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:

$$\begin{array}{c}\text{Differential equations}\\ \Downarrow \\ \text{Functions}\\ \Downarrow \\ \text{Variables}\\ \Downarrow \\ \text{Numbers}\end{array}$$

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:

$$\begin{array}{c}\text{Hamiltonians}\\ \Downarrow\\ \text{Differential equations}\\ \Downarrow\\ \text{Functions}\\ \Downarrow\\ \text{Variables}\\ \Downarrow\\ \text{Numbers}\end{array}$$

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.)

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

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:

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

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:

$$\{x_3, p_3^2\} = \{x_3, p_3\}p_3 + p_3\{x_3, p_3\} = 1p_3 + p_3 1 = 2p_3 \qquad (2.15)$$

Here now is the first set of exercises. These exercises, and all exercises up until Sec.2.3.5 are to be solved using only the above axioms for the Poisson bracket. In Sec.2.3.5, we will introduce a representation of the Poisson bracket in terms of derivatives[5]. You can use this representation only for the exercises from Sec.2.3.5 onward.

Exercise 2.1 *Prove Eq.2.14.*

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

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

Exercise 2.4 *Verify Eq.2.5.*

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

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

[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] The abstract, axiomatically-defined Poisson bracket that we defined above may sometimes be a bit tedious to evaluate but it is very powerful because of its abstractness. As Dirac first showed, this abstract, axiomatically-defined Poisson bracket is the exact same in classical and quantum mechanics. As we will see in Sec.2.3.5, there is a representation of this abstract Poisson bracket in terms of derivatives, which is more convenient to work with but this representation only works in classical mechanics.

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