar} \]

Used: \[ X = \int_{-\infty}^{\infty} e^{-A x^2} \, dx = \frac{\sqrt{2\pi}}{\sqrt{A}} e^{-A x^2} \bigg|_0^\infty = \sqrt{\pi / A} \]

\[\Rightarrow X = \sqrt{\pi / A} \]

c. The partition function \( z_1 \) can be expressed as \( z_1 = L / \lambda_T \), where \( \lambda_T \) is a temperature-dependent length-scale. What is the value of \( \lambda_T \)? Can you give a physical interpretation for \( \lambda_T \)?

\[ z_1 = \frac{L}{\lambda_T} \quad \text{where} \quad \lambda_T = \frac{\hbar}{\sqrt{2m \tau}} \]

This is the de Broglie wavelength of particles with momentum \( \sqrt{2m \tau} \), a characteristic momentum scale at temperature \( \tau \).

d. Based on parts a-c., write down the full partition function for the gas of \( N \) indistinguishable, non-interacting particles in terms of \( L, \lambda_T, \) and \( N \).

\[
z = \frac{z_1^{3N}}{N!} = \left( \frac{L}{\lambda_T} \right)^{3N} / N!
\]
2. Quantum mechanical derivation of the ideal gas law.

a. Calculate the Helmholtz free energy \( F \) of the ideal gas of \( N \) particles in a three-dimensional box of volume \( V \).

\[
F = -T \ln Z = -3N \ln \left( \frac{L}{\sqrt{2\pi}} \right) + N \ln N - NT,
\]

where we have used Stirling's approximation:

\[\ln N! \approx N \ln N - N\]

\[\Rightarrow F = -NT \ln \left( \frac{V}{2\pi} \right) + N \ln N - NT \] (in terms of \( V \))

b. Calculate the pressure \( p(N, T, V) \).

\[
dF = -p \, dV + \tau \, dt \Rightarrow p = -\frac{\partial F}{\partial V} = \frac{N \tau}{V}
\]

\[
\left[ p = -\frac{\partial F}{\partial V} \right] = \frac{2}{\sqrt{2\pi}} \left[ N \tau \ln V + \text{const} \right]
\]

\[\Rightarrow \boxed{pV = N\tau}\]

c. While we are at it, let's also calculate the energy \( E \) of the ideal gas from the partition function \( Z \).

\[
E = -\frac{\partial \ln Z}{\partial \beta} = -\frac{2}{\sqrt{2\pi}} \left[ 3N \ln \left( \frac{L}{\sqrt{2\pi}} \right) - N \ln N + N \right]
\]

\[
= -\frac{2}{\sqrt{2\pi}} \left[ 3N \ln \left( \frac{V}{2\pi} \right) + \text{const} \right]
\]

\[= -\frac{2}{\sqrt{2\pi}} \left[ \frac{3}{2} N \ln \tau + \text{const} \right] = +\frac{2}{\sqrt{2\pi}} \left( \frac{3}{2} N \ln \beta \right)
\]

\[\boxed{E = \frac{3}{2} \frac{N}{\beta} = \frac{3}{2} N\tau}\]

Comment: equipartition theorem, \( \frac{1}{2} k \) of kinetic energy for each translational d.o.f. We will come back to proving and seeing conditions for validity.
3. Entropy of the ideal gas.

a. Express the entropy of a generic system at known particle number $N$, temperature $\tau$, and volume $V$ as a derivative of the Helmholtz free energy $F(N, \tau, V)$.

\[
\frac{dF}{d\tau} = -\tau \frac{\partial F}{\partial \tau} - \sigma \frac{\partial F}{\partial \sigma}
\]

\[
\Rightarrow \sigma = -\left(\frac{\partial F}{\partial \tau}\right)_V
\]

b. Calculate the entropy of a gas of $N$ non-interacting identical particles. (You already wrote down the Helmholtz free energy $F$ in problem 2.)

Express your results in terms of the \textbf{thermal de Broglie wavelength}

\[
\lambda_\tau = \frac{\hbar}{\sqrt{2\pi m\tau}}
\]  

(2)

\[
F = -N\tau \ln \left(\frac{V}{\lambda_\tau^3}\right) + N\tau \ln N - N\tau
\]

\[
= -N\tau \left[ -\ln V + \ln N - 1 \right] + N\tau \ln \lambda_\tau^3
\]

\[
= -N\tau \ln \left(\frac{V}{\lambda_\tau^3}\right) + N\ln \lambda_\tau^3 + 3N\tau \frac{\partial \ln \lambda_\tau^3}{\partial \tau}
\]

\[
\left(\frac{\partial F}{\partial \tau}\right)_V = N \left[ \ln (N/\lambda_\tau^3) - 1 \right] + N\ln \lambda_\tau^3 + 3N\tau \frac{\partial \ln \lambda_\tau^3}{\partial \tau}
\]

\[
\left(\frac{\partial F}{\partial \tau}\right)_V = N \ln \left(\frac{N\lambda_\tau^3}{V}\right) - N + 3N\tau \left( \frac{\partial \ln \lambda_\tau^3}{\partial \tau} + \text{const} \right)
\]

\[
= N \ln \left(\frac{N\lambda_\tau^3}{V}\right) - N + 3N\tau \left( \frac{1}{2\tau} \ln \tau + \text{const} \right)
\]

\[
= N \ln \left(\frac{N\lambda_\tau^3}{V}\right) - \frac{3}{2} N\tau + \frac{1}{6} N
\]

\[
\therefore \sigma = -\left(\frac{\partial F}{\partial \tau}\right)_V = N \ln \left(\frac{V}{N\lambda_\tau^3}\right) + \frac{5}{2} N
\]
c. Your result from 3.b. is the Sackur-Tetrode equation for the entropy of a monatomic ideal gas. Subject it to another sanity check: does it violate any laws of thermodynamics? Explain.

Let's check the 3rd law:

As $\tau \to 0$, \( \frac{1}{N^3} \propto 2^{3/2} \to 0 \Rightarrow \ln\left(\frac{V}{N^{3/2}}\right) \to -\infty \)

So $\tau \to N(-\infty) + \frac{5}{2} N \to -\infty$

This is inconsistent with the 3rd law and even with our definition of entropy, which is always positive.

d. In what parameter regime must the Sackur-Tetrode equation (3.b.) break down? Why does it break down?

The Sackur-Tetrode equation must break down for

\[ \frac{V}{N^{3/2}} \leq 1 \quad \text{where} \quad N \ln\left(\frac{V}{N^{3/2}}\right) < 0 \]

We can think of \( \frac{V}{N^{3/2}} \) as the smallest volume within which a particle can be localized. (Defining a smaller wave-packet would require a larger spread of momenta than compatible with temperature \( T \).)

Thus, for \( V \leq N^{3/2} \), we start to have multiple particles in the same microstate. Here, our approximation \( Z \approx \frac{2^N N!}{N^N} \) breaks down.

We will come back to this regime later in the term.

e. In summary: what conditions must be satisfied for a gas to be considered ideal?

* Classical regime: low density \( \frac{N^{3/2}}{V} \ll 1 \) (or high temperature)

* Non-interacting \( \Rightarrow \) also a good approximation at sufficiently low density
f. Estimate the thermal de Broglie wavelength for...
   i. ...nitrogen gas (28 amu) at room temperature. Compare your result with the typical intermolecular spacing at 1 atmosphere of pressure.

\[ \lambda_T = \frac{\hbar}{\sqrt{2\pi mk_bT}} \]

\[ h = 6.63 \times 10^{-34} \text{ J s}, \quad k_b = 1.38 \times 10^{-23} \text{ J/K}, \quad m = 28 \times 1.67 \times 10^{-27} \text{ kg} \]

\( \Rightarrow \lambda_T = 19 \text{ pm} \)

Compare: \( pV = Nt \Rightarrow \frac{V}{N} = \frac{k_bT}{\rho} \Rightarrow (V/N)^{1/3} \approx 3 \text{ nm} \)

So \( \lambda_T < \text{ intermolecular spacing} \Rightarrow \text{ classical gas} \).

ii. ...atomic sodium (23 amu) at a temperature of 100 nK.

\( \lambda_T \approx 1 \text{ nm} \)

\( \therefore \) can be in quantum regime even for a comparatively dilute gas (under vacuum — collected from vapor pressure & laser-cooled...).

iii. ...an electron at room temperature. Compare your result with the lattice constant of a typical metal.

\[ \lambda_T = 4.3 \text{ nm} \at \ T = 300 \text{ K} \] (using \( m_e = 9.1 \times 10^{-31} \text{ kg} \))

Well actually using Wolfram alpha...

For comparison, copper has a lattice constant of \( d = 0.4 \text{ nm} \approx \lambda_e \)

\( \Rightarrow \) Electrons at room temperature can be viewed as a quantum gas. Crucial foundation for solid-state physics.

We’ll come back to this later in the term.
Course: Physics 152/252 — Introduction to Particle Physics

Instructor: Lauren Tompkins

Target Audience: Junior/seniors and first year graduate students

Brief summary of examples:

This lecture shows a wrap up of the topic of neutrino oscillations as well as the start of learning about isospin. I chose it because a) it was my most recent class and b) it shows how we work in current experimental data into the course.

Notes are in black/red
Student exercises are in blue
Solutions (which the students do not see) are in cyan
Wrapping up Neutrinos

Reminder from last class: the neutrino mixing equations for two-neutrino equation are:

\[ P_{\nu_\mu \to \nu_e} = \left\{ \sin(2\theta) \sin\left[ \frac{1.27\Delta m_{23}^2 L}{E} \right] \right\}^2 \] (1)

for atmospheric neutrinos and for solar neutrinos:

\[ P_{\nu_e \to \nu_\mu} = \left\{ \sin(2\theta) \sin\left[ \frac{1.27\Delta m_{12}^2 L}{E} \right] \right\}^2 \] (2)

where \( E \) is in MeV, \( \Delta m^2_{32(12)} = m^2_{2(1)} - m^2_{3(2)} \) is in eV/\( c^2 \), and \( L \) is in m. Experimentally it has been observed that

\[
\begin{align*}
\sin^2 2\theta_{\text{atm}} & \approx 0.97 \\
\sin^2 2\theta_{\text{solar}} & \approx 0.86 \\
\sin^2 2\theta_{13} & \approx 0.1 \\
\Delta m_{12}^2 & \approx 7.59 \times 10^{-5} \text{eV}^2 \\
\Delta m_{23}^2 & \approx 2.32 \times 10^{-3} \text{eV}^2
\end{align*}
\]

where \( \theta_{\text{atm}} = \theta_{23} \) and \( \theta_{\text{solar}} = \theta_{12} \).

The Kamland experiment in Japan used reactor neutrinos (an \( \nu_e \) source!) to test the same parameter space as the solar neutrinos. Neutrinos from nuclear reactors have a characteristic energy of \( \lesssim 8 \text{ MeV} \).

Given that they are targeting a solar neutrino mass splitting, and assuming a monochromatic energy spectrum of neutrinos, how far away do they need to be from the detector to make the best measurement of the oscillation parameters?

The oscillation is at a maximum at \( \frac{1.27\Delta m^2 L}{E} = \frac{\pi}{2} \). For \( \Delta m \approx 8 \times 10^{-5} \text{ eV}^2 \) and \( E = 8 \text{ MeV} \), \( L \) is optimal at \( \approx 1000\text{km} \). Due to the physical constraints of the reactor and detector, the average distance was actually more like 100-200km!

From the mass and mixing parameters above, qualitatively discuss the following plots. Relate them to the neutrino mixing fractions we had from last class. Then we’ll discuss a summary of all of the neutrino data from 2016:
Figure 1: Kamland data from http://demonstrations.wolfram.com/NeutrinoOscillations/.
Quick Recap of the course

- 3 fundamental interactions of the Standard Model
- QED: A single massless photon interacts with all electrically charged particles; the only effectively long-range force (other than gravity).
- Weak Interactions: 3 massive bosons interact with all particles (except gluons); allows flavor changing.
- QCD: 8 massless gluons interact with all color charged particles (quarks and gluons); allows color changing; binds quarks into a zoo of hadrons.
- Three generations of quarks and leptons, each have “mixing” due to mismatch between weak and mass eigenstates.

Where we are going

- We’ll learn how to add angular momentum in composite states.
- The symmetry of angular momentum will also apply to the flavor symmetry of isospin.
- We can generalize this type of symmetry to apply to forming mesons and baryons and even to the three fundamental interactions themselves.
- We’ll see how a local SU(3)×SU(2)×U(1) gauge symmetry describes these three interactions in a more unified way.
- That symmetry determines the number of gluons (8), weak bosons (3), and photons (1).
- The Higgs field plays a vital role in giving fermions and the W and Z bosons their masses.
- We’ll hear a survey of interesting and relevant topics presented by you!

Learning goals

- Refresh ourselves with addition of angular momentum.
- Learn (or re-learn) to use the Clebsch-Gordan coefficients.
- Determine whether angular momentum is conserved in a reaction.
- Learn how spinors are represented in SU(2)
Review of angular momentum

Let’s remind ourselves of why we care about angular momentum in particle reactions. In the “original” beta-decay reaction, \( n \rightarrow p + e^- \), the electron was expected to have a fixed energy in every decay. Conservation of four-momentum (set \( c = 1 \))

\[
(m_n, 0) = (m_p, p_p) + (m_e, p_e)
\]

gives us two equations (energy and three-momentum) and two unknowns (the momentum for each daughter particle). This yields a fixed energy or momentum for each particle, but the observed spectrum always looks like Fig. 3.

![Energy spectrum of beta decay electrons from Bi-210](https://github.com/tdwiser/physhandout)

**Figure 3**: Bi-210 has a decay energy (or Q value) of 1.426 MeV.

There must be an antineutrino to carry away some of that extra energy. What about angular momentum?

*Although you can just look it up in your particle table, what possible values of spin could you assign the antineutrino from this experiment?*

Combining the spins of the proton and electron can get us \( s = 1 \) or \( s = 0 \) but not \( s = \frac{1}{2} \) of the neutron. Adding orbital angular momentum to the system could only achieve \( s \geq 1 \).

If we add the neutrino with \( s = \frac{1}{2} \) to the electron and proton, we can get \( s = \frac{1}{2} \) or \( \frac{3}{2} \). The neutrino could also be \( s = \frac{3}{2} \) to yield \( \frac{5}{2}, \frac{3}{2}, \) or \( \frac{1}{2} \).

With angular momentum added to the final system, we could add or subtract any multiple of \( \hbar \) so that *any* half-integer spin would work.
Let’s go through one important example from the book: if we have three quarks bound with no orbital angular momentum, what are the possible spins of the baryons?

First we combine two of the quarks to get a total spin state of 0 or 1, then we can add a third quark to this, so we are adding $(0,1)\times(1/2)$. From this we either get $1/2$ or $3/2$.

The generic angular momentum $\mathbf{J}$ could be orbital ($\mathbf{L}$), spin ($\mathbf{S}$), or some combined quantity. $j$ is the quantum number that parametrizes the eigenvalues for the magnitude of generic angular momentum,

$$J^2 = j(j+1)\hbar^2.$$

Remember that we could replace $J^2$ with $S^2$ or $L^2$ and replace $j$ with $\ell$ or $s$, respectively.

Eq. (4.12) in Griffiths tells us how two particles (or composite particles) with $j_1$ and $j_2$ add to form a new $j$’s.

$$j = |j_1 - j_2|, |j_1 - j_2| + 1, \ldots, (j_1 + j_2) - 1, (j_1 + j_2)$$

**Clebsch-Gordan coefficients**

When two specific states $|j_1m_1\rangle \, |j_2m_2\rangle$ combine to form a composite state $|jm\rangle$, we may want to know the probability of finding the composite state in any one of the allowed $j$’s and $m$’s.

$$|j_1m_1\rangle \, |j_2m_2\rangle = \sum_{j=|j_1 - j_2|}^{(j_1+j_2)} C_{mnm_1m_2}^{j_1j_2} |jm\rangle$$  \hspace{1cm} (3)

$$|jm\rangle = \sum_{m_1m_2} C_{mnm_1m_2}^{j_1j_2} |j_1m_1\rangle \, |j_2m_2\rangle$$  \hspace{1cm} (4)

**How to read the table of Clebsch-Gordan coefficients**

- *Every* coefficient includes an understood square-root sign applied before the minus sign: $-1/2$ really means $-\sqrt{1/2}$.
- First, pick a sub-table based on the $j$’s of the particles you’re adding. If you have a spin-1/2 and spin-1 particle, go straight to the $1 \times 1/2$ table.
- The key at the top-right corner tells you that the coefficients in the shaded region correspond to the state $|jm\rangle$ in a single column. Note that the table uses $J$ and $M$ instead of lower-case letters.
- The rows correspond to states $|m_1\rangle \, |m_2\rangle = |j_1m_1\rangle \, |j_2m_2\rangle$. We dropped the $j$’s because we already know that from the sub-table at which we’re looking, but you may want to put them back in when working with them to know which particle is which. You can tell the rows involve $m$’s because both numbers have $+$ or $-$ signs.
• You can always check yourself with the known-by-heart (right???) coefficients for singlet state of two spin 1/2 particles: $|0, 0\rangle = \frac{1}{\sqrt{2}} |+1/2\rangle - \frac{1}{\sqrt{2}} |-1/2\rangle$.

• Each row and each column sums to 1 when the minus signs are properly squared away. This enforces that each state conserves probability when changing basis.

• Each sub-table is actually a big matrix that is collapsed to hide the many zero entries.

| $m_l$ | $m_s$ | $|j, m\rangle$ |
|-------|-------|----------------|
| 1     | 1/2   | $|3/2, 3/2\rangle$ |
| 1     | -1/2  | $|3/2, 1/2\rangle$ |
| 0     | 1/2   | $|1/2, 1/2\rangle$ |
| 0     | -1/2  | $|3/2, -1/2\rangle$ |
| -1    | 1/2   | $|1/2, -1/2\rangle$ |
| -1    | -1/2  | $|3/2, -3/2\rangle$ |

Figure 4: This is the expanded version of the $1 \times 1/2$ sub-table of Clebsch-Gordan coefficients. Every blank space is a 0.

Suppose you have a spin-$\frac{3}{2}$ particle and a spin-2 particle. If you knew that the orbital angular momentum was zero and that the total spin of the composite system was $\frac{5}{2}$ and its $z$ component was $-\frac{1}{2}$, what values might you get for a measurement of $S_z$ on the spin-2 particle? What is the probability of each? Check that they add up to 1.

From the $2 \times \frac{3}{2}$ table,

$$|\frac{5}{2}, -\frac{1}{2}\rangle = \sqrt{\frac{27}{70}} |2, 1\rangle + \sqrt{\frac{3}{35}} |2, 0\rangle - \sqrt{\frac{5}{14}} |2, -1\rangle - \sqrt{\frac{6}{35}} |2, -2\rangle$$

The LHS is $|jm\rangle$, and the RHS is $|j_1m_1\rangle |j_2m_2\rangle$.

So you could get

$m_1 = 1$, probability 27/70; or
$m_1 = 0$, probability 3/35; or
$m_1 = -1$, probability 5/14; or
$m_1 = -2$, probability 6/35.

Those probabilities do add up to one.

$$27/70 + 6/70 + 25/70 + 12/70 = 1$$

With the same $\frac{3}{2} \times 2$ system, if you know the $z$-components of the two particles are +1 and +1/2, what is the total state in terms of $j$ and $m$? Do the probabilities add to 1?
\[ |+1, +1/2 \rangle = |2, +1 \rangle |3/2, +1/2 \rangle = \sqrt{\frac{4}{7}} |7/2, +3/2 \rangle + \sqrt{\frac{1}{35}} |5/2, +3/2 \rangle - \sqrt{\frac{2}{5}} |3/2, +3/2 \rangle \]

Yes, the probabilities add to one.

More decays

We’ve had lots of practice with determining if decays are allowed or forbidden. Now there is one more conservation law we can add to the list: conservation of angular momentum.

Determine if this reaction is allowed or forbidden. If forbidden, why is it forbidden?

\( \omega^0 \rightarrow \pi^0 \pi^0 \)

Not allowed. \( \omega \) is spin 1, and the pions are spin 0.

You might guess that some orbital angular momentum could pick up the slack, but it’s actually impossible for two identical bosons to have \( \ell = 1 \) by the Pauli exclusion principle. The \( \omega^0 \) is a boson, so it has a symmetric total wavefunction. \( \ell = 1 \) is antisymmetric, and the spatial wavefunction for two \( \pi^0 \)’s is symmetric under interchange. The combination would be antisymmetric.

The two most common decay modes are

\( \omega^0 \rightarrow \pi^+ \pi^- \pi^0 \) and \( \omega^0 \rightarrow \pi^0 \gamma \)

We won’t deal with orbital angular momentum in decays much, but just know it’s there.
Representations of SU(2)

Spin-$\frac{1}{2}$ particles, such as electrons or protons, have two eigenstates $|s, s_z\rangle$ states for the $S_z$ operator.

$$\left| \frac{1}{2}, +\frac{1}{2} \right\rangle = |\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (5)$$

$$\left| \frac{1}{2}, -\frac{1}{2} \right\rangle = |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (6)$$

These two-component complex column matrices are known as spinors. Any arbitrary state can be represented as a linear combination of these two states, which form a complete basis,

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

where $|\alpha|^2 + |\beta|^2 = 1$ for normalization. So far, this is just a different way to represent the states instead of using $|j, j_z\rangle$ as we were doing before.

This is the two-dimensional representation of SU(2), the group of all $2 \times 2$ unitary matrices with determinant 1. The purpose of this is that there is a lot of group theory one could use (we'll only see a little) with this notation.

Remember when we had a rotation about the $z$ axis for a 3-vector, we had

$$R_z = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{or just } R = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad (7)$$

for a two-dimensional vector. Spinors transform in a similar but different way.

$$\begin{pmatrix} \alpha' \\ \beta' \end{pmatrix} = U(\theta) \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad (8)$$

$$U(\theta) \equiv e^{-i(\theta \cdot S)/\hbar} = e^{-i(\theta \cdot \sigma)/2} = \cos \frac{\theta}{2} I - i(\theta \cdot \sigma) \sin \frac{\theta}{2} \quad (9)$$

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (10)$$

Show that a rotation about the $y$ axis with Eqs. (8, 10) is exactly

$$U_y(\theta) = \begin{pmatrix} \cos \theta/2 & -\sin \theta/2 \\ \sin \theta/2 & \cos \theta/2 \end{pmatrix}.$$
\[ \begin{align*} 
\theta &= \hat{y} \\
\sigma &= (\sigma_x, \sigma_y, \sigma_z) \\
\theta \cdot \sigma &= \sigma_y \\
\cos \frac{\theta}{2} I - i(\theta \cdot \sigma) \sin \frac{\theta}{2} &= \cos \frac{\theta}{2} - i \sigma_y \sin \frac{\theta}{2} \\
&= \begin{pmatrix} 
\cos \frac{\theta}{2} & -\sin \frac{\theta}{2} \\
\sin \frac{\theta}{2} & \cos \frac{\theta}{2}
\end{pmatrix}
\end{align*} \]

It looks the same our familiar rotation matrix from Eq. 7 except that the angles are all strangely divided by 2! This reflects our notation where a \( \pi \) rotation in the familiar \( x-y \) plane from high school rotates \( \begin{pmatrix} 1 \\ 0 \end{pmatrix} \) to \( \begin{pmatrix} -1 \\ 0 \end{pmatrix} \), but a \( \pi \) rotation in our spinor space rotates \( \begin{pmatrix} 1 \\ 0 \end{pmatrix} \) to \( \begin{pmatrix} 0 \\ 1 \end{pmatrix} \).

At the end of the day, we just want you to know that SU(2) is very similar to rotations in three-dimensional space (described by the SO(3) group), except that rotating \( \begin{pmatrix} 1 \\ 0 \end{pmatrix} \) by \( 2\pi \) gets \( \begin{pmatrix} -1 \\ 0 \end{pmatrix} \).

For spin-1 particles, you need to go to the three-dimensional representation of SU(2) because you now have three eigenstates.

\[
|1, +1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1, 0\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad |1, -1\rangle = \begin{pmatrix} 0 \\ 0 \end{pmatrix}
\]

The eigenvalues are \( m = \hbar, 0, -\hbar \), respectively. The analog of the Pauli spin matrices \( (\sigma_i) \) are the following:

\[
\lambda_x = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \lambda_y = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}
\]

If you had particles with higher spin, you’d need a higher-dimensional representation of SU(2). I wrote this out to show later that this is not the same as SU(3).

When you deal with SU(2), think of the eigenstates as a lever for some old-fashioned machine. For \( s = \frac{1}{2} \), the lever eigenstates can be either up or down. For \( s = 1 \), the lever has three positions: up, down, and neutral, but it’s all along the same line. As we’ll see later, SU(3) is like a lever that has 3 separate directions in a Y shape.

\(^1\)The main reason I’m belaboring all of this is because I (Paul) have always been confused about this, and I haven’t seen a book that spells it out like this. Before you think we’ve just changed this notation to make things more confusing, remember that experiments have verified that this notation is essential because the phase of the wave function really does work this way.
Course: Physics 14 — Freshman Seminar on Quantum Information

Instructor: Monika Schleier-Smith

Target Audience: Freshmen

Brief summary of examples:

Quantum Key Distribution: lecture 3 (week 2) of Physics 14N

Note — this is a class for freshman and not necessarily physics majors.
Last time we introduced polarized photons as model qubits:

\[ \hat{\mathcal{E}} = \epsilon_x \hat{\mathbf{X}} + \epsilon_y \hat{\mathbf{Y}} \quad \Rightarrow \quad |\epsilon\rangle = \epsilon_x |\leftrightarrow\rangle + \epsilon_y |\uparrow\rangle \]

\[ \sim -\epsilon_x |\leftrightarrow\rangle - \epsilon_y |\uparrow\rangle \quad \text{same} \]

\[ \sim -\epsilon_x |\leftrightarrow\rangle + \epsilon_y |\uparrow\rangle \quad \text{different!} \]

More generally,

\[ e^{i\varphi} (|0\rangle + e^{i\varphi} |1\rangle) = |0\rangle + e^{i\varphi} |1\rangle \]

Same physical state

But \(|0\rangle + e^{i\varphi} |1\rangle\) is not the same.

Next time: how to visualize generic qubit states

Today: application of what we've learned so far — cryptography!

\[ \Rightarrow \text{States} \]

Recall: \(|\uparrow\rangle = \frac{1}{\sqrt{2}} (|1\rangle + |\leftrightarrow\rangle) \]

[Called \(1\rangle\) in book.]

"Hadamard basis?"
EXERCISE 2A: QUANTUM KEY DISTRIBUTION

Objective:

- Understand the BB84 protocol for quantum key distribution.
- Simulate the protocol in a simple optical setup of lasers, wave plates, and polarizers.

Many cryptographic protocols, both classical and quantum, involve a secret key that is shared between two people (Alice and Bob). This allows them to encode/decode messages sent in a public way without anyone else being able to decipher them. But how do Alice and Bob securely agree on a secret key when they are in different locations? Quantum mechanics provides a way to do this, such that Alice and Bob are able to guarantee that no one has intercepted their secret key. The use of quantum mechanics to securely transmit information is a called quantum cryptography, and we'll begin to investigate it with this worksheet.

1. Transmitting information with photons. Alice and Bob have two methods of communication: they can communicate classically (by text messages) and they can communicate quantumly (Alice can send Bob single photons over an optical fiber). It is possible for other people to hack into both of these methods of communication. They agree that Alice will encode a string of '0's and '1's in the polarization state of photons. If they can verify no one intercepted this string, they will use it as their secret key. They agree to use the following mapping:

\[
\begin{align*}
|\leftrightarrow\rangle & \mapsto 1 \\
|\uparrow\downarrow\rangle & \mapsto 0.
\end{align*}
\]  

(1)

Alice will send a photon with its polarization oriented either horizontally or vertically. She will choose the polarization randomly, with equal probability for either. She will also text Bob telling him that she sent a photon, without telling which polarization she sent. Bob will measure the photon and text her back to confirm that he received it, without telling her anything about how he measured the photon or the measurement outcome.

a. How can Bob measure the polarization of the photon that Alice sends? List two ways and explain why one would be better than the other.

- Polarizing filter - not good because he can't distinguish whether photon was absorbed by filter or lost inside (in the fiber)

- Polarizing beam splitter

\[ \xrightarrow{\text{detector (\uparrow)}} \xrightarrow{\text{detector (\leftrightarrow)}} \]
b. If Bob measures a vertically polarized photon, with what percent certainty does he know which polarization Alice sent?

100% (assuming a good polarizing beam splitter)

* Note that photon polarization can be maintained in an optical fiber, even if the fiber twists...

c. Now suppose there’s an eavesdropper named Eve. She is able to intercept the photon Alice sends to Bob, detect it, and send Bob a replacement photon. If she overhears all of the discussions Bob and Alice have about their plan to share the secret key, is it possible for her to know the polarization state of every photon Alice sends with 100% certainty?

Yes: she can measure in the $\mathbb{I} \leftrightarrow \sigma_\perp$ basis, in which case she is certain to detect the polarization that Alice sent.

d. Assuming that Eve is able to send the replacement photon instantaneously, would Alice and Bob know that she has eavesdropped? Why or why not?

* No — the replacement photon is identical to the original photon.

* I mentioned "no cloning" theorem, but this only is relevant to superpositions — we can't clone an arbitrary superposition $c_1 \left| \psi \right\rangle + c_2 \left| \phi \right\rangle$.

* Until we deal with superpositions, everything we are doing is classical.
e. Explain why it is not secure for Alice to tell Bob in which basis she is sending photons.

See (d). If Eve overhears the measurement basis, she can eavesdrop & go unnoticed.

2. Transmitting a secure secret key: BB84 protocol. The protocol in the previous problem is not secure; however there are various secure protocols to generate secret keys. One of the most well known is called the BB84 protocol, discovered in 1984 by C. Bennett and G. Brassard. In the BB84 protocol, Alice and Bob still have the same two ways of communicating (text messages and single photons), but Alice now has an additional choice of sending light polarized horizontally, vertically, or with one of the two diagonal polarizations. They agree to use the same mapping as in Equation 1 for horizontally or vertically polarized photons, and the following mapping for the diagonally polarized photons:

\[ |\uparrow\rangle \mapsto 0 \]
\[ |\downarrow\rangle \mapsto 1. \]  

Alice now randomly sends any of the four polarization states with equal probability. Bob doesn’t know which basis Alice used, so he randomly chooses a basis (horizontal/vertical or diagonal) in which to measure the photon.

a. Suppose Alice sends a photon in the state \( |\uparrow\rangle \) and Bob measures in the horizontal/vertical basis.

i. What digit ('0' or '1') does Alice’s photon represent?

ii. With what probability does Bob detect the same digit that Alice sent?

i) 0

\[ 50\% \text{ probability since} \]

\[ |\uparrow\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \]
b. For the transmission of Alice’s bit to be successful, Alice and Bob must agree on the bit value and know that they agree. What information can Bob publicly send Alice after his measurement to establish whether the transmission was a success?

If they agree, then they have a shared secret that nobody else knows.

Alice & Bob announce their transmission & measurement bases.

Would it be equivalent for Bob to send the same information before he performs his measurement? Why or why not?

If Bob announced his basis in advance, Eve could overhear and measure in the same basis, then send copy of outcome to Bob. Bob wouldn’t know the difference, so this would be problematic. So, not equivalent.

d. What percentage of the bits that Bob receives will be useful for forming the shared key? Explain.

50% — all the ones where Alice & Bob use the same basis.

e. Fill in the empty spaces in the following table. What is the secret key that Alice and Bob generate?

<table>
<thead>
<tr>
<th>Alice’s bit</th>
<th>1 1 0 1 0 0 0 1 0</th>
<th>Secret key is 1101</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alice’s basis</td>
<td>+ x x + + x + x</td>
<td></td>
</tr>
<tr>
<td>Alice’s polarization</td>
<td>← ↑ ↑</td>
<td>← ↑ ↑</td>
</tr>
<tr>
<td>Bob’s basis</td>
<td>x x + + x x</td>
<td></td>
</tr>
<tr>
<td>Bob’s measurement</td>
<td>← ↑ ↑</td>
<td>← ↑ ↑</td>
</tr>
<tr>
<td>Same basis?</td>
<td>No Yes N Y N Y Y N</td>
<td></td>
</tr>
<tr>
<td>Shared secret key</td>
<td>1 1 0</td>
<td></td>
</tr>
</tbody>
</table>
3. *An eavesdropper.* Suppose that Eve intercepts Alice’s photons and measures them just as Bob does, randomly picking one of the two bases each time. She then sends a replacement photon—polarized in the same way as her measurement result—to Bob.

a. What percentage of the time will Eve be 100% certain of the polarization of the photon Alice sent? Explain.

* 0% of the time will she be certain if she does not know Alice’s basis.

* However, if she *listens* over the classical channel to find out Alice’s basis after the fact, she will have measured in the correct basis 50% of the time.

So, Eve can be 100% certain of Alice’s polarization 50% of the time.

b. Fill in the spaces in the table below for the case of an eavesdropper. Fill in the last row (“same bit?”) only for the columns in which Alice & Bob used the same basis.

<table>
<thead>
<tr>
<th>Alice’s bit</th>
<th>1</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>0</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alice’s basis</td>
<td>+</td>
<td>x</td>
<td>x</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Alice’s polarization</td>
<td>↔</td>
<td>→</td>
<td>→</td>
<td>↔</td>
<td>↓</td>
<td>$\rightarrow$</td>
<td>$\rightarrow$</td>
<td>$\rightarrow$</td>
</tr>
<tr>
<td>Eve’s basis</td>
<td>+</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Eve’s measurement</td>
<td>↔</td>
<td>↓</td>
<td>Δ</td>
<td>$\rightarrow$</td>
<td>↓</td>
<td>$\rightarrow$</td>
<td>↔</td>
<td>↔</td>
</tr>
<tr>
<td>Bob’s basis</td>
<td>x</td>
<td>x</td>
<td>+</td>
<td>+</td>
<td>x</td>
<td>x</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Bob’s measurement</td>
<td>$\rightarrow$</td>
<td>y</td>
<td>y</td>
<td>↔</td>
<td>θ</td>
<td>↓</td>
<td>θ</td>
<td>θ</td>
</tr>
</tbody>
</table>

| Alice/Bob same basis? | N | Y | N | Y | N | Y | Y | N |
| Alice/Bob same bit? | Y | N | Y | Y |

c. Suppose that Alice and Bob try to construct a shared key with Eve eavesdropping as in part b. Of the bits that they keep, what percentage will be in error due to Eve’s presence?

* No error if Eve chose same basis as Alice & Bob.

* 25% will be in error

* 50% chance of error if Eve chose opposite basis.

Page 5
d. Can Alice and Bob use a classical communication channel to determine whether there is an eavesdropper, if they are willing to sacrifice a few of their shared bits? Explain.

   Yes. Even though Eve won’t affect all bits if they share enough bits they will notice the errors induced by Eve.

e. How many bits do Alice and Bob need to compare if they want to be 99% sure that there is no eavesdropper?

\[
\text{After } N \text{ bits with Eve eavesdropping, chance of no error is}
\]

\[
\left(\frac{3}{4}\right)^N = 0.01
\]

\[
N = \frac{\log(0.01)}{\log(3/4)} = \frac{\log(100)}{\log(4/3)} \approx 16
\]

f. Summarize the properties of quantum states and measurements that ensure that Alice and Bob know if somebody eavesdropped.

   \* Quantum states cannot be cloned (except special case of basis state in known basis = effectively classical).

   \* A measurement affects the quantum state, projecting it into one of the basis states.

We stopped here.

g. Does this protocol guarantee that Alice and Bob can establish a secure key? Why or why not?

   \* The protocol guarantees that Alice & Bob will know if the security of their quantum channel is compromised. It does not give them a way of circumventing Eve if she is there. \(\Rightarrow\) Need privacy amplification.

   \* Also susceptible to “woman-in-the-middle” attack where Eve impersonates Alice & Bob to one another...