Chapter 11

Physical Systems

Until now we have ignored most aspects of physical systems by dealing only with abstract ideas such as information. Although we assumed that each bit stored or transmitted was manifested in some physical object, we focused on the abstract bits, and ignored any limitations caused by the laws of physics. This is the fundamental mantra of the information age.

It has not always been that way, and it will not be that way in the future. In past centuries, the physical manifestation of information was of great importance because of its great cost. To preserve or communicate information, books had to be written or even words cut into stone. For example, think of the process of creating a medieval manuscript during the middle ages. Pages were laboriously copied and illustrated. The results may be viewed today with great admiration for their artistry and cultural importance, in part because they were so expensive to create – society could only afford to deal with the most important information, and the cost of superb art work was not high compared with the other costs of production.

Advances over the years have improved the efficiency of information storage and transmission – think of the printing press, telegraph, telephone, radio, television, digital signal processing, semiconductors, fiber optics. These have enabled complicated systems such as computers, data networks, and even economic systems for the creation and distribution of entertainment. As the cost of processing data drops, it is relevant to consider the case where the cost is negligible compared to the cost of creating, maintaining, and making use of information. It is in this domain that the abstract ideas of information theory, bits, coding, and indeed all of computer science are dominant. All aspects of modern society are coping with the increasing amount of information that is available. Even the ideas of intellectual property, copyrights, patents, and trade secrets are being redefined in light of the changing economics of information processing. This is the information age.

This model of information separated from its physical embodiment is, of course, an approximation of reality. Eventually, as we make microelectronic systems more and more complicated, using smaller and smaller components, we will need to face the fundamental limits imposed not so much by our ability to fabricate small structures, but by the basic laws of physics. The basic law which governs all physical systems is quantum mechanics.

Quantum mechanics is often believed to be of importance only for small structures, say the size of an atom. Although it is unavoidable at that length scale, it also governs everyday objects. When dealing with information processing in physical systems, it is pertinent to consider both very small systems with a small number of bits of information, and large systems with large amounts of information.

The key ideas we have used thus far that need to be re-interpreted in the regime where quantum mechanics is important include

- The **digital abstraction** made practical by devices that can restore data with small perturbations

Author: Paul Penfield, Jr.

- Use of **probability** to express our knowledge in the face of uncertainty
- The **Principle of Maximum Entropy** as a technique to estimate probabilities without bias

11.1 Nature of Quantum Mechanics

Quantum mechanics is weird. There seems to be no way to make it appear otherwise. Many of its predictions are not consistent with expectations that arise from everyday experience.

Quantum mechanics is mysterious, even to very good physicists. The underlying philosophy and interpretation of its equations and techniques are controversial.

Quantum mechanics is difficult to use. Relatively advanced mathematical skills are needed. The basic equation, although linear, is a partial differential equation that cannot be solved analytically except in a very few simple situations. Usually numerical solutions are necessary.

Quantum mechanics, like other physical theories, requires skill and judgement both in modelling and in mathematics. It is not generally taught in any depth before the graduate or advanced undergraduate level.

Quantum mechanics comes in different forms. It has many alternate formulations. Generally these are equivalent in the sense that they predict the same results of experiments, but are not equally easy to learn or to use for particular purposes.

In light of these attributes, why is quantum mechanics important? Because it works. It is the ONLY fundamental physical theory that works over such a wide range of situations. Its predictions have been verified experimentally time after time. It applies to everyday size objects, and to astronomical objects (although it is usually not necessary for them). It applies to atomic-size objects, to electromagnetic waves, and to sub-atomic objects. There is a version that is compatible with the theory of special relativity. About the only physical phenomenon not handled well at this time is gravity; quantum mechanics has not yet been extended to be compatible with the theory of general relativity.

In these notes we cannot cover quantum mechanics in much depth. For the purpose of examining information processing in physical systems, we only need to understand a few of the general features such systems must have. In particular, we need a model of physical systems in which there are many possible states, each with its own probability of being the one the system is actually in (i.e., the state "occupied"). These states all have physical properties associated with them, and energy is one of these. Quantum mechanics justifies this model.

We will use this model in two situations. The first (below) is one with many states, where the objective is to understand how the information associated with the occupancy of these states affects the flow of energy. The second (in a later chapter of these notes) is one with a very small number of states, where information is represented using the occupancy of these states, and the objective is to understand both the limits and opportunities afforded by quantum mechanics.

The next two sections, entitled "Introduction to Quantum Mechanics" and "Stationary States," may be skipped by readers who are prepared to accept the state model without justification. They may proceed directly to the section "Multi-State Model." Other readers may glean from these two sections some indication of how quantum considerations lead to that model, and in the process may find some aspects of quantum mechanics less mysterious.

11.2 Introduction to Quantum Mechanics

Perhaps the first question to ask about a physical object is, "where is it?" In everyday experience, it is possible to answer that question with great precision, limited only by the quality of measurement apparatus. In the realm of very small objects, there are some fundamental limitations and quantum mechanics must be used to address that question.

At its heart, quantum mechanics deals with energy. Because of the equivalence of mass and energy (remember Einstein's famous formula $E = mc^2$ where c is the speed of light, 2.998×10^8 meters per second) quantum mechanics also deals with particles with mass. And because of the relationship between energy of

a photon and its frequency ($E = hf$ where h is the Planck constant, 6.626×10^{-34} Joule-seconds) quantum mechanics deals with photons.

According to quantum mechanics, the question "where is it" cannot be answered with certainty. How do we deal with uncertainty? By assigning probabilities. It is a little more complicated because of the continuous nature of space, and because space is considered to be infinite in extent, but the idea is the same as for probabilities of a finite set of events. The probability density is nonnegative, and integrates over all space to 1 (this is like the sum of the probabilities of all events that are mutually exclusive and exhaustive adding up to 1).

Thus in quantum mechanics, an object is represented as a "probability blob" which evolves over time. How does it evolve? The underlying equation is not written in terms of the probability density, but rather in terms of another function of space and time from which the probability density can be found.

Consider the square root of the probability density, as a function of space and time. Then, for added generality, let the square root be either positive or negative – when you square it to get the probability density, either one will do. Next, for even more generality, allow this square root to have an arbitrary phase in the complex plane, so that it has both a real and an imaginary part. We will no longer call this the square root, but instead the "wave function" $\psi(r,t)$ which is a function of space r and time t. The probability density is then the magnitude of the wave function squared

$$| \psi(r,t) |^2 = \psi^*(r,t)\psi(r,t) \tag{11.1}$$

where the asterisk * denotes the complex conjugate.

In dealing with probabilities earlier, we never expressed them in terms of any more primitive concept. Why do we need to now? Because the fundamental equation of quantum mechanics deals with $\psi(r,t)$. Why is this? Don't ask. It's just one of many bizarre features of quantum mechanics.

The fundamental equation of quantum mechanics is the Schrödinger equation, published in 1926 by the Austrian physicist Erwin Schrödinger (1887 - 1961).[1]

$$i\hbar\frac{\partial\psi(r,t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi^2(r,t) + V(r)\psi(r,t) \tag{11.2}$$

where i is the (imaginary) square root of -1, m is the mass of this object, $V(r)$ is the potential energy function, whose spatial gradient is the negative of the force on the object, and $\hbar = \frac{h}{2\pi} = 1.054 \times 10^{-34}$ Joule-seconds. Note that this equation contains partial derivatives in both space and time. The derivative with respect to time is first order, and the spatial derivatives are second order. The Laplacian ∇^2 is defined as

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \tag{11.3}$$

where x, y, and z are the three spacial dimensions.

This equation is frequently interpreted by multiplying it by $\psi^*(r,t)$ and integrating over space. Then the left-hand side is identified as the total energy, and the right-hand side as the sum of the kinetic and potential energies (assuming the wave function is normalized so that the space integral of $| \psi(r,t) |^2$ is 1, a property required for the interpretation in terms of a probability density).

This equation is deceptively simple. It is a linear equation in $\psi(r,t)$ in the sense that if both 1 and 2 are solutions then so is any linear combination of them

$$\psi_{\text{total}} = \alpha_1\psi_1 + \alpha_2\psi_2 \tag{11.4}$$

where α_1 and α_2 are complex constants (if the linear combination is to lead to a valid probability distribution then the values of α_1 and α_2 must be such that the integral over all space of $| \psi_{\text{total}} |^2$ is 1). However, except for the simplest cases of $V(r)$ the equation cannot be solved in closed form.

[1] See a biography of Schrödinger at http://www-groups.dcs.st-andrews.ac.uk/~history/Mathematicians/Schrodinger.html

Strictly speaking, this equation is really only correct if the object being described is the entire universe, in which case the equation is useless because it is so complicated. However, it is often used as an approximation in the case where the universe is considered in two pieces – a small one (the object) whose wave function is being calculated, and the rest of the universe (the "environment") whose influence on the object is assumed to be represented by the $V(r)$ term. Note that the object may be a single photon, a single electron, or two or more particles, i.e., it need not correspond to the normal concept of a single particle.

An object might interact with its environment. Naturally, if the object changes its environment (as would happen if a measurement were made of some property of the object) then the environment in turn would change the object. One of the interesting results of quantum mechanics is the fact that after a measurement of some property, an object will generally have a different wave function, and as a consequence some properties of the object as it had previously been may not be able to be determined.

11.3 Stationary States

Even though, for a given $V(r)$ term, the Schrödinger equation may be impossible to solve in closed form, much can be said about the nature of the solutions without knowing them in detail. Consider those solutions that are of a particular form, a product of a function of space times another function of time. It can be easily shown from the Schrödinger equation that the most general form this class of wave function can have is

$$\psi(r,t) = \phi(r)e^{-iEt/\hbar} \tag{11.5}$$

for some real constant E (real because otherwise (r, t) would be grow without bound for very large or very small time), where $\phi(r)$ obeys the equation (not involving time)

$$E\phi(r) = -\frac{\hbar^2}{2m}\nabla^2\phi(r) + V(r)\phi(r) \tag{11.6}$$

Nonzero solutions for $\phi(r)$ cannot be obtained for arbitrary values of E. There may be some ranges in which any value of E is OK and other ranges in which only specific discrete values of E lead to nonzero wave functions. Generally speaking, those solutions corresponding to discrete values of E become small far away (i.e., they "vanish at infinity") and are therefore localized in space, although they might have their "probability blob" split over more than one location.

These solutions are called "stationary states" because the magnitude of the wave function (and therefore the probability density) does not change in time; it is only a function of space.

For these stationary states, E has an interesting interpretation. If we multiply this equation by $\phi^*(r)$ and integrate over space, we see that (just as in the previous section) E is the sum of two terms from the right-hand side, namely the kinetic and potential energies of the object. Thus E is the total energy associated with that solution.

Of course most solutions to the Schrödinger equation with this potential $V(r)$ do not have this form. But remember that any linear combination of solutions to the Schrödinger equation is also a solution. We can use these stationary states as building blocks to generate more general solutions.

We are most interested in stationary states that are localized in space, so that the allowed values of E are discrete, although there could be many of them (perhaps even a countable infinite number). If we let j be an index over the stationary states, then it is possible to define the resulting wave functions $\psi_j(r,t)$ so that they are both "normalized" in the sense that the space integral of the magnitude of each squared is 1 and "orthogonal" in the sense that the product of any one with the complex conjugate of another is zero when integrated over all space. We can then denote the values of E, which we have interpreted as the energy associated with that state, by e_j.

Then the general solutions to the Schrödinger equation are written as a linear combination of stationary states

$$\psi(r,t) = \sum_j a_j \phi_j(r) e^{-ie_j t/\hbar} \tag{11.7}$$

where a_j are the expansion coefficients, which may be complex. If the wave function $\psi(r,t)$ is normalized then it is easily shown that

$$1 = \sum_j \mid a_j \mid^2 \tag{11.8}$$

and that the energy associated with the function can be written in terms of the e_j as

$$\sum_j e_j \mid a_j \mid^2 \tag{11.9}$$

From these relationships we observe that $\mid a_j \mid^2$ behaves like a probability distribution over the events consisting of the various states being occupied, and that this distribution can be used to calculate the average energy associated with the object.

The conclusion of our brief excursion into quantum mechanics is to justify the multi-state model given in the next section. Those readers who were willing to accept this model without any explanation have skipped over the past two sections and are now rejoining us.

11.4 Multi-State Model

Our model of a physical object, justified by the brief discussion of quantum mechanics in the previous two sections, is as follows. The object has a wave function ψ which, in principle, characterizes its behavior over time. This wave function may be difficult or impossible to calculate, and it can change, perhaps in unpredictable ways, when the object interacts with its environment.

The object has a finite (or perhaps countable infinite) number of "stationary states" that are easier to calculate (although for complicated objects finding them may still be impossible). Each of the stationary states has its own wave function ψ_j where j is an index over the stationary states. If the actual wave function is one of these stationary states (i.e., if this state is "occupied") then the object stays in that state indefinitely (or until it interacts with its environment). Each stationary state has its own energy e_j and possibly its own values of other physical quantities of interest.

The object's wave function can be expressed as a linear combination of the stationary states, in the form

$$\psi = \sum_j a_j \psi_j \tag{11.10}$$

where the a_j are complex numbers called expansion coefficients. If the object occupies one of the stationary state then all a_j are 0 except one of them. Without loss of generality the expansion coefficients can be defined so that the sum of their magnitudes squared is one:

$$1 = \sum_j \mid a_j \mid^2 \tag{11.11}$$

Measurement an object's property, such as its energy, involves an interaction with the object's environment, and a change in the environment (if for no other reason than to record the answer). It is a consequence of quantum mechanics that if the object is in one of its stationary states and its energy is measured, the result of the measurement is simply the energy of that state, and the state does not change (i.e., the expansion coefficients, all of which are 0 except one, are not changed by the measurement). On the other hand, if the object is not in one of the stationary states, then the result of the measurement is the energy of one of the

stationary states, and the object immediately assumes that stationary state. Thus after each measurement the object ends up in a stationary state. Which state? The probability that state j is the one selected is $\mid a_j \mid^2$. Thus the expected value of the energy measured by an experiment is

$$\sum_j e_j \mid a_j \mid^2 \tag{11.12}$$

where e_j is the energy associated with the stationary state j. Measurement in quantum mechanics is thus not like measurement of everyday objects, where it is assumed that the energy or other physical properties can be measured with arbitrary accuracy, and that such measurements need not perturb the object. The nature of quantum measurement is one more of those aspects of quantum mechanics that must be accepted even though it may not conform to intuition developed in everyday life.

11.5 Energy Systems

An object intended to perform storage, transmission, or conversion of energy must have possible states. Such an object typically might consist of a large number (say Avogadro's number $N_A = 6.02 \times 10^{23}$) of similar or identical particles and therefore a huge number of stationary states. The Schrödinger equation cannot be solved in such circumstances. Interactions with the environment would occur often in order to transfer energy to and from the environment. It is impossible to know whether the system is in a stationary state, and even if it is known, unpredictable interactions with the environment make such knowledge irrelevant rapidly.

The most that can be done with such systems is to deal with the probabilities p_j of occupancy of the various stationary states

$$p_j = \mid a_j \mid^2 \tag{11.13}$$

The expected value of the energy E would then be

$$E = \sum_j e_j p_j \tag{11.14}$$

This model is set up in a way that is perfectly suited for the use of the Principle of Maximum Entropy to estimate the occupation probability distribution p_j. This topic will be pursued in the next chapter of these notes.

11.6 Information Systems

An object intended to perform information storage, transmission, or processing should avoid the errors that are inherent in unpredictable interactions with the environment. The simplest such object that could process information would need two states. One bit of information could be associated with the knowledge of which state is occupied. More complex objects, with more than two states, could represent more than one bit of information.

Quantum information systems, including computers and communication systems, will be the topic of a later chapter of these notes.