Advanced Quantum Theory

AMATH473/673, PHYS454

Achim Kempf

Department of Applied Mathematics
University of Waterloo
Canada

Contents

Introduction

Quantum theory, together with general relativity, represents humanity's so-far deepest understanding of the laws of nature. And quantum phenomena are not rare or difficult to observe. In fact, we experience quantum phenomena constantly! For example, the very stability of the desk at which you are sitting now has its origin in a quantum phenomenon. This is because atoms are mostly empty space and the only reason why atoms don't collapse is due to the uncertainty relations. Namely, the uncertainty relations imply that it costs plenty of momentum (and therefore energy) to compress atoms. Also, for example, the spectrum of sunlight is shaped by quantum effects - if Planck's constant were smaller, the sun would be bluer.

Over the past century, the understanding of quantum phenomena has led to a number of applications which have profoundly impacted society, applications ranging from nuclear power, lasers, transistors and photovoltaic cells, to the use of MRI in medicine. Ever new sophisticated applications of quantum phenomena are being developed, among them, for example, quantum computers which have the potential to revolutionize information processing.

Also on the level of pure discovery, significant progress is currently being made, for example, in the field of cosmology, where both quantum effects and general relativistic effects are important: high-precision astronomical data obtained by satellite telescopes over the past 15 years show that the statistical distribution of matter in the universe agrees with great precision with the distribution which quantum theory predicts to have arisen from quantum fluctuations shortly after the big bang. We appear to originate in quantum fluctuations. New satellite-based telescopes are being planned.

The aim of this course is to explain the mathematical structure of all quantum theories and to apply it to nonrelativistic quantum mechanics. Nonrelativistic quantum mechanics is the quantum theory that replaces Newton's mechanics and it is the simplest quantum theory. The more advanced quantum theory of fields, which is necessary for example to describe the ubiquitous particle creation and annihilation processes, is beyond the scope of this course, though of course I can't help but describe some of it. For example, the first chapter of these notes, up to section 1.5, describes the history of quantum theory as far as we will cover it in this course. The introduction goes on, however, with a historical overview that outlines the further developments, from relativistic quantum mechanics to quantum field theory and on to the modern day quest

for a theory of quantum gravity with applications in quantum cosmology. Quantum theory is still very much a work in progress and original ideas are needed as much as ever!

Note: This course prepares for a number of graduate courses, for example, the graduate course Quantum Field Theory for Cosmology (AMATH872/PHYS785) that I normally teach every other year. I will next teach that course in the term Winter 2018.

Chapter 1

A brief history of quantum theory

1.1 The classical period

At the end of the 19th century, it seemed that the basic laws of nature had been found. The world appeared to be a mechanical clockwork running according to Newton's laws of mechanics. Light appeared to be fully explained by the Faraday-Maxwell theory of electromagnetism which held that light was a wave phenomenon. In addition, heat had been understood as a form of energy. Together, these theories constituted "Classical Physics". Classical physics was so successful that it appeared that theoretical physics was almost complete, the only task left being to add more digits of precision. And so, one of Max Planck's teachers, von Jolly, advised him against a career in physics. But a few years later, classical physics was overthrown.

1.2 Planck and the "Ultraviolet Catastrophe"

The limits to the validity of classical physics first became apparent in measurements of the spectrum of heat radiation. It had been known that very hot objects, such as a smith's hot iron, are emitting light. They do because matter consists of charged particles which can act like little antennas that emit and absorb electromagnetic waves. This means that even relatively cold objects emit and absorb electromagnetic radiation. Their heat radiation is measurable but it is not visible because it too weak and too red for our eyes to see. Black objects are those that absorb electromagnetic radiation (of whichever frequency range under consideration) most easily and by time reversal symmetry they are therefore also the objects that emit electromagnetic heat radiation of that frequency range most readily. For that reason, for example, tea in a black tea pot cools down faster than tea in a white or reflecting tea pot.

Now in the late 1800s, researchers were ready to apply the laws of classical physics to a precise calculation of the radiation spectrum emitted by black bodies. To everybody's surprise the calculations, first performed by Rayleigh and Jeans, predicted far more

emission of waves of short wavelengths (such as ultraviolet) than what experimental measurements seemed to indicate. This was not a subtle discrepancy: the laws of classical physics were found to predict that any object would actually emit an infinite amount of heat radiation in an arbitrarily short time, especially at very short wave lengths.

At first, this was not seen as a reason to doubt the laws of classical physics. It seemed obvious that this nonsensical prediction could only be due to an error in the calculation. Eventually, however, as time passed and nobody appeared to be able to find a flaw in the calculation, the problem became considered serious enough to be called the "ultraviolet catastrophe". Planck decided to look into this problem.

1.3 Discovery of h

From about 1890 to 1900, Planck dedicated himself to thoroughly analyzing all assumptions and steps in the calculations of Rayleigh and Jeans. To his great disappointment and confusion he too did not find an error. In the year 1900, Planck then learned of a new precision measurement of the heat radiation spectrum. Those measurements were precise enough to allow curve fitting. Planck had so much experience with the calculations of the heat radiation that on the same day that he first saw the curve of the heat radiation spectrum he correctly *guessed* the formula for the frequency spectrum of heat radiation, i.e., the formula that is today called Planck's formula. After two further months of trying he was able to derive his formula from a simple but rather radical hypothesis. Planck's hypothesis was that matter cannot radiate energy continually, but only in discrete portions of energy which he called "quanta".

Concretely, Planck postulated that light of frequency f could only be emitted in packets of energy $E_q = hf$, as if light was consisting of particles. He found that the value of this constant, h, must be about 6.6 10^{-34}Kg m^2/s for the prediction of the heat radiation spectrum to come out right. Planck's quantum hypothesis was in clear contradiction to classical physics: light was supposed to consist of continuous waves - after all, light was known to be able to produce interference patterns[1]. Nevertheless, most researchers, including Planck himself, still expected to find an explanation of his quantum hypothesis within classical physics.

[1]It is easy to see these interference patterns: in a dark room, have a candle burning on a desk, then sit a few meters away from it. Close one of your eyes and hold a hair in front of your other eye, about 1cm in front of the eye, vertically. Align the hair with the flame of the candle. Do you see an interference pattern, i.e., the flame plus copies of it to its left and right? From the apparent distance between the copies of the flame and the distance of the hair to the flame you can work out the ratio of the thickness of the hair to the wavelength of the light.

1.4 Mounting evidence for the fundamental importance of h

The significance of Planck's constant was at first rather controversial and even Planck was hesitant to view his discretization of energy hypothesis as anything more than a mathematical trick that would presumably find some explanation within a continuum theory eventually. Einstein, however, was prepared to take Planck's finding at face value. In 1906, Einstein succeeded in quantitatively explaining the photoelectric effect[2]. Then, he reasoned, the light's energy packets must be of high enough energy and therefore of high enough frequency to be able to free electrons from the metal. For irrational reasons, Einstein's explanation of the photoelectric effect is the only result for which he was awarded a Nobel prize.

At about the same time, work by Rutherford and others had shown that atoms consist of charged particles which had to be assumed to be orbiting another. This had led to another deep crisis for classical physics: If matter consisted of charged particles that orbit another, how could matter ever be stable? When a duck swims in circles in a pond, it continually makes waves and the production of those waves costs the duck some energy. Similarly, an electron that orbits a nucleus should continually create electromagnetic waves. Just like the duck, also the electron should lose energy as it radiates off electromagnetic waves. A quick calculation showed that any orbiting electron should rather quickly lose its energy and therefore fall into the nucleus.

Finally, in 1913, Bohr was able to start explaining the stability of atoms. However, to this end he too had to make a radical hypothesis involving Planck's constant h: Bohr hypothesized that, in addition to Newton's laws, the orbiting particles should obey a strange new equation. The new equation says that a certain quantity calculated from the particle's motion (the so called "action"), can occur only in integer multiples of h. In this way, only certain orbits would be allowed. In particular, there would be a smallest orbit of some finite size, and this would be the explanation of the stability of atoms. Bohr's hypothesis also helped to explain another observation which had been made, namely that atoms absorb and emit light preferably at certain discrete frequencies.

1.5 The discovery of quantum theory

Planck's quantum hypothesis, Einstein's light quanta hypothesis and Bohr's new equation for the hydrogen atom all contained Planck's h in an essential way, and none of this could be explained within the laws of classical physics. Physicists, therefore, came

[2]Under certain circumstances light can kick electrons out of a metal's surface. Classical physics predicted that this ability depends on the brightness of the light. Einstein's quantum physics correctly explained that it instead depends on the color of the light: Einstein's radical idea was that light of frequency ω comes in quanta, i.e., in packets of energy $\hbar\omega$

to suspect that the laws of classical physics might have to be changed according to some overarching new principle, in which h would play a crucial role. The new physics would be called quantum physics. The theoretical task at hand was enormous: One would need to find a successor to Newton's mechanics, which would be called quantum mechanics. And, one would need to find a successor to Faraday and Maxwell's electromagnetism, which would be called quantum electrodynamics. The new quantum theory would have to reproduce all the successes of classical physics while at the same time explaining in a unified way all the quantum phenomena, from Planck's heat radiation formula, to the stability and the absorbtion and emission spectra of atoms.

The task took more than twenty years of intense experimental and theoretical research by numerous researchers. Finally, in 1925, it was Heisenberg who first found "quantum mechanics", the successor to Newton's mechanics. (At the time, Heisenberg was a 23 year old postdoctoral fellow with a Rockefeller grant at Bohr's institute in Copenhagen). Soon after, Schrödinger found a seemingly simpler formulation of quantum mechanics which turned out to be equivalent. Shortly after, Dirac was able to fully clarify the mathematical structure of quantum mechanics, thereby revealing the deep principles that underlie quantum theory. Dirac's textbook "Principles of Quantum Mechanics" is a key classic.

The new theory of "Quantum Mechanics", being the successor to Newton's mechanics, correctly described how objects move under the influence of electromagnetic forces. For example, it described how electrons and protons move under the influence of their mutual attraction. Thereby, quantum mechanics explained the stability of atoms and the details of their energy spectra. In fact, quantum mechanics was soon applied to explain the periodic table and the chemical bonds.

What was still needed, however, was the quantum theory of those electromagnetic forces, i.e., the quantum theoretic successor to Faraday and Maxwell's electromagnetism. Planck's heat radiation formula was still not explained from first principles! Fortunately, the discovery of quantum mechanics had already revealed most of the deep principles that underlie quantum theory. Following those principles, Maxwell's theory of electromagnetism was "quantized" to arrive at quantum electrodynamics so that Planck's formula for the heat radiation spectrum could be derived.

It then became clear that quantum mechanics, i.e., the quantization of classical mechanics, was merely the starting point. Somehow, quantum mechanics would have to be upgraded to become consistent with the brand new theory of relativity which Einstein had discovered! And then it would have to be covariantly combined with the quantization of electrodynamics in order to be able to describe both matter and radiation and their interactions.

1.6 Relativistic quantum mechanics

Already by around 1900, Lorentz, Einstein and others had realized that Newton's mechanics was in fact incompatible with Faraday and Maxwell's theory of electromagnetism, for reasons unrelated to quantum theory, thereby contributing to the crisis of classical physics. In a daring move, Einstein accepted Faraday and Maxwell's theory of electromagnetism as correct and questioned the validity of Newton's notion of absolute space and time:

Maxwell was able to calculate the speed of electromagnetic waves from first principles, and found it to match with the measured speed of light. His calculations also showed, however, that a traveller would with some large constant velocity would find the same speed of light. (Today we would say that this is because the Maxwell equations are covariant).

At the time, this was rather surprising as it clearly contradicted Newton's classical mechanics which says that velocities are simply additive. For example, according to Newton, a passenger who walks forward at $v_1 = 5$km/h in a train travelling at $v_2 = 100$km/h has a speed of $v_3 = v_1 + v_2 = 105$km/h relative to the ground. In fact, he does not. His speed to the ground is $v_3 = (v_1 + v_2)/(1 + v_1 v_2/c^2) = 104.9999994...$km/h. Today, the nonadditivity of velocities is an easy-to-measure everyday phenomenon. At the time, the nonadditivity of velocities was first confirmed experimentally by Michelson and Moreley, who compared the speed of two light rays travelling parallel and orthogonal to the motion of the earth around the sun. The new theory that explained it all was of course Einstein's special relativity. By 1916, he developed it into general relativity, which supersedes Newton's laws of gravity. General relativity very elegantly explains gravity as curvature of space-time.

Historically, the discovery of relativity therefore happened more or less simultaneously with the discovery of quantum theory. Yet, the two theories were developed virtually independently of another. In actual experiments, special relativity effects seemed of little importance to quantum mechanical effects and vice versa. For example, it was easy to estimate that an electron which orbits the nucleus of a hydrogen atom would travel at most at speeds smaller than one percent of the speed of light. Also, since gravity is extremely weak compared to the electromagnetic forces that rule the atom it was clear that general relativity would be even less important than special relativity for those early quantum mechanical studies. Conversely, the uncertainty principle appeared irrelevant at the astrophysical scales where general relativity was applied.

Nevertheless, soon after quantum mechanics had been found in 1925 it became apparent that at least the tiny special relativistic effect of the speed of an electron orbiting a nucleus was indeed measurable. This meant that there was experimental guidance for the development of an improved version of quantum mechanics that would be compatible with special relativity. Indeed, Klein, Gordon, Dirac and others soon

developed "relativistic quantum mechanics[3]". Dirac's analysis, in particular, led him to correctly predict surprising magnetic properties of electrons, and it led him to correctly predict the existence and properties of antiparticles such as the positron!

However, the fact that particles are able to create and annihilate another in collisions, which had clearly been observed, was beyond the power of even relativistic quantum mechanics. It was clear that a significant enlargement of the framework of quantum theory was needed.

1.7 Quantum field theory

The way forward was called "second quantization". The starting observation was that, in quantum mechanics, the wave functions behave completely deterministically, namely according to the Schrödinger equation. Given the initial wave function, one can calculate its evolution with absolute certainty. It was felt that to be able to predict the evolution of something, here the wavefunction, with absolute certainty was unusual for a quantum theory. The idea of second quantization was, therefore, to apply quantum theory to quantum theory itself. To this end, the quantum mechanical wave functions were to be treated as classical fields, much like the classical electromagnetic fields. Then, the aim was to find the quantum version of those fields. Since quantum theory was to be applied to the wave functions themselves, the amplitudes of wave functions would no longer be numbers but they would be operators instead. (An operator is a linear map on an infinite dimensional vector space). As a consequence, in quantum field theory, the amplitudes of the wave functions would be subject to uncertainty relations. One should not be able to be sure of the values of the wave function, nor should one be able to be sure of the norm of the wave function. Since in quantum mechanics the normalization of the wave function to norm one means that there is exactly one particle, somewhere, i.e., one would with second quantization not necessarily be sure how many particles there are. Roughly speaking, it is in this way that the quantum fluctuations of the wave functions themselves would then account for the creation and annihilation of particles[4].

The problem of finding a quantum theory for fields had of course already been encountered when one had first tried to find the quantum theoretic successor to Faraday and Maxwell's electrodynamics (which was consistent with special relativity from the start). As it turned out, guided by the general principles underlying quantum mechanics the quantum theory of the electromagnetic fields alone was not too hard to find. Following these lines, one was eventually able to write down a unifying quantum theory both of charged particles and their antiparticles, and also of their interaction through

[3]This relativistic quantum mechanics is an improvement of quantum mechanics which is consistent merely with *special* relativity. The search for a quantum theory that is consistent also with general relativity is still on today.

[4]Yes, third and higher quantization has been considered, but with no particular successes so far.

electromagnetic quanta, i.e., photons. While this theory succeeded well in describing all the interactions, including annihilation and creation processes, it did yield much more than one had bargained for. The reason was that, since now particle number was no longer conserved, the time-energy uncertainty principle made it possible for short time intervals that energy (and therefore all kinds of particles) could be virtually "borrowed" from the vacuum.

As a consequence, the new quantum field theory, called quantum electrodynamics, necessarily predicted that, for example, that an electron would sometimes spontaneously borrow energy from the vacuum to emit a photon which it then usually quickly reabsorbs. During its brief existence, this so-called "virtual" photon even has a chance to split into a virtual electron-positron pair which shortly after annihilates to become the virtual photon again. In fact, the virtual electron (or the positron) during its short existence, might actually emit and quickly reabsorb a virtual photon. That photon, might briefly split into an electron positron pair, etc etc ad infinitum. Even more intriguing is that even without a real electron to start with, the vacuum alone is predicted to have virtual particles continually appearing and disappearing!

In spite of all this new complexity, it turned out that the theory's predictions for the very simplest interactions and creation-annihilation processes were in very good agreement with experimental results. However, the calculation of those predicted endless chain reactions of virtual processes typically yielded divergent integrals! To take those virtual processes into account should have increased the precision of predictions. Instead, one only obtained seemingly meaningless predictions of infinite numbers. It took the combined efforts of numerous scientists, such as Feynman, Tomanaga, Weisskopf, Dyson and others, over about twenty years, to solve this problem.

It turned out that those calculations that had yielded infinities did make sense after all, if one suitably recalibrated the parameters of the theory, such as the fundamental masses and charges. This process of recalibration, called renormalization, also occurs in condensed matter physics, where it is easier to understand intuitively: Consider an electron that is traveling through a crystal. It has the usual mass and charge. But if you want to influence the electron's motion you will find that the traveling electron behaves as if it had a several times larger mass and a smaller charge. That's because the electron slightly deforms the crystal by slightly displacing the positive and negative charges that it passes by. It is these deformations of the crystal, which travel with the electron, which make the electron behave as if it were heavier and they also shield its charge. Also, the closer we get to the electron with our measurement device, the less is its charge shielded, i.e., the more we see of the bare charge of the electron.

The key lesson here is that the masses and charges that one observes in a crystal are generally not the "bare" masses and charges that the particles fundamentally possesses. The observed masses and charges even depend on how closely one looks at the electron.

Now when fundamental particles travel through the vacuum, then they deform the distribution of those virtual particles that pop in and out of existence due to the time-energy uncertainty principle. Again, this makes particles behave as if they had

a different mass and a different charge. The masses and charges that are observed are not the "bare" masses and charges that the particles fundamentally possess. The observed masses and charges actually depend again on how closely one looks at the particles, i.e., at what energy one observes them, say with an accelerator. In quantum field theory, it turns out that the bare masses and charges may formally even tend to zero or be divergent. This is OK, as long as the predicted measured values come out right.

Technically, if you would like to know the gist of it already, renormalization consists of the following steps: First, artificially render all predictions finite, say by cutting of the divergent integrals. It turned out that this can be achieved by postulating the existence of a smallest possible distance ϵ between any two particles and by calculating virtual processes accordingly. Next, adjust the parameters of the theory (charges, masses etc) such that a handful of predictions come out in agreement with experiment (namely as many as there are free parameters such as masses and charges in the theory). Now let $\epsilon \to 0$, while at the same time letting the bare parameters of the theory run so that the same handful of predictions comes out right. (The parameters of the theory will thereby usually tend to 0 or ∞.) Crucially, all other (infinitely many!) possible predictions of the theory will now also come out finite in the limit $\epsilon \to 0$ - and they can be compared to experiment. Indeed, predictions so-obtained through renormalization, for example for the energy levels of the hydrogen atom, match the experimental results to more than a dozen digits behind the comma!

Of course, renormalization has always been seen as mathematically and conceptually unsatisfactory. Nevertheless, it did open the door to the successful application of quantum field theory for the description of all the many species of particles that have been discovered since, from neutrinos and muons to quarks.

It is important also to mention two developments related to quantum field theory: First, on the applied side, it turned out that quantum field theoretic methods can also be used for the description of wave phenomena in solids. These are crucial, for example, for the understanding of superconductivity. Second, on the theoretical side, Feynman in his work on quantum electrodynamics, found an equivalent but very insightful and mathematically powerful new formulation for the principles of quantum theory, called the path integral formulation. I will briefly outline the path integral formulation of quantum mechanics later in this course.

1.8 Beyond quantum field theory?

Today, quantum field theory has served as the basis of elementary particle physics (and therefore as the basis for the description of all that we are made of) for about fifty years. Even though numerous new particles and even new forces have been discovered over the years, quantum field theory itself never needed to undergo any fundamental changes. Similarly successful has been Einstein's general relativity, which has now served as

the basis of all gravitational physics for over 80 years. Even the most sophisticated experiments seem to merely confirm the validity of quantum field theory and general relativity with more and more precision.

Could it be, therefore, that these two theories constitute the final laws of nature and that this is all there is? Should one discourage students from a career in the subject? Certainly not! In fact, the situation resembles in many ways the situation at the time Planck was a student. We have two highly successful theories - but they are inconsistent! As long as we consider gravity to be a fixed background for quantum theory some calculations can be performed. Hawking's prediction of black hole radiation is of this kind. However, once we fully take into account the dynamics of general relativity, we face a problem: The predictions of infinities in quantum field theory appear to persist. In the renormalization procedure, the limit $\epsilon \to 0$ does no longer seem to work (not for lack of trying!).

This problem is very deep. Many believe that this indicates that there actually exists a finite shortest length, ϵ, in nature, much like there is a finite fastest speed. Indeed, if we put together what we know from general relativity and what we know from quantum theory, we can conclude that we cannot even in principle devise an experimental operation that would allow us to resolve distances as small as about $10^{-35}m$, which is the so-called Planck scale:

Consider the task of resolving some very small structure. To this end, we need to shine on it some probing particles of very short wavelength. Due to quantum theory, the shorter the wavelength, the higher is the energy uncertainty of the probing particle. According to general relativity, energy gravitates and curves space. Thus, the probing particles will randomly curve space to the extent of their energy uncertainty. Assume now that a distance of $10^{-35}m$ or smaller is to be resolved. A short calculation shows that to this end the probing particles would have to be of such short wavelength, i.e., of such high energy uncertainty that they would significantly curve and thereby randomly disturb the region that they are meant to probe. It therefore appears that the very notion of distance loses operational meaning at distances of $10^{-35}m$ or so.

In order to describe the structure of space-time and matter at such small scales we will need a unifying theory of quantum gravity. Much effort is being put into this. In this field of research, it is widely expected that within the unified quantum gravity theory there will be a need for renormalization, but not for infinite renormalization. This yet-to-be found theory of quantum gravity may also solve several other major problems of quantum theory. In particular, it could yield an explanation for the particular masses and charges of the elementary particles, and perhaps even an explanation for the statistical nature of quantum theoretical predictions.

A very concrete major problem awaiting resolution in the theory of quantum gravity is the derivation of the cosmological constant, which represents the energy of the vacuum. Quantum field theory predicts the vacuum to possess significant amounts of energy due to vacuum fluctuations: Each field can be mathematically decomposed into a collection of quantum theoretical harmonic oscillators, each of which contributes a

finite ground state energy of $\hbar\omega/2$. General relativity predicts that the vacuum energy should gravitate, just like any other form of energy.

Evidence from recent astronomical observations of the expansion rate of the universe indicates that the cosmological constant has a small but nonzero value. How much vacuum energy does quantum field theory predict? Straightforwardly, quantum field theory predicts the vacuum energy density to be infinite. If we augment quantum field theory by the assumption that the Planck length is the shortest length in nature, then quantum field theory predicts a very large vacuum energy. In fact, it is by a factor of about 10^{120} larger than what is experimentally observed. This is the today's "ultraviolet catastrophe". It appears that whoever tries to reconcile quantum theory with general relativity must be prepared to question the very foundations of all we know of the laws of nature. Original ideas are needed that may be no less radical than those of Planck or Einstein. Current attempts are, for example, string theory and loop quantum gravity.

1.9 Experiment and theory

In the past, progress in the search for the theory of quantum gravity has been severely hampered by the fact that one cannot actually build a microscope with sufficiently strong resolving power to probe Planck scale physics. Even the best microscopes today, namely particle accelerators, can resolve distances only down to at most 10^{-20}m, which is still very far from the Planck scale of $10^{-35}m$. Of course, guidance from experiments is not strictly necessary, as Einstein demonstrated when he first developed general relativity. Nevertheless, any candidate theory must be tested experimentally before it can be given any credence.

In this context, an important recent realization was that there are possibilities for experimental access to the Planck scale other than through accelerators! One possibility could be the study of the very highly energetic cosmic rays that occasionally hit and locally briefly light up the earth's atmosphere. Another recently much discussed possibility arises from the simple fact that the universe itself was once very small and has dramatically expanded since. The idea is, roughly speaking, that if the early expansion was rapid enough then the universe might have acted as a microscope by stretching out everything to a much larger size. Astronomical evidence obtained over the past few years indicate that this did happen.

The statistical distribution of matter in the universe is currently being measured with great precision, both by direct observation of the distribution of galaxies, and through the measurement of the cosmic microwave background. Experimental evidence is mounting for the theory that the matter distribution in the universe agrees with what one would expect if it originated as tiny primordial quantum fluctuations - which were inflated to cosmic size by a very rapid initial expansion of the universe! It appears that the universe itself has acted as a giant microscope that enlarged initially small

quantum phenomena into an image on our night sky! It is just possible that even the as yet unknown quantum phenomena of Planck length size have left an imprint in this image. Some of my own research is in this area. New satellite based telescopes are currently further exploring these phenomena.

Chapter 2

Classical mechanics in Hamiltonian form

2.1 Newton's laws for classical mechanics cannot be upgraded

When physicists first tried to find the laws of quantum mechanics they knew from experiments that Planck's constant h would have to play an important role in those laws. Imagine yourself in the situation of these physicists. How would you go about guessing the laws of quantum mechanics? Clearly, quantum mechanics would have to strongly resemble classical mechanics. After all, quantum mechanics should be an improvement over classical mechanics. Thus, it would have to reproduce all the successful predictions of classical mechanics, from the motion of the planets to the forces in a car's transmission. So how if we try to carefully improve one or several Newton's three axioms of classical mechanics by suitably introducing Planck's constant?

For example, could it be that $F = ma$ should really be $F = ma + h$ instead? After all, h is such a small number that one could imagine that this correction term might have been overlooked for a long time. However, this attempt surely can't be right on the trivial grounds that h does not have the right units: F and ma have the units Kgm/s^2 while the units of h are Kgm^2/s. But then, could the correct second law perhaps be $F = ma(1 + h/xp)$? The units would match. Also this attempt can't be right because whenever x or p are small, the term h/xp would be enormous, and we could therefore not have overlooked this term for all the hundreds of years since Newton. Similarly, also $F = ma(1 + xp/h)$ can't be right because for the values of x and p that we encounter in daily life the term xp/h would usually be big enough to have been seen.

In fact, no attempt to improve on Newton's laws in such a manner works. This is why historically this search for the laws of quantum mechanics actually took a quarter century! When the first formulations of quantum mechanics were eventually found by

Heisenberg and Schrödinger, they did not at all look similar to classical mechanics.

It was Dirac who first clarified the mathematical structure of quantum mechanics and thereby its relation to classical mechanics. Dirac remembered that a more abstract formulation of classical mechanics than Newton's had long been developed, namely Hamiltonian's formulation of classical mechanics. Hamilton's formulation of classical mechanics made use of a mathematical tool called Poisson brackets. Dirac showed that the laws of classical mechanics, once formulated in their Hamiltonian form, can be repaired by suitably introducing h into its equations, thereby yielding quantum mechanics correctly. In this way, Dirac was able to show how quantum mechanics naturally supersedes classical mechanics while reproducing the successes of classical mechanics. We will follow Dirac in this course[1].

2.2 Levels of abstraction

In order to follow Dirac's thinking, let us consider the levels of abstraction in mathematical physics: Ideally, one starts from abstract laws of nature and at the end one obtains concrete number predictions for measurement outcomes. In the middle, there is usually a hierarchy of mathematical problems that one has to solve.

In particular, in Newton's formulation of classical mechanics one starts by writing down the equations of motion for the system at hand. The equations of motion will generally contain terms of the type $m\ddot{x}$ and will therefore of the type of differential equations. We begin our calculation by solving those differential equations, to obtain functions. These functions we then solve for variables. From those variables we eventually obtain some concrete numbers that we can compare with a measurement value. The hierarchy of abstraction is, therefore:

Differential equations
$\Downarrow$
Functions
$\Downarrow$
Variables
$\Downarrow$
Numbers

This begs the question if there is a level of abstraction above that of differential equations? Namely, can the differential equations of motion be obtained as the *solution* of

[1]In his paper introducing the Schrödinger equation, Schrödinger already tried to motivate his equation by an analogy with some aspect of Hamilton's work (the so called Hamilton Jacobi theory). This argument did not hold up. But, another part of Schrödinger's intuition was right on: His intuition was that the discreteness in quantum mechanics (e.g., of the energy levels of atoms and molecules) has its mathematical origin in the usual discreteness of the resonance frequencies of wave phenomena.

some higher level mathematical problem? The answer is yes, as Dirac remembered: Already in the first half of the 19th century, Lagrange, Hamilton and others had found this higher level formulation of classical mechanics. Their methods had proven useful for solving the dynamics of complicated systems, and some of those methods are still being used, for example, for the calculation of satellite trajectories. Dirac thought that if Newton's formulation of classical mechanics was not upgradable, it might be worth investigating if the higher level formulation of Hamilton might be upgradable to obtain quantum mechanics. Dirac succeeded and was thereby able to clearly display the similarities and differences between classical mechanics and the quantum mechanics of Heisenberg and Schrödinger. To see this is our first goal in this course.

Remark: For completeness, I should mention that there are two equivalent ways to present classical mechanics on this higher level of abstraction: One is due to Hamilton and one is due to Lagrange. Lagrange's formulation of classical mechanics is also upgradable, i.e., that there is a simple way to introduce h to obtain quantum mechanics from it, as Feynman first realized in the 1940s. In this way, Feynman discovered a whole new formulation of quantum mechanics, which is called the path integral formulation. I will explain Feynman's formulation of quantum mechanics later in the course.

2.3 Classical mechanics in Hamiltonian formulation

2.3.1 The energy function H contains all information

What was Hamilton's higher level of abstraction? How can classical mechanics be formulated so that Newton's differential equations of motion are themselves the solution of a higher level mathematical problem? Hamilton's crucial observation was the following: the expression for the total energy of a system already contains the complete information about that system! In particular, if we know a system's energy function, then we can derive from it the differential equations of motion of that system. In Hamilton's formulation of classical mechanics the highest level description of a system is therefore through its energy function. The expression for the total energy of a system is also called the Hamiltonian. The hierarchy of abstraction is now:

Hamiltonians
⇓
Differential equations
⇓
Functions
⇓
Variables
⇓
Numbers

As a very simple example, let us consider a system of two point[2] masses, m_1 and m_2, which are connected by a spring with spring constant k. We write their respective position vectors as $\vec{x}^{(r)} = (x_1^{(r)}, x_2^{(r)}, x_3^{(r)})$ and their momentum vectors as $\vec{p}^{(r)} = (p_1^{(r)}, p_2^{(r)}, p_3^{(r)})$, where r is 1 or 2 respectively (we will omit the superscript $^{(r)}$ when we talk about one mass only). The positions and momenta are of course functions of time. Let us, therefore, keep in mind that for example $x_3^{(1)}$ is just a short hand notation for the function $x_3^{(1)}(t)$. Since this is a simple system, it is easy to write down its equations of motion:

$$\frac{d}{dt}\, x_i^{(r)} = \frac{p_i^{(r)}}{m_r} \tag{2.1}$$

$$\frac{d}{dt}\, p_i^{(1)} = -k(x_i^{(1)} - x_i^{(2)}) \tag{2.2}$$

$$\frac{d}{dt}\, p_i^{(2)} = -k(x_i^{(2)} - x_i^{(1)}) \tag{2.3}$$

Here, $r \in \{1, 2\}$ labels the objects and $i \in \{1, 2, 3\}$ labels their coordinates. Hamilton's great insight was that these equations of motion (as well as those of arbitrarily complicated systems) can all be *derived* from just one piece of information, namely the expression for the system's total energy H alone! This is to say that Hamilton discovered that the expression for the total energy is what we now call the generator of the time evolution. The Hamiltonian H, i.e., the total energy of the system, is the kinetic energy plus the potential energy. In our example:

$$H = \frac{\left(\vec{p}^{(1)}\right)^2}{2m_1} + \frac{\left(\vec{p}^{(2)}\right)^2}{2m_2} + \frac{k}{2}\left(\vec{x}^{(1)} - \vec{x}^{(2)}\right)^2 \tag{2.4}$$

Here, $\left(\vec{p}^{(1)}\right)^2 = \sum_{i=1}^{3} \left(p_i^{(1)}\right)^2$ etc. Now imagine that the system in question is instead a complicated contraption with plenty of wheels, gears, discs, levers, weights, strings, masses, bells and whistles. Using Newton's laws it is possible to determine the equations of motion for that system but it will be complicated and will typically involve drawing lots of diagrams with forces. Hamilton's method promises a lot of simplification here. We just write down the sum of all kinetic and potential energies, which is generally not so difficult, and then Hamilton's methods should yield the equations of motion straightforwardly. In practice we won't be interested in complicated contraptions. We'll be interested in systems such as molecules, quantum computers or quantum fields, which all can be quite complicated too.

[2] In this course, we will always restrict attention to point masses: all known noncomposite particles, namely the three types of electrons and neutrinos, six types of quarks, the W and Z particles (which transmit the weak force responsible for radioactivity), the gluons (which transmit the strong force responsible for the nuclear force) and the photon are all point-like as far as we know.

But what is the technique with which one can derive the equations of motion from a Hamiltonian, for example, Eqs.2.1-2.3 from Eq.2.4? Exactly how does the generator, H, of the time evolution generate the time evolution equations Eqs.2.1-2.3?

2.3.2 The Poisson bracket

The general procedure by which the equations of motion can be derived from a Hamiltonian H requires the use of a powerful mathematical operation, called "Poisson bracket"[3]:

The Poisson bracket is a particular kind of multiplication: Assume that f and g are polynomials in terms of the positions and momenta of the system, say $f = -2p_1$ and $g = 3x_1^2 + 7p_3^4 - 2x_2^3p_1^3 + 6$. Then, the Poisson bracket of f and g is written as $\{f, g\}$ and the evaluation of the bracket will yield another polynomial in terms of the position and momenta of the system. In this case:

$$\{-2p_1\ ,\ 3x_1^2 + 7p_3^4 - 2x_2^3p_1^3 + 6\} = 12x_1 \tag{2.5}$$

But how does one evaluate such a Poisson bracket to obtain this answer? The rules for evaluating Poisson brackets are tailor-made for mechanics. There are two sets of rules:

A) By definition, for each particle, the Poisson brackets of the positions and momenta are:

$$\{x_i, p_j\} = \delta_{i,j} \tag{2.6}$$
$$\{x_i, x_j\} = 0 \tag{2.7}$$
$$\{p_i, p_j\} = 0 \tag{2.8}$$

for all $i, j \in \{1, 2, 3\}$. Here, $\delta_{i,j}$ is the Kronecker delta, which is 1 if $i = j$ and is 0 if $i \neq j$. But these are only the Poisson brackets between linear terms. How to evaluate then the Poisson bracket between two polynomials? The second set of rules allow us to reduce this general case to the case of the Poisson brackets between linear terms:

B) By definition, the Poisson bracket of two arbitrary expressions in the positions and momenta, $f(x, p)$ and $g(x, p)$, obey the following rules:

$$\{f, g\} = -\{g, f\} \qquad \text{antisymmetry} \tag{2.9}$$
$$\{cf, g\} = c\,\{f, g\}, \quad \text{for any number } c \qquad \text{linearity} \tag{2.10}$$
$$\{f, g + h\} = \{f, g\} + \{f, h\} \qquad \text{addition rule} \tag{2.11}$$

[3]Remark: In elementary particle physics there is a yet higher level of abstraction, which allows one to *derive* Hamiltonians. The new level is that of so-called "symmetry groups". The Poisson bracket operation plays an essential role also in the definition of symmetry groups. (Candidate quantum gravity theories such as string theory aim to derive these symmetry groups from a yet higher level of abstraction which is hoped to be the top level.)

$$\{f, gh\} = \{f, g\}h + g\{f, h\} \qquad \text{product rule} \tag{2.12}$$
$$0 = \{f, \{g, h\}\} + \{h, \{f, g\}\} + \{g, \{h, f\}\} \qquad \text{Jacobi id.} \tag{2.13}$$

Let us postpone the explanation for why these definitions had to be chosen in exactly this way[4]. For now, note that an immediate consequence of these rules is that the Poisson bracket of a number always vanishes:

$$\{c, f\} = 0 \qquad \text{if } c \text{ is a number} \tag{2.14}$$

The point of the second set of rules is that we can use them to successively break down the evaluation of a Poisson bracket like that of Eq.2.5 into sums and products of expressions that can be evaluated by using the first set of rules, Eqs.2.6,2.7,2.8. Using the product rule we immediately obtain, for example:

$$\{x_3, p_3^2\} = \{x_3, p_3\}p_3 + p_3\{x_3, p_3\} = 1p_3 + p_3 1 = 2p_3 \tag{2.15}$$

Exercise 2.1 *Prove*[5] *Eq.2.14.*

Exercise 2.2 *Show that* $\{f, f\} = 0$ *for any* f*.*

Exercise 2.3 *Assume that* n *is a positive integer.*
a) Evaluate $\{x_1, p_1^n\}$
b) Evaluate $\{x_2^n, p_2\}$

Exercise 2.4 *Verify Eq.2.5.*

Exercise 2.5 *Show that the Poisson bracket is not associative by giving a counter example.*

[4]If the product rule already reminds you of the product rule for derivatives (i.e., the Leibniz rule) this is not an accident. As we will see, the Poisson bracket can in fact be viewed as a sophisticated generalization of the notion of derivative.

[5]This exercise and all exercises up until Sec.2.3.5 are to be solved using only the above axioms for the Poisson bracket. You may have heard of the Poisson bracket before and you may be tempted, for example, to use its representation in terms of derivatives to solve these exercises. Don't do that, use the axioms above. There is an important point here: The abstract, axiomatically-defined Poisson bracket that we defined here using axioms is rather powerful. As Dirac first showed, this abstract, axiomatically-defined Poisson bracket is the exact same in classical and quantum mechanics. As we will see later, the representation of this abstract Poisson bracket in terms of derivatives represents the abstract Poisson bracket only in classical mechanics. We will find that the abstract axiomatically-defined Poisson bracket has a different representation in quantum mechanics that has nothing to do with derivatives. As I said though, the above axioms of the abstract Poisson bracket stay exactly the same also in quantum mechanics.

So far, we defined the Poisson brackets of polynomials in the positions and momenta of one point mass only. Let us now consider the general case of a system of n point masses, $m^{(r)}$ with position vectors $\vec{x}^{(r)} = (x_1^{(r)}, x_2^{(r)}, x_3^{(r)})$ and momentum vectors $\vec{p}^{(r)} = (p_1^{(r)}, p_2^{(r)}, p_3^{(r)})$, where $r \in \{1, 2, ..., n\}$. How can we evaluate the Poisson brackets of expressions that involve all those positions and momentum variables? To this end, we need to define what the Poisson brackets in between positions and momenta of different particles should be. They are defined to be simply zero. Therefore, to summarize, we define the basic Poisson brackets of n masses as

$$\{x_i^{(r)}, p_j^{(s)}\} = \delta_{i,j}\ \delta_{r,s} \tag{2.16}$$

$$\{x_i^{(r)}, x_j^{(s)}\} = 0 \tag{2.17}$$

$$\{p_i^{(r)}, p_j^{(s)}\} = 0 \tag{2.18}$$

where $r, s \in \{1, 2, ..., n\}$ and $i, j \in \{1, 2, 3\}$. The evaluation rules of Eqs.2.9-2.13 are defined to stay just the same.

Exercise 2.6 *Mathematically, the set of polynomials in positions and momenta is an example of what is called a Poisson algebra. A general Poisson algebra is a vector space with two extra multiplications: One multiplication which makes the vector space into an associative algebra, and one (non-associative) multiplication $\{,\}$, called the Lie bracket, which makes the vector space into what is called a Lie algebra. If the two multiplications are in a certain sense compatible then the set is said to be a Poisson algebra. Look up and state the axioms of a) a Lie algebra, b) an associative algebra and c) a Poisson algebra.*

2.3.3 The Hamilton equations

Let us recall why we introduced the Poisson bracket: A technique that uses the Poisson bracket is supposed to allow us to *derive* all the differential equations of motion of a system from the just one piece of information, namely from the expression of the total energy of the system, i.e., from its Hamiltonian.

To see how this works, let us consider an arbitrary polynomial f in terms of the positions and momentum variables $x_i^{(r)}, p_j^{(s)}$ of the system in question, for example, something like $f = 7x_2^{(3)}\left(x_3^{(1)}\right)^3 - 2\cos(4t^2)(p_1^{(1)})^7 + 3/2$. This f depends on time for two reasons: There is an explicit dependence on time through the cosine term, and there is an implicit dependence on time because the positions and momenta generally depend on time. According to Hamilton's formalism, the equation of motion for f is then given by:

$$\frac{df}{dt} = \{f, H\} + \frac{\partial f}{\partial t} \tag{2.19}$$

Here, the notation $\partial f/\partial t$ denotes differentiation of f with respect to only its explicit time dependence. In the example above it is the derivative of the time dependence in the term $\cos(4t^2)$.

Eq.2.19 is a famous equation which is called the Hamilton equation. Why is it famous? If you know how to evaluate Poisson brackets then the Hamilton equation Eq.2.19 encodes for you all of classical mechanics! Namely, given H, equation Eq.2.19 yields the differential equation of motion for any entity f by the simple procedure of evaluating the Poisson bracket on its right hand side.

If f is dependent on time only through x and p (say if we choose for f a polynomial in x and p's with constant coefficients) then $\partial f/\partial t = 0$ and Hamilton's equation simplifies to:

$$\frac{d}{dt}\, f \;=\; \{f, H\} \tag{2.20}$$

In the remainder of this book, unless otherwise specified, we will always choose our functions f, g, h to have no explicit dependence on time, i.e., they will depend on time only implicitly, namely through the time dependence of the x and p's. In particular, the most important choices for f are of this kind: $f = x_i^{(r)}$ or $f = p_i^{(r)}$. For these choices of f we immediately obtain the fundamental equations of motion:

$$\frac{d}{dt}\, x_i^{(r)} \;=\; \{x_i^{(r)}, H\} \tag{2.21}$$

$$\frac{d}{dt}\, p_i^{(r)} \;=\; \{p_i^{(r)}, H\} \tag{2.22}$$

Here is a concrete example: A single free particle of mass m possesses only kinetic energy. Its Hamiltonian is:

$$H = \sum_{j=1}^{3} \frac{p_j^2}{2m} \tag{2.23}$$

By using this H in Eqs.2.21,2.22, we obtain the following equations of motion for the positions and momenta:

$$\frac{d}{dt}\, x_i \;=\; \left\{ x_i \;,\; \sum_{j=1}^{3} \frac{p_j^2}{2m} \right\} \;=\; \frac{p_i}{m} \tag{2.24}$$

and

$$\frac{d}{dt}\, p_i \;=\; \left\{ p_i \;,\; \sum_{j=1}^{3} \frac{p_j^2}{2m} \right\} \;=\; 0 \tag{2.25}$$

They agree with what was expected: $p_i = m\dot{x}_i$ and $\ddot{x}_i = 0$, where the dot indicates the time derivative. For another example, consider again the system of two point masses

m_1, m_2 which are connected by a spring with spring constant k. Its Hamiltonian H was given in Eq.2.4. By using this H in Eqs.2.21,2.22 we should now be able to derive the system's equations of motion (as given in Eqs.2.1-2.3). Indeed:

$$\frac{d}{dt}\, x_i^{(r)} = \{x_i^{(r)}, H\} \tag{2.26}$$

$$= \frac{p_i^{(r)}}{m_r} \tag{2.27}$$

$$\frac{d}{dt}\, p_i^{(1)} = \{p_i^{(1)}, H\} \tag{2.28}$$

$$= -k(x_i^{(1)} - x_i^{(2)}) \tag{2.29}$$

$$\frac{d}{dt}\, p_i^{(2)} = \{p_i^{(2)}, H\} \tag{2.30}$$

$$= -k(x_i^{(2)} - x_i^{(1)}) \tag{2.31}$$

Let us omit the proof that Hamilton's formulation of classical mechanics always yields the same equations of motion as Newton's.

Exercise 2.7 *Consider $f = gh$, where g and h are some polynomial expressions in the position and momentum variables. There are two ways to calculate df/dt: Either we use the Leibnitz rule, i.e., $\dot{f} = \dot{g}h + g\dot{h}$, and apply Eq.2.20 to both $\dot{g}$ and $\dot{h}$, or we apply Eq.2.20 directly to gh and use the product rule (Eq.2.12) for Poisson brackets. Prove that both methods yield the same result.*

This exercise shows that a property of the derivative on the left hand side of Eq.2.20 determines a rule for how the Poisson bracket had to be defined. In fact, such requirements of consistency are the main reason why the Poisson bracket is defined the way it is.

Exercise 2.8 *Use Eq.2.13 to prove that:*

$$\frac{d}{dt}\, \{f, g\} = \{\dot{f}, g\} + \{f, \dot{g}\} \tag{2.32}$$

2.3.4 Symmetries and Conservation laws

Our reason for reviewing the Hamiltonian formulation of mechanics is that it will be useful for the study of quantum mechanics. Before we get to that, however, let us ask why the Hamiltonian formulation of mechanics was useful for classical mechanics. It was, after all, developed more than half a century before quantum mechanics.

Sure, it was fine to be able to derive all the differential equations of motion from the one unifying equation. Namely, if for simplicity, we are now assuming that $f(x,p)$ is chosen not to have any explicit dependence on time, the universal Hamilton equation which expresses the differential equation of all variables f, is:

$$\dot{f} \;=\; \{f,H\} \tag{2.33}$$

Ultimately, however, one obtains this way just the same equations of motion as Newton's methods would yield. Was there any practical advantage to using Hamilton's formulation of mechanics? Indeed, there is an important practical advantage: The main advantage of Hamilton's formulation of mechanics is that it gives us powerful methods for studying conserved quantities, such as the energy or angular momentum. In general, for complicated systems (such as planetary systems, for example), there can be much more complicated conserved quantities than these. To know conserved quantities usually significantly helps in solving the dynamics of complicated systems. This feature of Hamilton's formulation of mechanics will carry over to quantum mechanics, so studying it here will later help us also in quantum mechanics.

Consider a polynomial f in x and p's with constant coefficients. Then, $\partial f/\partial t = 0$ and Eq.2.33 applies. We can easily read off from Eq.2.33 that *any such f is conserved in time if and only if its Poisson bracket with the Hamiltonian vanishes:*

$$\{f,H\} = 0 \quad \Rightarrow \quad \dot{f} = 0 \tag{2.34}$$

Consider, for example, a free particle. Its Hamiltonian is given in Eq.2.23. We expect of course that its momenta p_i are conserved. Indeed:

$$\dot{p}_i = \left\{ p_i, \sum_{j=1}^{3} p_j^2/2m \right\} = 0 \tag{2.35}$$

For another example, consider a system whose Hamiltonian is any polynomial in x's and p's with constant coefficients. The proof that this system's energy is conserved is now fairly trivial to see (using Axiom Eq.2.9):

$$\dot{H} = \{H,H\} = 0 \tag{2.36}$$

I mentioned that in order to be able to find solutions to the equations of motion of complicated real-life systems it is often crucial to find as many conserved quantities as possible. Here is an example. Consider a 3-dimensional isotropic (i.e., rotation invariant) harmonic oscillator. Because of its symmetry under rotations, it angular momentum is conserved. But this oscillator has actually a much larger symmetry and therefore more conserved quantities. This is because a harmonic oscillator, being of the form x^2+p^2 also possesses rotation symmetry in phase space. I will here only remark that this means that the 3-dimensional isotropic harmonic oscillator possesses $SO(3)$ rotational symmetry as well as a larger $SU(3)$ symmetry.

Powerful methods for discovering symmetries and constructing the implied conserved quantities for arbitrary systems have been developed on the basis of Eq.2.33 and the Poisson bracket. A key technique is that of so-called canonical transformations, i.e., of changes variables for which the Poisson brackets remain the same. You can find these methods in classical mechanics texts under the keywords "canonical transformations" and "Hamilton Jacobi theory".

In fact, Poisson bracket methods reveal a very deep one-to-one correspondence between conserved quantities and so-called symmetries. For example, the statement that an experiment on a system gives the same result no matter when we perform the experiment, is the statement of a "symmetry" which is called time-translation invariance symmetry. In practice, it means that the Hamiltonian of the system does not explicitly depend on time: $\partial H/\partial t = 0$. As we just saw, this implies energy conservation: $dH/dt = 0$.

Similarly, the statement that an experiment on a system gives the same result wherever we perform the experiment is the statement of space-translation symmetry. It implies and is implied by momentum conservation. Further, the statement that an experiment on a system gives the same result whatever the angular orientation of the experiment is the statement of rotation symmetry. It implies and is implied by angular momentum conservation.

These are examples of the so-called "Noether theorem", of Emmy Noether (1882-1935). Noether's theorem plays a crucial role both in practical applications, and in fundamental physics[6]. We will later come back to Noether's theorem.

Exercise 2.9 *Show that in general, $dH/dt = \partial H/\partial t$ and give a nonzero example.*

Exercise 2.10 *Consider the system with the Hamiltonian of Eq.2.4. Show that the total momentum is conserved.*

2.3.5 A representation of the Poisson bracket

In principle, we can evaluate any Poisson bracket $\{f, g\}$ by using the rules Eqs.2.6-2.12 if, as we assume, f and g are polynomials or well-behaved power series in the position and momentum variables. This is because the product rule allows us to break Poisson brackets that contain polynomials into factors of Poisson brackets that contain polynomials of lower degree. Repeating the process, we are eventually left with having to evaluate only Poisson brackets of linear terms, which can easily be evaluated using the first set of rules.

[6]Mathematically, symmetries are described as groups (for example, the composition of two rotations yields a rotation and to every rotation there is an inverse rotation). In elementary particle physics, symmetry groups are one abstraction level higher than Hamiltonians: It has turned out that the complicated Hamiltonians which describe the fundamental forces, i.e., the electromagnetic, weak and strong force, are essentially derivable as being the simplest Hamiltonians associated with with three elementary symmetry groups.

This is all good and fine but when f or g contain high or even infinite powers of the position and momentum variables, then the evaluation of the Poisson bracket $\{f, g\}$ can become rather tedious and cumbersome.

For practical purposes it is of interest, therefore, to have a shortcut for the evaluation of Poisson brackets. Indeed, for complicated f and g, the Poisson bracket $\{f, g\}$ can be evaluated usually faster by the following formula:

$$\{f, g\} = \sum_{r=1}^{n} \sum_{i=1}^{3} \left(\frac{\partial f}{\partial x_i^{(r)}} \frac{\partial g}{\partial p_i^{(r)}} - \frac{\partial f}{\partial p_i^{(r)}} \frac{\partial g}{\partial x_i^{(r)}} \right) \tag{2.37}$$

Exercise 2.11 *Use Eq.2.37 to evaluate* $\{x^8 p^6, x^3 p^4\}$.

Exercise 2.12 *Show that Eq.2.37 is indeed a representation of the Poisson bracket, i.e., that it always yields the correct answer. To this end, check that it obeys Eqs.2.9-2.13 and Eqs.2.16-2.18.*

Exercise 2.13 *Find the representation of the Hamilton equations Eq.2.19 and Eqs.2.21, 2.22 obtained by using Eq.2.37.*

Remark: Some textbooks start with these representations of the Hamilton equations, along with the representation Eq.2.37 of the Poisson bracket - without reference to the Hamilton equations' more abstract origin in Eq.2.19 and Eqs.2.21, 2.22. This is unfortunate because those representations using Eq.2.37 do not carry over to quantum mechanics, while the more abstract equations Eq.2.19 and Eqs.2.21, 2.22 will carry over to quantum mechanics unchanged, as we will see!

2.4 Summary: The laws of classical mechanics

We already discussed that quantum mechanics must have strong similarities with classical mechanics, since it must reproduce all the successes of classical mechanics. This suggested that the laws of quantum mechanics might be a slight modification of Newton's laws which would somehow contain Planck's constant h. Since this did not work, we reformulated the laws of classical mechanics on a higher level of abstraction, namely in Hamilton's form. Before we now try to guess the laws of quantum mechanics, let us restate Hamilton's formulation of classical mechanics very carefully:

The starting point is the energy function H of the system in question. It is called the Hamiltonian, and it is an expression in terms of the position and momentum variables of the system. Then, assume we are interested in the time evolution of some quantity

f which is also a polynomial in the x and p's (say with constant coefficients). Then we can derive the equation of motion for f through:

$$\frac{d}{dt} f \;=\; \{f, H\} \tag{2.38}$$

In particular, f can be chosen to be any one of the position and momentum variables of the system, and we obtain their equations of motion as Eqs.2.21,2.22. In order to obtain explicit differential equations from Eqs.2.38,2.21,2.22 we evaluate the Poisson bracket on its right hand side. To this end, we use the definitions Eqs.2.6-2.13. The so-obtained differential equations are then solved to obtain the positions $x_i^{(r)}(t)$ and momenta $p_i^{(r)}(t)$ as functions of time.

We note that the Poisson bracket which is defined by the axioms Eqs.2.6-2.12 possesses an often convenient explicit representation through Eq.2.37. We need to keep in mind, however, that Eq.2.37 merely provides a convenient shortcut for evaluating the Poisson bracket. This shortcut only works in classical mechanics. In quantum mechanics, there will also be a representation of the Poisson bracket but it will look very different from Eq.2.37.

2.5 Classical field theory

This section is a mere comment. In classical mechanics, the dynamical variables are the positions of particles, together with their velocities or momenta. For each particle there are three degrees of freedom of position and momenta.

In a field theory, such as Maxwell's theory, positions (and momenta) are not dynamical variables. After all, unlike a particle that moves around, a field can be everywhere at the same time. In the case of a field theory, what is dynamical is its amplitude.

Consider say a scalar field ϕ. At every point x in space it has an amplitude $\phi(x,t)$ that changes in time with a 'velocity' of $\dot{\phi}(x,t)$ which we may call the canonically conjugate momentum field: $\pi(x,t) := \dot{\phi}(x,t)$. Unlike the three degrees of freedom that particle possesses, a field therefore possesses uncountably many degrees of freedom, one at each position x. Now one can define the Poisson brackets of the first kind for them in analogy to the Poisson brackets for particles:

$$\{\phi(x,t), \pi(x',t)\} \;=\; \delta^3(x-x') \tag{2.39}$$
$$\{\phi(x,t), \phi(x',t)\} \;=\; 0 \tag{2.40}$$
$$\{\pi(x,t), \pi(x',t)\} \;=\; 0 \tag{2.41}$$

Here, $\delta^3(x-x')$ is the three dimanional Dirac delta distribution. The second set of Poisson brackets is unchanged, i.e., it is still given by Eqs.2.9-2.13. The energy of the

classical field, i.e., its Hamiltonian, is:

$$H(\phi,\pi) = \int d^3x \, \frac{1}{2}\left(\pi(x,t)^2 + \sum_{i=1}^{3}(\partial_i\phi(x,t))^2 + m^2\phi(x,t)^2\right) \tag{2.42}$$

The Hamilton equation Eq.2.38 is unchanged.

Exercise 2.14 *(Bonus question) Derive the equations of motion for $\phi(x,t)$ and $\pi(x,t)$. Combine the two equations by eliminating $\pi(x,t)$.*

The combined equation is the so-called Klein Gordon equation. The Dirac equation and the Maxwell equations can be treated similarly, although with some small extra complications because the amplitudes of these fields are not scalar but are vectorial and spinorial respectively.

Chapter 3

Quantum mechanics in Hamiltonian form

We formulated the laws of classical mechanics on a higher level of abstraction, as summarized in Sec.2.4 because classical mechanics appeared to be not upgradeable when written in Newton's formulation. We are now ready to upgrade the more abstract Hamiltonian laws of classical mechanics to obtain quantum mechanics. A modification is needed which is a small enough to preserve all the successes of classical mechanics while it must also introduce h in order to correctly predict quantum mechanical effects.

For example, could it be that Eq.2.38 needs to be modified to obtain quantum mechanics? Could the correct equation be, say, $\frac{d}{dt}\, f = \{f, H\} + h$ or $\frac{d}{dt}\, f = \{f + h^2/f, H\}$, where h is Planck's constant? Those two equations can't be right, of course, already because the units generally don't match. Could it be then that to obtain quantum mechanics we will have to change the definitions for the Poisson bracket? Could it be that the definition Eq.2.12 needs to be changed? This, of course, is unlikely too because the definitions for the Poisson bracket were fixed by consistency conditions (recall, e.g., Exercise 2.7). The structure of our Poisson algebra is quite tightly constrained.

No, the necessary upgrade of the Hamiltonian formalism is actually much more subtle! Let us remember that when we defined the Poisson algebra structure in the previous section we did not make any assumptions about the mathematical nature of the functions $x(t)$ and $p(t)$ (let us omit writing out the indices). In particular, we did not make the assumption that these functions are number-valued. We can start with a Hamiltonian, i.e., a polynomial in the x and p and then by using the rules of the Poisson bracket we can derive the differential equations of motion. In the process, we never need to assume that the functions $x(t)$ and $p(t)$ are number valued. Could it be that the $x(t)$ and $p(t)$ need not be number valued and that this holds the key to upgrading classical mechanics to obtain quantum mechanics? Actually yes!

Before we get to this, we have to consider though that we actually did assume the x and p to be number valued at one specific point at the very end of the previous chapter. There, we wrote down a convenient representation of the Poisson bracket in

Eq.2.37, and there we needed the x and p to be number-valued - because to use this convenient representation we needed to be able to differentiate with respect to the x and p. We can conclude from this that if allowing the $x(t)$ and $p(t)$ to be something else than number valued is the key to upgrading to quantum mechanics, then the Poisson bracket will not be representable any more through Eq.2.37.

In fact, as we will see, this is how it will play out. Everything we did in the previous chapter, except for the representation Eq.2.37 will still exactly hold true in quantum mechanics. In particular, the differential equations of motion derived from the Hamiltonian will look exactly the same in quantum and classical mechanics. That's because they are derived from the same Hamiltonian polynomial in the x and p by using the same rules for the Poisson bracket. But then, if not in the equations of motion, how does the upgrade involve h at all?

3.1 Reconsidering the nature of observables

At this point, let us reconsider the very basics: How do the symbols we write on paper relate to real systems? We measure a system with concrete measurement devices in the lab, for example, devices for the measurement of positions and devices for the measurement of momenta. As usual, we invent for each kind of measurement a symbol, say $x_i^{(r)}$ and $p_i^{(r)}$. At this stage we need to be careful not to over-interpret these symbols. At this stage, these symbols have nothing to do (yet) with numbers, vectors, matrices, operators or bananas. Instead, these symbols are merely names for kinds of measurement. We need to find out more about the nature of these $x_i^{(r)}$ and $p_i^{(r)}$.

Now according to our everyday experience, the operation of a position measurement device does not interfere with the operation of a momentum measurement device: it seems that we can always measure both, positions and momenta. For example, GPS units are able to tell both position and velocity at the same time to considerable accuracy. It is tempting to assume, therefore, that there is no limit, in principle, to how accurately positions and velocities can be determined. And that would mean that we can let each of the symbols $x_i^{(r)}(t)$ and $p_i^{(r)}(t)$ stand for its measurement devices's output number at time t.

It is at this very point, namely when we make the assumption that positions and momenta can be accurately measured simultaneously, that we make the assumption that the symbols $x_i^{(r)}(t)$ and $p_i^{(r)}(t)$ can be represented mathematically as number-valued functions of time. And number-valued functions have the property of being commutative:

$$x_i^{(r)} p_j^{(s)} - p_j^{(s)} x_i^{(r)} \;=\; 0 \tag{3.1}$$

Since measurement values cannot be just any number but always come out real, we also have the law:

$$\left(x_i^{(r)}\right)^* = x_i^{(r)} \qquad \text{and} \qquad \left(p_j^{(s)}\right)^* = p_j^{(s)} \tag{3.2}$$

Similarly, we have $H^* = H$. Technically, the *-operation is an example of was is called an involution.

Exercise 3.1 *Find and list the defining property of an involution.*

The statements above, namely that position and momentum measurements are compatible and come out as real numbers are indeed a nontrivial part of the laws of classical mechanics. For completeness we should have included them in the summary of classical mechanics in Sec.2.4.

A reality property of the form of Eq.3.2 will still be true in quantum mechanics. But the commutativity property expressed in Eq.3.1 and its underlying assumption that the operation of position and momentum measurement devices do not interfere with another needs to be abandoned and upgraded. It turns out that position and momentum measurements are like taking a shower and working out at the gym. It matters in which sequence one does them:

$$\text{gym}\cdot\text{shower - shower}\cdot\text{gym} = \text{sweat} \tag{3.3}$$

3.2 The canonical commutation relations

For the remainder of this course, we will need a way to make it transparent in every equation whether a variable is number valued or not. To this end, we will decorate variables that may not be number valued with a hat, for example, $\hat{H}, \hat{p}, \hat{x}$, or more specifically $\hat{x}_i^{(r)}$ and $\hat{p}_i^{(r)}$ for each position and momentum measurement device. Now how can the interference of the measurement devices mathematically be expressed as properties of the symbols $\hat{x}_i^{(r)}$ and $\hat{p}_i^{(r)}$?

According to classical mechanics one would be able to operate all measurement devices all the time and they would not interfere with another. We could therefore choose the $\hat{x}_i^{(r)}(t)$ and $\hat{p}_i^{(r)}(t)$ to stand for the number-valued outcomes of those measurements as functions of time. Crucially, the fact that we can't actually know positions and momenta simultaneously means that we can no longer choose the $\hat{x}_i^{(r)}(t)$ and $\hat{p}_i^{(r)}(t)$ to stand simply for number-valued outcomes of those measurements as functions of time.

Mathematically, it was the commutativity law of Eq.3.1 which expressed that in classical mechanics the symbols $\hat{x}_i^{(r)}(t)$ and $\hat{p}_i^{(r)}(t)$ can be represented as number valued functions. Could it be that Eq.3.1 has to be modified to include h so that the $\hat{x}_i^{(r)}(t)$ and $\hat{p}_i^{(r)}(t)$ become non-commutative and therefore can no longer be number-valued functions?

Are position and momentum measurements noncommuting similar to how doing sports and having a shower don't commute?

It was Dirac who first realized that all of the Poisson algebra structure that we defined above can be kept (and therefore the ability to derive the equations of motion),

while changing just one little thing: allowing the symbols $\hat{x}_i^{(r)}(t)$ and $\hat{p}_i^{(r)}(t)$ to be noncommutative in a very particular way.

Before we can follow Dirac's argument, let us first reconsider the product rule for the Poisson bracket:

$$\{f, gh\} = \{f, g\}h + g\{f, h\} \tag{3.4}$$

Using the antisymmetry rule, $\{f, g\} = -\{g, f\}$, the product rule can be rewritten in this form:

$$\{gh, f\} = \{g, f\}h + g\{h, f\} \tag{3.5}$$

Using Eqs.3.4,3.5, we can now follow Dirac's argument for why the Poisson algebra structure imposes strict conditions on the form that any noncommutativity can take. Dirac considered the Poisson bracket

$$\{\hat{u}_1\hat{u}_2, \hat{v}_1\hat{v}_2\} \tag{3.6}$$

where $\hat{u}_1, \hat{u}_2, \hat{v}_1, \hat{v}_2$ are arbitrary polynomials in the variables $\hat{x}_i^{(r)}$ and $\hat{p}_j^{(s)}$. Expression Eq.3.6 can be decomposed into simpler Poisson brackets in two ways, namely using first Eq.3.4 and then Eq.3.5, or vice versa. And, of course, any noncommutativity of the $\hat{x}_i^{(r)}$ and $\hat{p}_j^{(s)}$ has to be such that both ways yield the same outcome:

$$\begin{aligned} \{\hat{u}_1\hat{u}_2, \hat{v}_1\hat{v}_2\} &= \hat{u}_1\{\hat{u}_2, \hat{v}_1\hat{v}_2\} + \{\hat{u}_1, \hat{v}_1\hat{v}_2\}\hat{u}_2 \\ &= \hat{u}_1(\hat{v}_1\{\hat{u}_2, \hat{v}_2\} + \{\hat{u}_2, \hat{v}_1\}\hat{v}_2) + (\hat{v}_1\{\hat{u}_1, \hat{v}_2\} + \{\hat{u}_1, \hat{v}_1\}\hat{v}_2)\hat{u}_2 \end{aligned} \tag{3.7}$$

must agree with:

$$\begin{aligned} \{\hat{u}_1\hat{u}_2, \hat{v}_1\hat{v}_2\} &= \hat{v}_1\{\hat{u}_1\hat{u}_2, \hat{v}_2\} + \{\hat{u}_1\hat{u}_2, \hat{v}_1\}\hat{v}_2 \\ &= \hat{v}_1(\hat{u}_1\{\hat{u}_2, \hat{v}_2\} + \{\hat{u}_1, \hat{v}_2\}\hat{u}_2) + (\hat{u}_1\{\hat{u}_2, \hat{v}_1\} + \{\hat{u}_1, \hat{v}_1\}\hat{u}_2)\hat{v}_2 \end{aligned} \tag{3.8}$$

We can, therefore, conclude that, independently of whether or not we have commutativity, it must always be true that:

$$\{\hat{u}_1, \hat{v}_1\}(\hat{v}_2\hat{u}_2 - \hat{u}_2\hat{v}_2) = (\hat{v}_1\hat{u}_1 - \hat{u}_1\hat{v}_1)\{\hat{u}_2, \hat{v}_2\} \tag{3.9}$$

And this equation has to be true for all possible choices of $\hat{u}_1, \hat{u}_2, \hat{v}_1, \hat{v}_2$. How can we ensure this? As is easy to check, Eq.3.9 will be true[1] if we require all expressions $\hat{u}, \hat{v}$ in the position and momentum variables obey:

$$\hat{v}\hat{u} - \hat{u}\hat{v} = k\{\hat{u}, \hat{v}\} \tag{3.10}$$

with k being some constant that commutes with everything. This is because in this case, the left and right hand sides of Eq.3.9 are automatically identical. But what

[1] We will not show here that, vice versa, *only* the condition Eq.3.10 ensures that Eq.3.9 holds true. If you are interested, there is plenty of literature on the topic of "quantization".

value does k take? Of course, the case $k = 0$ would be classical mechanics, because it implies that all expressions in the positions and momenta commute.

However, it turns out that in order to eventually yield the correct experimental predictions (we will later see how), we have to set $k = -ih/2\pi$, i.e., we have

$$\hat{u}\hat{v} - \hat{v}\hat{u} = i\hbar\{\hat{u}, \hat{v}\} \tag{3.11}$$

where we used the convenient definition:

$$\hbar = \frac{h}{2\pi} \tag{3.12}$$

In particular, choosing for $\hat{u}$ and $\hat{v}$ the variables $x_i^{(r)}$ and $p_j^{(s)}$, and using Eqs.2.6-2.8, we now obtain the quantum mechanical commutation relations for n particles:

$$\hat{x}_i^{(r)}\hat{p}_j^{(s)} - \hat{p}_j^{(s)}\hat{x}_i^{(r)} = i\hbar\ \delta_{i,j}\delta_{r,s} \tag{3.13}$$

$$\hat{x}_i^{(r)}\hat{x}_j^{(s)} - \hat{x}_j^{(s)}\hat{x}_i^{(r)} = 0 \tag{3.14}$$

$$\hat{p}_i^{(r)}\hat{p}_j^{(s)} - \hat{p}_j^{(s)}\hat{p}_i^{(r)} = 0 \tag{3.15}$$

Let us keep in mind that we did not modify the rules of the Poisson bracket. We still have:

$$\{\hat{x}_i, \hat{p}_j\} = \delta_{i,j} \tag{3.16}$$

$$\{\hat{x}_i, \hat{x}_j\} = 0 \tag{3.17}$$

$$\{\hat{p}_i, \hat{p}_j\} = 0 \tag{3.18}$$

$$\{f, g\} = -\{g, f\} \qquad \text{antisymmetry} \tag{3.19}$$

$$\{cf, g\} = c\,\{f, g\}, \quad \text{for any number } c \qquad \text{linearity} \tag{3.20}$$

$$\{f, g + h\} = \{f, g\} + \{f, h\} \qquad \text{addition rule} \tag{3.21}$$

$$\{f, gh\} = \{f, g\}h + g\{f, h\} \qquad \text{product rule} \tag{3.22}$$

$$0 = \{f, \{g, h\}\} + \{h, \{f, g\}\} + \{g, \{h, f\}\} \qquad \text{Jacobi id.} \tag{3.23}$$

Because the rules for the Poisson bracket did not change with the upgrade to quantum mechanics, one arrives in quantum mechanics at the same equations of motion as in classical mechanics. This is as long as one does not unnecessarily commute any variables.

The equations Eqs.3.13-3.15 are called the "Canonical Commutation Relations" (CCRs). The appearance of the imaginary unit i will be necessary to ensure that measurements are predicted as real numbers, as we will see below. Eqs.3.14,3.15 express that position measurements among another and momentum measurements among another do not interfere. Only positions and momenta of the same particle and in the same direction, i.e., for $i = j$ and $r = s$, are noncommutative.

In conclusion, we upgrade classical mechanics to quantum mechanics by first formulating classical mechanics in Hamiltonian form to identify the Poisson algebra structure. Then, we realize that while keeping all the rules for the Poisson bracket intact, there is still the freedom to make the associative multiplication in the Poisson algebra noncommutative, parametrized by some constant k. Nature chose the modulus of k to be nonzero though very small, namely $\hbar$. The fact that the Poisson bracket stays the same when quantizing explains why quantum mechanics has the same equation of motion as does classical mechanics. The fact that $\hbar$ is so small explains why it took long to discover quantum mechanics.

In spite of the tremendous similarity between classical and quantum mechanics from this perspective, quantum mechanical calculations will in practise look rather different from classical calculations. This is because they will require representations of the $\hat{x}_i^{(r)}(t)$ and $\hat{p}_i^{(r)}(t)$ variables as explicit non-number valued mathematical entities that obey the commutation relations. Even though there is only a slight noncommutativity in the Poisson algebra of quantum mechanics its representations will necessarily look quite different from the representation of the classical commutative Poisson algebra. This will explain why the Schrödinger equation looks rather different from Newton's equations.

3.3 From the Hamiltonian to the equations of motion

In quantum mechanics, as in classical mechanics, the energy function $\hat{H}$ encodes all information about the system. It is still called the Hamiltonian and it is in general some polynomial (or well-behaved power series) in the positions and momenta $\hat{x}_i^{(r)}$ and $\hat{p}_i^{(r)}$ of the system. In quantum mechanics, the sequence of steps that lead from the Hamiltonian down to concrete number predictions for experiments can be drawn schematically in this form:

Hamiltonian
$\Downarrow$
Equations of motion
$\Downarrow$
Differential equations
$\Downarrow$
Non-number-valued functions
$\Downarrow$
Number-valued functions
$\Downarrow$
Number predictions

So far, we can perform the first step, namely the derivation of the equations of motion from the Hamiltonian: Assume that we are interested in the time evolution of some $\hat{f}$ which is a polynomial in the $\hat{x}$ and $\hat{p}$'s (say with constant coefficients). Then we can derive the equation of motion for $\hat{f}$ through:

$$\frac{d}{dt}\,\hat{f} \;=\; \{\hat{f},\hat{H}\} \tag{3.24}$$

where $\{,\}$ is the usual Poisson bracket, as defined in Eqs.2.6-2.12. In particular, $\hat{f}$ can be chosen to be any one of the position and momentum variables of the system, so that we obtain for their equations of motion, exactly as in Eqs.2.21,2.22:

$$\frac{d}{dt}\,\hat{x}_i^{(r)} \;=\; \{\hat{x}_i^{(r)},\hat{H}\} \tag{3.25}$$

$$\frac{d}{dt}\,\hat{p}_i^{(r)} \;=\; \{\hat{p}_i^{(r)},\hat{H}\} \tag{3.26}$$

By evaluating the Poisson bracket on the right hand side of Eqs.3.25,3.26 these equations of motion then become differential equations for the entities $\hat{x}_i^{(r)}(t)$ and $\hat{p}_i^{(r)}(t)$. Clearly, the resulting equations of motion will be analogous to those of classical mechanics. The entities $\hat{x}_i^{(r)}(t)$ and $\hat{p}_i^{(r)}(t)$ must also still obey Eq.3.2, which in quantum mechanics is usually written as:

$$\left(\hat{x}_i^{(r)}\right)^\dagger = \hat{x}_i^{(r)} \qquad \text{and} \qquad \left(\hat{p}_j^{(s)}\right)^\dagger = \hat{p}_j^{(s)} \tag{3.27}$$

We will call any polynomial or well-behaved power series $\hat{f}$ in the $\hat{x}$ and $\hat{p}$ an "observable", if it obeys $\hat{f}^\dagger = \hat{f}$. As we will see later, the condition $\hat{f}^\dagger = \hat{f}$ will indeed imply that measurement outcomes are predicted as real numbers. In addition to the position variables $\hat{x}_i^{(r)}(t)$ and momentum variables $\hat{p}_j^{(s)}(t)$ also, e.g., the energy $\hat{H}(t)$ and the angular momentum variables $\hat{L}_i(t)$ are observables.

While classical mechanics requires the Poisson algebra to be commutative, quantum mechanics requires that the equations of motion be solved by entities $\hat{x}_i^{(r)}(t)$ and $\hat{p}_i^{(r)}(t)$ which are noncommutative:

$$\hat{x}_i^{(r)}\hat{p}_j^{(s)} - \hat{p}_j^{(s)}\hat{x}_i^{(r)} \;=\; i\hbar\;\delta_{i,j}\delta_{r,s} \tag{3.28}$$

$$\hat{x}_i^{(r)}\hat{x}_j^{(s)} - \hat{x}_j^{(s)}\hat{x}_i^{(r)} \;=\; 0 \tag{3.29}$$

$$\hat{p}_i^{(r)}\hat{p}_j^{(s)} - \hat{p}_j^{(s)}\hat{p}_i^{(r)} \;=\; 0 \tag{3.30}$$

Technically, we will, therefore, need to solve differential equations of motion with noncommutative entities. In practice, the task is then to start from the top level of

abstraction, the Hamiltonian of a system, then working one's way down by calculating the equations of motion, and then solving them to obtain something from which eventually predictions can be made of numbers that can be measured in experiments on the system. In the next section, we will investigate what kind of noncommutative mathematical objects, such as, for example, matrices, may represent the position and momentum variables.

Exercise 3.2 *For classical mechanics, formula Eq.2.37 provided a convenient representation of the Poisson bracket. However, Eq.2.37 is not a valid representation of the Poisson bracket in the case of quantum mechanics. In quantum mechanics, we have a (not so convenient) representation of the Poisson bracket through Eq.3.11:*

$$\{\hat{u}, \hat{v}\} = \frac{1}{i\hbar}(\hat{u}\hat{v} - \hat{v}\hat{u}) \tag{3.31}$$

Use this representation to evaluate the Poisson bracket $\{\hat{x}^2, \hat{p}\}$.

Let us introduce an often-used notation, called "the commutator":

$$[A, B] \;:=\; A\;B \;-\; B\;A \tag{3.32}$$

For simplicity, assume that $\hat{H}$ and $\hat{f}$ are polynomials in the positions and momenta which depend on time only through their dependence on the $\hat{x}$ and $\hat{p}$. Then the Hamilton equation Eq.3.24 holds and takes the form:

$$i\hbar\,\frac{d}{dt}\,\hat{f}(t) \;=\; [\hat{f}(t), \hat{H}] \tag{3.33}$$

$$\tag{3.34}$$

and, in particular:

$$i\hbar\,\frac{d}{dt}\,\hat{x}_i^{(r)}(t) \;=\; [\hat{x}_i^{(r)}(t), \hat{H}]$$

$$i\hbar\,\frac{d}{dt}\,\hat{p}_i^{(r)}(t) \;=\; [\hat{p}_i^{(r)}(t), \hat{H}] \tag{3.35}$$

These equations are called the Heisenberg equations of motion.

Remark: The particular method by which in the past few sections we upgraded classical mechanics to quantum mechanics is called canonical quantization. I covered it in some detail because of its importance: Essentially the same method was used to find quantum electrodynamics starting from Faraday and Maxwell's electromagnetism. All the quantum field theories of elementary particles can be derived this way. Even string

theory and most other modern attempts at finding the unifying theory of quantum gravity try to employ canonical quantization. I should mention too that the problem of canonical quantization for constrained classical systems was also pioneered by Dirac but is still not fully understood. A simple example of a constrained system would be a particle that is constrained to move on a curved surface. The most important constrained system is general relativity.

Exercise 3.3 *Reconsider the system with the Hamiltonian Eq.2.4, which consists of two particles that are attracted to another through a harmonic force (a force which is proportional to their distance). In practice, for example the force that binds diatomic molecules and the force that keeps nucleons (i.e., neutrons and protons) inside a nucleus are essentially harmonic. In those cases the effect of $\hbar$ cannot be neglected. One obtains the correct quantum theoretic Hamiltonian from the classical Hamiltonian of Eq.2.4 by simply placing hats on the x and p's. Find explicitly all the equations which the $\hat{x}_i^{(r)}$ and $\hat{p}_j^{(r)}$ (where $r \in \{1, 2\}$) of this system must obey.*

Exercise 3.4 *To obtain the quantum Hamiltonian from the classical Hamiltonian and vice versa by placing or removing hats on the x and p's is generally not as straightforward as in the previous exercise! Namely, there can occur so-called "ordering ambiguities": Consider the two Hamiltonians $\hat{H}_1 = \hat{p}^2/2m + a(\hat{x}^2\hat{p}\hat{x} - \hat{x}\hat{p}\hat{x}^2)$ and $\hat{H}_2 = \hat{p}^2/2m + b(\hat{p}\hat{x}\hat{p}^2 - \hat{p}^2\hat{x}\hat{p})$ where a and b are nonzero constants with appropriate units. Check whether or not these two Hamiltonians are the same in classical mechanics. Also check whether or not they can be the same in quantum mechanics for some nonzero choices of a and b.*

Exercise 3.5 *Find a Hamiltonian which contains at least one thousand powers of $\hat{x}$ and which also agrees with the Hamiltonian $\hat{H}_1$ of the previous exercise in classical mechanics. Make sure that your Hamiltonian is formally hermitean, i.e., that it obeys $\hat{H}^\dagger = \hat{H}$. Help: To ensure hermiticity, you can symmetrize. For example, $\hat{x}\hat{p}^2$ is not hermitean but $(\hat{x}\hat{p}^2 + \hat{p}^2\hat{x})/2$ is hermitean.*

Remark: In quantum theory, the choice of Hamiltonian always has an ordering ambiguity because one could always add to the Hamiltonian any extra terms that are proportional to $(\hat{x}\hat{p} - \hat{p}\hat{x})$ because those terms don't affect what the Hamiltonian is in classical mechanics. In principle, experiments are needed to decide which Hamiltonian is the correct one. In practice, the simplest choice is usually the correct choice. The simplest choice is obtained by symmetrizing the given classical Hamiltonian and then not adding any extra terms that are proportional to $\hat{x}\hat{p} - \hat{p}\hat{x}$. This is called the Weyl ordered Hamiltonian.

3.4 From the Hamiltonian to predictions of numbers

In the framework of classical mechanics we know how to descend from the most abstract level, where the system is described simply by giving its Hamiltonian H, down to the concrete level of predicting numbers for measurement outcomes. Now we will have to develop methods for descending in quantum mechanics from the level of the Hamiltonian down to the concrete predictions of numbers in experiments.

In the previous section, we already took the first step: we found that we can use the Hamiltonian to derive the differential equations of motion of the system. Since the Poisson brackets have not changed as we went from classical to quantum mechanics, the equations of motion are the same as those of classical mechanics[2].

The big change compared to classical mechanics, is that now the position and momentum variables $\hat{x}_i^{(r)}(t)$ and $\hat{p}_j^{(s)}(t)$ are noncommutative according to Eqs.3.28-3.30, and that they, therefore, can no longer be represented by number-valued functions of time. This means that the equations of motion can no longer be interpreted as differential equations for number-valued functions!

But we need to find a way to descend the ladder of abstractions all the way from the Hamiltonian on the top of the ladder down to concrete predictions of numbers at the bottom of the ladder. To that end, in order now to be able to solve the equations of motion as explicit differential equations, the $\hat{x}_i^{(r)}(t)$ and $\hat{p}_j^{(s)}(t)$ must be viewed as functions of time whose values are some kind of noncommutative mathematical objects. What kind of mathematical objects could these be?

3.4.1 Linear maps

Actually, *every* mathematical object can be viewed as a map, if need be, as a trivial map. For example the number 5 can be identified with the map that maps everything to 5. So let us look at maps. Let's try to represent the symbols $\hat{x}_i^{(r)}(t)$ and $\hat{p}_j^{(s)}(t)$ as some kind of explicit map-valued functions of time. A simple kind of maps is the linear maps. And they can be noncommutative! So this looks promising.

For example, any square matrix that acts on a finite-dimensional vector space of column vectors represents a linear map. And square matrices are generally noncommutative! In principle, we need the matrices to be square matrices so that they map back into the same vector space, so that we can multiply any two matrices on the same vector space. As we will see later, with some precautions, we can also consider infinite-by-infinite matrices that act as linear maps on infinite-dimensional vector spaces.

[2]Except for the ordering ambiguity: when going from the classical to the quantum Hamiltonian we could (if we had any experimental reason to do so) add to the quantum Hamiltonian any hermitean terms that are proportional to $\hbar$, such as terms like $ig(\hat{x})(\hat{x}\hat{p} - \hat{p}\hat{x})\hat{g}(\hat{x})$ where g is some polynomial in $\hat{x}$.

There are actually many kinds of linear maps and they may not act on vector spaces of column vectors at all!

Let us consider, for example, the infinite-dimensional vector space $V := C^7(\mathbb{R})$ of seven times continuously differentiable functions on the real line. The set V forms a vector space because it obeys the defining axioms of a vector space: in brief, one can suitably add any two elements of V and get an element of V and one can suitably multiply any element of V with a number and get an element of V.

Exercise 3.6 *Find and list the precise axioms that a set has to obey to be called a vector space.*

Definition: *A map on a vector space that is infinite dimensional is called an operator.*

For example, the derivative operator, D, acts on functions in V in this way:

$$D : g(\lambda) \to \frac{d}{d\lambda} g(\lambda) \tag{3.36}$$

The operator D is a linear operator, i.e., it is a linear map, because it obeys $\partial_\lambda\,(c\,g(\lambda)) = c\;\partial_\lambda g(\lambda)$ for all numbers c and because $\partial_\lambda(g_1(\lambda) + g_2(\lambda)) = \partial_\lambda g_1(\lambda) + \partial_\lambda g_2(\lambda)$. Here, in order to simplify the notation, we introduced the notation: $\partial_\lambda := \frac{d}{d\lambda}$

Exercise 3.7 *Check whether or not the multiplication operator, M, which maps $M : g(\lambda) \to \lambda g(\lambda)$ is a linear operator.*

Exercise 3.8 *Show that the two operators D and M on V do not commute, namely by calculating $(DM - MD)g(\lambda)$.*

Exercise 3.9 *Check whether or not the operator Q which acts on functions in V as $Q : g(\lambda) \to \lambda^5 g(\lambda)$ is a linear operator.*

3.4.2 Choices of representation

We have just seen examples of linear maps and, since they generally do not commute, they may be useful for representing the variables $\hat{x}_i^{(r)}(t)$ and $\hat{p}_j^{(s)}(t)$ as explicit mathematical objects. And this is what we need to be able to descend further down the ladder of abstractions, down to predictions of numbers for measurement outcomes.

But could it be that one should use representations of the $\hat{x}_i^{(r)}(t)$ and $\hat{p}_j^{(s)}(t)$ as nonlinear maps instead? Non-linear representations have been considered in the literature. There are articles by Steven Weinberg, for example, on this topic. This work has shown, however, that any attempt at using nonlinear spaces or nonlinear operators to define quantum theories generally leads to physically incorrect predictions. We will, therefore, here only consider linear representations.

Now that we have settled on representations of the variables $\hat{x}_i^{(r)}(t)$ and $\hat{p}_j^{(s)}(t)$ as linear operators, we still have plenty of choice, because there are so many vector spaces and so many linear operators on them. And this leads to a worry: could it happen that we invest great effort in developing one particular kind of representation of the variables $\hat{x}_i^{(r)}(t)$ and $\hat{p}_j^{(s)}(t)$ as operators, say as matrices, and then it turns out that we have bet on the wrong horse? Maybe, we should have instead developed a representation of the $\hat{x}_i^{(r)}(t)$ and $\hat{p}_j^{(s)}(t)$ as, for example, multiplication and differentiation operators?

Fortunately, essentially[3] all linear representations of variables $\hat{x}_i^{(r)}(t)$ and $\hat{p}_j^{(s)}(t)$ that obey the canonical commutation relations are equivalent, i.e., they lead to the exact same predictions! This is the content of the Stone von Neumann theorem, which we will later cover more precisely. Technically, as we will see, all linear representations are the same, up to a change of basis in the vector space. It may seem strange that, for example, a space of column vectors with countably infinitely many entries could be isomorphic to some space of functions on the real line. But this is what will turn out to be the case[4]!

So to recapitulate: our task is to solve the equations of motion, the hermiticity conditions and the canonical commutation relations for $\hat{x}_i^{(r)}(t)$ and $\hat{p}_j^{(s)}(t)$ as linear-map-valued (instead of number-valued) functions of time.

We now know that the choice of which kind of linear representation we use will ultimately not matter when calculating physical predictions.

As our first choice, let us, therefore, use the most concrete kind of linear maps to represent the $\hat{x}_i^{(r)}(t)$ and $\hat{p}_j^{(s)}(t)$, namely, let us try to represent tham as matrix-valued functions in time. Historically, quantum mechanics was actually first written down in terms of matrix-valued functions, back in June 1925 when the young Heisenberg had some quiet time while escaping his hay fever on the island of Helgoland in the North Sea.

3.4.3 A matrix representation

Let us now find out how the variables $\hat{x}_i^{(r)}(t)$ and $\hat{p}_j^{(s)}(t)$ can be represented as matrix-valued functions in time, and how, therefore, the abstract equations of motion can be represented as explicit matrix differential equations for matrix-valued functions of time. To keep the number of indices in check, we will restrict ourselves here to the case of just one $\hat{x}(t)$ and one $\hat{p}(t)$ operator.

[3] There is a small subtlety, arising from the fact that, as we'll see later, the $\hat{x}$ and $\hat{p}$ are what are called unbounded operators. This leaves some loopholes, in principle, but nature does not appear to make use of those.

[4] If you want to know the essence already: the space of functions will be the set of equivalence classes of square-integrable functions, two functions being in the same equivalence class if their difference has vanishing integral over its norm squared. In this space of equivalence classes one can find bases of countably infinitely many basis vectors.

The canonical commutation relations are of course to hold at all times. To begin with, let us ask whether it is possible to find two $N \times N$ matrices $\hat{x}(t_0)$ and $\hat{p}(t_0)$ so that at the starting time, t_0, of the experiment the canonical commutation relations hold:

$$\hat{x}(t_0)\,\hat{p}(t_0) \;-\hat{p}(t_0)\,\hat{x}(t_0) \;=\; i\hbar\,\mathbf{1} \tag{3.37}$$

Here, $\mathbf{1}$ is the identity matrix. At this point it is useful to remember that the trace of matrices $\mathrm{Tr}(A) = \sum_n A_{n,n}$ is linear and cyclic:

$$\mathrm{Tr}(A+B) = \mathrm{Tr}(A) + \mathrm{Tr}(B) \quad \text{and} \quad \mathrm{Tr}(AB) = \mathrm{Tr}(BA) \tag{3.38}$$

Exercise 3.10 *Verify Eqs.3.38.*

We see that the trace of the left hand side of Eq.3.37 vanishes, while the trace of the right hand side is $i\hbar N$. Thus, there are in fact no $N \times N$ matrices, i.e., there are no finite-dimensional matrices $\hat{x}(t_0)$ and $\hat{p}(t_0)$ that obey the commutation relation Eq.3.37! For infinite dimensional matrices, however, the trace may be ill-defined on both sides, and our argument then does not apply. In fact, there exist infinite-dimensional matrices which do obey the commutation relation.

In order to find such matrices we start by defining the $\infty \times \infty$ dimensional matrix:

$$a_{n,m} = \begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & 0 & \\ 0 & 0 & \sqrt{2} & 0 & 0 & \\ 0 & 0 & 0 & \sqrt{3} & 0 & \\ 0 & 0 & 0 & 0 & \sqrt{4} & \\ 0 & 0 & 0 & 0 & 0 & \\ & & & & & \ddots \end{pmatrix} \tag{3.39}$$

The hermitean conjugate is:

$$a^\dagger_{n,m} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \\ \sqrt{1} & 0 & 0 & 0 & 0 & \\ 0 & \sqrt{2} & 0 & 0 & 0 & \\ 0 & 0 & \sqrt{3} & 0 & 0 & \\ 0 & 0 & 0 & \sqrt{4} & 0 & \\ & & & & & \ddots \end{pmatrix} \tag{3.40}$$

Their commutation commutation relation is:

$$a a^\dagger - a^\dagger a \;=\; \mathbf{1} \tag{3.41}$$

Since they are not numbers, we should decorate a and $a^\dagger$ with hats but traditionally one doesn't put hats on these operators.

Remark: *In case you are wondering because you feel that you have seen similar things before: fundamentally, these operators a and $a^\dagger$ have absolutely nothing to do with harmonic oscillators. What we are currently doing will be good for any choice of system, not just harmonic oscillators. We are currently developing a representation of the variables $\hat{x}(t)$ and $\hat{p}(t)$ as matrices and this representation will, of course, be good for any arbitrary choice of Hamiltonian*[5]*.*

Exercise 3.11 *Verify Eq.3.41.*

Using a and $a^\dagger$, we can now represent $\hat{x}(t_0)$ and $\hat{p}(t_0)$ as matrices that obey the canonical commutation relation, namely by defining:

$$\hat{x}(t_0) = L(a^\dagger + a) \tag{3.42}$$

and

$$\hat{p}(t_0) = \frac{i\hbar}{2L}(a^\dagger - a) \tag{3.43}$$

Here, L is some arbitrary real number with units of length, which we need because $\hat{x}$ has a unit of length while a and $a^\dagger$ do not have units. The definitions are such that the realness conditions Eqs.3.27 are obeyed, i.e., such that the matrices are formally[6] hermitean: $\hat{x}^\dagger(t_0) = \hat{x}(t_0)$ and $\hat{p}^\dagger(t_0) = \hat{p}(t_0)$.

Exercise 3.12 *Verify that the two matrices defined in Eqs.3.42,3.43 are formally hermitean. I am using the term "formally" here to indicate that for the purposes of this exercise you should not worry about potential subtleties that may arise because these matrices are infinite dimensional.*

Exercise 3.13 *Show that the hermitean conjugation of matrices reverses the order, i.e., that if A and B are linear maps, then $(AB)^\dagger = B^\dagger A^\dagger$.*

Technically, † is a map from the Poisson algebra into itself which is called an involution because it is its own inverse. Because it also reverses the order it is called and "anti" algebra mapping: First multiplying and then applying † is the same as first applying † and then multiplying, up to the reversal of the order.

We can now understand the appearance of the imaginary unit i in the canonical commutation relations: If we apply † to the commutation relations $\hat{x}\hat{p} - \hat{p}\hat{x} = k\mathbf{1}$ we obtain $\hat{p}\hat{x} - \hat{p}\hat{x} = k^*\mathbf{1}$, i.e., we obtain $k = -k^*$. Thus, k has to be imaginary. And of course it is: $k = i\hbar$.

Exercise 3.14 *Verify that the two matrices defined in Eqs.3.42,3.43 do obey the commutation relation Eq.3.37. You may use the results of Exercise 3.11.*

[5]Still, it is true also that, when compared to other linear representations, the use of the a and $a^\dagger$ will be particularly convenient when considering the special case of harmonic oscillators.

[6]I am writing here "formally" hermitean, because the issue of whether a matrix is hermitean, symmetric or self-adjoint is quite subtle for infinite-dimensional matrices, as we will see later.

3.4.4 Example: Solving the equations of motion for a free particle with matrix-valued functions

In the case of the free particle which moves in one dimension, the Hamiltonian is $\hat{H} = \hat{p}^2/2m$. The Hamilton equations or, equivalently, the Heisenberg equations, yield the abstract equations of motion:

$$\frac{d}{dt}\hat{x}(t) = \frac{1}{m}\hat{p}(t) \tag{3.44}$$

$$\frac{d}{dt}\hat{p}(t) = 0 \tag{3.45}$$

Let us view these equations as matrix equations. Using the results of the previous section, it becomes clear that these equations are solved through

$$\hat{x}(t) = \hat{x}(t_0) + \frac{(t-t_0)}{m}\hat{p}(t_0) \tag{3.46}$$

and

$$\hat{p}(t) = \hat{p}(t_0), \tag{3.47}$$

where $\hat{x}(t_0)$ and $\hat{p}(t_0)$ are the matrices of Eqs.3.42,3.43. Concretely, by substituting in the matrices a and $a^\dagger$, we have:

$$\hat{x}(t)_{n,m} = \begin{pmatrix} 0 & \sqrt{1}\left(L - \frac{i\hbar(t-t_0)}{2Lm}\right) & 0 & \\ \sqrt{1}\left(L + \frac{i\hbar(t-t_0)}{2Lm}\right) & 0 & \sqrt{2}\left(L - \frac{i\hbar(t-t_0)}{2Lm}\right) & \\ 0 & \sqrt{2}\left(L + \frac{i\hbar(t-t_0)}{2Lm}\right) & 0 & \\ & & & \ddots \end{pmatrix} \tag{3.48}$$

$$\hat{p}(t)_{n,m} = \begin{pmatrix} 0 & -\sqrt{1}\frac{i\hbar}{2L} & 0 & \\ \sqrt{1}\frac{i\hbar}{2L} & 0 & -\sqrt{2}\frac{i\hbar}{2L} & \\ 0 & \sqrt{2}\frac{i\hbar}{2L} & 0 & \\ & & & \ddots \end{pmatrix} \tag{3.49}$$

Exercise 3.15 *Show that the matrices $\hat{x}(t)$ and $\hat{p}(t)$ obey at all times $t > t_0$ all the quantum mechanical conditions, i.e., the equations of motion, the hermiticity condition, and the commutation relation.*

Remark: We had constructed the representation in such a way that the commutation relation and the hermiticity condition hold at the initial time t_0. Having solved

the equations of motion we found that the commutation relation and the hermiticity conditions continue to hold at all times t. This is nontrivial but it is not a coincidence. As we will soon see, the quantum mechanical time evolution of all systems[7] preserves the commutation relations and hermiticity. The preservation of the commutation relations is of course the preservation of the Poisson bracket. And we have in classical and quantum mechanics that the Poisson brackets between the positions and momenta are preserved by the dynamics through the Hamilton equation: $d/dt\ \{\hat{x}, \hat{p}\} = \{\{\hat{x}, \hat{p}\}, \hat{H}\} = \{1, \hat{H}\} = 0$. We can also turn the logic around. Assume we know nothing about Hamiltonians and about the dynamics of quantum systems. Except, we may want to assume that, whatever the time evolution is, it must preserve the Poisson algebra structure, i.e., we require that the Poisson brackets be conserved in time. The structure of the Poisson algebra then demands (we don't show this explicitly here) that the time evolution must be generated through an equation of the type of the Hamilton equation, by some generator which we may call H, and which we may then as well call the Hamiltonian.

[7]With the possible exception of systems that involve black hole horizons or other gravitational horizons or singularities.