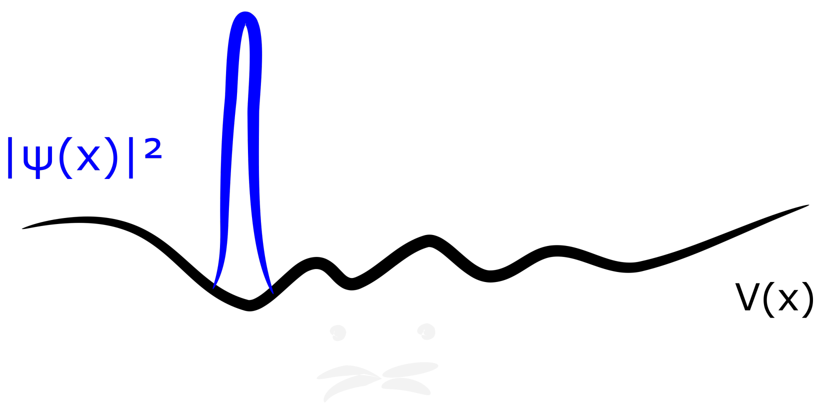
Gravity and the flea

A gravitational approach to the measurement problem

Loek van Rossem



Bachelor Thesis in Mathematics and Physics & Astronomy
Supervisor: Prof. dr. N.P. Landsman
Radboud University Nijmegen
The Netherlands
August 20, 2018

Abstract

In this text we study some possible solutions to the measurement problem in quantum mechanics. This problem is about an inherent contradiction between the Schrödinger equation and the Born postulate: the Schrödinger equation tells us how a system evolves, yet when it is applied to a measurement it does not reproduce the Born postulate. Although many solutions to the measurement problem have been proposed, to this day the problem remains unsolved.

We first look at "the flea on Schrödinger's cat", which is based on the idea that small perturbations can have a large influence on macroscopic quantum mechanical systems. Therefore, the randomness in the Born postulate could be explained as a lack of knowledge about these perturbations. Some problems with this method are discussed, and an attempt will be made to address them. Then we will have a look at the Penrose approach, which tries to use gravity to solve the measurement problem. Inspired by the problems of these two solutions, we will try to combine them into a single hybrid solution, in an attempt to solve the issues.

Contents

Chapter 1

Introduction

The measurement problem is one of the oldest problems in quantum mechanics, and has been around since its discovery. Records of it go as far back as 1926, when Born questioned whether or not the outcomes of measurements in quantum mechanics are determined by hidden properties. The measurement problem is about a discrepancy between quantum mechanics and classical mechanics. Quantum mechanics predicts superpositions, even at the macroscopic scale, yet such things are completely absent in classical mechanics. The most famous formulation of the measurement problem is probably Schrödinger's cat, where quantum mechanics paradoxically predicts a cat to be in a superposition of being both alive and dead at the same time. Many solutions have been proposed over the years, such as the many-worlds interpretation, pilot wave theory, and Ghirardi-Rimini-Weber theory. Still, to this day, no completely satisfying solution has been found, despite the fundamental nature of the measurement problem.

The aim of this text is to explore and discuss some of the solutions to the measurement problem. In chapter 2 we provide the mathematical details of the measurement problem and give a few examples of well known solutions. In chapter 3 we take a look at a more recent solution to the measurement problem: "the flea on Schrödinger's cat". We also discuss some of the issues it has and possible approaches to solving them. In chapter 4 we will look at a solution proposed by Penrose which claims that gravity is the crucial component causing wave function collapse. Finally, in chapter 5 we will propose a solution which makes use of ideas from both "the flea on Schrödinger's cat" and Penrose's solution.

Chapter 2

The measurement problem

The measurement problem may be described as the failure of quantum theory to reproduce the macroscopic classical world.[1] To put it briefly, quantum mechanics predicts macroscopic superpositions, yet none are ever observed. Despite the measurement problem being almost as old as quantum mechanics itself, no satisfying solution has been found.

2.1 A contradiction in quantum mechanics

In quantum mechanics, states evolve according to the Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}\left|\psi(t)\right\rangle = H\left|\psi(t)\right\rangle. \tag{2.1}$$

On the other hand, measurements occur according to the Born postulate, which says that the possible measurement outcomes for measuring an observable A are its eigenvectors, and the probability p_ϕ of obtaining such an eigenvector $\left|\phi\right\rangle$ as an outcome is

$$p_\phi = |\left\langle\phi|\psi\right\rangle|^2, \tag{2.2}$$

where $\left|\psi\right\rangle$ is the state before the measurement. After the measurement, the state of the system has been reduced to $\left|\phi\right\rangle$, and we say the wave function has collapsed.

Using these postulates, most of modern day physics can be derived, yet they also lead to a contradiction. After all, measurement apparatuses are physical objects, and so they too will evolve according to the Schrödinger equation. Thus we can also use the Schrödinger equation to see what the outcome of a measurement is. In general, this does not lead to the same anwser as the Born postulate.

For instance, consider the spin of an electron in the following superposition:

$$\frac{1}{\sqrt{2}}(\left|\uparrow\right\rangle + \left|\downarrow\right\rangle). \tag{2.3}$$

Here $\left|\uparrow\right\rangle$ represents the state where the spin is in the positive z direction (i.e. $\sigma_z\left|\uparrow\right\rangle = \left|\uparrow\right\rangle$), and is $\left|\downarrow\right\rangle$ the state where the spin is in the negative z direction ($\sigma_z\left|\downarrow\right\rangle = -\left|\downarrow\right\rangle$). If we measure the spin in the z direction, then according to the Born postulate there is a 50% chance of getting the outcome spin up and a 50% chance of getting spin down.

Now we investigate what should happen according to the Schrödinger equation. Let $\left|0\right\rangle_A$ be the state initially describing the measurement apparatus. Then our total initial state is:

$$\left|\psi(0)\right\rangle = \frac{1}{\sqrt{2}}(\left|\uparrow\right\rangle + \left|\downarrow\right\rangle) \otimes \left|0\right\rangle_A. \tag{2.4}$$

[1] Many different formulations of the measurement problem are possible [5].

The Schrödinger equation says that time evolution is a unitary operator $U(t)$, so

$$\begin{aligned} |\psi(t)\rangle &= U(t)\left(\frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle) \otimes |0\rangle_A\right) \\ &= \frac{1}{\sqrt{2}}U(t)\left(|\uparrow\rangle \otimes |0\rangle_A\right) + \frac{1}{\sqrt{2}}U(t)\left(|\downarrow\rangle \otimes |0\rangle_A\right). \end{aligned} \tag{2.5}$$

Since the measurement apparatus measures the spin, we have

$$\begin{aligned} U(T)\left(|\uparrow\rangle \otimes |0\rangle_A\right) &= |\uparrow\rangle \otimes |\uparrow\rangle_A\,, \\ U(T)\left(|\downarrow\rangle \otimes |0\rangle_A\right) &= |\downarrow\rangle \otimes |\downarrow\rangle_A\,, \end{aligned} \tag{2.6}$$

where $|\uparrow\rangle_A$ is the state of the measurement apparatus indicating spin up, $|\downarrow\rangle_A$ is the state of the measurement apparatus indicating spin down, and T is a point in time large enough such that the measurement can be considered complete. Combining (2.5) and (2.6), we obtain

$$|\psi(T)\rangle = \frac{1}{\sqrt{2}}\,|\uparrow\rangle \otimes |\uparrow\rangle_A + \frac{1}{\sqrt{2}}\,|\downarrow\rangle \otimes |\downarrow\rangle_A\,. \tag{2.7}$$

So instead of collapsing the superposition of the electron, the measurement apparatus joins the electron in the superposition. This is in contradiction with the Born postulate; we now have a 100% chance of getting the state $\frac{1}{\sqrt{2}}\,|\uparrow\rangle \otimes |\uparrow\rangle_A + \frac{1}{\sqrt{2}}\,|\downarrow\rangle \otimes |\downarrow\rangle_A$ instead of a 50% chance of getting $|\uparrow\rangle \otimes |\uparrow\rangle_A$ and a 50% chance of getting $|\downarrow\rangle \otimes |\downarrow\rangle_A$.

This paradox is probably best known in the form of Schrödinger's cat. There, a scientist puts a cat in a box, along with a radioactive atom. The box is set up so that it will release a deadly poison if the atom decays. Since the atom is a quantum system, it will enter a superposition of being decayed and not decayed. This will result in the cat being in a superposition of dead and alive, which is not the expected macroscopic behavior. Here, the atom plays the role of the electron and the cat plays the role of the measurement apparatus.

2.2 Interpreting measurements

This paradox is known as the measurement problem. The different ways of resolving it correspond to the interpretations of quantum mechanics. We will now give a few examples of these interpretations.

One of the better known interpretations is the many-worlds interpretation, also known as the Everett interpretation. According to this, the Schrödinger equation always gives the correct outcome. Thus, when a measurement takes place, a macroscopic superposition like (2.7) really does occur physically. The interpretation says that the two parts of the wave function each represent their own branch of the universe, one in which the outcome was spin up and the device registered spin up, and one in which the outcome was spin down and the device registered spin down. An observer would read off spin up or spin down, so it would appear to him that the wave function collapsed.

Another interpretation is the pilot wave theory, or de Broglie–Bohm theory. In pilot wave theory the particle and its wave function are separate entities. The wave function guides the behavior of the particle according to the so called guiding equation. This equation is constructed in such a way as to recover standard quantum mechanics. The particle always has a well-defined position (as opposed to a superposition), so no collapse happens during a measurement.

Finally, there is Ghirardi–Rimini–Weber theory. Ghirardi–Rimini–Weber theory is an example of an objective collapse theory, which means that wave function collapse is a physical process, independent of the observer. This is unlike the previous two examples, where wave function collapse is an effect caused by the perspective of the observer, or where wave function collapse does not happen at all, respectively. Ghirardi–Rimini–Weber theory postulates that the wave function of a particle has a chance to spontaneously collapse at any moment. This probability is very small, so that we do not notice it for a single particle. A macroscopic superposition on the other hand, is an entangled mess of many particles, each with a probability to collapse. So it has a much higher chance of collapsing.

In the next chapter we will study another interpretation called the flea on Schrödinger's cat, in which small perturbations cause wave function collapse. After that we will have a short look at the Penrose interpertation, in which collapse takes place due to gravity. Finally, we will combine these two interpretations into yet another interpretation. All three of these interpretations are objective collapse theories.

Chapter 3

The flea on Schrödinger's cat

The flea on Schrödinger's cat is a relatively new approach to the measurement problem [5] [6]. The basic idea is that some macroscopic systems are incredibly sensitive to small perturbations. So the outcome of a measurement might be determined by these perturbations, thus giving the illusion of a random outcome.[1] The name "the flea on Schrödinger's cat" comes from the fact that a tiny perturbation (flea) is responsible for collapsing the Schrödinger's cat state (2.7).

3.1 Instability of large quantum systems

To demonstrate the effect of a small perturbation on a large quantum system we consider the double well potential (figure 3.1).

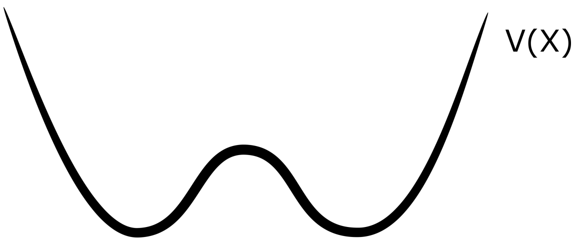


Figure 3.1: A symmetrical double well potential.

Suppose that there is a slight perturbation, such that the the potential is not perfectly symmetrical. For instance, suppose that the left well is ΔE lower in energy than the right well.

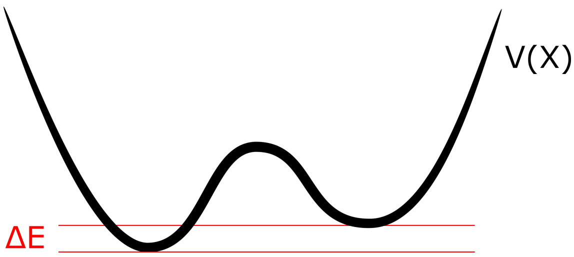


Figure 3.2: A biased double well potential with energy difference ΔE.

Consider now a single quantum particle present in this well. At first sight, we would expect the ground state to be entirely in the left well, since the potential is lower.

[1] Unless the perturbations themselves are non-deterministic, in which case the outcome will still be truly random.

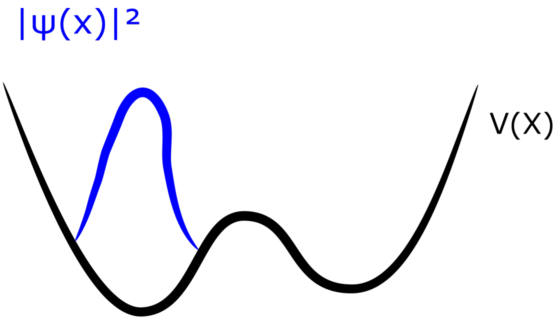


Figure 3.3: A quantum object in the left well.

However, due to the wavy nature of quantum mechanics, the wave function cannot fully localize in the minimum of the potential. After all, if the entire wave function were located at the minimum, the uncertainty principle would give infinite uncertainty in momentum. Generally in quantum mechanics, the wave function spreads out to a characteristic wave length λ as given by the de Broglie relation:

$$\lambda = \frac{h}{p}, \tag{3.1}$$

where h is the Planck constant, and p the momentum. The wave function will therefore occupy a region around the minimum, thus forcing parts of it to be at a relatively high potential (see figure 3.3). Since the center of the right well has a lower potential than these parts, the wave function can lower its energy by tunneling to the right. Thus the ground state occupies both wells (although the part of the wave function in the left well is slightly larger).

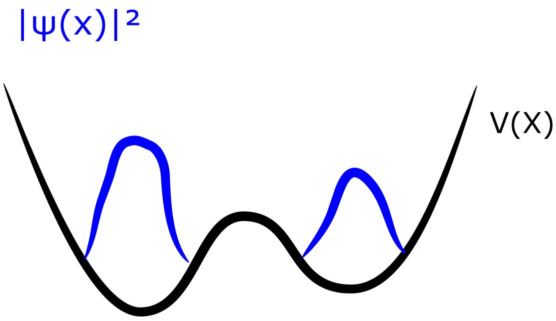


Figure 3.4: A quantum object in the ground state. It occupies both wells.

Now consider a more classical object (for instance one consisting of a large number of quantum particles). The constant h in (3.1) represents the significance of quantum mechanics[2], so taking the limit h to 0 corresponds to the classical limit. Thus we can see that for a more classical system, the characteristic wavelength will be smaller, representing a more localized wave function. Now, the entire wave function fits inside the part of the left well that is lower than the lowest point in the right well (see figure 3.5) [3]. Thus the ground state will almost be contained in the left well.

[2]This can seen in the formula $E = hf$, where E is the energy of a photon and f its frequency. So at fixed frequency f, h gives the minimum energy quantum (one photon). Therefore, setting h to zero corresponds to the classical limit. Alternatively, by looking at the Schrödinger equation, it can be seen that small h corresponds to large m, so it can also be interpreted as the limit of large mass, which is again a classical limit.

[3]Wave functions in quantum mechanics always have a non-zero amplitude everywhere in space, so there should still be a tiny bit of wave function in the right well. Nevertheless, for simplicity, these small parts are ignored in the drawings here.

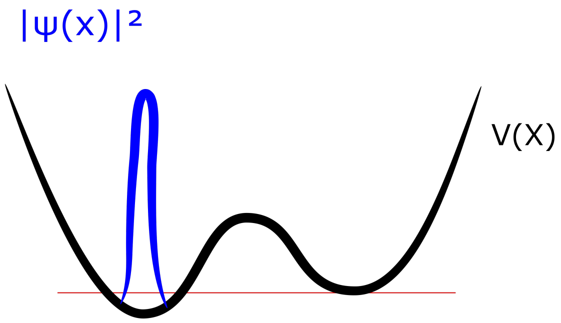


Figure 3.5: An almost classical object in the ground state. It is localized in the left well, since it fits almost completely inside the part of the left well that is completely below the lowest point of the right well (red line).

As we can see from figure 3.4, for large h a small perturbation ΔE has very little effect on a quantum system. However, in figure we can see 3.5 that it has a much larger effect when the system becomes classical. This is actually quite counter-intuitive, as larger systems are apparently *more* affected by perturbations than smaller systems. This property for large quantum systems to localize has been studied both numerically and analytically [5] [6] [3] [11].

We can repeat the same argument for an arbitrary macroscopic physical potential, since in practice potentials have only one global minimum (perfect symmetry is rare when you have many particles).

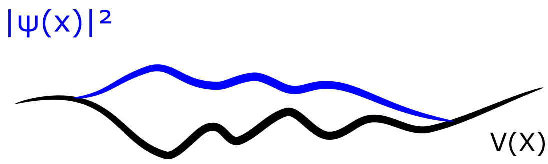


Figure 3.6: The ground state of a quantum system in some arbitrary potential. It has peaks at multiple minima.

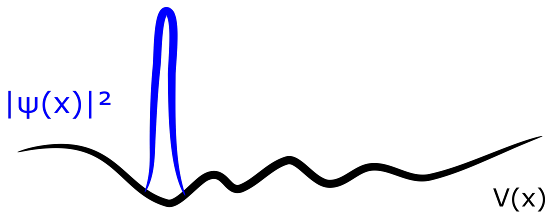


Figure 3.7: The ground state of a classical system in some arbitrary potential. It is almost completely localized at the global minimum.

The fact that classical systems localize at a single point is of great interest to the measurement problem. After all, measurement apparatuses have to be classical since they need to be large enough for us to read the result, and this localization to a single point is rather reminiscent of wave function collapse.

3.2 The measurement process

Let us try to apply this instability property to a measurement. We consider the simplest possible setup, in order to understand best what is going on. Suppose we want to measure the spin of an electron. We will assume our measurement apparatus to be in a double well potential, initially with the wave function located in the center. We also assume an interaction term in the Hamiltonian of the form $-\mu \cdot S \otimes X$, where S is the spin operator of the electron, X is the position operator of the measurement apparatus, and μ is a coupling constant [4]. This term is chosen such that different measurement outcomes correspond to different values for the electron spin. When the spin is up, this term becomes a positive constant times the position operator, so the measurement apparatus can lower its energy by moving left. Similarly, it will move right when the electron has spin up. Thus the two wells represent the two different measurement outcomes (think of a needle on a measurement apparatus moving left to point at the label spin down, or moving right to point at the label spin up).

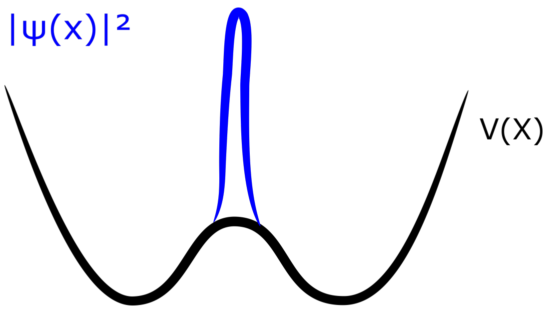


Figure 3.8: The initial state of a measurement apparatus.

Let us now consider the case of an electron in a superposition

$$\frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle). \tag{3.2}$$

Because of the interaction term in the Hamiltonian, half of the wave function will move left, while the other half will go to the right. This will put the measurement apparatus in a macroscopic superposition:

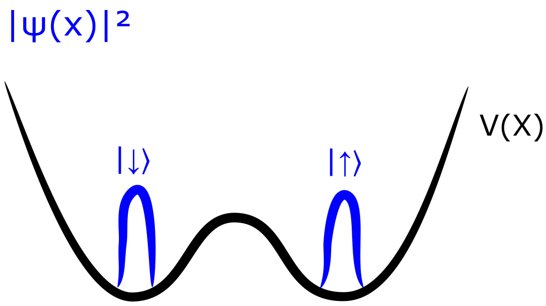


Figure 3.9: The measurement device coupled to an electron in superposition. The left peak is coupled to spin down, while the right peak is coupled to spin up.

This is precisely the Schrödinger cat state from the measurement problem that we wish to avoid. However, suppose that there is a small perturbation of the potential in one of the wells. Now the potential has a single global minimum, so we can use the instability argument such that the apparatus will collapse to a single outcome (figure 3.10). Here we use that any measurement apparatus is a classical object.

[4] An alternative would be a term of the form $\mu \cdot S \otimes P$, where P is the momentum operator. However, we will stick with the $-\mu \cdot S \otimes X$ term, since it will make some concepts easier to explain.

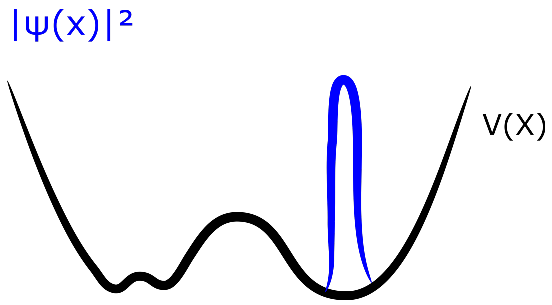


Figure 3.10: A small perturbation in the left well. The global minimum of the potential is now in the right well, so the entire wave function localizes there.

We thus read a single outcome on the measurement device. Since the perturbation is too small to be observed, the outcome appears random. However, this is an illusion; as long as the perturbation is deterministic, so is the outcome. We simply lack the information to observe it as such.

Since the measurement apparatus is macroscopic, perturbations are to be expected. In particular, particles are constantly moving around due to thermal fluctuations, and since the apparatus consists of a large number of particles, we cannot expect the potential to be perfectly symmetrical.

3.3 Born probabilities

So far this method only works in the 50% spin up 50% spin down case. After all, the perturbation completely determines the outcome of the measurement, and it does not depend on the electron spin. There is always a 50% chance for the outcome spin up and a 50% chance for spin down, regardless of the initial superposition of the electron.

So how do we get the Born probabilities (2.2)? We need perturbations that in some way depend on the electron spin. But even that is not enough, as shown by the following argument, which is just an extension of the argument leading to the measurement problem to include all possible initial conditions.

Suppose we do have some source of perturbations that gives rise to the Born probabilities. Let $|0(p)\rangle_A$ be the initial state of the measurement apparatus, including the perturbation p. This state can be different at each measurement, since the perturbation typically differs. Now suppose we measure an electron in the state $|\uparrow\rangle$. To accommodate the Born probabilities, this state must always give the outcome spin up. So for *any* initial state $|0(p)\rangle_A$, we must have

$$U(T)(|\uparrow\rangle \otimes |0(p)\rangle_A) = |\uparrow\rangle \otimes |\uparrow(p)\rangle_A \tag{3.3}$$

where $|\uparrow(p)\rangle_A$ is a state possibly varying depending on p, but always such that the apparatus indicates spin up. Similarly, we have

$$U(T)(|\downarrow\rangle \otimes |0(p)\rangle_A) = |\downarrow\rangle \otimes |\downarrow(p)\rangle_A\,. \tag{3.4}$$

Applying these two relations to the state $\frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$, we find for *any* perturbation p

$$\begin{aligned}
|\psi(T)\rangle &= U(T)\left(\frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle) \otimes |0(p)\rangle_A\right) \\
&= \frac{1}{\sqrt{2}}U(T)\left(|\uparrow\rangle \otimes |0(p)\rangle_A\right) + \frac{1}{\sqrt{2}}U(T)\left(|\downarrow\rangle \otimes |0(p)\rangle_A\right) \\
&= \frac{1}{\sqrt{2}}|\uparrow\rangle \otimes |\uparrow(p)\rangle_A + \frac{1}{\sqrt{2}}|\downarrow\rangle \otimes |\downarrow(p)\rangle_A \qquad (3.5)
\end{aligned}$$

which is clearly a macroscopic superposition rather than a definite outcome. We are therefore back to the measurement problem and conclude that Born probabilities do not come out. By only assuming that the 100% spin up and the 100% spin down cases work, we automatically find that the 50% spin up 50% spin down case fails.

The problem here is that the outcome of the measurement is entirely determined by the perturbation, while the initial superposition of the wave function has no impact whatsoever. If the part of the wave function coupled to spin up somehow influenced the perturbation, then per linear time evolution that influence would be entangled to that part of the wave function, and thus have no influence on the part that is coupled to spin down.

In order to recover the Born probabilities, we must add something that breaks the linearity of time evolution. Since time evolution is always linear in quantum mechanics, we need to look outside standard quantum mechanics for the solution. Because the origin of the perturbation is not set, it seems reasonable to look for non-linear perturbations (possibly following from some underlying theory).

Another problem with recovering the Born probabilities is that since the collapse is caused by the measurement apparatus, the probabilities are likely to be dependent on the properties of the measurement apparatus. However, when measuring the same state, we always want to get the same probabilities, namely the Born probabilities. A possible solution could be that the perturbations are caused by the measurement apparatus, so that they scale with the properties of the device. If the perturbations remain in proportion to the measurement device, the chance of tunneling right instead of left will remain the same, such that the probabilities are the same at all scales.

3.4 Tunneling times

The Born postulate says that when a measurement is made, the measured state collapses *instantly* to the outcome state. In the flea method, however, the collapse happens due to a tunneling process. So in order to match the observed behavior, this tunneling process must happen faster than we can measure. As it turns out, the tunneling times actually get slower the more classical the device is.

To derive the tunneling times, we will approximate the double well potential as a two-level system [6]. Let $\Delta = E_1 - E_0$ be the difference between the first excited state and the ground state. If we ignore all higher energy states and keep only the first two energy states, the Hamiltonian is reduced to

$$H = \frac{1}{2}\begin{pmatrix} 0 & -\Delta \\ -\Delta & 0 \end{pmatrix}. \tag{3.6}$$

The energy eigenstates of H are given by

$$\begin{aligned} \phi_0 &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \\ \phi_1 &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}. \end{aligned} \tag{3.7}$$

From these, the localized states can be constructed as

$$\begin{aligned} \phi_l &= \frac{1}{\sqrt{2}}(\phi_0 + \phi_1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ \phi_r &= \frac{1}{\sqrt{2}}(\phi_0 - \phi_1) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned} \tag{3.8}$$

Suppose that we now add a perturbation of size δ to the left well:

$$H = \frac{1}{2}\begin{pmatrix} \delta & -\Delta \\ -\Delta & 0 \end{pmatrix}. \tag{3.9}$$

The eigenstates now become:

$$\begin{aligned}\phi_0 =& \frac{1}{\sqrt{2}}\left(\delta^2+\Delta^2+\delta\sqrt{\delta^2+\Delta^2}\right)^{-1/2}\begin{pmatrix}\Delta \\ \delta+\sqrt{\delta^2+\Delta^2}\end{pmatrix}, \\ \phi_1 =& \frac{1}{\sqrt{2}}\left(\delta^2+\Delta^2+\delta\sqrt{\delta^2+\Delta^2}\right)^{-1/2}\begin{pmatrix}\Delta \\ \delta-\sqrt{\delta^2+\Delta^2}\end{pmatrix}.\end{aligned} \tag{3.10}$$

For a double well potential $V(x) = \frac{1}{2}\omega^2 x^2 + \frac{1}{4}\lambda x^4$, we can use the WKB approximation to get [1]

$$\Delta = \frac{\hbar\omega}{\sqrt{\frac{1}{2}e\pi}} e^{-d_V/\hbar}, \tag{3.11}$$

where the WKB factor d_V is given by

$$d_V = \int_{-a}^{a} dx\sqrt{V(x)}. \tag{3.12}$$

Therefore, in the classical limit ($\hbar \to 0$), Δ goes to zero exponentially. Applying $\Delta \ll \delta$ to 3.10, we can see that $\phi_0 \to \phi_r$ and $\phi_1 \to \phi_r$.[5] This means that the localized states become stationary states, and thus no tunneling will take place. This is a serious problem, since we want the measurement device to correspond to a classical system.

We can calculate the tunneling times more explicitly by approximating them with half the oscillation time of the two level system

$$T = \frac{\pi\hbar}{\Delta}, \tag{3.13}$$

since this is the time it takes for the system to go from one localized state to the other. If $\hbar$ goes to zero, Δ goes to zero and so the tunneling times get arbitrarily large, which is consistent with the previous argument.

3.5 Energy conservation

The flea on Schrödinger's cat is an objective collapse theory, which means that it describes wave function collapse as a physical process independent of the observer. Objective collapse theories typically have issues with energy conservation, since a wave function in a superposition of different energies has an (expected) energy different from the states it might collapse to. For instance, $\frac{1}{\sqrt{2}}(|E=0\rangle + |E=1\rangle)$ has expected energy $\frac{1}{2}$. If you were to measure its energy, there is a probability $\frac{1}{2}$ for it to collapse to $|E=1\rangle$. However, this state has energy 1, so energy is not conserved.

The flea method deals with this problem by taking into account the environment. Since the measurement device is macroscopic, there are internal degrees of freedom that act as an energy reservoir. One possible model for this is the Spehner-Haake model [12]. Here it is assumed that the apparatus consists of N particles. The Schrödinger equation is then transformed to center of mass coordinates, given by

$$h(P,Q,\pi,\rho) = h_A(P,Q) + h_{AE}(\pi) + h_E(Q,\rho), \tag{3.14}$$

[5] The limit $\phi_0 \to \phi_r$ corresponds to what we have seen in section 3.1 , where we have argued that the ground state localizes to the deepest well.

where

$$\begin{aligned} h_A(P,Q) &= \frac{P^2}{2M} + NV(Q) \\ h_E(\pi) &= \frac{1}{2M}\left(\sum_{n=1}^{N-1}\pi_n^2 + \left(\sum_{n=1}^{N-1}\pi_n\right)^2\right) \\ h_{AE}(Q,\rho) &= \sum_{k=1}^{\infty}\frac{1}{k!}f_k(\rho)V^k(Q), \end{aligned} \tag{3.15}$$

where Q is the center of mass position operator, P is the center of mass momentum operator, ρ_n is the relative position operators with corresponding momentum operators π_n, M is the total mass of the system, V the potential, and finally $f_k(\rho)$ is some function of ρ.

As you can see, the Hamiltonian for the center of mass $h_A(P,Q)$ is just the standard Hamiltonian for a single particle of mass M. The second term $h_E(\pi)$ describes the environment, which consists of relative movement between the particles. This means there can be energy hidden in these internal degrees of freedom. Finally, we have some interaction term $h_{AE}(Q,\rho)$ between the center of mass and the environment. This allows energy to move between the two systems.

When the center of mass wave function collapses, the energy it loses can move to the internal degrees of freedom, so that energy is still conserved. This process is entirely described within standard quantum mechanics, which guarantees that it will conserve energy.[6] This process is not much different from friction, where energy seems to disappear but really just transfers to relative motion between particles.

3.6 Stability of outcomes

Suppose a measurement has taken place, and due to a perturbation the wave function has collapsed to a single outcome. We assume the perturbations to be constantly changing, since different measurements with the same apparatus are supposed to give different outcomes in quantum mechanics. So suppose that after some amount of time, the perturbation has changed location. Then, the wave function will collapse to the new minimum, thus changing the displayed outcome.

This instability of the outcome is undesirable, as it would make measurements impossible (possibly even undefined). Luckily, a solution is already in place, namely the interaction term $-\mu \cdot S \otimes X$. When the wave function has collapsed to the spin up outcome (as in figure 3.10), the position is positive for the entire wave function. The electron can thus lower the interaction energy by becoming entirely spin up. If some kind of environment is present, providing a source of energy and spin, it will try to minimize energy and thus become entirely spin up. We thus find that measuring spin up will collapse the electron itself into the spin up state, not just the measurement device. This is a property that any interpretation of quantum mechanics needs to explain, since it is part of the Born postulate, so it is nice that it already follows from the interaction term.

Given now that the electron is in the spin up state, we can again look at the interaction term $-\mu \cdot S \otimes X$ and see that this lowers the energy at the right well. This in turn keeps the wave function at that outcome, making it stable.

There are however, some issues with this approach. It depends on the value of μ being large enough to outclass the perturbations. This may seem relatively unproblematic at

[6] The potential is time independent between the moment the perturbation has appeared and the wave function has fully collapsed

first, since the perturbations are assumed to be small, while μ is part of a macroscopic effect. However, if the interaction energy $\mu S a$ (where a is the distance from the center to the minima) were larger than the perturbations, the spin up part would always move to the right and the spin down part would always move to the left. So no collapse would occur in a spin superposition.

3.7 Entanglement and locality

According to the flea model, the outcomes of measurements are determined purely by local properties, namely the perturbation in the measurement apparatus. However, quantum mechanics is a non local theory, due to the phenomenon of entanglement. To demonstrate this, consider the following two electron entangled state:

$$\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \tag{3.16}$$

According to the Born postulate, a measurement of both spins of this state can only have outcomes $|\uparrow\downarrow\rangle$ and $|\downarrow\uparrow\rangle$, each with a 50% chance. This is true even if the electrons are separated by a large distance.

Now consider the same situation in the flea model. When we measure one of the electrons, the outcome is determined by the perturbation in the measurement device. Since the $|\uparrow\downarrow\rangle$ part of the wave function behaves like $|\uparrow\rangle$ locally, we have a 50% chance to get spin up. Similarly, if we measure the second electron, the outcome is determined by the perturbation in the measurement device used. At first sight, there is no reason to think that the perturbations in the two measurement devices are correlated, especially if the electrons are separated by a large distance. So we again have a 50% chance of measuring spin up. This means we have a 25% chance to get the outcome $|\uparrow\uparrow\rangle$, which is in contradiction with the Born postulate.

However, a closer look shows that a possible solution would be to again make use of the interaction term in the Hamiltonian. If we make a measurement of the spin of one electron, a random perturbation will collapse the wave function of the apparatus to an outcome. The interaction term will then collapse the superposition of the electron spin to that outcome, as mentioned in section 3.6. Since the electrons *are* correlated (since they are entangled), the collapse of one electron might induce a collapse of the other. Then, due to the interaction term, the other measurement device will collapse to the corresponding outcome.

It is not exactly clear how the electron will induce the collapse of the other electron. But at least the electrons are correlated, in contrast to the perturbations, which really have nothing to do with each other.

Unfortunately, it turns out that such an induced collapse is not possible. Suppose that the issues mentioned in section 3.3 have been solved and p is a perturbation for the first electron is such that it will collapse from the superposition $\frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$ to spin up. This means that

$$\begin{aligned} U(T)(|\downarrow\uparrow\rangle \otimes |0\rangle_A) &= |\uparrow\uparrow\rangle \otimes |\uparrow\rangle_A\,, \\ U(T)(|\uparrow\downarrow\rangle \otimes |0\rangle_A) &= |\uparrow\downarrow\rangle \otimes |\uparrow\rangle_A\,, \end{aligned} \tag{3.17}$$

where A is the measurement apparatus measuring the first electron. Here we use the fact that the states $|\downarrow\uparrow\rangle$ and $|\uparrow\downarrow\rangle$ are product states, so the first electron has no influence on the second electron. We can now use linear evolution to get

$$\begin{aligned}
U(T)\left(\frac{1}{\sqrt{2}}(|{\downarrow\uparrow}\rangle + |{\uparrow\downarrow}\rangle) \otimes |0\rangle_A\right) &= \frac{1}{\sqrt{2}}\left(U(T)(|{\downarrow\uparrow}\rangle \otimes |A\rangle) + U(T)(|{\uparrow\downarrow}\rangle \otimes |0\rangle_A)\right) \\
&= \frac{1}{\sqrt{2}}\left(|{\uparrow\uparrow}\rangle \otimes |{\uparrow}\rangle_A + |{\uparrow\downarrow}\rangle \otimes |{\uparrow}\rangle_A\right) \\
&= \frac{1}{\sqrt{2}}\left(|{\uparrow\uparrow}\rangle + |{\uparrow\downarrow}\rangle\right) \otimes |{\uparrow}\rangle_A \\
&= |{\uparrow}\rangle \otimes \frac{1}{\sqrt{2}}(|{\uparrow}\rangle + |{\downarrow}\rangle) \otimes |{\uparrow}\rangle_A\,. \qquad (3.18)
\end{aligned}$$

So the second electron remains in the superposition $\frac{1}{\sqrt{2}}(|{\uparrow}\rangle + |{\downarrow}\rangle)$, and no induced collapse has taken place. Just as with the Born probabilities, it is the linearity of time evolution that causes problems. By the same token, if a solution to the Born probability problem is found, it might also invalidate this argument.

Chapter 4

The Penrose interpretation

Penrose suggested another approach to the measurement problem [7] [8]. There are compelling reasons to think that gravity and quantum superpositions are incompatible. Therefore, it might be gravity that causes the wave function to collapse.

4.1 Superpositions in space-time

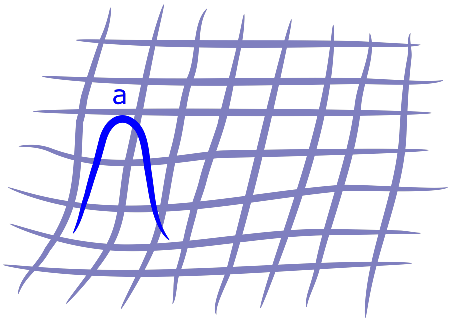
Gravity has a very interesting property with respect to the measurement problem, namely, bigger objects feel more of it. This is interesting, because measurement devices have larger masses than the small quantum system we have observed to have superpositions. So if gravity were to cause wave function collapse, it might do so just for superpositions in measurement devices but not for superpositions in these small quantum systems. Thus keeping the Schrödinger equation for small system, while recovering the Born postulate for measurement devices.

The idea in the Penrose interpretation is that large quantum superpositions are incompatible with general relativity, so a theory of quantum mechanics must somehow prevent large quantum superposition from occurring. This means some kind of wave function collapse must occur, once a measurement device enters superposition.

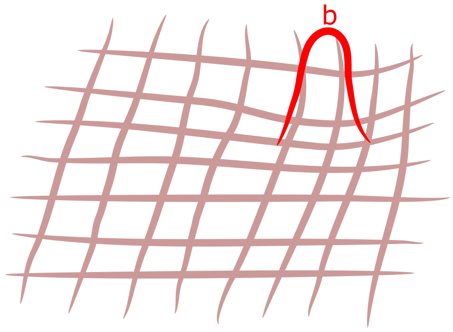
To illustrate the argument, suppose we have a large mass object in a superposition

$$\frac{1}{\sqrt{2}}(|a\rangle + |b\rangle), \tag{4.1}$$

where $|a\rangle$ indicates a state localized around position a, and $|b\rangle$ around b. From general relativity we know that mass curves space-time. We do not know exactly how this happens in quantum mechanics, since no complete theory of quantum gravity exists (yet). However, we do know what space-time looks like when the state is just $|a\rangle$ (see figure 4.1), so let us call this space-time G_a. Similarly, we write G_b for the space-time corresponding to $|b\rangle$.



(a) The system localized around a and its induced space-time G_a.



(b) The system localized around b and its induced space-time G_b.

Figure 4.1: Wave functions in space-time.

For the state $\frac{1}{\sqrt{2}}(|a\rangle + |b\rangle)$ we do not know the space-time, however, a reasonably guess would be to just apply linearity:

$$\frac{1}{\sqrt{2}}(|a\rangle \otimes G_a + |b\rangle \otimes G_b) \tag{4.2}$$

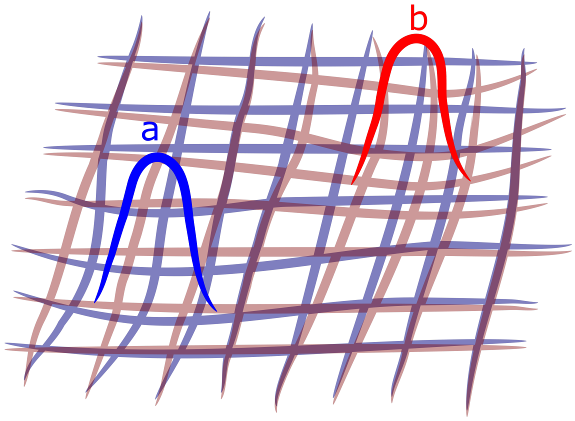


Figure 4.2: The system in a superposition $\frac{1}{\sqrt{2}}(|a\rangle + |b\rangle)$, with its induced space-time.

We now no longer have a single space-time, but a superposition of two space-times. But this leads to problems. The quantum states $|a\rangle$ and $|b\rangle$ are wave functions in space, however, it is no longer clear in *which* space, since there are now two different space-times. So these states no longer have a clear meaning. Even worse, there is no single direction of time anymore. Thus time evolution is no longer defined.

According to Penrose, the degree of ill-definedness can be measured in terms of an uncertainty in energy. This uncertainty in energy corresponds to a decay time $T = \hbar\, \Delta E$, representing the time it takes for the superposition to collapse. This is similar to calculating the decay time of unstable particles, using the uncertainty in energy.

4.2 The Schrödinger-Newton equation

There does not yet exist a complete description of quantum gravity, In order to approximate gravity at a semi-classical scale, we introduce the Schrödinger-Newton equation:

$$i\hbar\frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi + V\psi + m\Phi\psi. \tag{4.3}$$

Here V is the ordinary potential and Φ is the gravitational potential, which satisfies the Poisson equation:

$$\nabla^2\Phi = 4\pi Gm|\psi|^2. \tag{4.4}$$

Equation (4.3) can be derived [10] as the effective behavior of a many particle system. Solving (4.4), we obtain

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi + V\psi - Gm^2\int\frac{|\psi(t,\vec{y})|^2}{|\vec{x}-\vec{y}|}d^3\vec{y}\psi. \tag{4.5}$$

This is the equation you would find when you interpret $m|\psi|^2$ as a mass density. This means there is a self-interaction of the particle, since different parts of the wave function treat each other as a mass density, and hence as a source of gravity. The self-interaction comes from the equation representing a many particle system, and these particles interact with each other.

The Schrödinger-Newton equation can tell us something about the states that the wave function might collapse to. This is because we expect the final state to be stable, meaning it must be a stationary solution of the Schrödinger-Newton equation. However, because of the self-interaction, the Schrödinger-Newton equation is distinctly non-linear, making it even harder to solve than the already difficult ordinary Schrödinger equation.

We can, however, find the gravitational potential for a Gaussian wave function [9]:

$$|\psi(\vec{x})|^2 = \frac{1}{\pi^{3/2}\sigma^3}e^{-r^2/\sigma^2}, \tag{4.6}$$

where A is a normalization constant, σ^2 is the variance and $r^2 = x^2 + y^2 + z^2$. Solving (4.4) for a Gaussian density, we obtain the gravitational potential:

$$\Phi(\vec{x}) = -\frac{Gm}{r}\,\mathrm{erf}\left(\frac{r}{\sigma}\right), \tag{4.7}$$

where $\mathrm{erf}(x)$ is the error function, given by

$$\mathrm{erf}(x) = \frac{1}{\sqrt{\pi}}\int_{-x}^{x} e^{-t^2}dt. \tag{4.8}$$

In figure 4.3, the gravitational potential Φ is plotted for $\sigma = 1$.

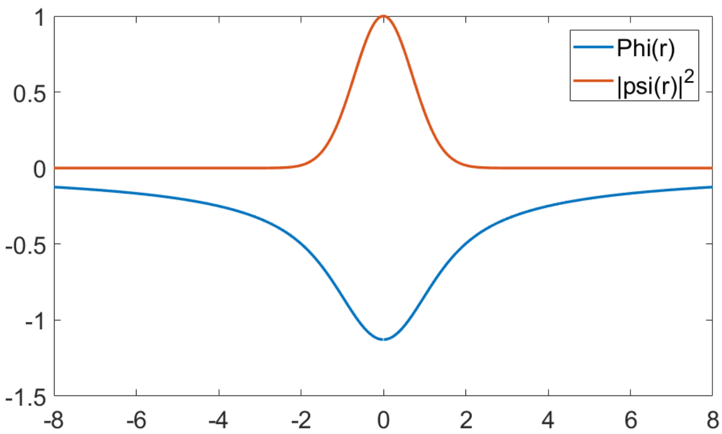


Figure 4.3: A plot of a Gaussian wave function density $|\psi(r)|^2$ with $\sigma = G = m = 1$, centered around 0. The corresponding gravitational potential $V(r)$ is also plotted.

The potential shows that particles are attracted to the center of the Gaussian, which coincides with what you would expect from gravity. For large r, we have $\mathrm{erf}(r) \approx 1$, and so we recover the classical gravitational potential:

$$\Phi(\vec{x}) = -\frac{Gm}{r}. \tag{4.9}$$

Although the stationary solutions might not be Gaussians, we can still try to approximate a stationary solution using a Gaussian in order to better understand the behavior of such solutions. The time-independent Schrödinger-Newton equation in vacuum is

$$E\psi = -\frac{\hbar^2}{2m}\nabla^2\psi - Gm^2 \int \frac{|\psi(t,\vec{y})|^2}{|\vec{x}-\vec{y}|} d^3\vec{y}\psi. \tag{4.10}$$

We thus need the right hand side to be approximately a constant times ψ. Plugging in $\psi(\vec{x}) = \frac{1}{\pi^{3/4}\sigma^{3/2}} e^{-r^2/2\sigma^2}$, and using 4.7, we find

$$E = -\frac{\hbar^2}{2m}\left(\frac{r^2}{\sigma^4} - \frac{3}{\sigma^2}\right) - \frac{Gm^2}{r}\,\mathrm{erf}\left(\frac{r}{\sigma}\right). \tag{4.11}$$

We now apply the Taylor expansion of erf around zero

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\left(x - \frac{x^3}{3} + \frac{x^5}{10} - \frac{x^7}{42} + \frac{x^9}{216} - \dots\right) \tag{4.12}$$

to obtain

$$E = -\frac{\hbar^2}{2m}\left(\frac{r^2}{\sigma^4} - \frac{3}{\sigma^2}\right) - \frac{2Gm^2}{\sqrt{\pi}\sigma}\left(1 - \frac{r^2}{3\sigma^2}\right) + \mathcal{O}(r^4). \tag{4.13}$$

This is approximately a constant if

$$\frac{\hbar^2}{2m\sigma^4} = \frac{2Gm^2}{3\sqrt{\pi}\sigma^3}, \tag{4.14}$$

which can be solved as

$$\sigma = \frac{3\sqrt{\pi}\hbar^2}{4Gm^3}. \tag{4.15}$$

We can see that as m increases, σ decreases. This is to be expected, since a larger mass increases the gravitational attraction between different parts of the wave function, thus narrowing the width of the wave function. However, σ is never completely zero, since the wave nature of quantum mechanics prevents this. Hence the value of σ arises as a balance between gravity and quantum mechanics.

Similarly to section 3.1 we get the behavior that more classical (larger mass) systems become increasingly localized. For similar reasons, this is an interesting property with respect to the measurement problem.

4.3 Collapse from the Schrödinger-Newton equation

We want to use the Schrödinger-Newton equation to explain wave function collapse. Suppose that we have a wave function in a superposition of two locations:[1]

$$|\psi(\vec{x})|^2 = \frac{1}{\pi^{3/2}\sigma^3} e^{-(\vec{x}-\vec{a})^2/\sigma^2} + \frac{1}{\pi^{3/2}\sigma^3} e^{-(\vec{x}-\vec{b})^2/\sigma^2}. \tag{4.16}$$

[1] We can use a sum of delta functions, however Gaussians are more natural as they have no singularities in the potential.

The Poisson equation is linear, so the gravitational potential is

$$\Phi(\vec{x}) = -\frac{Gm}{|\vec{x}-\vec{a}|}\operatorname{erf}\left(\frac{|\vec{x}-\vec{a}|}{\sigma}\right) - \frac{Gm}{|\vec{x}-\vec{b}|}\operatorname{erf}\left(\frac{|\vec{x}-\vec{b}|}{\sigma}\right). \tag{4.17}$$

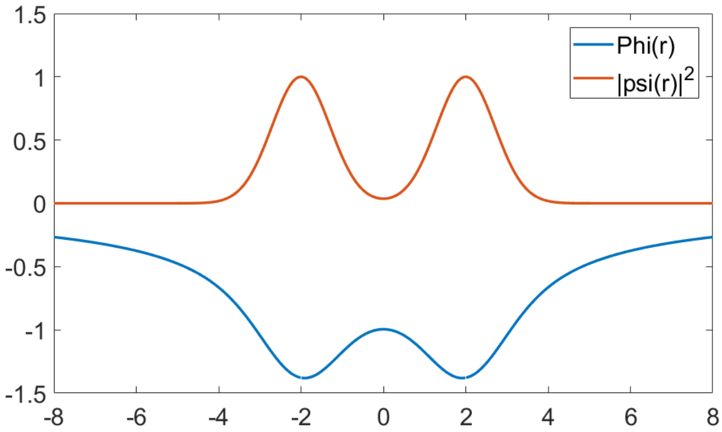


Figure 4.4: A plot of a wave function in a superpostion of two Gaussians with $\sigma = G = m = 1$, centered around -2 and 2. The corresponding gravitational potential $\Phi(r)$ is also plotted.

The left Gaussian can lower its energy by moving right, and the right Gaussian can lower its energy by moving left. By symmetry, the system will collapse to the center. This might not appear to be the case at first, since there is a bump in the middle, however, the potential generated by the left Gaussian moves with the left Gaussian, so only the potential generated by the right Gaussian should be considered for the left Gaussian. Alternatively, you can think of two classical particles which attract each other and end up in the center, which is approximately correct because of (4.9).

So we do have wave function collapse. However, the outcome is problematic. The wave function ends up in the center of the original two Gaussians, while in order to recover the Born postulate, we need the outcome to be near a or b. In fact, if we were measuring the electron spin, the center would not even be a valid outcome, for only spin up and down would be, which are represented by a and b. So even though the Schrödinger-Newton equation might cause wave function collapse, it does not explain the Born postulate.

Chapter 5

Gravity and the flea

In this chapter we combine the ideas of the flea method and the Penrose interpretation. It will turn out that this solves some of the problems in these interpretations. However, other problems still remain unsolved.

5.1 The measurement process 2.0

There are multiple reasons to consider applying the Schrödinger-Newton equation to the measurement process in section 3.2. Most of the issues in the flea method seem to arise from the linearity of time evolution. This makes the Schrödinger-Newton equation interesting, since it is a non-linear equation. But the specific way in which it is non-linear is interesting too. In particular, it gives wave functions a tendency to collapse to a single point (or a narrow Gaussian), since different parts of the wave function attract each other. This means that superpositions are unstable, so we are pretty much guaranteed not to have a superposition at the end of the measurement. Also, as argued in section 4.1, gravity might be related to the measurement problem.

We will try to make as few assumptions as possible in the derivation of the measurement process. Suppose we have a measurement device, capable of measuring the spin of an electron. This must mean that some part of the apparatus does something different depending on whether the spin is up or down. This difference must be something we can read off, so it must be macroscopic. So we will assume this part moves left if the spin is down, and moves right if it is up, and also that it will stay at that location for some amount of time.

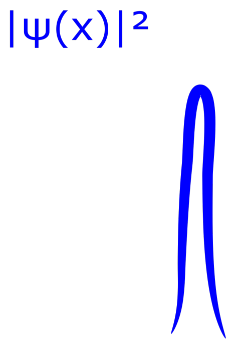


Figure 5.1: The device in its initial state.

Suppose that we try to measure the spin of an electron in a superposition

$$\frac{1}{\sqrt{2}}\left(|\uparrow\rangle + |\downarrow\rangle\right). \tag{5.1}$$

If we wish to use the Schrödinger-Newton equation, time evolution is no longer linear. However, since the effects of gravity are much smaller than those of other forces, the effect of the force pushing the system to the left or right (most likely electromagnetic) is still approximately linear. So the wave function still ends up in a superposition:

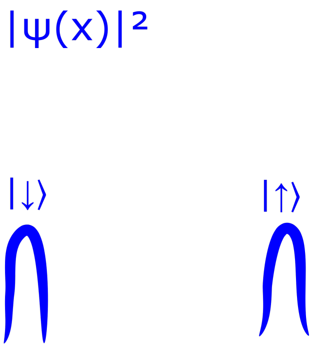


Figure 5.2: The device after coupling to an electron with spin $\frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$.

The measurement device is macroscopic, meaning it consists of many particles. Since the Schrödinger-Newton equation can be derived as an effective many particle equation, it does not seem unreasonable to use it here. This gives us a gravitational potential, which is similar to the one in figure 4.4, since it is generated by a wave function localized at two locations. But this is just the double well potential. So here we get it for free, instead of assuming it[1].

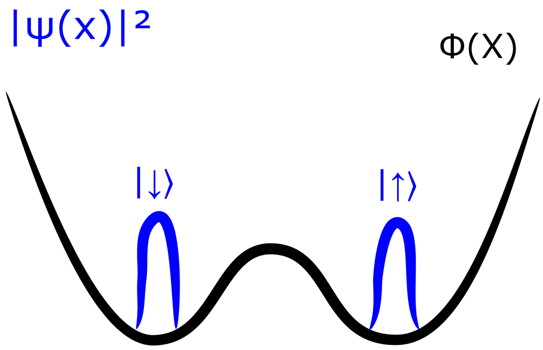


Figure 5.3: The gravitational potential corresponding to the superposition state.

Suppose that we have some small perturbation in one of the wells.

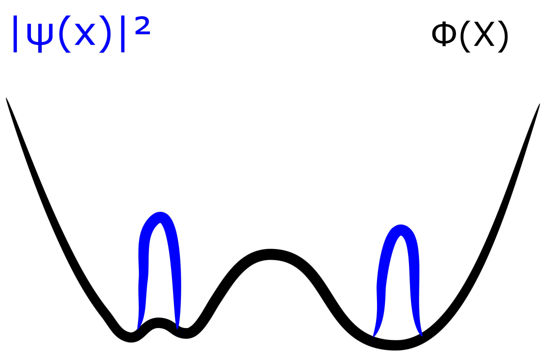


Figure 5.4: The gravitational potential plus some small perturbation.

[1]Assuming a double well potential for measurement devices might be reasonable, however wave function collapse happens for any macroscopic object, not just measurement devices, which makes the assumption a lot less obvious.

This will cause a small part of the wave function to tunnel to the other well. This in turn will lower the potential at that well, since the potential is generated by the mass density of the wave function.

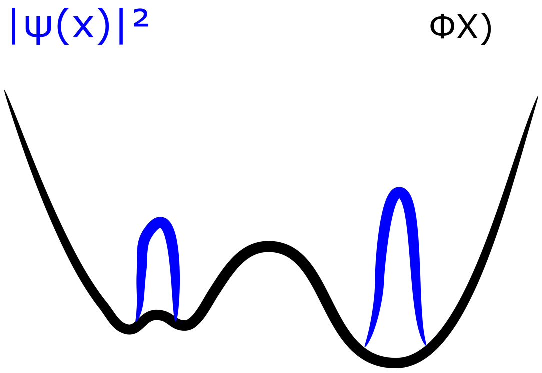


Figure 5.5: The gravitational potential after some part of the wave function has tunneled to the lower well.

Now the wave function will start to tunnel even faster. So there is a self-reinforcing effect causing the wave function to tunnel faster and faster to the lower well, until the entire wave function has collapsed to a single outcome.

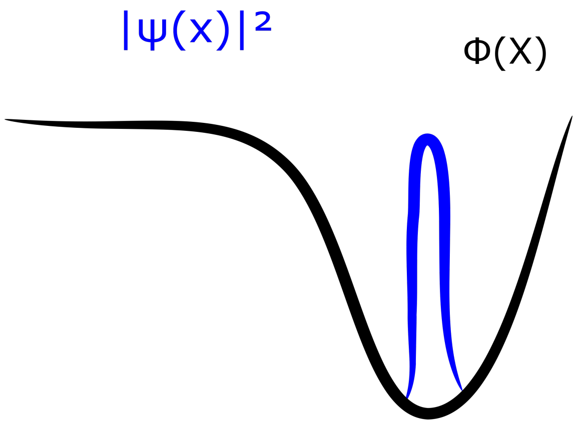


Figure 5.6: The gravitational potential after the wave function has collapsed to a single well.

If a perturbation were to appear in the other well, it would not have any effect, since it would never be larger than the gravitational potential. The outcome state is thus stable, solving the issues mentioned in section 3.6. Another advantage of using the Schrödinger-Newton equation, is that a collapse mechanism is already built in. So we no longer need the argument from section 3.1 to work. And finally, since the Schrödinger-Newton equation already breaks the linearity, we are free in the choice of perturbation (they no longer need to come from some non-linear theory).

5.2 Applicability of the Schrödinger-Newton equation

At first sight the Schrödinger-Newton equation seems like a reasonable assumption. After all, the measurement apparatus is a classical system, so it must consist of many particles. This also makes the self interaction of the wave function seem fairly reasonable, since it is really just an approximation of the interaction between the different particles that the apparatus consists.

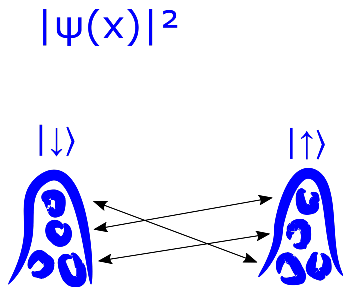


Figure 5.7: A wave function consisting of many particles. The particles have been drawn as circles. There is self-interaction, since different particles in the different parts of the wave function interact with each other.

However, we still run into problems. Consider again the apparatus in its initial state.

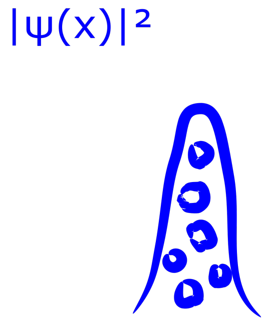


Figure 5.8: The initial state of a measurement apparatus consisting of many particles.

When the apparatus couples to the electron in superposition, the apparatus takes over the superposition. However, each individual particle making up the apparatus also gets in a superposition. Thus, we end up in a state that appears microscopically similar to figure 5.7, but which at the individual particle level looks very different.

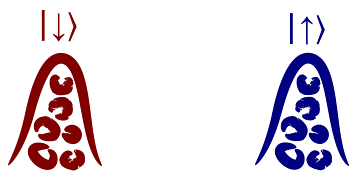


Figure 5.9: The many particle measurement apparatus after coupling to an electron in superposition.

Now there is no longer a reason to expect a self-interaction. Particles in one peak of the wave function are entangled to a different part of the electron spin from particles in the other peak. So quantum linearity suggest the peaks should not interact.

We can still use the Schrödinger-Newton equation, but we have to assume it at the single particle level, meaning we assume that there is something fundamentally non-linear about gravity. This is not particularly strange, since it is still unclear whether or not gravity can be incorporated in quantum mechanics. Also, as we have seen in section 4.1, linearity and general relativity do not go well together.

5.3 Born probabilities

Similarly to the flea on Schrödinger's cat, a priori this process only seems to work in the fifty-fifty case. Since the perturbation determines to which side the wave function tunnels, the outcome is just random. However, unlike the flea model, the 100% spin up and the 100% spin down cases also seem to work. If the entire wave function is on the right side, there will be a deep well there and so no small perturbation will be large enough to make the wave function tunnel to the left side. This is important, since for the flea model it was proven that if the 100% cases work, then the fifty-fifty case cannot work. This was because of the linearity of time evolution, which the Schrödinger-Newton equation breaks.

It remains unclear how this gravitational method would produce the Born probabilities for other spin distributions. Suppose that the electron is in such a superposition, say $\frac{1}{\sqrt{3}}\left|\uparrow\right\rangle + \sqrt{\frac{2}{3}}\left|\downarrow\right\rangle$. The well corresponding to spin down will now be quite a bit deeper than the one corresponding to spin up. So you would expect the wave function to start tunneling towards the spin down outcome, unless there is a perturbation large enough to make the spin up well the deeper. However, the wells are generated by gravity, so they are of the order Gm^2. This means that the perturbation must be of a similar order in order to change the outcome. So it must scale with Gm^2 as well. It is not quite clear if such perturbations even exist.

A possible source of such perturbations might be fluctuations in the many particle wave function. As seen in section 5.2, the wave function of the measurement device is really an average over many particles. These particles move around due to temperature, so the wave function tends to fluctuate. This will lead to fluctuations in the gravitational potential, which can be interpreted as perturbations. If there are more particles, there will be more fluctuations, and since the mass scales with the number of particles, the perturbations will scale with the mass.

However, it turns out that these perturbations do not scale quadratically with the mass. To illustrate this, suppose we do have such perturbations for all m (or at least some range of m). We will show that this assumption contradicts itself. If we have some apparatus with large mass M, we know that it generates perturbations P_M of the order GM in the gravitational field,[2] so P_M is a random variable with standard deviation of order GM. Since M is large, we know the apparatus must consist of smaller objects, so we can for instance interpret it as a collection of N objects, each with mass $m = M/N$. For these objects we have perturbations P_m of the order Gm in the gravitational field. Since gravitational fields add, we get $P_M = \sum_{i=1}^{N} P_{m_i}$, which, by the central limit theorem, has a standard deviation of order $GN^{1/2}m = GM^{1/2}m^{1/2}$, thus giving a contradiction. So perturbations of such nature only exists when they scale as $Gm^{1/2}$ in the gravitational field ($Gm^{3/2}$ in the potential), and not for Gm (Gm^2 in the potential).

The fact that such perturbations cannot scale as Gm^2 should not come as a surprise. We know that such fluctuations become less relevant at larger scales, since they must produce the classical behavior we observe. This leads to a more fundamental problem with recovering the Born probabilities. We need perturbations of the order Gm^2 to get the Born probabilities. These perturbations have to be fundamental, in the sense that we cannot remove them by taking more accurate measurements, since we must always recover the Born probabilities, even for incredibly accurate measurements. Such perturbations would be observable at the macroscopic level, yet all perturbations we observe at this scale can be removed by taking more accurate measurements.

To get around this issue, we propose two possible approaches to a solution. First of

[2] Perturbations of order GM^2 in the gravitational potential correspond to perturbations of order GM in the gravitational field, if an object of mass M is used to probe the field. This is true here, since we are considering the effects of the perturbations on the measurement device which has mass M.

all, we can get around the problem if we only need collapse at a fixed mass scale. Then there would be no need for the perturbations to scale as Gm^2. When the measurement apparatus couples to the electron, it is not instantly brought entirely into a superposition. Instead, this happens gradually, where the electron first couples to a few atoms, which then in turn couple to more atoms, until eventually the entire device is in a superposition state. So the mass of the part that is in superposition effectively slowly increases, and so does the relevance of gravity. There might be some mass scale at which gravity is important enough to induce collapse, so when the superposition reaches this scale, it collapses. This always happens at the same scale, so we only need the perturbations to produce the Born probabilities at that scale.

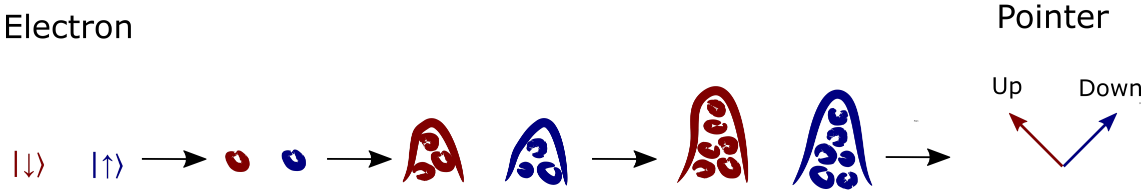


Figure 5.10: The superposition of the electron couples gradually to increasingly larger objects, until finally it couples to the entire pointer.

Another approach would be perturbations that exist for a very short amount of time. Suppose for instance, that the gravity between two objects is not a deterministic variable, but a very quickly varying stochastic variable with the classical value as an average, and a standard deviation of order Gm^2. This would not make a difference for the macroscopic behavior of gravity, but if the fluctuations times are at least comparable in size to tunneling times, they might produce Born probabilities. This basically reduces the problem to the problem of getting fast enough tunneling times.

To demonstrate how Born probabilities might appear, we consider normally distributed fluctuations. The cumulative distribution function for a normal distribution with standard deviation σ, centered around μ, is given by

$$F(x) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right]. \tag{5.2}$$

We take the fluctuations at each point to be normally distributed independently in space, with σ proportional to $|\psi(x,t)|^2$. The motivation is that they are fluctuations of gravity, and gravity is proportional to the mass density (so to $m|\psi(x,t)|^2$). We will also take the mean to be 0, since we want to recover gravity at the classical scale. For convenience, we will assume the tunneling times to be much shorter than the fluctuation times.[3] Suppose we have an electron in spin superposition $x\,|{\uparrow}\rangle + \sqrt{1-x^2}\,|{\downarrow}\rangle$, and perturbation P_l in the left well and P_r in the right well. To figure out what the outcome will be, we need to calculate which well is the lowest one, since that is where the wave function will tunnel to. We approximate the superposition as consisting of two Gaussians, one on the left with mass x^2m, and one on the right with mass $(1-x^2)m$. Taking the limit $r \to 0$ of equation (4.7) using (4.12), we have the following expression for the potential at the center of a Gaussian wave function

$$V(0) = -\frac{2Gm^2}{\sqrt{\pi}\sigma}. \tag{5.3}$$

Pluggin in (4.15), we get

$$V(0) = -\frac{8G^2m^5}{3\pi\hbar^2}. \tag{5.4}$$

[3]If the tunneling times are not significantly shorter than the fluctuation times, the calculation becomes much harder.

We thus find that the wave function will collapse to the left if and only if

$$-\frac{8G^2x^{10}m^5}{3\pi\hbar^2} + P_l < -\frac{8G^2(1-x^2)^5m^5}{3\pi\hbar^2} + P_r. \tag{5.5}$$

Rewriting this, we get the condition

$$P_l - P_r < A\left(x^{10} - (1-x^2)^5\right), \tag{5.6}$$

where $A = \frac{8G^2m^5}{3\pi\hbar^2}$.

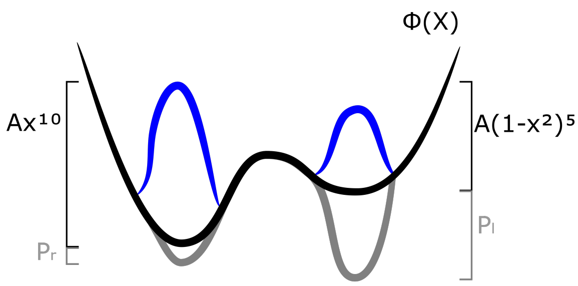


Figure 5.11: The double well with perturbations in each of the wells. Here the wave function would tunnel to the left.

The sum of two normal distributions is again a normal distribution with $\mu = \mu_1 + \mu_2$ and $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}$. Using this and (5.2), we see that (5.6) happens with probability

$$\begin{aligned} F(x^{10} - (1-x^2)^5) &= \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{A(x^{10}-(1-x^2)^5)}{\sqrt{c^2x^4 + c^2(1-x^2)^2}\sqrt{2}}\right)\right] \\ &= \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{A}{c}\frac{x^{10}-(1-x^2)^5}{\sqrt{1-2x^2+2x^4}\sqrt{2}}\right)\right], \end{aligned} \tag{5.7}$$

where c is the proportionality constant $\sigma = c|\psi|^2$.

A plot of these probabilities can be seen in figure 5.12 for $c = 0.464A$. Since c had to be chosen to give a good result, it is essentially a one parameter fit. For a one-parameter fit, the result is relatively good, however there are still some slight deviations from the Born probabilities.

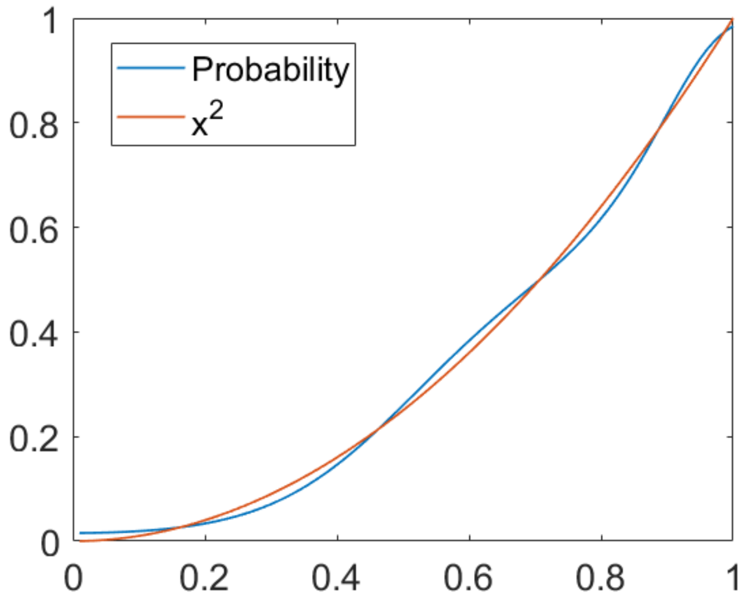


Figure 5.12: The derived probabilities (blue line) plotted against the Born probabilities (red line). The proportionality constant c is taken to be $0.464A$, since this gives the best match.

Interestingly, most formulas that look similar to (5.7) do not give as good a result. In fact, for any possible power of m in equation (5.4), 5 gives the best result (see figure 5.13).

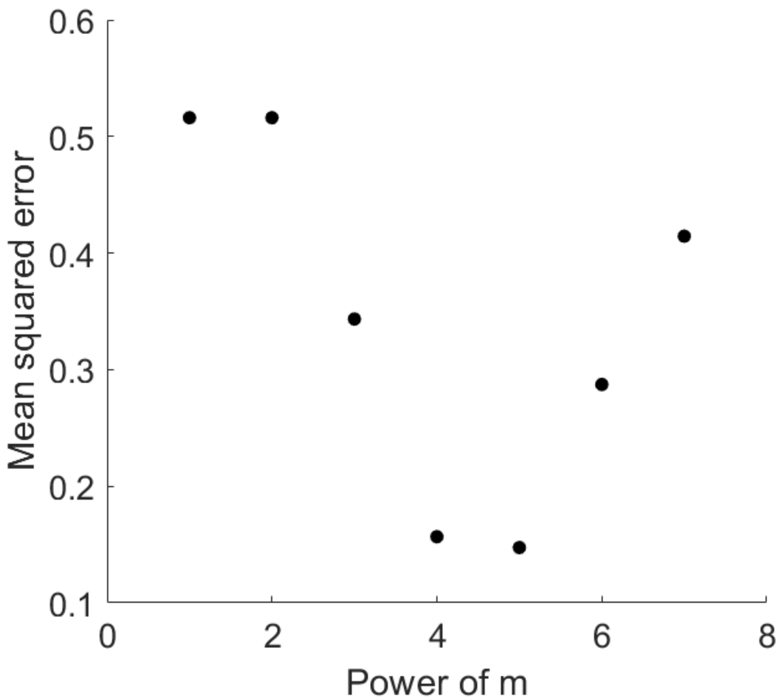


Figure 5.13: The minimal mean squared error of the derived probabilities compared to the Born probabilities for each possible power of m in equation (5.4). The mean squared error for a power 5 is 0.1474 which is slightly lower than for a power 4, which has 0.1567.

For instance, if we look at the best fit for m^6, it looks noticeable worse:

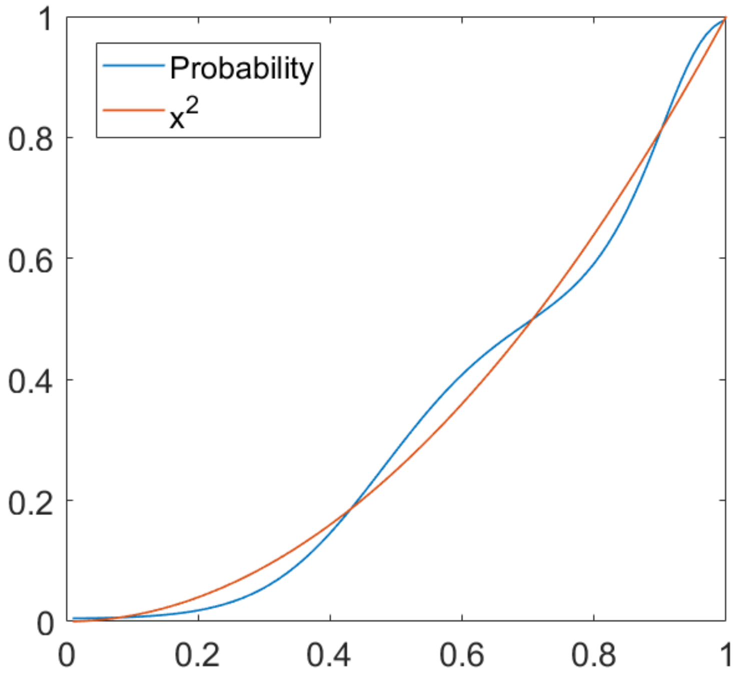


Figure 5.14: The derived probabilities for $V(0)$ proportional to m^6 (blue line) plotted against the Born probabilities (red line). The proportionality constant c is taken to be $0.39A$, since this gives the best match.

There is a rather subtle point to be made about this derivation. By taking the electron in a superposition $x\left|\uparrow\right\rangle + \sqrt{1-x^2}\left|\downarrow\right\rangle$, we implicitly assume this state to be normalized. However, the motivation for normalization comes from the Born postulate, which is precisely what we are trying to derive. Therefore, it is not completely obvious that we can assume the spin wave function to be normalized. This is not a problem here, since we do not use the wave function directly, but instead only use the mass density which we assume to be distributed according to the wave function squared. The normalization condition therefore reduces to the condition that the total mass is m, which is clearly true by assumption.

5.4 Tunneling times

In this case the classical limit corresponds to the limit of large mass. Since the double well potential is generated by gravity, increasing the mass increases the potential. So according to (3.11) and (3.12), the energy difference Δ will go to 0 for large masses. We thus again find that in the classical limit the tunneling times decrease, similar to section 3.4.

This argument is not completely watertight. There are reasons to believe that the WKB-method used to derive (3.11) only work if the double well potential is symmetric [2]. But in order to get Born probabilities, we need large perturbations, meaning the well will not be approximately symmetric. There is also an intuitive explanation why (3.11) might fail in the asymmetric case. The potential contributes in two ways to the tunneling times. First of all, the potential in between the two wells acts as a barrier which slows down the tunneling time. This can be seen in (3.12), as it integrates over the potential between the two wells. The potential also contributes in the sense that, at least intuitively, tunneling should go faster from a higher well to a lower, since energy is lowered. This is only relevant to the asymmetric case, since the otherwise there is no difference between the wells. This effect cannot be seen in (3.12), and so it is not surprising that it might fail in the asymmetric case.

It is also important to note that in the flea method the potential was static, whereas here it is generated by the wave function itself and thus is changing. Since the wave function tunnels to one well, it also lowers the potential even more at that well, and so the tunneling speed should increase. This could mean that the actual tunneling times are smaller than the initial static tunneling times.

Using an expression for Δ the tunneling times can be approximated with (3.13). These tunneling times will depend on the mass m will tell us at which mass scales tunneling (and thus collapse) will take place within a reasonable time frame. For single particles the tunneling should take very long, since we know particles can be in a superposition without the superposition spontaneously decaying. For macroscopic objects the tunneling time should be incredibly small, as we have never observed a large scale superposition. Calculating these tunneling times could give us the means to test this theory, as it will tell us at which scales superposition become fundamentally unstable.

To demonstrate this, assume that $\Delta \approx Gm^2$, which is a typical gravitational energy for an object with mass m.[4] From (3.13) we find that the tunneling time is $5.98 \cdot 10^{36} s$ for an electron, or approximately 10^{19} times the age of the universe. With such long tunneling times, spontaneous collapse would practically never occur for single particles, which is consistent with our current observations. For an object with a mass of $1kg$, the tunneling time would be $4.97 \cdot 10^{-24} s$, which would appear instantaneous to us. This would mean that there would be no macroscopic superpositions, which again is consistent with our current observations. Finally, the mass required for the tunneling time to be one second is $2.23 \cdot 10^{-12} kg$, or approximately the mass of a human cell. This would give us a testable prediction, as it would mean that superpositions at such mass scales are fundamentally unstable.

5.5 Entanglement and locality

The same problem with entanglement as mentioned in section 3.7 is present here as well. There seems to be no reason one measurement would influence the outcome of another measurement a large distance away. However, the counterargument in section 3.7 now fails, because time evolution is no longer linear under the Schrödinger-Newton equation. So the solution mentioned there could in principle work, although it is still not clear how exactly the collapse of one electron would induce a collapse in the other one. Since it is gravity that breaks the linearity, it must be gravity that somehow influences the other electron at long distance, if this argument were to hold. Either way, entanglement seems like a though problem since the collapse at least appears to be completely local.

.

[4] In fact, the self-interaction energy of a sphere of radius R with uniform mass density is $U = -\frac{3}{5R} Gm^2$ [4], which is not unlike the type of energy we are considering.

Chapter 6

Conclusion

In this text we considered three possible solutions to the measurement problem. We first looked at "The flea on Schrödinger's cat", which turned out to not be able to produce the correct Born probabilities because it has linear time evolution. It also has issues with tunneling times being too long, unstable measurement outcomes, and reproducing entanglement. On the other hand, problems with energy conservation could be avoided.

Next, we looked at the Penrose interpretation. This led to the Schrödinger-Newton equation as a way of making quantum mechanics non-linear. The Penrose interpretation itself however, lacks the ability to explain concretely how wave function collapse takes place.

By combining the two interpretations, we get both the non-linearity from the Penrose interpretation, as well as a concrete explanation of wave function collapse from "The flea on Schrödinger's cat". So unlike "The flea on Schrödinger's cat" it is not impossible to get Born probabilities, at least in principle. It also solves the problem of unstable outcomes, as the self-interaction keeps a collapsed state localized.

There are still some unanswered questions. The fact that perturbations similar in size to the potential are needed to derive Born probabilities still seems like a weak point. The two possible solutions mentioned in section 5.3 could work in principle, but more work is needed to confirm this. Also, even though it has been shown that reasonably behaving perturbations can produce the Born probabilities, it is not clear if such perturbations actually physically exist and are always present during a measurement. More research could go into searching for possible physical sources of such perturbations, and figuring out which ones generate Born probabilities. The tunneling times have a better shot at being fast enough than in the flea method, however, explicit calculations are still necessary to confirm this. Using this, it can be calculated at which scales the collapse takes place in a reasonable time frame, which could possibly be compared to experiments. Finally, an explanation is still missing for entanglement. Although entanglement is generally hard to explain due to its strange non-local properties, doing so is still necessary for reproducing quantum mechanics.

References

[1] Anupam Garg. "Tunnel splittings for one-dimensional potential wells revisited". *American Journal of Physics* 68 (2000), pp. 430–437.

[2] Ricard Gelabert, Miquel Moreno, and José M. Lluch. "Applicability of the WKB method in Asymmetric Double Wells with Degenerate and Nondegenerate Minima". *Journal of Computational Chemistry* 15 (1994).

[3] G. Jona-Lasinio, F. Martinelli, and E. Scoppola. "New approach to the semiclassical limit of quantum mechanics". *Communications in Mathematical Physics* 80 (1981), pp. 223–254.

[4] C. Kittel, W. D. Knight, and M. A. Ruderman. *Mechanics 2nd edition*. US: McGraw-Hill Inc., 1973, pp. 268–269.

[5] N.P. (Klaas) Landsman. *Foundations of Quantum Theory*. Cham, Switzerland: Springer, 2017. DOI: `https://link.springer.com/book/10.1007%2F978-3-319-51777-3`.

[6] N.P. (Klaas) Landsman and Robin Reuvers. "A Flea on Schrödinger's Cat". *Foundations of Physics* 43 (2013), pp. 373–407. URL: `https://arxiv.org/abs/1210.2353`.

[7] Roger Penrose. "On Gravity's role in Quantum State Reduction". *General Relativity and Gravitation* 28 (1996).

[8] Roger Penrose. "On the Gravitization of Quantum Mechanics 1: Quantum State Reduction". *Foundations of Physics* 44 (2014). DOI: `https://link.springer.com/article/10.1007/s10701-013-9770-0`.

[9] Carlos Castro Perelman. "Exact Solutions of the Newton-Schrödinger Equation, Infinite Derivative Gravity and Schwarzschild Atoms". *Physics and Astronomy International Journal* 1 (2017).

[10] Aaron Schaal. "Derivation and Meaning of the Newton-Schrödinger Equation". MA thesis. Ludwig-Maximilians-Universität München, Technische Universität München, 2016.

[11] Barry Simon. "Semiclassical analysis of low lying eigenvalues. IV. The flea on the elephant". *Journal of Functional Analysis* 63 (1985), pp. 123–136.

[12] Dominique Spehner and Fritz Haake. "Decoherence bypass of macroscopic superpositions in quantum measurement". *Journal of Physics A: Mathematical and Theoretical* 41 (2008).