

*Lecture 1**Relevant sections in text: §1.1***What is a theory?**

Students often consider quantum mechanics to be rather strange compared to other theories they have encountered (*e.g.*, classical mechanics, electromagnetic theory). And I guess it *is* a little unsettling to find position and momentum being represented as differential operators, particles being described by complex waves, dynamics being defined by the Schrödinger equation, measurements characterized by probabilities, and so forth. In particular, the rules of quantum mechanics, at first sight, seem rather strange compared to the rules of Newtonian mechanics (essentially just Newton's 3 laws). But a lot of the intuitive appeal of Newton's laws comes from years of familiarity. Unfortunately, the "strangeness" of quantum mechanics, largely bred from lack of familiarity, often leads to the feeling that the subject is very difficult. Of course, the subject is not easy; nor should it be - it is one of our most advanced descriptions of nature! But even if the basics seem a little difficult, it is only because one is having to use new mathematical models for familiar physical structures. After all, wasn't it a little unsettling when you first started using vectors in a systematic way to describe displacements, velocities, and so forth? (Can you remember that far back?) So, I would like to begin by emphasizing that, in many ways, quantum mechanics is a theory like any other. To do this I must give a very coarse grained description of what a theory is in general, and quantum mechanics in particular. Of course, the devil - and the physics - is in the details.

Essentially, a physical theory is a set of rules (*i.e.*, postulates) that can be used to "explain" and/or predict the behavior of the world around us - in particular, the outcome of experiments. Which experiments can be explained, how to characterize the physical ingredients in these experiments, what information needs to be specified in advance, *etc.* are part of the rules of the theory. By the way, keep in mind that a theory can never be said to be "true", but only that it "agrees with experiment". It is always possible that the next experiment will falsify the theory. And it is possible that more than one theory will be able to explain a given set of results. Insofar as these results are all we have experimental access to, the theories are equally good. Usually, though, there is one theory that explains the most results with the "simplest" postulates. This theory is usually considered the best. So, while classical mechanics can be used to explain a wide variety of macroscopic observations, quantum mechanics can also explain these results* and a host of other results from microscopic physics (atomic physics, nuclear physics, *etc.*) that cannot be explained using classical mechanics.

* This is essentially because quantum mechanics *includes* classical mechanics as an approximation which is valid for macroscopic phenomena.

It should be emphasized that the word “theory” has a number of connotations, and this can be confusing. For example, quantum mechanics is a “theory”, but the use of quantum mechanics to model the hydrogen atom as, say, a non-relativistic electron moving in a fixed Coulomb field is also a “theory” in some sense — a theory of the hydrogen atom. Clearly these two notions of “theory” have somewhat different logical standings in the sense that one can build a variety of “theories” of the hydrogen atom (by adding, *e.g.*, spin-orbit coupling, special relativity, finite-size nucleus, *etc.*) within the framework of *the* theory of quantum mechanics. Given this slightly confusing state of affairs, I will try (but may fail) to call quantum mechanics a “theory”, and I will call “models” the various theories built from quantum mechanics (of, *e.g.*, the hydrogen atom). Thus we can discuss the physical “theory” called “quantum mechanics” and using the framework defined by this theory we can build “models” of various physical phenomena (*e.g.*, crystalline solids) using various external considerations (*e.g.*, periodic potentials, various approaches to conductivity, *etc.*). The model of a given phenomenon can be wrong without the theory of quantum mechanics itself being wrong. On the other hand, if one is unable to build a satisfactory model of the phenomenon owing to a fundamental failure of the parent theory (quantum mechanics), then one might try to improve the theory of quantum mechanics itself. Similar comments will apply to other theories/models, of course.

What are some successful physical theories? There are many, of course. Some examples are: Newton’s theory of matter and its interactions, valid at large length scales, weak gravitational fields, and small velocities (“classical mechanics”); Maxwell’s theory of the electromagnetic field and its interaction with electrically charged “sources”; Einstein’s theories of relativity and gravitation; and, of course, the theory of matter and its interactions at small length scales, which we call quantum mechanics, along with its descendant, quantum field theory. One confusing feature of all this is that theories really come to us in overlapping hierarchies. For example, using the classical Maxwell theory of electrodynamics in conjunction with classical mechanics we can successfully model all macroscopic electrodynamic systems. Such models are, ultimately, incorrect (being classical) and must be incorporated into a more correct theory when dealing with microscopic phenomena. A “correct” theory of electromagnetic phenomena in general, and resultant models of atoms in particular, arises via quantum electrodynamics in which the theory of quantum mechanics (better: quantum field theory) is melded with Maxwell’s theory and then used to build models of interacting charges, atoms, *etc.*

After all these general, philosophical-sounding statements it is time to get down to business. What does it mean to have a set of “rules that can be used to ‘explain’ and/or predict the outcome of experiments”? Of course useful theories are necessarily somewhat intricate, but taking a very coarse-grained view of things, the structure of a generic theory can be characterized in a pretty simple way. Basically, one can view a theory as a way of

describing *observables, states, and dynamics*. Below we will describe each of these three. As you know, mathematics has, for some centuries now, been the language and tool of choice in building a physical theory. Consequently, one of our principal tasks in this course will be to give a mathematical representation for these 3 basic ingredients.

Observables

Measurable aspects of the experimentally accessible world are the *observables*. Any theory is to provide a means of assigning a mathematical representation to the observables. By specifying the observables, we are going a long way toward specifying the kinds of physical situations that our theory is meant to cover. Quantum mechanics postulates a universal ground rule for observables: they must be self-adjoint operators on a Hilbert space. The way in which we implement this rule may vary from physical model to physical model. In fact, one physical model may admit several equivalent ways of representing observables as operators.

As another example of observables, recall that using the theory of Newtonian mechanics we can build a model of a “particle” in which all observables can be viewed as functions of position and momentum, which we can call the *basic observables*. Mathematically, the basic observables for a single particle in Newtonian theory are represented by 6 numbers, (x, y, z, p_x, p_y, p_z) , which can be represented as a pair of vectors in Euclidean space. The observables are then functions on the six-dimensional *phase space*. More generally, given any dynamical system one can view the observables as being functions on the phase space. Other quantities, like mass, electric charge, and time are in some sense “observable” too, but in Newtonian mechanics these quantities appear as parameters in various equations, not as quantities which one measures to document the behavior of the system. In other words, while we may consider the way in which the position of a particle changes, we normally don’t include in our model of a “particle” a time-varying mass or electric charge. Of course, more sophisticated models of matter may attempt to give a better, or more “fundamental” description of mass and electric charge in which these quantities become observables in the sense described above.

As another example, consider the electromagnetic field as it is described in Maxwell’s theory. Observables are, of course, the electric and magnetic field strengths. Also we have, polarization of waves, energy density, momentum density, *etc.* All electromagnetic observables (in Maxwell’s theory) are built from electric and magnetic field strengths, which are the basic observables of the theory. Mathematically, they are represented as vector *fields*. Another measurable quantity is the speed of light c . Like mass and charge in Newtonian mechanics, this quantity appears as a parameter in Maxwell’s theory and is not something that can change with the configuration of the system. In the context of classical electrodynamics we don’t use the term “observable” for the speed of light in the

language we are developing.

As my final example, consider the theory of bulk matter known as “statistical mechanics”. This theory has a number of similarities with quantum mechanics. For now, let us note that some of the observables are things like free energy, entropy, critical exponents, *etc.* All of these quantities can be computed from the partition function, which is, in some sense, the basic observable. Of course, statistical mechanics is normally built from classical and/or quantum mechanics in which case the partition function itself is built from more basic observables. By the way, temperature is, of course, a measurable quantity. But it plays the role of a parameter in statistical mechanics in a canonical ensemble and so in this case temperature is treated as a parameter just like mass and electric charge in classical electrodynamics.

So, we see that different physical phenomena require different kinds of observables, and different theories use different mathematical representations for the observables. One of our two main goals this semester is to get a solid understanding of how quantum mechanics represents observables.

Let us remark that in all of our examples — indeed, in most theories — the observable called “time” enters as an adjustable parameter. It is normally *not* modeled as an observable in the same sense as we model, say, the energy of a particle. In quantum mechanics time is not an observable in the sense described above.

States

A physical system can be in a variety of “configurations”, that is, it can display a variety of possible values for its observables. When we speak of the *state* of the system we are referring to a mathematical object which can completely characterize the outcomes of all possible measurements of the observables of the system (at a given time). Thus the mathematical representation of state is intimately tied to the representation of observables. Let us illustrate the notion of “state” with our previous examples.

In quantum mechanics the state of the system is usually characterized by a (unit) vector from the Hilbert space of vectors upon which the operator-observables are acting. Using this vector one can compute the probabilities for the outcomes of any measurement via various mathematical manipulations involving the operator-observables. Of course, we will be spending a lot of time developing this framework.

In the Newtonian mechanics of a particle, the state of the system can be specified by giving a point in phase space. Indeed, all other observables are functions of these observables (and the time) and so once one knows the point in phase space the system occupies, one can say what any measurement will yield.

In electromagnetic theory, the state of affairs is quite similar. In source-free electrodynamics, the state of the system is equivalent to specifying the electric and magnetic fields (at a given time).

In statistical mechanics we have a more subtle way of defining the state, which is similar in spirit to what happens in quantum mechanics. To be explicit, let us focus on “classical statistical mechanics” of a system of particles. From the point of view of classical mechanics, the underlying, basic observables can be viewed as the coordinates and momenta of all the particles making up the system (the phase space) and specifying a point in phase space determines the state of the system. In statistical mechanics the state is specified by giving a probability distribution on the phase space (which is a single function), rather than a point in phase space as is done in Newtonian mechanics. From the probability distribution one can compute all observables of statistical mechanics.

To get a better handle on the distinction between states and observables, you can think as follows. The observables are the aspects of the system which, in principle, are experimentally accessible. The state of a system characterizes the status of the observables. Often time the state reflects the way the system has been “prepared” (think “initial conditions”). Given a particular preparation procedure (performed by various measurements and/or filtering processes) the system will behave in a particular – indeed, unique – way, as reflected in the behavior of its observables. A physical model for a system principally involves an identification of the observables needed to describe the system. This is done once and for all. The states of the system represent various ways the system can be “started off” and can be adjusted by experimental procedures.

Dynamics

The measured values of observables of a system will usually change in time; this is *dynamics*. Normally, a theory will contain a means of describing time evolution of the system, that is, a “dynamical law”, or a “law of motion”. Assuming that we use a time-independent mathematical model for the observables, we can view dynamics as a continuous change (in time) of the state of the system according to some system of equations. This way of formulating dynamics is what is often called the *Schrödinger picture* of dynamics, and a famous example of the dynamical law is provided by the Schrödinger equation. In statistical mechanics, the state of the system is determined by a probability distribution on phase space. In the Schrödinger picture this distribution evolves in time according to the Liouville equation.

In classical mechanics the state of the system is specified by selecting a point in phase space. The dynamical law is then determined, *e.g.*, by a choice of Hamiltonian via Hamilton’s equations. By solving Hamilton’s equations (with some choice of initial state) one

determines a curve in phase space which can be viewed as the state at a given time. Similarly, in electrodynamics, one can view the Maxwell equations as determining a one parameter family of electric and magnetic fields and hence the state as a function of time.

One important aspect of dynamics that is usually incorporated in any theory is a very basic notion of causality. In the Schrödinger picture, given the state of the system at one time, the dynamical law should determine the state uniquely at any other time. Granted this, you see that the state at a given time (along with the dynamical law - which is part of the specification of the theory/model) will determine the outcomes of all measurements *at any time*.

In general, the only thing that ultimately matters is what I will call the “physical output” of the theory – the combination of states and observables which produce a verifiable prediction about a measurement of the system. The physical output will, in general, vary with time and one can mathematically represent this time variation by a change in the state – this is the Schrödinger picture. One can also represent time evolution by changing the representation of the observables in time, keeping the state fixed; this is the Heisenberg picture. Other pictures are possible.

To summarize: A theory requires (1) A mathematical representation of observables; (2) A mathematical representation of states and a prescription for determining the values of the observables — the physical output of the theory — from any given state; (3) A specification of a dynamical law, which tells us how to extract physical output as a function of time. Our goal this semester will be to see how quantum mechanics takes care of (1), (2), and (3) and uses them to build models of a number of physical systems.

Stern-Gerlach experiment

We now describe an experiment conducted by Stern and Gerlach in the early 1920’s. It gives us a valuable demonstration of the kind of phenomenon that needs quantum mechanics to explain it. It also provides an example of what is probably the simplest possible quantum mechanical model. As you probably know, this experiment involves the property of particles known (perhaps misleadingly) as their *spin*, which is an intrinsic angular momentum possessed by the particles. Note, though, at the time of the experiment, neither intrinsic spin nor quantum mechanics was very well understood! Our goal in studying this important experiment is to introduce the basic rules of quantum mechanics in what is probably the simplest possible mathematical setting.

The Stern-Gerlach experiment amounts to passing a beam of particles through a region with a magnetic field which has the same direction everywhere, but a varying magnitude. Recall that the classical potential energy of interaction of a magnetic moment $\vec{\mu}$ with a magnetic field \vec{B} is $-\vec{\mu} \cdot \vec{B}$. Thus the force (as opposed to the torque) exerted on the

magnetic moment is non-zero if the magnetic field varies in space. We have

$$\vec{F} = \nabla(\vec{\mu} \cdot \vec{B}).$$

If the magnetic field varies only in one direction, then the force is parallel or anti-parallel to that direction, depending upon the orientation of the magnetic moment relative to the fixed direction of \vec{B} . Thus the force experienced by the magnetic moment allows us to measure the component of $\vec{\mu}$ along the direction of \vec{B} ; we simply have to observe which way the particles are deflected. Of course, this reasoning was purely classical, so we are initially confident it will work with macroscopic objects, but it can be justified/explained quantum mechanically. For now we just note that the correct quantum model of the interaction of a magnetic moment with a magnetic field does indeed use the potential energy function shown above, and since the atom is sufficiently massive its motion can be well approximated using a classical mechanics description. Here we have a nice example of how classical, macroscopic reasoning gives us important clues to modeling a quantum mechanical, microscopic system.

Stern and Gerlach passed a beam of silver atoms through such an apparatus. These atoms have a magnetic moment (thanks to the electron spin) and so are deflected by the apparatus. Based upon a classical model of the atom's magnetic moment as coming from motion of the charge distribution, and given that the atoms in the beam have random orientations of their magnetic moments, one expects a continuous deflection of the beam, reflecting a continuous spread in the projections of the magnetic moment along the inhomogeneity axis. Instead, what was observed was that the beam splits into two parts. The explanation for this phenomenon is that the magnetic moment vector $\vec{\mu}$ of the atom is due to the presence of an electron which carries intrinsic angular momentum \vec{S} – its “spin” — and hence a magnetic moment. (Here $\vec{\mu}$ is proportional to \vec{S} : $\mu = \frac{q}{mc}S$) The electron is in an atomic state which does not have any orbital angular momentum, so its motion about the atom does not provide a contribution to the magnetic moment.* In this (initially bizarre) explanation, the electron can be in one of two “spin states” relative to any given direction. More precisely, the projection of the spin of an electron along any axis can only take two values ($\pm\hbar/2$). Such particles are said to have “spin 1/2”, an intrinsic (state-independent) property of the electron. If each atom is randomly selected, one expects that these two alternatives occur with 50-50 probability, and this is what is observed. Half the beam gets deflected in each direction. To explain this discreteness – indeed, two-valued nature – of an angular momentum is one challenge faced by any putative theory. But there is much more...

Using Stern-Gerlach apparatuses, we can measure any component of the magnetic moment – equivalently, the spin vector – of a spin 1/2 particle by aligning the SG magnetic

* Of course, there are 47 electrons in a silver atom. However 46 of them are in a state with no net angular momentum and hence no net contribution to the magnetic moment.

field direction along the axis of interest, then passing the beam of particles through and seeing which way the particles are deflected, corresponding to spin “up” or “down” along that direction. Let us therefore try to model the behavior of the spin vector \vec{S} in such an experiment, ignoring all the other “degrees of freedom” that the atoms might have. Thus the atom is modeled as a “spin 1/2 particle”. Let us call an SG apparatus that measures the spin along an axis characterized by a unit vector \hat{n} “ SG_n ”. Thus the apparatus SG_n measures $\vec{S} \cdot \hat{n}$. The empirical fact is that if you measure $\vec{S} \cdot \hat{n}$ you always get $\pm\hbar/2$. Let us pass a beam of spin 1/2 particles through SG_n and keep, say, only the particles that deflect according to $\vec{S} \cdot \hat{n}$ having the value $+\hbar/2$. If we pass this *filtered* beam through another such SG_n /filter device we see that 100% of the beam passes through. We say that we have “determined the spin along \hat{n} with certainty” for all the particles in the filtered beam. We model this situation by saying that all the particles in the (filtered) beam are in the state $|\vec{S} \cdot \hat{n}, +\rangle$. We say that we have “prepared” many particles all in the same state by passing a beam of particles through an SG apparatus and only keeping those deflected up or down.

Suppose we pass a beam through the apparatus SG_z and only keep one spin projection: “spin up”. We now have many electrons prepared in the state $|S_z, +\rangle$. Let us try to pin down the value of S_x that these electrons possess. Pass the beam (all particles in the state $|S_z, +\rangle$) through another Stern-Gerlach apparatus SG_x . Particles are now deflected according to the projection of their magnetic moments (or spin vectors) along the x direction. What you find in this experiment is that the beam splits in half. This is perfectly reasonable; we have already decided that any component of the spin has just two projections along any given axis. Since there is nothing special about the x or z directions; we should get similar behavior for both. In the SG_z filtered beam we did not “prepare” S_x in any special way, so it is not too surprising that we get the beam to split in half.

Let us continue our investigation as follows. We have passed our beam through SG_z and kept the “spin up” particles. We then pass these spin up particles through SG_x ; let us focus on the beam that gave $\hbar/2$ for the S_x measurement. Therefore, roughly half of the beam that entered the SG_x apparatus is kept, and we now have 1/4 of the original particles left in our prepared beam. After this filtering process we can, if we like, verify the value of S_x using SG_x again; and find all the particles do indeed have $S_x = \hbar/2$. Thus we can represent the state of the particles by $|S_x, +\rangle$.*

Now we have a beam of electrons that have been measured to have the following properties: first, S_z is $+\hbar/2$; second S_x is $+\hbar/2$.[†] Given (1) and (2) above, it is reasonable to believe that the electrons we have kept now have definite values for S_z and S_x since we

* Of course, one naturally prefers to write the state as something like $|S_z, +; S_x, +\rangle$, but we shall see that this is not appropriate.

† We could now go and measure S_y in this doubly filtered beam; you will find that half the beam has spin up along y , half has spin down (exercise). But let us not even bother with this.

have filtered out the only other possibilities. This belief does not agree with experiment! Suppose you go back to check on the value of S_z . Take the beam that came out of SG_z with value $\hbar/2$ and then SG_x with the value $+\hbar/2$ and pass it through SG_z again. You may expect that all of the beam is found to have a value $+\hbar/2$ for S_z , but instead you will find that beam splits in two! This is despite the fact that we supposedly filtered out the spin down components along z .

So, if you measure S_z and get, say, $\hbar/2$, and then you measure it again, you will get $\hbar/2$ with probability one (assuming no other interactions have taken place). If you measure S_z and get, say, $\hbar/2$, then measure S_x and then measure S_z , the final measurement will be $\pm\hbar/2$ with a 50-50 probability. This should get your attention: the values that you can get for the observable S_z in two measurements depends upon whether or not you have determined the value of S_x in between the S_z measurements.

Given this state of affairs, it is hard to make sense of the classical picture in which one imagines the electron to have given, definite values of all its observables, *e.g.*, S_x and S_z . One sometimes says that the measurement of S_x has somehow “disturbed” the value of S_z . This point of view is not incorrect, but is not a perfect description of what is going on. For example, as we shall see, the quantum mechanical prediction is unambiguously independent of the way in which we make the measurements. Nowhere do we really need to know how the SG devices worked. Moreover, the “disturbance” in S_z due to the S_x measurement is not a function of how carefully we make the S_x measurement, that is, one cannot blame the strange behavior as coming from some “experimental error”, the measurements can, ideally, be perfect and we still get the same result. The fact of the matter is that one shouldn’t think of observables (such as S_z and S_x) as having given, fixed, values that “exist” in the object of interest. This may be philosophically a bit sticky (and psychologically a bit disturbing), but it seems to be quite alright as a description of how nature actually works.

If all this seems perfectly reasonable to you, then you probably don’t understand it too well. Our macroscopic experience with matter just doesn’t give any hint that this is the way nature works.

Electrons (and other elementary particles) are not like tiny baseballs following classical trajectories with tiny spin angular momentum arrows attached to them, and there is no reason (experimentally) to believe that they are. It is a purely classical prejudice that a particle has definite values for all observables that we can measure. Try to think this way: what is a particle? It has mass, (total) spin, charge, *etc.* and other intrinsic, “real” properties that do not change with the state of the particle. Based upon experiment, one may want to assign other observable properties such as position, energy, orbital angular momentum, spin component along an axis to the particle. But according to experiment, these properties change with the state of the particle and cannot be viewed as “existing”

in the particle independently of the measuring process (which changes the state). As it turns out, according to the quantum mechanical explanation of this sort of phenomenon, all you are guaranteed to be able to “assign” to a particle is probability distributions for its various observables. Our next task is to build up the quantum mechanical model of the spin $1/2$ system using the rules of quantum mechanics.