I. Collapse

Recall: 1. In general, states change via the Schrödinger dynamics:

\[ |\psi(t_1)\rangle \xrightarrow{\text{Schrödinger evolution}} |\psi(t_2)\rangle \]

2. When a measurement occurs, states change via the Projection Postulate:

When a measurement of a macroscopic system interacts with a microscopic system, states change via the Projection Postulate:

\[ |\psi\rangle \xrightarrow{\text{collapse}} |b_i\rangle \]

Problem: What is a measurement? When is the Projection Postulate supposed to take over from the Schrödinger dynamics?

Possible Responses:

1. When a conscious observer looks at a measuring device.

   Consequence: Dualism -- 2 fundamentally different types of physical system.

   A. Purely physical systems: Always evolve via Schrödinger dynamics.

   B. Conscious systems: Interact with physical systems in certain situations to cause collapse via Projection Postulate.

2. When a macroscopic system interacts with a microscopic system.

   Consequence: Dualism again -- 2 fundamentally different types of physical system.

   A’. Microscopic systems: Always evolve via Schrödinger dynamics.

   B’. Macroscopic systems: Interact with microscopic systems in certain situations to cause collapse via Projection Postulate.

Not much progress!
Can experiments determine *when* a collapse occurs?
In principle, **Yes!** But in practice, **No!**
Let’s see why:

![Diagram](image)

**initial \( t_0 \) state:**
\[
\frac{1}{\sqrt{2}} \lvert \text{ready} \rangle_m \lvert \text{hard} \rangle_e + \frac{1}{\sqrt{2}} \lvert \text{ready} \rangle_m \lvert \text{soft} \rangle_e
\]

**\( t_1 \) state:**
\[
\frac{1}{\sqrt{2}} \lvert \text{“hard”} \rangle_m \lvert \text{hard} \rangle_e + \frac{1}{\sqrt{2}} \lvert \text{“soft”} \rangle_m \lvert \text{soft} \rangle_e
\]

**Theory 1:** Collapse occurs at \( t_1 \).

- **Prediction 1:** At \( t_1 \), state will be:
  - **either** \( \frac{1}{\sqrt{2}} \lvert \text{“hard”} \rangle_m \lvert \text{hard} \rangle_e \) with prob = \( \frac{1}{2} \)
  - **or** \( \frac{1}{\sqrt{2}} \lvert \text{“soft”} \rangle_m \lvert \text{soft} \rangle_e \) with prob = \( \frac{1}{2} \)

**Theory 2:** Collapse occurs sometime after \( t_1 \).

- **Prediction 2:** At \( t_1 \), state will be:
  \[
  \frac{1}{\sqrt{2}} \lvert \text{“hard”} \rangle_m \lvert \text{hard} \rangle_e + \frac{1}{\sqrt{2}} \lvert \text{“soft”} \rangle_m \lvert \text{soft} \rangle_e
  \]

Two theories, two different predictions about when collapse occurs. Are their predictions *experimentally distinguishable?*

**Note:** Prediction 2 is an entangled state, while Prediction 1 is one or the other of the components of the entangled state (Prediction 1, in either form, is a separable state).

**Recall:** The entangled “2-particle” state of measuring device + electron can have a well-defined 2-particle property when neither the measuring device nor the electron separately have well-defined properties (remember the example of the positions of a 2-particle entangled state from Chap 2). If this 2-particle property is one that neither of the separable states of Prediction 2 possess, this would let us distinguish between Prediction 1 and Prediction 2. Let’s find such a 2-particle property.
We have *Hardness* $H^e$ and *Color* $C^e$ operators for the electron (see Chap 2):

\[
H^e |\text{hard}\rangle_e = +1 |\text{hard}\rangle_e \\
H^e |\text{soft}\rangle_e = -1 |\text{soft}\rangle_e \\
C^e |\text{black}\rangle_e = +1 |\text{black}\rangle_e \\
C^e |\text{white}\rangle_e = -1 |\text{white}\rangle_e \\
\]

\[
|\text{black}\rangle_e = \frac{1}{\sqrt{2}} |\text{hard}\rangle_e + \frac{1}{\sqrt{2}} |\text{soft}\rangle_e \\
|\text{white}\rangle_e = \frac{1}{\sqrt{2}} |\text{hard}\rangle_e - \frac{1}{\sqrt{2}} |\text{soft}\rangle_e \\
|\text{hard}\rangle_e = \frac{1}{\sqrt{2}} |\text{black}\rangle_e + \frac{1}{\sqrt{2}} |\text{white}\rangle_e \\
|\text{soft}\rangle_e = \frac{1}{\sqrt{2}} |\text{black}\rangle_e - \frac{1}{\sqrt{2}} |\text{white}\rangle_e \\
\]

Let’s define operators that represent pointer properties of the measuring device. We want a “*Hardness*” pointer operator, call it $H^m$, and a “*Color*” pointer operator, call it $C^m$ (Albert calls it *zip*). Consider:

\[
|\text{ready}\rangle_m = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\
|\text{“hard”}\rangle_m = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\
|\text{“black”}\rangle_m = \begin{pmatrix} 0 \\ \sqrt{2} \\ \sqrt{2} \end{pmatrix} \\
|\text{“white”}\rangle_m = \begin{pmatrix} 0 \\ \sqrt{2} \\ -\sqrt{2} \end{pmatrix} \\

H^m = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \\
C^m = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\
\]

**SO:**

\[
H^m |\text{ready}\rangle_m = +1 |\text{ready}\rangle_m \\
H^m |\text{“hard”}\rangle_m = +1 |\text{“hard”}\rangle_m \\
H^m |\text{“soft”}\rangle_m = -1 |\text{“soft”}\rangle_m \\
C^m |\text{ready}\rangle_m = +1 |\text{ready}\rangle_m \\
C^m |\text{“black”}\rangle_m = +1 |\text{“black”}\rangle_m \\
C^m |\text{“white”}\rangle_m = -1 |\text{“white”}\rangle_m \\
\]

\[
|\text{“black”}\rangle_m = \frac{1}{\sqrt{2}} |\text{“hard”}\rangle_m + \frac{1}{\sqrt{2}} |\text{“soft”}\rangle_m \\
|\text{“white”}\rangle_m = \frac{1}{\sqrt{2}} |\text{“hard”}\rangle_m - \frac{1}{\sqrt{2}} |\text{“soft”}\rangle_m \\
|\text{“hard”}\rangle_m = \frac{1}{\sqrt{2}} |\text{“black”}\rangle_m + \frac{1}{\sqrt{2}} |\text{“white”}\rangle_m \\
|\text{“soft”}\rangle_m = \frac{1}{\sqrt{2}} |\text{“black”}\rangle_m - \frac{1}{\sqrt{2}} |\text{“white”}\rangle_m \\
\]

We can interpret this operator in this way, even though our measuring device doesn’t have *Color* settings. (That’s why Albert calls it “*zip*”.) We can imagine another measuring device that *does* have *Color* settings (just as we can imagine an electron that has *Hardness* and another that has *Color*).
Now call Theory 1’s prediction $|T1\rangle$, and call Theory 2’s prediction $|T2\rangle$:

$$
|T1\rangle = \begin{cases} 
|\text{“hard”}\rangle_m |\text{hard}\rangle_e & \text{(prob} = \frac{1}{2}) \\
|\text{“soft”}\rangle_m |\text{soft}\rangle_e & \text{(prob} = \frac{1}{2})
\end{cases}
$$

$$
|T2\rangle = \sqrt{\frac{1}{2}} |\text{“hard”}\rangle_m |\text{hard}\rangle_e + \sqrt{\frac{1}{2}} |\text{“soft”}\rangle_m |\text{soft}\rangle_e
$$

- If we only want to measure $m$’s “Color” pointer-reading, we must use the 2-particle operator $C^m \otimes I^e$ (where $I^e$ is the identity operator on the electron’s state space).

- If we only want to measure $e$’s Color, we must use the 2-particle operator $I^m \otimes C^e$ (where $I^m$ is the identity operator on $m$’s state space).

- So the difference (“Color” − Color) is a property of the 2-particle system represented by the 2-particle operator $(C^m \otimes I^e) - (I^m \otimes C^e)$. (Albert calls this property zip − color.)

**Claim:** $|T2\rangle$ is an eigenstate of $(C^m \otimes I^e) - (I^m \otimes C^e)$ with eigenvalue 0, but $|T1\rangle$ in either form is not!

**ASIDE:** Two questions concerning zero eigenvalues:

1. **What is the physical significance of a zero eigenvalue?** It just indicates that the system has the value 0 of the property in question. For instance, if the property in question is angular momentum, then when the system is in a state represented by an eigenvector of the operator representing angular momentum with eigenvalue 0, the system has zero angular momentum. Note that this is different from saying the system has no definite value of angular momentum. Zero is a definite value! (So we can’t say, for instance, "A black electron has zero as its value for Hardness", which would mean that it is capable of possessing a value of Hardness. So not possessing a property and possessing the value zero of that property are, conceptually, distinct.)

2. **What is the mathematical significance of a zero eigenvalue?** Let $A$ be an operator, and $|x\rangle$ be an eigenvector of $A$ with eigenvalue 0. This means $A|x\rangle = 0|x\rangle = |0\rangle$, where $|0\rangle$ is the zero vector. Note that $|x\rangle$ can, in general, have non-zero components. Here’s one possible matrix representation of $A$ and $|x\rangle$:

$$
A = \begin{pmatrix} 6 & 3 \\
-2 & -1 \end{pmatrix} \quad |x\rangle = \begin{pmatrix} 1 \\
-2 \end{pmatrix}
$$

**SO:** To see which prediction is correct, at time $t_1$ measure the property (“Color” − Color). If Theory 2 is correct, each time you conduct this experiment, the value will always be 0. If Theory 1 is correct, after a number of experiments, you will get a range of values, some differing from 0.
ASIDE: Proof of Claim:

First note: $|Tl\rangle = \sqrt{\frac{1}{2}}|\text{“hard”}_m|\text{hard}_e\rangle + \sqrt{\frac{1}{2}}|\text{“soft”}_m|\text{soft}_e\rangle$

$= \sqrt{\frac{1}{2}}|\text{“hard”}_m|\{\sqrt{\frac{1}{2}}|\text{black}_e\rangle + \sqrt{\frac{1}{2}}|\text{white}_e\rangle\}\} + \sqrt{\frac{1}{2}}|\text{“soft”}_m|\{\sqrt{\frac{1}{2}}|\text{black}_e\rangle - \sqrt{\frac{1}{2}}|\text{white}_e\rangle\}\}$

$= \sqrt{\frac{1}{2}}|\text{black}_e\rangle\{\sqrt{\frac{1}{2}}|\text{“hard”}_m\rangle + \sqrt{\frac{1}{2}}|\text{“soft”}_m\rangle\} + \sqrt{\frac{1}{2}}|\text{white}_e\rangle\{\sqrt{\frac{1}{2}}|\text{“hard”}_m\rangle - \sqrt{\frac{1}{2}}|\text{“soft”}_m\rangle\}$

$= \sqrt{\frac{1}{2}}|\text{“black”}_m|\text{black}_e\rangle + \sqrt{\frac{1}{2}}|\text{“white”}_m|\text{white}_e\rangle$

Claim (a): $|Tl\rangle$ is an eigenstate of $(C^n \otimes F) - (I^n \otimes C^e)$ with eigenvalue 0

Proof: $( (C^n \otimes F) - (I^n \otimes C^e) )|Tl\rangle = (C^n \otimes F)|Tl\rangle - (I^n \otimes C^e)|Tl\rangle$

$= (C^n \otimes F)\{\sqrt{\frac{1}{2}}|\text{“black”}_m|\text{black}_e\rangle + \sqrt{\frac{1}{2}}|\text{“white”}_m|\text{white}_e\rangle\} - (I^n \otimes C^e)\{\sqrt{\frac{1}{2}}|\text{“black”}_m|\text{black}_e\rangle + \sqrt{\frac{1}{2}}|\text{“white”}_m|\text{white}_e\rangle\}$

$= \{\sqrt{\frac{1}{2}}|\text{“black”}_m|\text{black}_e\rangle - \sqrt{\frac{1}{2}}|\text{“white”}_m|\text{white}_e\rangle\} - \{\sqrt{\frac{1}{2}}|\text{“black”}_m|\text{black}_e\rangle - \sqrt{\frac{1}{2}}|\text{“white”}_m|\text{white}_e\rangle\}$

$= |0\rangle = 0|Tl\rangle$

Next note: $|\text{“hard”}_m|\text{hard}_e\rangle = \{\sqrt{\frac{1}{2}}|\text{“black”}_m|\text{black}_e\rangle + \sqrt{\frac{1}{2}}|\text{“white”}_m|\text{white}_e\rangle\} \{\sqrt{\frac{1}{2}}|\text{black}_e\rangle + \sqrt{\frac{1}{2}}|\text{white}_e\rangle\}$

$= \frac{1}{2}\{|\text{black”}_m|\text{black}_e\rangle + |\text{“white”}_m|\text{black}_e\rangle + |\text{“black”}_m|\text{white}_e\rangle + |\text{“white”}_m|\text{white}_e\rangle\}$

Claim (b): $|\text{“hard”}_m|\text{hard}_e\rangle$ is not an eigenstate of $(C^n \otimes F) - (I^n \otimes C^e)$

Proof: $( (C^n \otimes F) - (I^n \otimes C^e) )|\text{“hard”}_m|\text{hard}_e\rangle = (C^n \otimes F)|\text{“hard”}_m|\text{hard}_e\rangle - (I^n \otimes C^e)|\text{“hard”}_m|\text{hard}_e\rangle$

$= \frac{1}{2}(C^n \otimes F)\{\{\text{“black”}_m|\text{black}_e\rangle + \text{“white”}_m|\text{black}_e\rangle + \text{“black”}_m|\text{white}_e\rangle + \text{“white”}_m|\text{white}_e\rangle\}

- \frac{1}{2}(I^n \otimes C^e)\{\{\text{“black”}_m|\text{black}_e\rangle + \text{“white”}_m|\text{black}_e\rangle + \text{“black”}_m|\text{white}_e\rangle + \text{“white”}_m|\text{white}_e\rangle\}$

$= \frac{1}{2}\{\{\text{“black”}_m|\text{black}_e\rangle - \text{“white”}_m|\text{black}_e\rangle + \text{“black”}_m|\text{white}_e\rangle - \text{“white”}_m|\text{white}_e\rangle\}$

$- \frac{1}{2}\{\{\text{“black”}_m|\text{black}_e\rangle + \text{“white”}_m|\text{black}_e\rangle - \text{“black”}_m|\text{white}_e\rangle - \text{“white”}_m|\text{white}_e\rangle\}$

$= -|\text{“white”}_m|\text{black}_e\rangle + |\text{“black”}_m|\text{white}_e\rangle$

$\neq \lambda|\text{“hard”}_m|\text{hard}_e\rangle$, for any value of $\lambda$.

Similarly for $|\text{“soft”}_m|\text{soft}_e\rangle$. 
SO: In principle, can use the property ("Color" − Color) to experimentally distinguish between Theory 1 and Theory 2. But in practice, this is very difficult:

What happens if there is a single air molecule in the measuring device next to the pointer?

This will change the predictions of Theories 1 and 2!

<table>
<thead>
<tr>
<th>Theory 1 with air molecule</th>
<th>Collapse occurs at $t_1$.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prediction 1:</strong> At $t_1$, state will be:</td>
<td></td>
</tr>
<tr>
<td>either $</td>
<td>\text{&quot;hard&quot;}_m</td>
</tr>
<tr>
<td>or $</td>
<td>\text{&quot;soft&quot;}_m</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Theory 2 with air molecule</th>
<th>Collapse occurs sometime after $t_1$.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prediction 2:</strong> At $t_1$, state will be:</td>
<td></td>
</tr>
<tr>
<td>$\sqrt{\frac{1}{2}}</td>
<td>\text{&quot;hard&quot;}_m</td>
</tr>
</tbody>
</table>

These predictions are not eigenstates of ("Color" − Color)! To distinguish between them, we need a different, more complicated, 3-particle property of the m-e-a system. If there are other air molecules, or other microscopic systems, in the measuring device, we’ll need even more complicated multi-particle properties to distinguish between Theory 1 and Theory 2! Very difficult to implement in practice!

**Basic Problem:** To distinguish between Theory 1 and Theory 2 in practice, either we must have an ideally isolated system (impossible in the lab), or we have to take into account all subsystems that interact with the pointer (air molecules, etc.) (practically impossible).

**Note:** In the absence of a well-defined notion of measurement, any system that interacts with the pointer can serve as a measuring device that tells us where the pointer is located.
II. The GRW Collapse Theory  
Ghirardi, Rimini, Weber (1986)

**Motivation:** Most, if not all, properties can be correlated with the *position* of a pointer in an appropriate measuring device. So perhaps we can maintain Option (A1)* if we can modify the dynamics of states in such a way that always makes superpositions of *position states* collapse. According to Option (A1), this would allow the property of position to always be determinate.


**Additional GRW Dynamical Law (added under QM Principle C):**
During any time interval, there is a non-zero probability that the state of an elementary particle will collapse to a position eigenstate.

**SO:** Suppose $|\psi\rangle$ represents the state of an elementary particle.
We can always expand $|\psi\rangle$ in a basis of eigenvectors of position:

$$|\psi\rangle = a_1|x_1\rangle + a_2|x_2\rangle + \ldots + a_N|x_N\rangle$$

The GRW Collapse Theory says:
1. $|\psi\rangle$ generally evolves *via* the Schrödinger dynamics.
2. But, there is a non-zero chance that, during any time interval, $|\psi\rangle$ will collapse to one of the $|x_i\rangle$ eigenstates of position. If it does undergo such a collapse, the probability of it ending up in a specific position eigenstate $|x_i\rangle$ is given by the Born Rule:

$$\text{Pr}(\text{particle is located at position } x_i \text{ in state } |\psi\rangle) = |\langle\psi|x_i\rangle|^2 = |\psi(x_i)|^2$$

**Why this helps:**
Consider the entangled state of pointer particles and electron:

$$\sqrt{\frac{1}{2}}(|x_1\rangle_1|x_1\rangle_2|x_1\rangle_3\ldots)|\text{hard}\rangle_e + \sqrt{\frac{1}{2}}(|x_2\rangle_1|x_2\rangle_2|x_2\rangle_3\ldots)|\text{soft}\rangle_e$$

position states of pointer particles for a pointer located at $x_1$ (the position that registers “hard”)  
position states of pointer particles for a pointer located at $x_2$ (the position that registers “soft”)

If just *one* of all of these (trillions of) pointer particles has a definite location, then according to the Projection Postulate, the superposition will collapse to the term containing that position eigenstate. So even if the non-zero probability of collapse in the GRW Law is extremely small, since any measuring device has trillions of particles, any superposition containing them will have an extremely likely chance of spontaneously collapsing.
The GRW Law in terms of Wavefunctions:

**Recall:** \( \psi(x) = \langle \psi | x \rangle \) The wavefunction encodes all the values of the expansion coefficients. \( \text{Ex: } \psi(x_i) = a_i \)

When \( |\psi\rangle \xrightarrow{\text{collapse}} |x_i\rangle \) we want \( \psi(x) \xrightarrow{\text{collapse}} \psi(x_i) \)

Mathematically, the wavefunction collapse on the right is accomplished by multiplying the general wavefunction \( \psi(x) \) by a “Dirac delta function” \( \delta(x - x_i) \):

\[
\psi(x_i) = \delta(x - x_i)\psi(x)
\]

We can heuristically think of \( \delta(x - x_i) \) as a position “eigenfunction” (as Albert does). It is an infinite spike exactly at \( x_i \), and zero everywhere else:

![Diagram of Dirac delta function](Image)

**ASIDE:** We should really write \( \psi(x_i) \) in terms of \( \psi(x) \) in the form of an integral:

\[
\psi(x_i) = \int \delta(x - x_i)\psi(x) \, dx
\]

The Dirac delta “function” really isn’t a function. It’s only well-defined inside an integral sign (it’s really what’s called a “functional”). It has the properties:

\[
\delta(x - x') = 0, \text{ when } x \neq x', \quad \delta(x - x') = \infty, \text{ when } x = x', \quad \int \delta(x - x') \, dx' = 1
\]

It fixes the normalization of eigenvectors of an operator with a continuous range of eigenvalues:

\[
\langle x | x' \rangle = \delta(x - x')
\]

Again, this last line isn’t well-defined, mathematically. The physicist Paul Dirac invented these guys to be able to talk heuristically about position eigenvectors.

**Initial Problem:** If our state is represented by an eigenvector of position with associated eigenfunction \( \delta(x - x_i) \), this means (according to Option A1) that it has a *definite value* of position; namely, \( x_i \). Recall from an earlier lecture: This means that properties *incompatible* with position (like momentum or energy) will be “maximally” indeterminate.
When a GRW collapse occurs (and we get an exact value of position), we could potentially have violations of conservation of momentum and/or energy!

**Technical Solution:** In GRW collapses, instead of multiplying the wavefunction by a Dirac delta “function” \( \delta(x - x_i) \), use a “Gaussian” (or “Bell-shaped”) function \( g_L(x - x_i) \) spread out a finite width \( L \) about \( x_i \).

\[ L \text{ can be chosen in such a way that the uncertainty in momentum/energy is effectively cut-off.} \]

The price, thought, is that we’ve had to “smear” the position about \( x_i \).

**Essential Characteristics of GRW Collapse Theory:**

1. Modification of Schrödinger dynamics to make it consistent with the Projection Postulate.
2. Introduces two new constants of nature:
   a. Probability per time per particle of collapse.
   b. Width \( L \) of Gaussian position eigenfunctions.
Three Problems with the GRW Collapse Theory

1. The Problem of Wavefunction “tails”

We just saw that GRW needs Gaussian position functions $g_L(x - x_i)$ in order to avoid violations of energy/momentum conservation.

**Problem:** Mathematically, the Gaussian function $g_L(x - x_i)$ is *never* zero, no matter how far from $x_i$ you get! It has non-vanishing “tails”.

[Diagram: Gaussian function $g_L(x - x_i)$ with non-vanishing tails]

The ends asymptotically approach the $x$-axis, but never reach it, no matter how far away from $x_i$ you get. In other words, $g_L(x - x_i) \neq 0$ for all finite values of $x$.

**Upshot:** Even after a GRW collapse, an elementary particle is still in a superposition of position eigenvectors (there is still a non-zero probability of finding it at some location $x \neq x_i$): So, according to Option (1A), it still has no definite value of position! (When GRW agreed to use Gaussians instead of Dirac delta functions, they gave up exact position collapses.)

2. The Problem of “Positionless” Measurements

GRW assume that the position property is “fundamental” in so far as all other properties must be correlated with the positions of pointers in measuring devices in order to measure them.

**BUT:** Is this correct? Do all measurements of properties require reference to position?

**Albert’s example:** Measuring a particle’s Hardness by means of a fluorescent screen.

[Diagram: Set-up: To measure Hardness of $P$, insert it into Hardness box. If it’s hard, it will exit at $h$ and impact screen at $A$. If it’s soft, it will exit at $s$ and impact screen at $B$.]
3. **The Problem of Microscopic Measurements**

Due to the randomness of the GRW collapse, collapses will only occur in real times for macroscopic measuring devices (that have trillions of elementary particles).

**BUT:** What about the possibility of microscopic measuring devices? These would not be expected to have GRW collapses in real times. If they could be correlated with macroscopic measuring devices, we would have the measurement problem all over again!

**Albert’s example: Particle-Boy**

Albert wants us to imagine a being (“John”) who’s beliefs are correlated with a microscopic measuring device (think of a pointer Hardness device in which the pointer is a single elementary particle whose position gets correlated with the Hardness property of any other particle that enters the device). Then either John’s beliefs about the Hardness of any particle he measures are radically mistaken, or the GRW theory can’t be right!

In fact, if John’s beliefs are correct, not only is GRW wrong, but so is any collapse theory, since we know experimentally that isolated microscopic systems do exhibit superpositions.