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 them 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 by writing the matrices in an orthonormal basis, i.e., with indices, and by then evaluating the trace by summing up the matrix elements on the diagonal.*

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:

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

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 with the help of Eqs.3.39,3.40, are formally hermitean. I am using the term "formally" here to indicate that, for the purposes of this exercise, you need 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 $(AB)^\dagger = B^\dagger A^\dagger$. To this end, write out the matrices with indices and use that hermitean conjugating a matrix means transposing and complex conjugating it.*

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 an "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 see, therefore, why the imaginary unit i appeared 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: $k = i\hbar$.

[5]Still, it is true also that 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:

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

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

For simplicity, not all the many zeros in these matrices are shown. The only nonzero terms are immediately to the left and right of the diagonal.

Exercise 3.14 *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.

3.4.5 Example: Solving the equations of motion for a harmonic oscillator with matrix-valued functions

The vibrational degree of freedom of a diatomic molecule such as HF, CO or HCl can be described as a harmonic oscillator (as long as the oscillations are small). Now let x stand for the deviation from the equilibrium distance between the two nuclei. This distance oscillates harmonically and is described by this effective Hamiltonian of the form of a harmonic oscillator:

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m\omega^2}{2}\hat{x}^2 \tag{3.50}$$

The term "Effective Hamiltonian" expresses the fact that this Hamiltonian is not really the exact Hamiltonian but that it is a good approximation to the Hamiltonian in the regime of low energies (i.e., of small oscillations) that we are considering here. By the way, how do we know that the true Hamiltonian is not simply a harmonic oscillator? Easy: we know from experiments that diatomic molecules will, for example, split apart at sufficiently high temperatures, i.e., that they do not have infinite binding energy. A harmonic oscillator potential, however, just keeps going up faster with distance and therefore if you tried to pull apart the two particles in the diatomic molecule, they would just get pulled together more and more strongly. Diatomic molecules could never be split if they were truly harmonically bound.

So then if we know that the true potential is not harmonic, how do we know that a harmonic potential is a good approximation at low energies? That's because

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