6 Quantum mechanical systems with infinite-dimensional Hilbert spaces

We have discussed at length quantum mechanical systems with finite dimensional Hilbert spaces, like spin $\frac{1}{2}$ (single qbit) or multiple tensor products of them (multiple qbits). While they are mathematically simple, their physical realisation and experimental manipulation is certainly more recent than the early study of atomic spectra. But even the simple hydrogen atom has infinitely many energy levels meaning that the corresponding Hilbert space is infinite-dimensional.

The spin-$\frac{1}{2}$ systems we described so far had no classical analogue and we just deduced from the experimental results of the Stern-Gerlach experiment that we needed some abstract two-dimensional Hilbert space. For the ammonia molecule we made some contact with a classical picture by introducing the position operator X, although in this case this operator only had 2 eigenvalues.

Now we would like to see what is the quantum mechanical description of phenomena that have a classical counterpart, or said differently, that are such that we can retrieve the classical behaviour from the quantum mechanical description in an appropriate limit. In particular, this quantum mechanical description must be such that it takes into account the experimental facts of particle-wave duality and relations like $p \sim \frac{h}{\lambda} = \hbar k$ or $E \sim \hbar\omega$. We will see that this leads us to infinite-dimensional Hilbert spaces and, obviously, we want the resulting rules of quantum mechanics to be some generalisation of the rules that were valid in finite-dimensional Hilbert spaces, as given before.

6.1 Position

Our first task must be to give a quantum mechanical description of position and momentum. To begin with, let's just take a single spatial dimension with (classical) coordinate x. But even this will turn out to be quite subtle. The obvious difficulty is that the position x is a continuous quantity and that quantum mechanics should tell us that any measurement could only give a single discrete result x which is the eigenvalue of some position operator X, i.e. $X|x\rangle = x|x\rangle$ where we labelled the eigenvector by the corresponding eigenvalue. The difficulty here is not so much that we have infinitely many eigenvalues ($x \in \mathbf{R}$) and corresponding eigenvectors, but that this infinity is uncountable. Whenever, $x' \neq x$ we would like that $|x'\rangle$ and $|x\rangle$ are orthogonal, however close x' and x may be. The resulting Hilbert space would have a basis with uncountably many basis vectors. This is not only mathematically much more complicated, it also would not lead to a satisfactory quantum mechanical description, and eventually not correspond to experiment. Indeed, if x' and x are "infinitesimally close", say $|x' - x| < 10^{-30}\,\mathrm{m}$, we have absolutely no experimental possibility[31] to distinguish x' from x, nor $|x'\rangle$ from $|x\rangle$, although they should be orthogonal.

[31] The current resolution of very small scales in high-energy accelerators is of the order of $10^{-20}\,\mathrm{m}$.

We can evade this problem by declaring that the position eigenvalues can only take the discrete values

$$x_n = n\,\epsilon \ , \tag{6.1}$$

where ϵ is the "step size". To begin with, we choose ϵ to be the smallest resolution we can achieve experimentally, so that we can barely distinguish x_n and x_{n+1}, e.g. $\epsilon = 10^{-20}\,\mathrm{m}$. Then, as $n \in \mathbf{Z}$ takes any possible value, for almost all experimental purposes we can treat x_n as the usual continuous real position. We can also interpret this somewhat differently by saying that we have divided the infinite continuous space $(-\infty, +\infty)$ into infinitely many intervals as $\cup_{n\in\mathbf{Z}}\, I_n$ with $I_n = [x_n - \frac{\epsilon}{2}, x_n + \frac{\epsilon}{2})$, and that for each $x \in I_n$ we assimilate $x \simeq x_n$. The corresponding position operator X and eigenstates $|x_n\rangle$ then are such that

$$X\,|x_n\rangle = x_n\,|x_n\rangle \ \ , \tag{6.2}$$

and we can now assume that the $|x_n\rangle$ are orthogonal and form a countable basis for our Hilbert space. However, for reasons that will become clear, we do not choose to normalise them to one, but instead as

$$\langle x_n\,|x_m\rangle = \frac{1}{\epsilon}\,\delta_{nm} \ . \tag{6.3}$$

We also introduce the states

$$|\widetilde{x}_n\rangle = \sqrt{\epsilon}\ |x_n\rangle \ \ , \tag{6.4}$$

that are normalised to one and, hence, form an orthonormal basis.

Suppose now that we have a particle described by the normalised state $|\psi\rangle$. What should be the probability to find this particle at x_n ? More precisely, the question should be what is the probability to find it at x_n to within our smallest experimental uncertainty ϵ. This is the same as asking what is the probability to find the particle in the interval I_n. According to our rules, this should be

$$\mathcal{P}(x_n, \psi) = |\langle \widetilde{x}_n\,|\psi\rangle\,|^2 = \epsilon |\langle x_n\,|\psi\rangle\,|^2 \ . \tag{6.5}$$

We *define* the so-called *wave-function* associated to $|\psi\rangle$ as

$$\psi(x_n) = \langle x_n\,|\psi\rangle \ \ , \tag{6.6}$$

and then the probability to find the particle in the interval I_n of size ϵ around x_n is $\epsilon\,|\psi(x_n)|^2$. But this means that $|\psi(x_n)|^2$ is the probability to find the particle in a (small) interval divided by the size of that interval, i.e. $|\psi(x_n)|^2$ is the *probability density*.

Because $|\psi\rangle$ is normalised, and the states $|\widetilde{x}_n\rangle$ are also normalised, the fact that the total probability to find the particle somewhere is one is expressed as :

$$1 = \sum_n \mathcal{P}(x_n, \psi) = \sum_n \epsilon\,|\langle x_n\,|\psi\rangle\,|^2 = \sum_n \epsilon\,|\psi(x_n)|^2 \ , \tag{6.7}$$

where the sum over n is an infinite sum over the (countably) infinitely many discrete positions x_n. In general, we expect that the wave-function $\psi(x)$ does not vary much as x is changed from

x_n to x_{n+1} and that we may perfectly well consider $\psi(x)$ as a (smooth[32]) function defined for all real x. Then, since $\epsilon = x_{n+1} - x_n \simeq \mathrm{d}x$, we may (to a very good precision) replace $\sum_n \epsilon \to \int \mathrm{d}x$ and then the previous relation becomes

$$1 = \int \mathrm{d}x \, |\psi(x)|^2 \ , \tag{6.8}$$

where it is understood that the integral runs over all real values of x.

6.2 Infinite-dimensional Hilbert spaces and linear operators on $L^2(\mathbf{R})$

Mathematically, (6.8) is the statement that the wave-function must be square-integrable (not necessarily continuous though). The space of these (complex-valued) square-integrable functions with real argument is denoted as $L^2(\mathbf{R})$. More precisely,

$$\psi \in L^2(\mathbf{R}) \quad \Leftrightarrow \quad \int \mathrm{d}x \, |\psi(x)|^2 < \infty \ . \tag{6.9}$$

But whenever $\psi \neq 0$ satisfies this condition, we can always multiply it by an appropriate constant to achieve (6.8). One can show that $L^2(\mathbf{R})$ has the structure of an (infinite-dimensional) vector space, and one can define an inner product on this space as

$$(\psi, \chi) = \int \mathrm{d}x \, \psi^*(x)\chi(x) \ , \tag{6.10}$$

which makes it a pre-Hilbert space. The inner product always induces a norm as $||\psi||^2 = (\psi, \psi)$. A norm must be such that $||\psi|| = 0 \Rightarrow \psi = 0$. In particular also $||\psi - \chi|| = 0 \Rightarrow \psi = \chi$. This means that whenever $\int \mathrm{d}x |\psi(x) - \chi(x)|^2 = 0$ we identify ψ with χ in $L^2(\mathbf{R})$.[33]

Exercise 6.1 : *Show that* $0 \leq |a\psi - b\chi|^2$ *implies* $ab^*\psi\chi^* + c.c \leq |a\psi|^2 + |b\chi|^2$, *and then deduce* $|a\psi + b\chi|^2 \leq 2|a\psi|^2 + 2|b\chi|^2$. *Use this result to show that if* $\psi, \chi \in L^2(\mathbf{R})$ *then any linear combination also is in* $L^2(\mathbf{R})$, *so this is indeed a vector space over* $\mathbf{C}$.

A Hilbert space is a pre-Hilbert space that is complete, which means that for every Cauchy sequence there exists an element in this Hilbert space to which the sequence converges. Every finite dimensional pre-Hilbert space is a Hilbert space, but this is not necessarily true for infinite-dimensional spaces. However, one can show that every Cauchy sequence in $L^2(\mathbf{R})$ converges to an element in $L^2(\mathbf{R})$ (convergence being appreciated in the $L^2(\mathbf{R})$-norm $||\psi||^2 = \int \mathrm{d}x \, |\psi(x)|^2$), so that $L^2(\mathbf{R})$ is a Hilbert space. It is an infinite-dimensional Hilbert space but there exist (infinitely many) basis' that have countably infinitely many basis elements. Moreover, for each such basis

[32] A smooth function is usually meant to be a function of class C^∞, i.e. many times infinitely differentiable. This will turn out to be too strong, but for all reasonable physical situations we should expect that $\psi(x)$ is at least continuous.

[33] For continuous functions this is trivial, but $L^2(\mathbf{R})$ also contains non-continuous functions. In particular, the function that vanishes for all $x \in \mathbf{R}$ except at isolated points (on a set of measure 0) where it takes finite non-zero values, is identified with the function that is 0 everywhere. Similarly, if $f(x) = g(x)$ except on a set of measure 0 where they differ by finite constants, they get identified in $L^2(\mathbf{R})$.

one can choose the basis elements as orthonormal. This makes $L^2(\mathbf{R})$ a "nice" Hilbert space. Let $\{f_n\}_{n=1,2,...\infty}$ be one such orthonormal basis. This means that an element ψ of $L^2(\mathbf{R})$ can be decomposed on this basis with (complex) coefficients c_n as

$$\psi = \sum_{n=1}^{\infty} c_n f_n \quad \Leftrightarrow \quad \psi(x) = \sum_{n=1}^{\infty} c_n f_n(x) \ , \tag{6.11}$$

where the infinite sum over n converges in the L^2-norm, i.e.

$$\lim_{N\to\infty} \Big|\Big|\psi - \sum_{n=1}^{N} c_n f_n\Big|\Big|^2 \equiv \lim_{N\to\infty} \int \mathrm{d}x \, \big|\psi(x) - \sum_{n=1}^{N} c_n f_n(x)\big|^2 = 0 \ . \tag{6.12}$$

The coefficients c_n are obtained by taking the inner product of (6.11) with f_m :

$$(f_m, \psi) = (f_m, \sum_{n=1}^{\infty} c_n f_n) = \sum_{n=1}^{\infty} c_n (f_m, f_n) = \sum_{n=1}^{\infty} c_n \delta_{mn} = c_m \ , \tag{6.13}$$

where we used the linearity of the inner product defined by the integral (6.10), interchanged the infinite sum and the integral and used the orthonormality of the basis. The interchange can be justified by first considering the finite sums as in (6.12) and then taking the limit $N \to \infty$. We also have

$$||\psi||^2 = (\psi, \psi) = \big(\sum_n c_n f_n, \sum_m c_m f_m\big) = \sum_{n,m} c_n^* c_m (f_n, f_m) = \sum_{n,m} c_n^* c_m \delta_{nm} = \sum_n |c_n|^2 \ . \tag{6.14}$$

Giving the function ψ or giving the sequence of coefficients c_n with respect to some fixed basis is equivalent information. We see that $\psi \in L^2(\mathbf{R})$ is equivalent to the statement that $\sum_n |c_n|^2 < \infty$. The infinite sequences $\{c_n\}$ such that $\sum_n |c_n|^2 < \infty$ form another Hilbert space, called ℓ^2, which obviously is isomorphic to $L^2(\mathbf{R})$.

Exercise 6.2 : *Consider the periodic functions on the interval* $[0, 2\pi]$*, or equivalently the functions defined on the unit circle* S^1*. One defines the inner product of two functions* ψ *and* χ *as in* (6.10) *but with the integral going from 0 to* 2π*. What is the corresponding norm* $||\psi||$ *? Show that the set* $\{f_n\}_{n\in\mathbf{Z}}$ *with* $f_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}$ *forms an orthonormal set. Using the known results about Fourier series show that any continuous periodic function* ψ *can be written as* $\psi(x) = \sum_n c_n f_n(x)$ *where the sum converges to* $\psi(x)$ *at every point* $x \in [0, 2\pi]$*. Determine the coefficients* c_n *in terms of integrals involving* ψ *and* f_n*. Discuss how this is extended to arbitrary elements of* $L^2(S^1)$ *when convergence is appreciated in the corresponding norm.*

A linear functional on a Hilbert space $\mathcal{H}$ is a linear map that maps each element of the Hilbert space to a complex number. Clearly, for each fixed $\chi \in L^2(\mathbf{R})$ the map $\psi \to (\chi, \psi)$ is a linear functional. Conversely let $F(\psi)$ be a linear functional. Then $F(\psi) = F(\sum_n c_n f_n) = \sum_n c_n F(f_n)$. If we let $\chi_F = \sum_m (F(f_m))^* f_m$ then

$$\begin{aligned} (\chi_F, \psi) &= \big(\sum_m (F(f_m))^* f_m, \sum_n c_n f_n\big) = \sum_{n,m} F(f_m) c_n \, (f_m, f_n) = \sum_{n,m} F(f_m) c_n \delta_{mn} \\ &= \sum_n c_n F(f_n) = F(\psi) \ . \end{aligned} \tag{6.15}$$

Hence, the linear functional F determines a unique χ_F such that $(\chi_F, \psi) = F(\psi)$. The argument we have given does not rely on the properties of $L^2(\mathbf{R})$ but only on the existence of a orthonormal basis and the fact that we could interchange the sums and taking the inner products. We can then identify the space of linear functionals on $\mathcal{H}$, called the dual Hilbert space, and denoted as $\mathcal{H}^\dagger$ with the Hilbert space $\mathcal{H}$ itself : $\mathcal{H}^\dagger \simeq \mathcal{H}$.

Defining linear operators on infinite dimensional Hilbert spaces is much more subtle than on finite-dimensional Hilbert spaces. The linearity requirement is the same, but there are many interesting linear operators that cannot be defined on all elements of the Hilbert space. More precisely, this means that if we try to apply these operators on an arbitrary element of the Hilbert space, the resulting element no longer necessarily is in this Hilbert space. This can be easily seen for the Hilbert space ℓ^2. Consider the linear operator that consists of multiplying each c_n by n. Then the sequence $c_n = \frac{1}{n}$ is in ℓ^2 (since $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$) but obviously $c'_n = nc_n = 1$ is not. Similarly, the functions $\psi_1(x) = \frac{1}{\sqrt{1+x^2}}$ and $\psi_2(x) = \frac{1}{1+x^2}$ are in $L^2(\mathbf{R})$. The multiplication of $\psi(x)$ by x clearly is a linear operation, but $x\psi_2(x)$ is in $L^2(\mathbf{R})$, while $x\psi_1(x)$ is not. The operation D which maps a function to its derivative cannot even be defined on all elements of $L^2(\mathbf{R})$ but only on those functions that are differentiable. Hence each linear operator A comes with its corresponding domain of definition $\mathcal{D}(A) \subset L^2(\mathbf{R})$ such that for all $\psi \in \mathcal{D}(A)$ we have that $A\psi$ is well defined and is an element of $L^2(\mathbf{R})$. Note that the same linear operation with two different domains of definition should be considered as two different linear operators.

While differentiable functions only are a "small" subset of all functions in $L^2(\mathbf{R})$, they actually constitute a dense subset. A dense subset is a subset such that any element of $L^2(\mathbf{R})$ can be arbitrarily well approximated by an element in the dense subset. Said differently, for every $\psi \in L^2(\mathbf{R})$ and ϵ there is a differentiable function $f^\epsilon_\psi \in C^1(\mathbf{R})$ such that $||\psi - f^\epsilon_\psi|| < \epsilon$. Even the set $C^\infty_c(\mathbf{R})$ of infinitely differentiable functions with compact support is dense in $L^2(\mathbf{R})$! Another very useful subspace of $L^2(\mathbf{R})$ is $\mathcal{S}$ consisting of all infinitely differentiable functions f such that f and all its derivatives vanish faster than any (positive) power of $\frac{1}{|x|}$ as $|x| \to \infty$. An example of such a function is $P_n(x)e^{-\alpha x^2}$ with $\alpha > 0$ and P_n a polynomial in x. This subspace $\mathcal{S}$ is again dense in $L^2(\mathbf{R})$. Elements of $\mathcal{S}$ are sometimes called test-functions. Note that if $f \in \mathcal{S}$ then so are $g_{(k)}$ and $h_{(k)}$ with $g_{(k)} = \frac{\mathrm{d}^k f}{\mathrm{d}x^k}$ and $h_{(k)}(x) = x^k f(x)$.

A linear operator is called densely defined if its domain of definition is dense in $L^2(\mathbf{R})$. Since we can always define multiple differentiation and multiplication with x^k on any function in $\mathcal{S}$, these linear operators are densely defined.

We now want to define the operator $A^\dagger$ such that, as before, $(A^\dagger\chi, \psi) = (\chi, A\psi)$. The subtleties come from the domains of definition of A and $A^\dagger$. We suppose that A is densely defined on $\mathcal{H}$, i.e. $\mathcal{D}(A)$ is dense in $\mathcal{H}$. Then the adjoint operator $A^\dagger$ is defined to act on a domain $\mathcal{D}(A^\dagger) \subset \mathcal{H}$ that consists of those φ for which there is a $\chi \in \mathcal{H}$ such that $(\varphi, A\psi) = (\chi, \psi)$. Then we write $\chi = A^\dagger\varphi$:

$$(\varphi, A\psi) = (\chi, \psi) \quad , \quad \chi = A^\dagger\varphi \ . \tag{6.16}$$

Of course, this is just saying that

$$(\varphi, A\psi) = (A^\dagger\varphi, \psi) \quad , \quad A^\dagger\varphi \in \mathcal{H} \ . \tag{6.17}$$

Note that it is obvious from this definition that

$$(AB)^\dagger = B^\dagger\, A^\dagger \ . \tag{6.18}$$

The set $G(A) = \{(\psi, A\psi)|\psi \in \mathcal{D}(A)\}$ is called the graph of A, and similarly for $G(A^\dagger)$. If $G(A) \subseteq G(A^\dagger)$ and for all $\varphi, \psi \in \mathcal{D}(A)$ we have $(\varphi, A\psi) = (A\varphi, \psi)$ then A is called symmetric. This looks like $A^\dagger = A$ but actually $A^\dagger$ may have a larger domain of definition than A so that strictly speaking they are not the same. However, if $G(A) = G(A^\dagger)$ and for all $\varphi, \psi \in \mathcal{D}(A)$ we have $(\varphi, A\psi) = (A\varphi, \psi)$ then A is called self-adjoint, and we have indeed $A^\dagger = A$. Clearly, every self-adjoint operator is symmetric but the converse is not true.

The *spectral theorem* states that a self-adjoint linear operator A "can be diagonalised". This means that the set of all its eigenfunctions (eigenvectors) forms a basis of the Hilbert space which can be taken as orthonormal.[34] This is most important since this means that any element $\psi \in \mathcal{H}$ can be written as $\psi = \sum_n c_n f_n$ with the f_n being the eigenvectors of A. Also, for self-adjoint operators all eigenvalues are real. In physics, the term "hermitian" is also used and often refers to either symmetric or self-adjoint. Although the spectral theorem only applies to self-adjoint operators, there are certain (but not all !) operators that are symmetric without being self-adjoint, that nevertheless admit an orthonormal basis of eigenvectors.

Example 1 : Let $\mathcal{H} = L^2(S^1)$ where S^1 is the unit circle, i.e. the interval $[0, 2\pi]$ with 0 and 2π identified. The functions on S^1 are just the periodic functions of period 2π. Let $A = i\frac{\mathrm{d}}{\mathrm{d}x}$. Its domain of definition are just the differentiable functions on the circle, $\mathcal{D}(A) = C^1(S^1)$. Indeed, every differentiable function on the circle is bounded and hence square integrable and, if $f \in C^1(S^1)$, then if' is continuous and hence bounded and square integrable. We have

$$(g, Af) = \int_0^{2\pi} g^*(x) i f'(x) = -\int_0^{2\pi} g'^*(x) i f(x) = \int_0^{2\pi} (ig')^*(x) f(x) \ , \tag{6.19}$$

since both f and g are periodic and the integration by parts did not generate any boundary term. (Note that one gets two minus signs, one from integrating by parts and one from $i^* = -i$.) The domain of definition of $A^\dagger$ consists of all those $g \in L^2(S^1)$ that are differentiable, i.e. $\mathcal{D}(A^\dagger) = C^1(S^1) = \mathcal{D}(A)$ and obviously also $A^\dagger = i\frac{\mathrm{d}}{\mathrm{d}x} = A$. Hence A is self-adjoint.

Example 2 : Let $\mathcal{H} = L^2([0, 2\pi])$ be the space of square-integrable functions on the interval $[0, 2\pi]$ (without any periodicity requirement). Let again $A = i\frac{\mathrm{d}}{\mathrm{d}x}$. One must now specify the domain on which one wants to define A. Obviously $\mathcal{D}(A) \subseteq C^1([0, 2\pi])$, but in addition one can also impose boundary conditions. A first choice is $\mathcal{D}(A) = \{f \in C^1([0, 2\pi])|f(0) = f(2\pi) = 0\}$. Then as above

[34] As we will see below, for this to be true, we will have to slightly generalise the notion of eigenvectors and orthonormality. In particular, one will allow "non-normalisable" basis' depending on a continuous parameter ν, and then $\sum_n \to \int \mathrm{d}\nu$ and $\delta_{nm} \to \delta(\nu - \nu')$.

in (6.19) we can integrate by parts as long as $g \in C^1([0,2\pi])$, without generating boundary terms because now f vanishes at the boundaries. One concludes again that $A^\dagger = i\frac{\mathrm{d}}{\mathrm{d}x} = A$, but obviously $\mathcal{D}(A^\dagger) = C^1([0,2\pi])$ which is strictly larger than $\mathcal{D}(A)$. Hence, at present, A is symmetric but not self-adjoint.

Note that in both examples the eigenvalue equation $Af(x) \equiv if'(x) = af(x)$ leads to $f(x) \sim e^{-iax}$. In the first example, periodicity imposes $e^{-ia2\pi} = 1$ so that $a = n \in \mathbf{Z}$, and the eigenvalues are real (and discrete), while in the second example there are no boundary conditions for $A^\dagger$ and any $a \in \mathbf{C}$ would be an eigenvalue, while the boundary conditions for A (vanishing at $x = 0$ and at $x = 2\pi$) cannot be satisfied by ce^{iax} unless $c = 0$. We see that the distinction between symmetric and self-adjoint is not merely a question of mathematical rigour, but it is essential for the spectral theorem to apply and ensure that the eigenvalues are real.

Notwithstanding their importance, in the rest of this lecture we will not go any further into these mathematical subtleties.

6.3 Infinite-dimensional Hilbert spaces : kets, bras, position operator

The elements of the Hilbert space $L^2(\mathbf{R})$ appeared as the wave-functions $\psi(x)$ that where defined as "interpolating" functions from the inner products of the ket $|\psi\rangle \in \mathcal{H}$ with the discrete position eigenstates $|x_n\rangle \in \mathcal{H}$. The condition of square-integrability of the wave-function was just the translation of the fact that $|\psi\rangle$ was normalised. The vector space structure of $L^2(\mathbf{R})$ directly corresponds to the vector space structure of $\mathcal{H}$. Also the definition of the inner products translate as follows

$$\begin{aligned}(\chi,\psi) &= \int \mathrm{d}x\, \chi^*(x)\psi(x) \simeq \sum_n \epsilon\, \chi^*(x_n)\psi(x_n) = \sum_n \epsilon\, \langle\chi\,|x_n\rangle\,\langle x_n\,|\psi\rangle = \sum_n \langle\chi\,|\widetilde{x}_n\rangle\,\langle\widetilde{x}_n\,|\psi\rangle \\ &= \langle\chi|\,\big(\sum_n |\widetilde{x}_n\rangle\,\langle\widetilde{x}_n|\,\big)\,|\psi\rangle = \langle\chi\,|\psi\rangle \ , \end{aligned} \tag{6.20}$$

where we used that the $|\widetilde{x}_n\rangle$ constitute an orthonormal basis. We then have a perfect isomorphism between the Hilbert space $\mathcal{H}$ of the kets $|\psi\rangle$ and the Hilbert space $L^2(\mathbf{R})$ of the wave functions. In particular also, if $|f_n\rangle$ is an orthonormal basis of $\mathcal{H}$ (and then $f_n(x) = \langle x\,|f_n\rangle$ is an orthonormal basis of $L^2(\mathbf{R})$), we have the relations analogous to (6.11) :

$$|\psi\rangle = \sum_{n=1}^{\infty} c_n\,|f_n\rangle \quad , \quad c_n = \langle f_n\,|\psi\rangle \ , \tag{6.21}$$

where the infinite sum over n converges in the norm on $\mathcal{H}$ induced by the inner product, i.e.

$$\lim_{N\to\infty} \Big\langle \psi - \sum_{n=1}^{N} c_n f_n \Big| \psi - \sum_{m=1}^{N} c_m f_m \Big\rangle = 0 \ . \tag{6.22}$$

The orthonormal basis $|f_n\rangle$ just introduced of course satisfies $\langle f_n\,|f_m\rangle = \delta_{nm}$. The same was true for the discrets position eigenstates $|\widetilde{x}_n\rangle$ since we had $\langle\widetilde{x}_n\,|\widetilde{x}_m\rangle = \delta_{nm}$. However, the

wave-functions $\psi(x) = \langle x\,|\psi\rangle$ are defined with the differently normalized $|x\rangle$ and considering the step-size ϵ as infinitesimally small. Let's see what happens if we try to take the $\epsilon \to 0$ limit. We should have

$$\mathbf{1} = \sum_n |\widetilde{x}_n\rangle\,\langle\widetilde{x}_n| = \sum_n \epsilon\ |x_n\rangle\,\langle x_n| \sim_{\epsilon\to 0} \int \mathrm{d}x\ |x\rangle\,\langle x|\ , \tag{6.23}$$

as well as

$$\frac{1}{\epsilon}\,\delta_{nm} = \frac{1}{\epsilon}\,\langle\widetilde{x}_n\,|\widetilde{x}_m\rangle = \langle x_n\,|x_m\rangle\ . \tag{6.24}$$

In the limit $\epsilon \to 0$ this last relation should tell us that $\langle x\,|x'\rangle = 0$ whenever $x \neq x'$ and if $x = x'$ this inner product (which then is the norm squared of $|x\rangle$) becomes infinite. Stated this way, this is not very satisfactory. But let us apply (6.23) to the ket $|x'\rangle$. This gives in the $\epsilon \to 0$ limit

$$|x'\rangle = \int \mathrm{d}x\ |x\rangle\,\langle x\,|x'\rangle\ . \tag{6.25}$$

This tells us precisely that

$$\langle x\,|x'\rangle = \delta(x - x')\ , \tag{6.26}$$

where $\delta(x - x')$ is Dirac's "delta-function". The intuitive representation of this "delta-function" is that it is obtained as the $\epsilon \to 0$ limit of a function $\delta_{x',e}(x)$ that vanishes everywhere except in an interval of size ϵ around x' where it equals $\frac{1}{\epsilon}$. Then for any continuous funtion $f(x)$ one has

$$\int \mathrm{d}x\, f(x)\,\delta_{x',e}(x) = \frac{1}{\epsilon}\int_{x'-\epsilon/2}^{x'+\epsilon/2} \mathrm{d}x\, f(x) \sim_{\epsilon\to 0}\ f(x')\ , \tag{6.27}$$

which is written as

$$\int \mathrm{d}x\, f(x)\,\delta(x - x') = f(x')\ . \tag{6.28}$$

This notation suggests that we have linearity, i.e. $\int \mathrm{d}x\,\big(af(x) + bg(x)\big)\,\delta(x - x') = af(x') + bg(x')$. Mathematically, what Dirac's delta does is to act as a linear functional on an appropriate space of functions since it maps functions f to (complex) numbers $f(x')$, and it does so linearly. The linear functionals on an appropriate set of so-called test-functions (infinitely derivable and falling off sufficiently fast as $|x| \to \infty$) are called distributions and then Dirac's delta is just a particular distribution. One can then also define $\delta'(x - x')$ which is such that $\int \mathrm{d}x\,\delta'(x - x')f(x) = -f'(x')$ as one would expect from integrating by parts, etc. By definition, $\delta(x - x') = \delta(x' - x)$. One can then also easily derive relations like $x\delta(x - x') = x'\delta(x - x')$, etc. Here we will simply manipulate Dirac's delta according to these rules and remember that

$$\boxed{\langle x\,|x'\rangle = \delta(x - x') \quad , \quad \int \mathrm{d}x\ |x\rangle\,\langle x| = \mathbf{1}\ .} \tag{6.29}$$

These are almost the characteristics of an orthonormal basis except, of course, that these states $|x\rangle$ are of "infinite norm". Taking naively $x = x'$ in the first relation yields $\delta(0)$ which should be

interpreted as ∞. However, the correct thing to do is to interpret these relations as appropriate for distributions. This means that $\langle x\,|x'\rangle$ has no meaning as a function of x and x' and one cannot simply set $x = x'$. What makes sense, however, are manipulations as the following ones

$$|\psi\rangle = \int \mathrm{d}x\ |x\rangle\,\langle x\,|\psi\rangle = \int \mathrm{d}x\ |x\rangle\,\psi(x) \quad , \quad \langle\chi| = \int \mathrm{d}x'\,\langle\chi\,|x'\rangle\,\langle x'| = \int \mathrm{d}x'\,\chi(x')^*\,\langle x'| \ , \quad (6.30)$$

and then

$$\langle\chi\,|\psi\rangle = \int \mathrm{d}x\,\mathrm{d}x'\,\chi(x')^*\psi(x)\,\langle x'\,|x\rangle = \int \mathrm{d}x\,\mathrm{d}x'\,\chi(x')^*\psi(x)\,\delta(x'-x) = \int \mathrm{d}x\,\chi(x')^*\psi(x)\ . \quad (6.31)$$

In particular, this shows again the correspondence (6.20) between the inner product of two elements of $\mathcal{H}$ and the inner product between the corresponding wave-functions in $L^2(\mathbf{R})$.

Initially we introduced the $|x_n\rangle$ as eigenstates of the position operator as $X\,|x_n\rangle = x_n\,|x_n\rangle$. Obviously in the limit $\epsilon \to 0$ this becomes

$$X\,|x\rangle = x\,|x\rangle \ . \quad (6.32)$$

Then of course also $X\,|x'\rangle = x'\,|x'\rangle$ and

$$\begin{aligned} \langle x'|\,\big(X\,|x\rangle\,\big) &= \langle x'|\,\big(x\,|x\rangle\,\big) = x\langle x'\,|x\rangle = x\,\delta(x-x') = x'\,\delta(x-x') = x'\langle x\,|x'\rangle = \langle x|\,\big(x'\,|x'\rangle\,\big) \\ &= \langle x|\,\big(X\,|x'\rangle\,\big) \equiv \langle x\,|Xx'\rangle = \big(\langle Xx'\,|x\rangle\,\big)^* = \langle Xx'\,|x\rangle \ , \end{aligned} \quad (6.33)$$

where in the next to last step we used $\langle\chi\,|\psi\rangle^* = \langle\psi\,|\chi\rangle$ and in the last step that this inner product is real. We see that the linear operator X is symmetric, and if we appropriately adjust its domain of definition it is also self-adjoint. Then the action of this operator on a ket $|\psi\rangle$ is such that

$$\langle x|\,X\,|\psi\rangle = \langle Xx\,|\psi\rangle = x\langle x\,|\psi\rangle \ . \quad (6.34)$$

Le us insist that the eigenstates $|x\rangle$ of the position operator are not strictly speaking in the Hilbert space $\mathcal{H}$ since they are not normalisable. However, the states

$$|F\rangle = \int \mathrm{d}x\,F(x)\,|x\rangle \quad , \quad \int \mathrm{d}x\,|F(x)|^2 < \infty \ , \quad (6.35)$$

are in $\mathcal{H}$ since $\langle F\,|F\rangle = \int \mathrm{d}x\,|F(x)|^2 < \infty$. Of course, $F(x)$ is just the wave-function corresponding to the ket $|F\rangle$. By taking $F(x) = \frac{1}{\sqrt{\epsilon}}$ for $x \in [x_0 - \epsilon/2, x_0 + \epsilon/2]$ one gets states that are normalised to one and that are approximate eigenstates of X with approximate eigenvalue x_0. Taking such a superposition of exact eigenstates to obtain a normalisable state is referred to as making a wave packet. It is much easier to work with the non-normalisable eigenstates $|x\rangle$ but having in mind that, if needed, we can always replace them by such normalisable wave packets that are approximate eigenstates. Of course, for $x_0 = x_n$ these wave packets are just like the $|\widehat{x}_n\rangle$ we started from.

6.4 Momentum operator and eigenstates, Fourier transform

The next thing we need to do is to define the momentum operator and its eigenstates. At this point we need the experimental input of de Broglie's relation that a momentum p should correspond to a wave-vector k by the relation $p = \hbar k$. This should mean that a particle of (almost) well-defined momentum should somehow correspond to an (almost) plane wave of wave vector $k = p/\hbar$ i.e. to $e^{ikx} = e^{ipx/\hbar}$. This is a function of x and hence should represent the wave-function $\psi(x)$. Let us denote this wave-function as $f_p(x) \sim e^{ipx/\hbar}$. We did not indicate any normalisation since if we try to compute its normalisation we find $\int \mathrm{d}x\, |e^{ipx/\hbar}|^2 = \int \mathrm{d}x\, 1 = \infty$, so these "wave-functions" are not in $L^2(\mathbf{R})$, just as the $|x\rangle$ where not in $\mathcal{H}$, or the $\delta(x - x_0)$ are not in $L^2(\mathbf{R})$. But just as with the latter, these non-normalisable "wave-functions" of definite momentum will be a most useful computational tool. We then define the eigenstates of the momentum operator P and corresponding wave functions as

$$P\,|p\rangle = p\,|p\rangle \quad , \quad f_p(x) \equiv \langle x\,|p\rangle = \frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}}\ . \tag{6.36}$$

We will need to evaluate the inner product $(f_{p'}, f_p) = \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} \mathrm{d}x\, e^{i(p-p')x/\hbar}$. If $p \neq p'$, the integrand oscillates and the different contributions cancel, resulting in a vanishing integral. If $p = p'$, the integrand is 1 and the integral diverges. This ressembles much a Dirac delta $\delta(p - p')$. One can make things more precise by first working with a finite interval and periodic boundary conditions resulting in discrete values p_n of p, then the corresponding $f_n \equiv f_{p_n}$ are orthogonal for $n' \neq n$. Let us not go through the details and simply state the required relation:

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}x\, e^{i(k-k')x} = \delta(k - k') \quad , \quad \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}k\, e^{ik(x-x')} = \delta(x - x')\ . \tag{6.37}$$

By changing variables as $k = p/\hbar$ one also gets $\frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} \mathrm{d}x\, e^{i(p-p')x/\hbar} = \delta(p - p')$, as well as $\frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} \mathrm{d}p\, e^{ip(x-x')/\hbar} = \delta(x - x')$. It follows that

$$\begin{aligned} \langle p'\,|p\rangle &= \langle p'|\,\Big(\int \mathrm{d}x\ |x\rangle\,\langle x|\,\Big)\,|p\rangle = \int \mathrm{d}x\,\langle p'\,|x\rangle\,\langle x\,|p\rangle = \int \mathrm{d}x\, f^*_{p'}(x) f_p(x) = \frac{1}{2\pi\hbar}\int \mathrm{d}x\, e^{i(p-p')x/\hbar} \\ &= \delta(p - p')\ . \end{aligned} \tag{6.38}$$

These momentum eigenstates form a non-normalisable basis just as did the position eigenstates. Again, one can form superpositions (wave packets) that are normalisable states and approximate momentum eigenstates. We also have the corresponding completeness relation as one can see as follows (interchanging freely the orders of integrations)

$$\begin{aligned} \int \mathrm{d}p\ |p\rangle\,\langle p| &= \int \mathrm{d}p\, \mathbf{1}\,|p\rangle\,\langle p|\,\mathbf{1} = \int \mathrm{d}p\,\mathrm{d}x\,\mathrm{d}x'\ |x\rangle\,\langle x\,|p\rangle\,\langle p\,|x'\rangle\,\langle x'| = \int \mathrm{d}x\mathrm{d}x'\mathrm{d}p\ |x\rangle\,\frac{e^{ip(x-x')/\hbar}}{2\pi\hbar}\,\langle x'| \\ &= \int \mathrm{d}x\mathrm{d}x'\,\delta(x - x')\,|x\rangle\,\langle x'| = \int \mathrm{d}x\,|x\rangle\,\langle x| = \mathbf{1}\ . \end{aligned} \tag{6.39}$$

To summarise :

$$\boxed{\langle p' \,|p\rangle = \delta(p'-p) \quad , \quad \int \mathrm{d}p \;|p\rangle\,\langle p| = \mathbf{1} \quad , \quad \langle x\,|p\rangle = \frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}}\;.} \tag{6.40}$$

We can then decompose any state in $\mathcal{H}$ on these momentum eigenstates as

$$|\psi\rangle = \int \mathrm{d}p \;|p\rangle\,\widetilde{\psi}(p) \quad , \quad \widetilde{\psi}(p) = \langle p\,|\psi\rangle \;. \tag{6.41}$$

The relation between the wave-function $\psi(x)$ and the $\widetilde{\psi}(p)$ is given by the *Fourier transform*

$$\boxed{\widetilde{\psi}(p) = \langle p\,|\psi\rangle = \int \mathrm{d}x\,\langle p\,|x\rangle\,\langle x\,|\psi\rangle = \frac{1}{\sqrt{2\pi\hbar}}\int \mathrm{d}x\, e^{-ipx/\hbar}\,\psi(x)\;,} \tag{6.42}$$

and inverse Fourier transform

$$\boxed{\psi(x) = \langle x\,|\psi\rangle = \int \mathrm{d}p\,\langle x\,|p\rangle\,\langle p\,|\psi\rangle = \frac{1}{\sqrt{2\pi\hbar}}\int \mathrm{d}p\, e^{ipx/\hbar}\,\widetilde{\psi}(p)\;.} \tag{6.43}$$

The Fourier transform is an isometry which means that

$$\int \mathrm{d}x\,|\psi(x)|^2 = \int \mathrm{d}p\,|\widetilde{\psi}(p)|^2\;. \tag{6.44}$$

This follows from $\int \mathrm{d}x\,|\psi(x)|^2 = \int \mathrm{d}x\,\langle\psi\,|x\rangle\,\langle x\,|\psi\rangle = \langle\psi|\,\mathbf{1}\,|\psi\rangle = \int \mathrm{d}p\,\langle\psi\,|p\rangle\,\langle p\,|\psi\rangle = \int \mathrm{d}p\,|\widetilde{\psi}(p)|^2$.

Exercise 6.3 : *Rederive the isometry property* (6.44) *without using the bra and ket notation, but by directly working with* $\psi(x) = \frac{1}{\sqrt{2\pi\hbar}}\int \mathrm{d}p\, e^{ipx/\hbar}\,\widetilde{\psi}(p)$.

We may interpret the right-hand side of equation (6.43) as a superposition of plane waves $\sim e^{ipx/\hbar}$, corresponding to momentum eigenstates $|p\rangle$, with $\widetilde{\psi}(p) = \langle p\,|\psi\rangle$ being the coefficient (amplitude) for each of them. While a single plane wave is not normalisable, these superpositions are normalisable precisely if the coefficients satisfy $\int \mathrm{d}p\,|\widetilde{\psi}(p)|^2 < \infty$.

Let us work out the important example of a superposition of momentum eigenstates with p centered around some p_0. We suppose that $\widetilde{\psi}(p)$ is a Gaussian, centered at p_0 of width $\sqrt{2}\sigma$

$$\widetilde{\psi}_{p_0,\sigma}(p) = \frac{1}{\sqrt{\pi^{1/2}\sigma}}\, e^{-(p-p_0)^2/(2\sigma^2)}\;. \tag{6.45}$$

Then the probability density that the particle has a momentum p is $|\widetilde{\psi}(p)|^2 = \frac{1}{\pi^{1/2}\sigma}\, e^{-(p-p_0)^2/\sigma^2}$ which is again a Gaussian centered at p_0, now of width σ. The prefactor ensures that it is

correctly normalised. The Fourrier transform of $\widetilde{\psi}_{p_0,\sigma}$ is computed by "completing the square" in the exponent :

$$\begin{aligned}
\psi_{x_0,\sigma}(x) &= \frac{1}{\sqrt{2\pi^{3/2}\hbar\sigma}} \int \mathrm{d}p\, e^{ipx/\hbar}\, e^{-(p-p_0)^2/(2\sigma^2)} = \frac{e^{-p_0^2/(2\sigma^2)}}{\sqrt{2\pi^{3/2}\hbar\sigma}} \int \mathrm{d}p\, e^{-\left(p^2-2p(p_0+i\frac{x}{\hbar}\sigma^2)\right)/(2\sigma^2)} \\
&= \frac{e^{-p_0^2/(2\sigma^2)}}{\sqrt{2\pi^{3/2}\hbar\sigma}} \int \mathrm{d}p\, e^{-\left(p-(p_0+i\frac{x}{\hbar}\sigma^2)\right)^2/(2\sigma^2)}\, e^{(p_0+i\frac{x}{\hbar}\sigma^2)^2/(2\sigma^2)} \\
&= \frac{e^{-p_0^2/(2\sigma^2)}}{\sqrt{2\pi^{3/2}\hbar\sigma}}\, e^{(p_0+i\frac{x}{\hbar}\sigma^2)^2/(2\sigma^2)}\, \sqrt{2\pi}\sigma \\
&= \sqrt{\frac{\sigma}{\pi^{1/2}\hbar}}\, e^{-\frac{x^2}{2}\left(\frac{\sigma}{\hbar}\right)^2}\, e^{ip_0x/\hbar}\ . \qquad (6.46)
\end{aligned}$$

We get back a Gaussian (now in x). This can be interpreted as a wave with $k = p_0/\hbar$ but with a Gaussian envelope of width $\sqrt{2}\frac{\hbar}{\sigma}$. The corresponding probability density is also a Gaussian, centered at $x = 0$ with width $\frac{\hbar}{\sigma}$. We see that the smaller the width of the momentum distribution the larger is the spatial extension and vice versa.

Exercise 6.4 : *For these Gaussian wave functions $\psi(x)$ and $\widetilde{\psi}(p)$, compute the mean values $\langle X\rangle$ and $\langle P\rangle$ of the position and momentum and show that their variances equal $(\Delta X)^2 = \langle (X - \langle X\rangle)^2\rangle = \frac{\hbar^2}{2\sigma^2}$ and $(\Delta P)^2 = \langle (P - \langle P\rangle)^2\rangle = \frac{\sigma^2}{2}$, and conclude that their product satisfies $\Delta X\, \Delta P = \frac{\hbar}{2}$ which is the minimum value allowed by the uncertainty relation (see below).*

Let us now establish what is the wave function of $P\,|\psi\rangle$. Using the closure relation (6.40) and (6.36) one has

$$\begin{aligned}
\langle x|\left(P\,|\psi\rangle\right) &= \int \mathrm{d}p\, \langle x|\,P\,|p\rangle\,\langle p\,|\psi\rangle = \int \mathrm{d}p\, p\,\langle x\,|p\rangle\,\langle p\,|\psi\rangle = \int \mathrm{d}p\, p\,\frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}}\langle p\,|\psi\rangle \\
&= -i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\int \mathrm{d}p\,\frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}}\langle p\,|\psi\rangle = -i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\int \mathrm{d}p\,\langle x\,|p\rangle\,\langle p\,|\psi\rangle = -i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\langle x\,|\psi\rangle \\
&= -i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\psi(x)\ . \qquad (6.47)
\end{aligned}$$

This shows that the wave-function associated to $P\,|\psi\rangle$ is $-i\frac{\mathrm{d}}{\mathrm{d}x}\psi(x)$. The corresponding relation for the position operator was already obtained in (6.34). Let us summarise :

$$\boxed{\langle x|\,X\,|\psi\rangle = x\langle x\,|\psi\rangle = x\,\psi(x) \quad , \quad \langle x|\left(P\,|\psi\rangle\right) = -i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\langle x\,|\psi\rangle = -i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\psi(x)\ .} \qquad (6.48)$$

Sometimes one denotes by $\hat{X}$ and $\hat{P}$ the corresponding operators as acting on the wave-functions, and then the previous relations translate into

$$\boxed{\hat{X}\psi(x) = x\,\psi(x) \quad , \quad \hat{P}\psi(x) = -i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\psi(x)\ .} \qquad (6.49)$$

From either (6.48) or (6.49) one derives the commutation relations of X and P :

$$\boxed{[X,P] \equiv XP - PX = i\hbar\,\mathbf{1}\ .} \tag{6.50}$$

Let us demonstrate this using (6.48) :

$$\begin{aligned}\langle x|\,PX\,|\psi\rangle = \langle x|\,P\big(X\,|\psi\rangle\,\big) = -i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\,\langle x|\,X\,|\psi\rangle = -i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\big(x\,\langle x\,|\psi\rangle\,\big) \;\;&=\;\; x(-i\hbar)\frac{\mathrm{d}}{\mathrm{d}x}\langle x\,|\psi\rangle - i\hbar\langle x\,|\psi\rangle \\ \langle x|\,XP\,|\psi\rangle = \langle x|\,X\big(P\,|\psi\rangle\,\big) = x\,\langle x|\,\big(P\,|\psi\rangle\,\big) \;\;&=\;\; x(-i\hbar)\frac{\mathrm{d}}{\mathrm{d}x}\langle x\,|\psi\rangle\ .\end{aligned} \tag{6.51}$$

Subtracting the first from the second equation yields $\langle x|\,(XP-PX)\,|\psi\rangle = i\hbar\langle x\,|\psi\rangle$, for all $\langle x|$ and all $|\psi\rangle$, which proves (6.50) as a relation between operators. Of course, one should restrict the previous manipulations to those $|\psi\rangle$ or $\psi(x)$ such that both $XP\,|\psi\rangle$ and $PX\,|\psi\rangle$ are well-defined. A convenient dense subspace of $L^2(\mathbf{R})$ is $\mathcal{S}$, already defined before, consisting of all infinitely differentiable functions f such that f and all its derivatives vanish faster than any (positive) power of $\frac{1}{|x|}$ as $|x|\to\infty$. Then, on $\mathcal{S}$ any product of powers of X and P is well-defined and maps elements of $\mathcal{S}$ to elements of $\mathcal{S}$.

Translation operators

The translation operators $T(b)$ are defined as

$$T(b) = e^{-ibP/\hbar}\ , \tag{6.52}$$

and the goal of the following exercise is to show that they indeed act on position eigenstates as one would expect from translations, namely

$$T(b)\,|x\rangle = |x+b\rangle\ . \tag{6.53}$$

Exercise 6.5 : *Show that*

$$[P^n, X] = -i\hbar\,n\,P^{n-1} \quad\Rightarrow\quad [e^{-ibP/\hbar}, X] = -b\,e^{-ibP/\hbar}\ . \tag{6.54}$$

Deduce that

$$e^{-ibP/\hbar}X = (X-b)e^{-ibP/\hbar} \quad\Rightarrow\quad X\Big(e^{-ibP/\hbar}\,|x\rangle\Big) = (x+b)\Big(e^{-ibP/\hbar}\,|x\rangle\Big)\ , \tag{6.55}$$

so that $T(b)\,|x\rangle$ is a position eigenstate with eigenvalue $x+b$. Argue that $T(b)$ is unitary and hence norm-preserving and one can identify

$$|x+b\rangle = e^{-ibP/\hbar}\,|x\rangle \equiv T(b)\,|x\rangle \quad\Leftrightarrow\quad \langle x+b| = \langle x|\,e^{ibP/\hbar} = \langle x|\,T^{-1}(b)\ . \tag{6.56}$$

Show that one then consistently also has

$$\psi(x+b) = \langle x+b\,|\psi\rangle = \langle x|\,e^{ibP/\hbar}\,|\psi\rangle = \exp\Big(b\frac{\mathrm{d}}{\mathrm{d}x}\Big)\langle x\,|\psi\rangle = \exp\Big(b\frac{\mathrm{d}}{\mathrm{d}x}\Big)\psi(x) = \sum_{n=0}^{\infty}\frac{b^n}{n!}\frac{\mathrm{d}^n\psi}{\mathrm{d}x^n}(x)\ , \tag{6.57}$$

where we recover the Taylor expansion of ψ around x.

6.5 Observables

As before, we will assume that our Hilbert space $\mathcal{H}$ admits a basis $\{e_n\}_{n=1,2,...}$ containing countably infinitely many basis "vectors" e_n. (It is then said to be separable.) Then any other Hilbert space $\mathcal{H}'$ that also admits a countably infinite basis $\{f_n\}_{n=1,2,...}$ is isomorphic to $\mathcal{H}$ and the isomorphism is induced by the bijective mapping $e_n \leftrightarrow f_n$. In particular, one may think of $\mathcal{H}$ as $L^2(\mathbf{R})$ or just in terms of the coefficients as ℓ^2. We have also seen that it was *convenient* to use the continuous "basis" of the position or momentum eigenstates in this same Hilbert space.

We have defined symmetric and self-adjoint linear operators acting in this Hilbert space and we have cited the spectral theorem that states that a self-adjoint operator A admits a basis of eigenvectors. Let us now be more precise about this. The set of all eigenvalues of A is called the spectrum of A. This spectrum of A always is a subset of the real numbers, as we will show again shortly. It may contain a discrete part and a continuous part. The discrete part corresponds to eigenvalues with normalisable eigenvectors :

$$A\,|\varphi_{n,i}\rangle = a_n\,|\varphi_{n,i}\rangle \;\; , i = 1, \dots d_n \;, \quad ||\,|\varphi_{n.i}\rangle\,||^2 < \infty \tag{6.58}$$

where d_n is called the degeneracy (or degree of degeneracy) of the eigenvalue a_n. The continuous part of the spectrum corresponds to continuously varying eigenvalues (and non-normalisable eigenvectors) :

$$A\,|\nu\rangle = a(\nu)\,|\nu\rangle \qquad |\nu\rangle \text{ not normalisable} \;, \tag{6.59}$$

where ν is some (possibly multiple) continuous parameter.

The proof that the eigenvalues in the discrete part of the spectrum are real for a self-adjoint operator, is analogue to the one for finite-dimensional Hilbert spaces :

$$a_n\,\langle\varphi_{n,i}\,|\varphi_{n,i}\rangle = \langle\varphi_{n,i}|\,\big(A\,|\varphi_{n,i}\rangle\,\big) = \langle A\varphi_{n,i}\,|\varphi_{n,i}\rangle = \big(\langle\varphi_{n,i}\,|A\varphi_{n,i}\rangle\,\big)^* = a_n^*\,\langle\varphi_{n,i}\,|\varphi_{n,i}\rangle \;\; . \tag{6.60}$$

This cannot be immediately transposed to the continuous spectrum since $\langle\nu\,|\nu\rangle$ is infinite, but one can again start with superpositions (wave-packets) of nearby ν and then take a corresponding limit.

Exercise 6.6 : *Work out the details of this demonstration.*

Similarly also, eigenvectors with different eigenvalues are orthogonal :

$$a_m\,\langle\varphi_{n,i}\,|\varphi_{m,j}\rangle = \langle\varphi_{n,i}|\,\big(A\,|\varphi_{m,j}\rangle\,\big) = \langle A\varphi_{n,i}\,|\varphi_{m,j}\rangle = a_n^*\,\langle\varphi_{n,i}\,|\varphi_{m,j}\rangle = a_n\,\langle\varphi_{n,i}\,|\varphi_{m,j}\rangle \;\; . \tag{6.61}$$

Then $(a_n - a_m)\langle\varphi_{n,i}\,|\varphi_{m,j}\rangle = 0$ and if $a_n \neq a_m$ one necessarily has $\langle\varphi_{n,i}\,|\varphi_{m,j}\rangle = 0$. For the eigenvectors corresponding to the same eigenvalue one can always apply the orthonormalisation procedure so that the different $|\varphi_{n,i}\rangle$ are all orthogonal and normalised. This computation can be immediately adapted to the continuous spectrum also and then if $a(\nu) \neq a(\nu')$ then $\langle\nu\,|\nu'\rangle = 0$, so that

$$\langle\varphi_{n,i}\,|\varphi_{m,j}\rangle = \delta_{nm}\,\delta_{i,j} \quad , \quad \langle\nu\,|\nu'\rangle = \delta(\nu - \nu') \;. \tag{6.62}$$

The spectral theorem can than be stated as

$$\mathbf{1} = \sum_n \sum_{i=1}^{d_n} |\varphi_{n,i}\rangle \langle\varphi_{n,i}| + \int_{I_c} \mathrm{d}\nu \; |\nu\rangle \langle\nu| \tag{6.63}$$

where I_c denotes the (union of) intervals where one has continuous eigenvalues. This relation is called the completeness relation. In quantum mechanics linear operators for which (6.63) holds are called observables. In particular, self-adjoint operators are observables, but one may also encounter certain symmetric but not self-adjoint operators for which (6.63) nevertheless holds. Then they will also be called observables. One can then write the corresponding spectral decomposition of the observable A as

$$A = \sum_n a_n \sum_{i=1}^{d_n} |\varphi_{n,i}\rangle \langle\varphi_{n,i}| + \int_{I_c} \mathrm{d}\nu \, a(\nu) \; |\nu\rangle \langle\nu| \;\; . \tag{6.64}$$

We define the "projectors"

$$P(a_n) = \sum_{i=1}^{d_n} |\varphi_{n,i}\rangle \langle\varphi_{n,i}| \quad , \quad P(\nu_0, \epsilon) = \int_{\nu_0-\epsilon/2}^{\nu_0+\epsilon/2} \mathrm{d}\nu \; |\nu\rangle \langle\nu| \;\; . \tag{6.65}$$

It is straightforward to verify that they are projectors. For $P(a_n)$ this goes as in the finite-dimensional case and for $P(\nu_0, \epsilon)$ one has

$$\begin{aligned} P(\nu_0,\epsilon)P(\nu_0,\epsilon) &= \int_{\nu_0-\epsilon/2}^{\nu_0+\epsilon/2} \mathrm{d}\nu \int_{\nu_0-\epsilon/2}^{\nu_0+\epsilon/2} \mathrm{d}\nu' \; |\nu\rangle \langle\nu\,|\nu'\rangle \langle\nu'| = \int_{\nu_0-\epsilon/2}^{\nu_0+\epsilon/2} \mathrm{d}\nu \int_{\nu_0-\epsilon/2}^{\nu_0+\epsilon/2} \mathrm{d}\nu' \, \delta(\nu-\nu') \, |\nu\rangle \langle\nu'| \\ &= \int_{\nu_0-\epsilon/2}^{\nu_0+\epsilon/2} \mathrm{d}\nu \; |\nu\rangle \langle\nu| = P(\nu_0,\epsilon) \;\; . \end{aligned} \tag{6.66}$$

The spectral theorem states that the union of the $|\varphi_{n,i}\rangle$ and the $|\nu\rangle$ form a "generalised orthonormal basis" of $\mathcal{H}$ in the sense that every state in $\mathcal{H}$ can be uniquely written as

$$|\psi\rangle = \sum_n \sum_{i=1}^{d_n} c_{n,i} \, |\varphi_{n,i}\rangle + \int_{I_c} \mathrm{d}\nu \, c(\nu) \; |\nu\rangle \quad , \quad c_{n,i} = \langle\varphi_{n,i}\,|\psi\rangle \;\; , \quad c(\nu) = \langle\nu\,|\psi\rangle \;\; , \tag{6.67}$$

where the expressions for the coefficients c follow from writing $|\psi\rangle = \mathbf{1}\,|\psi\rangle$ and using (6.63). Of course, the state $|\psi\rangle$ must have finite norm to be really in $\mathcal{H}$. The norm is easily computed using (6.62) :

$$\langle\psi\,|\psi\rangle = \sum_n \sum_{i=1}^{d_n} |c_{n,i}|^2 + \int_{I_c} \mathrm{d}\nu \, |c(\nu)|^2 \quad , \quad \langle\psi\,|\psi\rangle < \infty \;\; . \tag{6.68}$$

An important result we have seen for finite-dimensional Hilbert spaces is that for commuting observables there is a basis of *common* eigenvectors, i.e. of eigenvectors of both observables. This also remains true for observables on infinite-dimensional Hilbert spaces. Of course, one has to be more careful, already to define the commutator $[A, B] = AB - BA$ of two observables. To make sense, when acting on a state $|\psi\rangle$, not only must $|\psi\rangle$ be in the domain of definition of A and of B,

but also $A\,|\psi\rangle$ must be in the domain of definition of B and vice versa. Without going into these details, we have the following result.

Lemma : If A and B are observables on $\mathcal{H}$ and $[A,B]=0$ on an appropriate domain that is dense in $\mathcal{H}$, then there exists a basis of common eigenvectors of A and B.

Let us only show this for the case that the spectra of A and B are discrete. Since A is an observable there exists an eigen-basis $\{|\varphi_{n,i}\rangle\}$ of A corresponding to the eigenvalues a_n with eigenvectors $|\varphi_{n,i}\rangle$. Let us call $\mathcal{E}(a_n)$ the eigen-space corresponding to a_n. Then the projector on this eigen-space is $P(a_n)=\sum_{i=1}^{d_n}|\varphi_{n,i}\rangle\langle\varphi_{n,i}|$ where the degree of degeneracy d_n can be finite or infinite. These projectors are symmetric and actually self-adjoint. We have

$$A\,B\;|\varphi_{n,i}\rangle = B\,A\;|\varphi_{n,i}\rangle = a_n B\;|\varphi_{n,i}\rangle\;\;, \tag{6.69}$$

which shows that $B\,|\varphi_{n,i}\rangle\in\mathcal{E}(a_n)$, i.e. B maps $\mathcal{E}(a_n)$ to itself. The restriction of B to $\mathcal{E}(a_n)$ can be written as $P(a_n)BP(a_n)\equiv B|_n$. This is again symmetric since $B|_n^\dagger=(P(a_n)BP(a_n))^\dagger = P(a_n)^\dagger B^\dagger P(a_n)^\dagger = P(a_n)BP(a_n)=B|_n$. If $d_n<\infty$, so that $\mathcal{E}(a_n)$ is finite-dimensional, $B|_n$ is automatically self-adjoint. If $d_n=\infty$, one has to be more careful, but one can still adjust the domain of definition such that $B|_n$ is self-adjoint or at least that the spectral theorem applies, and then

$$B|_n\,|\phi_{n,j}\rangle = b_{n,j}\,|\phi_{n,j}\rangle\quad,\quad \mathbf{1}\big|_{\mathcal{E}_n}=\sum_{j=1}^{d_n}|\phi_{n,j}\rangle\langle\phi_{n,j}|\;\;. \tag{6.70}$$

Now, the $|\phi_{n,j}\rangle$ are all in $\mathcal{E}_n$ and hence also eigenvectors of A with eigenvalue a_n. Hence, the $\{|\phi_{n,j}\rangle\}_{n,j}$ is the desired basis of common eigenvectors of A and B.

6.6 The postulates of quantum mechanics for infinite-dimensional (separable) Hilbert spaces

- Let $\mathcal{H}$ be a separable Hilbert space, i.e. a vector space over $\mathbf{C}$ for which an inner product is defined and which has a countably infinite basis that is orthonormal with respect to this inner product.

- A physical state is represented by a ray $\mathcal{R}_\psi$ in $\mathcal{H}$, i.e. a normed state vector $|\psi\rangle$, up to a phase: $|\psi\rangle\simeq e^{i\alpha}\,|\psi\rangle$ and $\langle\psi\,|\psi\rangle=1$.

- Physical observables $\mathcal{A}$ are represented by linear operators A on $\mathcal{H}$ that are observables. These are self-adjoint operators or symmetric operators for which (6.63) holds.

- Any measurement of the physical observable $\mathcal{A}$ can only give values a_n or $a(\nu)$ that are in the spectrum (are eigenvalues) of A, i.e. $A\,|\phi_n^i\rangle=a_n\,|\phi_n^i\rangle$ or $A\,|\nu\rangle=a(\nu)\,|\nu\rangle$, where we assume that the eigenvectors are normalised as in (6.62). Then the probability to obtain an eigenvalue a_n of the discrete spectrum is

$$\mathcal{P}_n(\psi)\equiv\mathcal{P}(a_n,\psi)=\sum_{i=1}^{d_n}|\langle\phi_n^i\,|\psi\rangle\,|^2=\langle\psi|\,P(a_n)\,|\psi\rangle\;\;. \tag{6.71}$$

The probability to obtain an eigenvalue $a(\widetilde{\nu})$ for $\widetilde{\nu}$ between $\nu - \epsilon/2$ and $\nu + \epsilon/2$ is

$$\mathcal{P}(\nu, \epsilon, \psi) = \int_{\nu-\epsilon/2}^{\nu+\epsilon/2} \mathrm{d}\nu' \, |\langle \nu' | \psi \rangle|^2 = \langle \psi | \, P(\nu, \epsilon) \, | \psi \rangle \ . \tag{6.72}$$

This means that

$$\mathcal{P}(\nu, \psi) = |\langle \nu | \psi \rangle|^2 \tag{6.73}$$

is the probability density to measure $a(\nu)$.

- If a measurement of a physical observable $\mathcal{A}$ on a state initially in ray $\mathcal{R}_\psi$ gives the result a_n of the discrete spectrum, then immediately *after* the measurement the state is an eigenstate of A corresponding to the measured eigenvalue a_n. This goes under the name of "reduction of the wave-packet". More precisely, up to a phase, the state is $P(a_n) \, |\psi\rangle \, / ||P(a_n) \, |\psi\rangle \, ||$.

  If the measurement results in an eigenvalue of the continuous spectrum, there is a corresponding reduction of the wave-packet but one must now take into account the finite resolution of the measuring apparatus and accordingly project on a superposition of nearby $|\nu\rangle$. More precisely, if one measures $a(\widetilde{\nu})$ for $\widetilde{\nu}$ between $\nu - \epsilon/2$ and $\nu + \epsilon/2$, then immediately *after* the measurement the state is, up to a phase, $P(\nu, \epsilon) \, |\psi\rangle \, / ||P(\nu, \epsilon) \, |\psi\rangle \, ||$.

- The physical observable "energy" is represented by a linear operator H called the Hamilton operator or simply Hamiltonian. H is an observable, and the time evolution of a physical state is governed by the Schrödinger equation

  $$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \, |\psi(t)\rangle = H(t) \, |\psi(t)\rangle \ , \tag{6.74}$$

  where the Hamiltonian possibly depends on time (unless $|\psi\rangle$ describes an isolated system).

Note that the Schrödinger equation relates a time-dependence (frequency) to the energy, and thus reflects Planck's relation $\hbar\omega \simeq E$. The "correct" Hamiltonian must be such that it's eigenvalues correspond to the possible energies of the system. Moreover, in some appropriate classical limit, the Schrödinger equation should give back the classical equations of motion. Of course, such a limit will "wash out" certain details of H and similar but different Hamiltonians may lead to the same classical equations. Ultimately it will be experiment (measurement of the energies, time evolution of transition probabilities as discussed with Rabi's formula, etc.) that will decide whether one has chosen the correct Hamitonian or not.

6.7 Some simple implications

All the statements we made in subsection 2.8 for quantum mechanics in finite-dimensional Hilbert spaces directly carry over to the present general case :

- We have the conjugate Schrödinger equation

  $$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \, \langle \psi(t)| = - \, \langle \psi(t)| \, H(t) \ . \tag{6.75}$$

- The mean value of the observable A in a state $|\psi\rangle$ is given by

$$\begin{aligned}\langle A\rangle_\psi &= \sum_n \mathcal{P}(a_n,\psi)a_n + \int_{I_c} \mathrm{d}\nu\, \mathcal{P}(\nu,\psi)a(\nu) = \sum_n \langle\psi|\, P(a_n)\,|\psi\rangle\, a_n + \int_{I_c} \mathrm{d}\nu \langle\psi\,|\nu\rangle\,\langle\nu|\,\psi\rangle a(\nu) \\ &= \langle\psi|\,\Big(\sum_n a_n P(a_n) + \int_{I_c} \mathrm{d}\nu\, a(\nu)\;|\nu\rangle\,\langle\nu|\Big)\,|\psi\rangle = \langle\psi|\,A\,|\psi\rangle\;\;, \end{aligned} \tag{6.76}$$

  where the last equality follows from the spectral decomposition (6.64) of A.

- The variance ΔA of A is

$$(\Delta A)^2_\psi \equiv \langle (A - \langle A\rangle_\psi)^2\rangle_\psi = \langle A^2\rangle_\psi - \langle A\rangle^2_\psi \geq 0\;. \tag{6.77}$$

- One can again show that $(\Delta A)^2_\psi$ vanishes if and only if $|\psi\rangle$ is an eigenstate of A. Since $|\psi\rangle$ must be normed it cannot be an eigenstate of the continuous spectrum. To get a normed state from the continuous spectrum, one must necessarily superpose nearby eigenstates and then $(\Delta A)^2_\psi > 0$.

- If A and B commute and $|\psi\rangle$ is a common eigenstate then, of course, $(\Delta A)^2_\psi = (\Delta B)^2_\psi = 0$. On the other hand, if $[A,B] = iC \neq 0$, one shows again the uncertainty relations

$$(\Delta A)_\psi (\Delta B)_\psi \geq \frac{1}{2}\,|\,\langle\,[A,B]\,\rangle_\psi\,|\;. \tag{6.78}$$

  In particular, if we take $A = X$ and $B = P$ so that $C = \hbar\,\mathbf{1}$, we get the famous Heisenberg uncertainty relation

$$\Delta X\,\Delta P \geq \frac{\hbar}{2}\;, \tag{6.79}$$

  where we suppressed the subscripts ψ.

- Finally, the definition of the density operator ρ and its properties, as given in subsection 2.10 carry over to the present case with the obvious modifications for the continuous part of the spectrum. In particular, for a pure state $|\psi\rangle$ one has $\rho = |\psi\rangle\,\langle\psi|$ and then

$$\begin{aligned}\mathrm{tr}\,\rho\,A &= \mathrm{tr}\,\left(\Big(\sum_n a_n P(a_n) + \int_{I_c} \mathrm{d}\nu\, a(\nu)\,|\nu\rangle\,\langle\nu|\Big)\,|\psi\rangle\,\langle\psi|\right) \\ &= \langle\psi|\,\Big(\sum_n a_n P(a_n) + \int_{I_c} \mathrm{d}\nu\, a(\nu)\,|\nu\rangle\,\langle\nu|\Big)\,|\psi\rangle \\ &= \sum_n a_n \mathcal{P}(a_n,\psi) + \int_{I_c} \mathrm{d}\nu\, a(\nu)\mathcal{P}(\nu,\psi) = \langle A\rangle_\psi\;. \end{aligned} \tag{6.80}$$

6.8 Classical limit and Ehrenfest's theorem

We will now discuss in how far the quantum mechanical principles and the choice of Hamiltonian give rise to the correct classical limit. This will show up as equations for the mean values of the

position and momentum operators that will take on the roles of the classical position and classical momentum.

To begin with let us recall the classical equations of motion in Hamiltonian form. Here we will denote by $\mathscr{H}(x,p)$ the classical Hamilton function. The value taken by this function is the energy of the system. Consider a classical point particle of mass m, moving only in a single direction with coordinate x. The particle is subject to a force $F_x(x)$ that derives from a potential $V(x)$ as $F_x(x) = -V'(x)$. Then the classical equation of motion is $m\ddot{x} = -V'(x)$. This can be rewritten as $\dot{x} = \frac{p}{m}$, $\dot{p} = -V'(x)$, which can be recast in Hamiltonian form as

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\partial \mathscr{H}(x,p)}{\partial p} \quad , \quad \frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial \mathscr{H}(x,p)}{\partial x} \quad , \quad \mathscr{H}(x,p) = \frac{p^2}{2m} + V(x) \ . \tag{6.81}$$

One also defines the antisymmetric Poisson bracket $\{F,G\} = -\{G,F\}$ of two functionals F and G of x and p as

$$\{F,G\} = \frac{\partial F}{\partial x}\frac{\partial G}{\partial p} - \frac{\partial F}{\partial p}\frac{\partial G}{\partial x} = -\{G,F\} \ , \tag{6.82}$$

and then obviously

$$\{x,p\} = 1 \quad , \quad \{x,G(p)\} = \frac{\partial G}{\partial p} \quad , \quad \{p,F(x)\} = -\frac{\partial F}{\partial x} \ . \tag{6.83}$$

The first relation is reminiscent of the commutator $[X,P] = i\hbar$ and we will come back to this. We can then rewrite (6.81) as

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \{x, \mathscr{H}(x,p)\} \quad , \quad \frac{\mathrm{d}p}{\mathrm{d}t} = \{p, \mathscr{H}(x,p)\} \ . \tag{6.84}$$

This is the general form of the equations of motion in Hamiltonian form, even with a more general Hamiltonian function than the one given above. It is then easy to show that for an arbitrary functional $F(x,p,t)$ of x and p, and possibly depending explicitly on t one has

$$\begin{aligned} \frac{\mathrm{d}F}{\mathrm{d}t} &= \frac{\partial F}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial F}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}t} + \frac{\partial F}{\partial t} = \frac{\partial F}{\partial x}\frac{\partial \mathscr{H}}{\mathrm{d}p} - \frac{\partial F}{\partial p}\frac{\partial \mathscr{H}}{\partial x} + \frac{\partial F}{\partial t} \\ &= \{F, \mathscr{H}\} + \frac{\partial F}{\partial t} \ . \end{aligned} \tag{6.85}$$

Let us insist that the variation $\frac{\mathrm{d}F}{\mathrm{d}t}$ is due to the explicit time dependence of F *and* due to the fact that x and p are time-dependent.

Let us now study, in quantum mechanics, the time evolution of the mean value of an observable A in a given state $|\psi(t)\rangle$. We also allow the observable to depend explicitly on time. Then

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle A(t)\rangle_{\psi(t)} = \frac{\mathrm{d}}{\mathrm{d}t}\Big(\langle\psi(t)|\, A(t)\, |\psi(t)\rangle \Big) = \frac{1}{i\hbar}\, \langle\psi(t)|\, \big(A(t)H - HA(t)\, |\psi(t)\rangle + \langle\psi(t)|\, \frac{\partial A(t)}{\partial t}\, |\psi(t)\rangle \ , \tag{6.86}$$

which yields Ehrenfest's theorem :

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\langle A(t)\rangle_{\psi(t)} = \frac{1}{i\hbar}\langle [A(t),H]\rangle_{\psi(t)} + \langle \frac{\partial A(t)}{\partial t}\rangle_{\psi(t)} \ .} \tag{6.87}$$

This looks very similar to (6.85). Of course, we need to take expectation values to get classical quantities from the quantum theory. Furthermore, (6.92) involves the commutator $[A, H]$, while (6.85) contains the Poisson bracket. It is often said that the Poisson bracket is the classical limit of the commutator (divided by $i\hbar$), or that a classical system is "quantised" by replacing Poisson brackets by commutators. Indeed, for X and P one has

$$\{x, p\} = 1 = \frac{1}{i\hbar}[X, P] \ , \tag{6.88}$$

and also

$$\begin{aligned} \{x, G(p)\} &= \frac{\partial G}{\partial p} \leftrightarrow \frac{1}{i\hbar}[X, G(P)] = \frac{\partial G(p)}{\partial p}\Big|_{p\to P} \\ \{p, F(x)\} &= -\frac{\partial F}{\partial x} \leftrightarrow \frac{1}{i\hbar}[P, F(X)] = -\frac{\partial F(x)}{\partial x}\Big|_{x\to X} \ . \end{aligned} \tag{6.89}$$

Let us check e.g. the first commutator relation for $G(p) = p^3$: $[X, P^3] = XP^3 - P^3X = (XP - PX)P^2 + P(XP - PX)P + P^2(XP - PX) = 3i\hbar\, P^2$. However this simple correspondence between Poisson bracket and commutator ceases when both factors are non-linear, like e.g. $[X^2, P^2]$.[35]

$$[X^2, P^2] = i\hbar\big(2XP + 2PX\big) = 4i\hbar\, XP + 4\hbar^2\mathbf{1} \ , \tag{6.91}$$

where the last term arises from a double commutator (i.e. using the commutation relation twice). There is no classical counterpart of such double commutators : there is no "double Poisson bracket". If we now apply Ehrenfest's theorem (6.92) with $A = X$ or $A = P$, there is no "double commutator" issue. Let us also assume that $H = \frac{1}{2m}P^2 + V(X)$. Then $\frac{1}{i\hbar}[X, H] = \frac{1}{2mi\hbar}[X, P^2] = \frac{P}{m}$ and $\frac{1}{i\hbar}[P, H] = -V'(X)$ so that

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle X\rangle = \frac{\langle P\rangle}{m} \quad , \quad \frac{\mathrm{d}}{\mathrm{d}t}\langle P\rangle = -\langle V'(X)\rangle \ . \tag{6.92}$$

Now *if* we could replace $\langle V'(X)\rangle$ by $V'(\langle X\rangle)$ then this would be the classical equations of motion for $x = \langle X\rangle$ and $p = \langle P\rangle$. But, of course, in general, $\langle V'(X)\rangle \neq V'(\langle X\rangle)$. We only have equality if V' is linear in X, i.e. for a harmonic oscillator potential $V(X) \sim X^2$, so that $V'(X) \sim X$. Thus, for the harmonic oscillator potential we expect that the mean position and mean momentum oscillate with the corresponding characteristic frequency. On the other hand, for $V(X) \sim X^3$ we have $V'(X) \sim X^2$ and $\langle X^2\rangle - \langle X\rangle^2 = (\Delta X)^2 > 0$ and then $\langle V'(X)\rangle \neq V'(\langle X\rangle)$ so that the equations for $\langle X\rangle$ and $\langle P\rangle$ differ from the classical equations of motion. But this also shows that if ΔX is small, i.e. $(\Delta X)^2 \ll \langle X^2\rangle$, we can expect that the mean position and mean momentum almost obey the classical equations. This is what happens in the "classical limit".

[35] To evaluate this, a useful formula for commutators is

$$[AB, C] = A[B, C] + [A, C]B \quad , \quad [B, CD] = C[B, D] + [B, C]D \ , \tag{6.90}$$

and then also $[AB, CD] = A[B, CD] + [A, CD]B = AC[B, D] + A[B, C]D + C[A, D]B + [A, C]DB$.

6.9 Several spatial directions

So far, we have only considered a single spatial direction with coordinate x and corresponding position operator X, and similarly for $p_x \equiv p$ with corresponding operator P. Although quite a few systems effectively behave as if there is only a single spatial direction, most of the interesting physics implies all 3 spatial directions. We could then redo our discussion about the eigenstates $|x_n\rangle$ of the position. Now one also need to specify the y and z-coordinates with corresponding operators Y and Z. From all experimental evidence it appears that a measurement of the x-position does not interfere with the measurement of the y- or z-position. This is not a priori obvious, and was not the case for the different components of the spin-operator. This was due to the non-commutativity of the latter. For the position operators X, Y and Z we then assume that they commute with each other, so that indeed the measurement of one of them is independent of the measurement of any other. One expects that the simple wave-function $\psi(x)$ now gets replaced by a $\psi(x, y, z)$ and that the operators X, Y and Z then simply act by multiplication with the argument x, y and z. As before, $\psi(x, y, z) = \langle (x, y, z) \, | \psi \rangle \equiv \langle \vec{r} \, | \psi \rangle$. The $|(x, y, z)\rangle$ are an improper basis of (non-normalisable) simultaneous eigenstates of X, Y and Z. The simplest way to construct such simultaneous eigenstates is as a tensor product. Define $|y\rangle$ and $|z\rangle$ in a way that is totally analogous to the way we defined the $|x\rangle$. Then

$$|(x, y, z)\rangle \equiv |\vec{r}\rangle = |x\rangle \otimes |y\rangle \otimes |z\rangle \ . \tag{6.93}$$

Then the operators X, Y and Z are in reality $X \otimes \mathbf{1} \otimes \mathbf{1}$, $\mathbf{1} \otimes Y \otimes \mathbf{1}$ and $\mathbf{1} \otimes \mathbf{1} \otimes Z$, and then they obviously commute with each other.

The corresponding Hilbert space is isomorphic[36] to $L^2(\mathbf{R^3}) \simeq L^2(\mathbf{R}) \otimes L^2(\mathbf{R}) \otimes L^2(\mathbf{R})$. In particular the operators P_x, P_y and P_z are defined in the obvious way in this tensor product structure :

$$|\vec{p}\rangle \equiv |(p_x, p_y, p_z)\rangle = |p_x\rangle \otimes |p_y\rangle \otimes |p_z\rangle \quad , \quad P_x \, |(p_x, p_y, p_z)\rangle = p_x \, |(p_x, p_y, p_z)\rangle \ \text{etc} \ , \tag{6.94}$$

and

$$\langle \vec{r} \, | \vec{p} \rangle \equiv \langle (x, y, z) \, | (p_x, p_y, p_z) \rangle = \langle x \, | p_x \rangle \ \langle y \, | p_y \rangle \ \langle z \, | p_z \rangle = \frac{e^{ip_x x/\hbar}}{\sqrt{2\pi\hbar}} \ \frac{e^{ip_y y/\hbar}}{\sqrt{2\pi\hbar}} \ \frac{e^{ip_z z/\hbar}}{\sqrt{2\pi\hbar}} = \frac{e^{i\vec{p}\cdot\vec{r}/\hbar}}{(2\pi\hbar)^{3/2}} \tag{6.95}$$

We see that mathematically, the description of the Hilbert space of a particle in the 3-dimensional space $\mathbf{R}^3$ is not any different from the description of the Hilbert space of 3 (distinguishable) particles, each living in a one-dimensional space, which would be given by the same three-fold tensor product.

But before exploring the quantum mechanics of a particle in 3 dimensions, we will study a few (simpler) systems of a particle living in just one spatial dimension.

[36] Actually, all Hilbert spaces with a countably infinite basis are isomorphic so that $L^2(\mathbf{R^3}) \simeq L^2(\mathbf{R})$. While mathematically true, this statement does not add much to our understanding of the structure of this Hilbert space, and it is more fruitful to continue thinking about $L^2(\mathbf{R^3})$ as $L^2(\mathbf{R}) \otimes L^2(\mathbf{R}) \otimes L^2(\mathbf{R})$.