Old Quantum Theory ends as Probability appears

In our historical development of quantum mechanics, we have reached 1924 where the “Old Quantum Theory” was established. It was postulated that strange experimental phenomena such as discrete atomic spectra, non-classical blackbody radiation, and the photoelectric effect could be explained by adding a restriction to classical mechanics in the form of quantization rules.

A system’s evolution with time is determined by the solutions of its classical EOMs (equations of motion), but for solutions that are periodic, the only ones that are realized in nature are those that satisfy the Bohr-Sommerfeld quantization rule:

\[ \int_{\text{one period}} p_{q} \cdot dq = n_{q} \hbar \quad \text{for integer values } n_{q} = 0, 1, 2, \ldots \]

In addition, Old Quantum Theory incorporated the idea of wave-particle duality. In 1905, Einstein said “radiation (EM waves) is composed of particles called photons”; in 1923, deBroglie completed the idea by stating that “particles (such as atomic electrons) behave like waves”. Waves are described by frequency \( f \) and wavelength \( \lambda \); particles are described by energy \( E \) and momentum \( p \). The relations between the two are:

\[ E = hf = \hbar \omega \quad \text{and} \quad p = \frac{\hbar}{\lambda} = \hbar k. \]

Quantum Mechanics itself was developed independently in 1925 and 1926, in two different forms that were quickly shown to be equivalent. The main new idea to enter the game is probability. Let’s take a moment to solidify the essential concepts of probability and statistics.

Basic Probability

Consider some observable \( x \) that can take on any value from \( x_{\text{min}} \) to \( x_{\text{max}} \), but where some values between these limits may have higher probability that others to occur. The classic example of such a situation is an experiment where \( x \) will be measured many times, and each measurement will be affected by some underlying randomness, caused by e.g. measurement uncertainty or a random physical process like radioactive decay. To provide a concrete example, imagine a champion darts player practicing, and let \( x \) be the horizontal distance that a dart lands from the center of the bullseye. The champion will usually hit close to the center (close to \( x = 0 \)) but sometimes they will hit a bit to the left (negative \( x \)) and sometimes a bit to the right (positive \( x \)). Perhaps they

---

1. To be exact, sinusoidal waves are described by frequency and wavelength, i.e. waves that oscillate sinusoidally in position and time as \( \cos(kx \pm \omega t) \) or \( \sin(kx \pm \omega t) \). The strict definition of a wave is much more general, and may not involve any oscillations at all. However, Fourier’s theorem tells us that any wave can be built as a superposition of sinusoidal waves. We can therefore happily restrict our wave discussions to oscillating things with frequency and wavelength, without any loss of generality.

2. English desperately needs a gender-neutral pronoun! “He/she” is too awkward, “it” is too inanimate, and inventing something new will never catch on. I have thus decided to bend English grammar slightly and employ “they” as both a singular and plural pronoun. It is already in fairly commonly use as such, as in “the UFO was caught on camera by an unidentified hiker; they immediately posted the photo on their Instagram account”. In German, the pronoun “sie” means both “she” and “they”, so the root language of English already offers support for a pronoun doing double-duty as singular and plural. Let’s change grammar!  ☺
have a leftward bias in their throwing and tend to hit a bit to the left of center. Occasionally they may throw a dart far off to the left or right, but this is less likely to occur if they are very skilled at the game. These features are described by three mathematical objects:

- The **probability distribution** $P(x)$ describes the likelihood of obtaining a particular value of $x$:

  
  
  \[
  \text{probability of obtaining an } x \text{ value in the range } a < x < b = \int_a^b P(x) \, dx
  \]

  
  The term “probability density” and the acronym “pdf = probability density function” are also used for $P(x)$; and Griffiths uses a different symbol, $p(x)$. Any sensible probability distribution must be **normalized**: if the full allowed range of $x$ is $x_{\text{min}} \leq x \leq x_{\text{max}}$, then

  \[
  1 = \int_{x_{\text{min}}}^{x_{\text{max}}} P(x) \, dx
  \]

  i.e. the probability of obtaining a value of $x$ somewhere within its allowed range is 100% = 1.

- The **mean** $\langle x \rangle$ is the average value of $x$ that you obtain if you measure $x$ a huge number of times. The formula for obtaining the mean $\langle x \rangle$ from a probability distribution is:

  \[
  \langle x \rangle \equiv \int_{x_{\text{min}}}^{x_{\text{max}}} x \, P(x) \, dx
  \]

  Very important: “average” is not the same as “most likely”. We’ll clarify that in one of our examples.

- The **standard deviation** $\sigma_x$ is a measure of the width of the distribution $P(x)$ around the mean. Its square, $\sigma_x^2$, is called the **variance** of the distribution. You calculate variance as follows:

  \[
  \sigma_x^2 \equiv \int_{x_{\text{min}}}^{x_{\text{max}}} (x - \langle x \rangle)^2 P(x) \, dx = \left\langle (x - \langle x \rangle)^2 \right\rangle
  \]

  In words: “variance is the average of the squared-deviation from the mean”. An important theorem is:

  \[
  \sigma_x^2 = \left\langle (x - \langle x \rangle)^2 \right\rangle = \langle x^2 \rangle - \langle x \rangle^2
  \]

  Some examples will clarify all this.

**Problem 1 : Robot Dart Player**

Some parts adapted from Griffiths 1.11, 1.12

A team of UIUC students build a robot dart player, HAL\(^3\), for their senior project. The robot is hitting the dartboard pretty consistently, but the team still has some bugs to work out … The team defines an $(x, y)$ coordinate system for the dartboard with the $x$-axis pointing to the right and the $y$-axis pointing upwards. Polar coordinates $(r, \phi)$ can be also be used, with $\phi = 0$ corresponding to the $+x$ axis as usual.

(a) HAL is throwing the darts in a strange pattern: a thin half-circle at a radial distance $d$ from the center of the board, and in the top half of the board only. The region HAL hits is thus $\{ \ r = d \ , \ 0 < \phi < \pi \ \}$. The students also measure that HAL hits all the angles between 0 and $\pi$ with equal likelihood, but never hits the angles between $\pi$ and $2\pi$. How strange! Write down a mathematical expression for the probability distribution $P(\phi)$ that HAL hits, remembering that the full allowed range of $\phi$ values is $0$ to $2\pi$ and that $P(\phi)$ must be normalized over this range. Also make a plot of $P(\phi)$, making sure to run your $\phi$ axis over the full allowed range of angles.

(b) What is the mean value $\langle \phi \rangle$ for HAL’s faulty dart-throwing?

---

\(^3\) Yes, it’s a sci fi reference. It’s also a famous UIUC reference! :-) Google “HAL 9000”.  

(c) Calculate the standard deviation $\sigma_\phi$ in two ways:

(i) Directly calculate the variance $\sigma_\phi^2 = \left(\langle \phi - \langle \phi \rangle \rangle \right)^2$ with an integral.

(ii) Calculate $\langle \phi^2 \rangle$ with an integral, then use the theorem $\sigma_x^2 = \langle x^2 \rangle - \langle x \rangle^2$ to get the variance.

FYI: You have just calculated an important result that comes up a lot: the **standard deviation** of a flat distribution = the range of the distribution over $\text{sqrt}(12)$. This $1/\text{sqrt}(12)$ actually comes up so much in statistics that it is worth memorizing (or being able to re-derive it super-quickly 😊).

(d) $\left(\left(x - \langle x \rangle\right)^2 \right) = \langle x^2 \rangle - \langle x \rangle^2$ is a truly useful theorem! You must prove it. Repeat the calculation (c)(i) you just did, only replace the particular value of $\langle \phi \rangle$ you used with a general expression for the mean. You will quickly find that the integral you set up in (c)(i) simplifies to the integrals you used in (c)(ii), which proves the theorem.

(e) We will now learn how to change variables in probability distributions. What is the probability density for HAL’s dart-throwing in the horizontal coordinate $x$? You must translate your distribution $P(\phi)$ into an equivalent one $P(x) ... how do we perform such a change? It is not hard at all as it is physically intuitive:

If the expressions $(u \pm du)$ and $(v \pm dv)$ correspond to the exact same range of physical situations (e.g. the same outcomes of an experiment), then the probability distributions $P_u(u)$ and $P_v(v)$ for the observables $u$ and $v$ respectively must be related as follows: $P_u(u) \ du = P_v(v) \ dv$.

You will need to change variables a lot in statistics, so this $P_u(u) \ du = P_v(v) \ dv$ formula needs to be intuitively obvious. Let’s write it again using our dart-throwing robot to provide a specific example:

The angle range $\phi \pm d\phi = 30^\circ \pm 1^\circ$ corresponds to a specific section of the dartboard given that the robot only hits on a ring of radius $d = 5$ cm. I can describe the exact same section of the dartboard using the horizontal position range $x \pm dx = 4.33 \pm 0.04$ cm. The probability $P_\phi(\phi) \ d\phi$ for the darts to land at $\phi = 30^\circ \pm 1^\circ$ is equal to the probability $P_x(x) \ dx$ for the darts to land at $x = 4.33 \pm 0.04$ cm because the angle and position ranges describe the exact same set of darts.

If the formula $P_u(u) \ du = P_v(v) \ dv$ is still not intuitively obvious, you must! ask! at office hours, lecture, etc. SO ... using the pdf $P_\phi(\phi)$ for HAL to hit angle $\phi$, calculate and plot the pdf $P_x(x)$ for HAL to hit horizontal coordinate $x$. So that you can plot $P_x(x)$ over the full range of allowed values, let the radius of the dartboard be $R > d$. If you get a negative probability distribution, just take the absolute value as probabilities are always positive (see footnote4). Finally, verify explicitly that your new distribution is properly normalized.

(f) Calculate the mean $\langle x \rangle$ and the standard deviation $\sigma_x$ for HAL’s faulty dart-throwing.

(g) As we mentioned earlier, “average” is NOT the same as “most likely”. HAL’s dart distribution in $x$ illustrates this very well. What is the average $x$ value that he hits, and what are the most likely $x$ values (there are two) for him to hit?

---

4 I guess it is pretty clear that a negative probability makes no physical sense. (The chance for something to happen can’t be less than zero! 😊) Minus signs sometimes show up when changing variables via $P_u(u) \ du = P_v(v) \ dv$, as the relation between $du$ and $dv$ may introduce a sign. For example, if HAL’s ring of darts has radius 5 cm, the angle-range $\phi = 30^\circ \rightarrow 31^\circ$ corresponds to the position-range $x = 4.33 \rightarrow 4.29$ cm. The $\phi$ variable increased while the $x$ variable decreased; that introduces a minus sign into the calculation of $P_x(x)$ from $P_\phi(\phi)$, but obviously $30^\circ \rightarrow 31^\circ$ and $4.33 \rightarrow 4.29$ cm describe the same physical ranges on the dartboard. Just take the absolute value of your probability distribution at the end of your calculation if a minus sign shows up.
Problem 2: Gaussian Probability Distribution

The most common probability distribution by far is the **Gaussian** or normal distribution. It has this form:

\[ P(x) = A e^{-(x-B)^2/2C^2} \]

(a) Figure out the value of \( A \) that is required to normalize the distribution (assuming that \( x \) can range from \(-\infty \) to \(+\infty \)), then show explicitly that the mean of the distribution is \( B \) and the standard deviation is \( C \).

➤ If you have trouble integrating \( \exp(-x^2) \) or \( x^n \exp(-x^2) \), please see the Appendix at the end of this homework.

(b) Using a mean \( \langle x \rangle = 4 \) and a standard deviation \( \sigma_x = 2 \), sketch this distribution.

FYI: The Gaussian distribution will be your constant companion no matter what area of STEM you end up in. It is worth memorizing. In brief, it is the distribution that controls random processes such as counting experiments (e.g. how many radium nuclei will decay within 1 second given a sample of 1 g) and combinatoric situations (e.g. the sum of 6 dice thrown at the same time). Also memorize the number “68”:

For any quantity with a Gaussian probability distribution — and there are many! — there is a **68% chance** that the quantity will land **within 1 standard deviation of the mean**, a 95% chance that the quantity will land within 2 standard deviations of the mean, & a 99.7% chance that the quantity will land within 3 standard deviations of the mean.

That 68% is particularly important because error bars on measured quantities show ±1 standard deviation on either side of the measured value, by long-standing convention. For example, the mass of the electron is

\[ m_e = (9.109\,383\,56 \pm 0.000\,000\,11) \times 10^{-31} \text{ kg according to CODATA 2014}. \]

That uncertainty of ±0.000 000 11 \( \times 10^{-31} \) kg is the **standard deviation** of the quoted value, taking all the statistical and systematic uncertainties of the contributing experiments into account. What it means is that, according to current world data, there is a 68% chance that the electron’s mass is in this range:

\[ 9.109\,383\,45 \times 10^{-31} \text{ kg} \,<\, m_e \,<\, 9.109\,383\,67 \times 10^{-31} \text{ kg} \quad \text{with 68% probability}. \]

So you know, there is a compact way of writing errors that is often used with quantities like \( m_e \) that are known to great precision:

\[ Q = 3.14159 \pm 0.00012 \quad \text{may equivalently be written} \quad Q = 3.14159(12). \]

**Wave Mechanics**

One of the two formulations of quantum mechanics developed in 1925-1926 is called “Wave Mechanics”. Below are the five principal axioms of Wave Mechanics (there is no standard way of counting, only the content matters). As we started in discussion 2, let’s work with these axioms and get used to their strange, probabilistic description of what a particle is.

**AXIOM 1**: A particle is described by a complex-valued wavefunction \( \psi(x, t) \) whose magnitude squared, \( |\psi|^2 \equiv \psi^* \psi \), represents the probability \( P(x, t) \) of finding the particle at location \( x \) at a particular time \( t \):

\[ P(x, t) \ dx = |\psi(x, t)|^2 \ dx \equiv \psi^*(x, t) \psi(x, t) \ dx \]

**AXIOM 2**: A particle’s wavefunction is normalized so that the probability of finding the particle somewhere is 1:

\[ 1 = \int_{-\infty}^{+\infty} |\psi(x, t)|^2 \ dx , \quad \text{which requires that} \quad \psi(x, t) \ \text{goes to zero at} \ x = \pm \infty \ (\text{see Griffiths §1.5}). \]
AXIOM 3: Each dynamical property \( Q \) of the particle is associated with an operator \( \hat{Q} \) that acts on \( \psi \):

\[
\hat{x} = x, \quad \hat{p} = -i\hbar \frac{\partial}{\partial x}, \quad \hat{E} = i\hbar \frac{\partial}{\partial t}
\]

We carry over from classical mechanics the fact that every dynamical property \( Q \) of a particle at a particular time \( t \) is determined by the particle’s state at that moment, i.e. by its position \( x \) and momentum \( p \). In quantum mechanics, each \( Q \) becomes an operator \( \hat{Q} \) that is constructed from the position operator \( \hat{x} = x \) and the momentum operator \( \hat{p} = -i\hbar \frac{\partial}{\partial x} \). We can thus write that, in general,

\[
\text{classical dynamical property } Q(x, p) \rightarrow \text{quantum operator } \hat{Q} = Q(\hat{x}, \hat{p}) = Q(x, -i\hbar \frac{\partial}{\partial x})
\]

AXIOM 4: When we measure a dynamical property \( Q \) of the particle at a particular time \( t \), the particle’s wavefunction tells us the expectation value \( \langle Q \rangle = \text{the average value of the quantity via this equation :} \)

\[
\langle Q(t) \rangle = \int_{-\infty}^{\infty} \psi^*(x, t) \hat{Q} \psi(x, t) \, dx
\]

AXIOM 5: The wavefunction of a non-relativistic particle of mass \( m \) in a potential-energy field \( V(x) \) evolves with time according to the Schrödinger Equation:

\[
i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} + V\psi \quad \text{... or using our operator symbols, } \hat{E} \psi = \left( \frac{\hat{p}^2}{2m} + V \right) \psi.
\]

The Schrödinger equation is the EOM (equation of motion) of non-relativistic wave mechanics.

Problem 3: A Wavefunction to Play With

A particle of mass \( m \) is in the state described by this wavefunction:

\[
\psi(x, t) = A e^{-a((mx^2)/\hbar + it)}
\]

where \( A \) and \( a \) are positive real constants.

(a) Find \( A \).

(b) Calculate the expectation values of \( x, x^2, p, \) and \( p^2 \).

(c) Find the standard deviations \( \sigma_x \) and \( \sigma_p \).

(d) Show that the wavefunction satisfies the Schrödinger equation for a particle of mass \( m \) in a potential-energy field \( V(x) \) ... what is this potential-energy field \( V(x) \)?

FYI: Since the wavefunction satisfies the Schrödinger equation (part d) and you can normalize it (part a), it is a legitimate wavefunction for a single particle of mass \( m \). We will soon prove a relation that you have no doubt encountered before: the Heisenberg Uncertainty Principle. One of its incarnations is that

---

5 A “dynamical property” of a particle means “a property that can change with time”. Examples are a particle’s kinetic energy or its speed. In classical mechanics, such dynamical properties are determined by a particle’s state = its {position & momentum}; in quantum mechanics, they become operators built from the position operator \( x \) and the momentum operator \( -i\hbar \partial/\partial x \).

In contrast, a particle also has intrinsic properties = those that cannot change with time. Examples are a particle’s rest mass and electrical charge. These properties are not determined by operators, they simply have values.
\[ \sigma_x \sigma_p \geq \frac{\hbar}{2} \]

We will prove it later on, but you can already see it in action: using the standard deviations in position and momentum you found in part (c), you can quickly verify that their product is indeed \( \geq \hbar / 2 \).

**Problem 4 : Another Wavefunction to Play With**

A particle is represented by this wavefunction at time \( t = 0 \):

\[
\psi(x,0) = \begin{cases} 
A(a^2 - x^2) & \text{for } -a \leq x \leq +a \\
0 & \text{for } |x| > a 
\end{cases}
\]

where \( a \) is a given (i.e. known) constant and \( A \) is a constant whose value you must determine. You should sketch it, otherwise what fun is this problem? We’re trying to build some intuition here! ☺

(a) Determine the normalization constant \( A \).
(b) Calculate the expectation values of \( x \), \( p \), \( x^2 \), and \( p^2 \) at time \( t = 0 \).
(c) Calculate the standard deviations \( \sigma_x \) and \( \sigma_p \) and check that your results are consistent with Heisenberg’s uncertainty principle.

**Problem 5 [no points] : Are you comfortable with complex numbers?**

If you feel the least bit uncertain about working with complex numbers in physics, please do work through these 225 unit-sections, you will find them in the homework folder:

- § 13.1 Essentials of complex numbers
- § 13.2 Using complex numbers to describe physical systems : AC circuits
- § 14.1 Common pitfalls = dangers of using complex numbers to describe physical systems

**Appendix : Gaussian Integrals**

The Gaussian probability distribution

\[
P(x; x_0, \sigma) = \frac{1}{\sqrt{2\pi} \sigma} e^{-|x-x_0|^2 / 2\sigma^2}
\]

describes a huge majority of random processes, such as the statistical uncertainties on measured quantities. Since it is so common, you should know how to perform integrals with such forms in the integrand. Since you can trivially change variables to get rid of the \( x_0 \) and \( \sigma \) constants in the exponent, the integrals we are talking about are of this form:

\[
\int_{-\infty}^{\infty} dx \; x^n e^{-x^2} = \text{the (n+1)th moment of } e^{-x^2} \quad \text{(Just jargon-busting the word “moment”. ☺).}
\]

- **n = 1**: The first thing to know is that when \( n = 1 \), the integral is really easy:

\[
\int dx \; x e^{-x^2} = \text{obvious} = -\frac{1}{2} \exp(-x^2).
\]

This integral is easily done because of that “x” factor in the integrand.
\( n = 0 \): The next thing to know is that, in contrast, when \( n = 0 \) the **indefinite integral** has no analytic solution:
\[
\int dx \ e^{-x^2} \text{ can only be evaluated numerically.}
\]

However, the **definite integral** can be evaluated if the bounds are 0 or \( \pm \infty \). Here is a trick for how to do it:

Define \( I \equiv \int_{-\infty}^{\infty} dx \ e^{-x^2} \), then make a double-integral by multiplying two of them with different integration var’s:
\[
I^2 = \left( \int_{-\infty}^{\infty} dx \ e^{-x^2} \right) \left( \int_{-\infty}^{\infty} dy \ e^{-y^2} \right) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx \ dy \ e^{-(x^2+y^2)}
\]

\( x \) and \( y \) are just arbitrarily chosen letters … but hah! If you *treat* them as **Cartesian Coordinates**, you recognize the product (\( dx \ dy \)) as the 2D differential of area, \( dA \) … and you can then switch to **Polar Coordinates** using \( dA = (r \ dr \ d\phi) \) and \( r^2 = x^2 + y^2 \):
\[
I^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx \ dy \ e^{-(x^2+y^2)} = 2\pi \int_{0}^{\infty} dr \ 2\pi \int_{0}^{\infty} r \ dr \ e^{-r^2} = 2\pi \left[ -\frac{1}{2} e^{-r^2} \right]_{0}^{\infty} = \pi
\]

\[
\therefore \quad I \equiv \int_{-\infty}^{\infty} dx \ e^{-x^2} = \sqrt{\pi}
\]

Neat trick eh! You can use this often-needed result in your calculations without redoing the proof.

\( n = 2 \): Finally, when \( n = 2 \), use **integration by parts** to reduce the integral to its \( n = 1 \) form:
\[
\int dx \ x^2 \ e^{-x^2} = \int dx \ (x e^{-x^2}) = \int f \ g' = f \ g - \int g \ f'.
\]