Chapter 1

Quantum Kinematics Reviewed

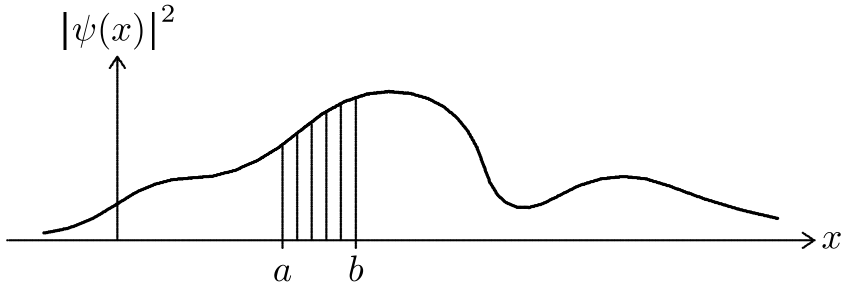
1.1 Schrödinger's wave function

A typical first course on quantum mechanics is likely to adopt the strategy of the typical textbooks for beginners, and will, therefore, focus predominantly on single objects moving along the x axis. In such an approach, Erwin Schrödinger's wave function $\psi(x)$ plays the central role in the mathematical description of the physical situation. It is taken for granted that the reader is somewhat familiar with the standard material of such a first course.

We remind ourselves of the significance of the wave function $\psi(x)$: by integrating the squared modulus of $\psi(x)$, you get probabilities. In particular, we recall that

$$\int_a^b \mathrm{d}x\, |\psi(x)|^2 \quad \text{is the probability of finding the object between } x = a \text{ and } x = b \text{ (whereby } a < b\text{)}, \tag{1.1.1}$$

which is graphically represented by



where the area is (proportional to) that probability. Therefore, the squared wave function $|\psi(x)|^2$ is a *probability density*, and one refers to the wave function itself as a *probability density amplitude*. Since we shall surely find the object somewhere, we have unit probability (= 100%) in the limit of $a \to -\infty$, $b \to \infty$, so that $\psi(x)$ is normalized in accordance with

$$\int_{-\infty}^{\infty} \mathrm{d}x\, |\psi(x)|^2 = 1\,. \tag{1.1.2}$$

We note that these probabilities are of a fundamental nature, they do not result from a lack of knowledge, as it would be typical for the probabilities in classical statistical physics. Also, one must remember that it is the *sole* role of $\psi(x)$ to supply the probabilistic predictions, it has no other significance beyond that. In particular, it would be wrong to think of $|\psi(x)|^2$ as a statement of how the object (electron, atom, ...) is spread out in space. Electrons, and atoms for the present matter as well, are point-like objects. You look for them, and you find them in one place and in one piece. It is only that we cannot predict with certainty the outcome of such a position measurement of an electron. What we can predict reliably are the probabilities of finding the electron in certain regions. And by repeating the measurement very often, we can verify such statistical predictions experimentally.

"Repeating the measurement" means "measure again on an equally prepared electron", it does not mean "measure again the position of the same electron". In the latter situation, the second measurement has probabilities different from the first measurement because the first measurement involves an interaction with the electron and thus a disturbance of the electron. In short, after the first position measurement, there is an altered wave function from which the probabilities for the second measurement are to be derived.

This last remark is a reminder that all probabilities are conditional probabilities. We make statistical predictions based on what we know on the conditions under which the experiment is performed. When speaking of "equally prepared electrons" we mean that the same conditions are realized. After the measurement has been carried out, the conditions are changed, and we must update our statistical predictions accordingly, because the altered conditions determine the probabilities of subsequent measurements.

We recall further that, in addition to the *position wave function* $\psi(x)$, there is also a *momentum wave function* $\psi(p)$, and the two are related to each other by Fourier transformations,

$$\psi(p) = \int \mathrm{d}x\, \frac{\mathrm{e}^{-\mathrm{i}px/\hbar}}{\sqrt{2\pi\hbar}}\psi(x)\,, \qquad \psi(x) = \int \mathrm{d}p\, \frac{\mathrm{e}^{\mathrm{i}xp/\hbar}}{\sqrt{2\pi\hbar}}\psi(p)\,, \tag{1.1.3}$$

where $\hbar$ is Planck's constant (Max K. E. L. Planck). We get probabilities for momentum measurements by integrating $|\psi(p)|^2$,

$$\int_q^r \mathrm{d}p\, |\psi(p)|^2 = \begin{array}{l}\text{probability of finding the object's} \\ \text{momentum in the range } q < p < r\,,\end{array} \tag{1.1.4}$$

and

$$\int_{-\infty}^{\infty} \mathrm{d}p\, |\psi(p)|^2 = 1 \tag{1.1.5}$$

is the appropriate normalization of $\psi(p)$.

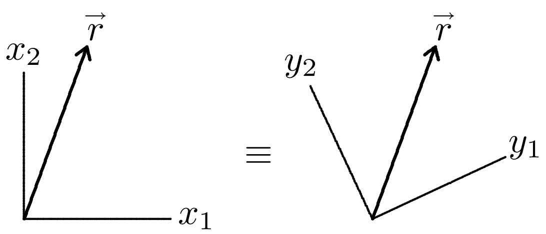
As Jean B. J. Fourier's taught us, the two transformations in (1.1.3) are inverses of each other, so that we can go back and forth between $\psi(x)$ and $\psi(p)$. Their one-to-one correspondence tells us that either one contains all the information of the other. And it does not stop here. For example, we could keep a record of all the gaussian moments of $\psi(x)$,

$$\psi_n = \int \mathrm{d}x\, \psi(x) x^n\, \mathrm{e}^{-(x/a)^2} \tag{1.1.6}$$

with an arbitrary length parameter a, and the set of $\psi_n, n = 0, 1, 2, \ldots$, would specify $\psi(x)$ uniquely. Clearly, then, the wave functions $\psi(x)$ and $\psi(p)$, and the moments ψ_n are just particular parameterizations of the given *physical state of affairs.* We are thus invited to look for a more abstract entity, a mathematical object that we shall call the *state of the system.*

1.2 Digression: Vectors, coordinates, and all that

It helps to build on an analogy that you might want to keep in mind because it will be useful for the visualization of some rather abstract quantum mechanical statements in terms of geometrical objects. We consider real n-component vectors and their numerical description in terms of coefficients (coordinates) that refer to agreed-upon coordinate systems:



$$\begin{aligned}\vec{r} &\mathrel{\widehat{=}} (x_1, x_2, \ldots, x_n) \\ &\mathrel{\widehat{=}} (y_1, y_2, \ldots, y_n)\,.\end{aligned} \tag{1.2.1}$$

One and the same vector has two (or more) numerical descriptions, the coordinates x_j and the coordinates y_k. These numbers, although not unrelated, can be quite different, but they mean *the same vector* $\vec{r}$. We make this explicit with the aid of the basis vectors $\vec{e}_j$ (for the x description) and $\vec{f}_k$ (for the y description),

$$\vec{r} = \sum_{j=1}^{n} x_j \vec{e}_j = \sum_{k=1}^{n} y_k \vec{f}_k \,. \tag{1.2.2}$$

We take for granted (this is a matter of convenient simplicity, not one of necessity) that the basis vectors of each set are orthonormal,

$$\begin{aligned} \vec{e}_j \cdot \vec{e}_k &= \delta_{jk} = \begin{cases} 1 & \text{if} \quad j = k\,, \\ 0 & \text{if} \quad j \neq k\,, \end{cases} \\ \vec{f}_j \cdot \vec{f}_k &= \delta_{jk}\,, \end{aligned} \tag{1.2.3}$$

where δ_{jk} is Leopold Kronecker's delta symbol. Then

$$x_j = \vec{e}_j \cdot \vec{r} \quad \text{and} \quad y_k = \vec{f}_k \cdot \vec{r} \tag{1.2.4}$$

tell us how we determine the coordinates of $\vec{r}$ if the basis vectors are given.

More implicitly than explicitly we have been thinking of $\vec{r}$, $\vec{e}_j$, $\vec{f}_k$ as being numerically represented by *rows* of coordinates. So let us regard the vectors themselves as *row-type vectors*. But just as well we could have arranged the coordinates in columns and would then regard the vectors themselves as *column-type vectors*. It is expedient to emphasize the row or column nature by the notation. We continue to write $\vec{r}, \vec{e}_j, \vec{f}_k$ for the row vectors, and denote the corresponding column vectors by $r^{\downarrow}$, $e_j^{\downarrow}$, $f_k^{\downarrow}$. Thus

$$\vec{r} \mathrel{\widehat{=}} (x_1, x_2, \ldots, x_n) \mathrel{\widehat{=}} (y_1, y_2, \ldots, y_n) \tag{1.2.5}$$

is paired with

$$r^{\downarrow} \mathrel{\widehat{=}} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mathrel{\widehat{=}} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \tag{1.2.6}$$

and the two kinds of vectors are related to each other by *transposition*,

$$r^{\downarrow} = \vec{r}^{\,\mathrm{T}}\,, \quad \vec{r} = r^{\downarrow\mathrm{T}}\,. \tag{1.2.7}$$

One immediate benefit of distinguishing between row vectors and column vectors is that we can write inner (scalar, dot) products as simple column-times-row products. This is illustrated by

$$\vec{r} \mathrel{\widehat{=}} (x_1, \ldots, x_n)\,, \quad \vec{s} \mathrel{\widehat{=}} (u_1, u_2, \ldots, u_n)\,, \tag{1.2.8}$$

$$\vec{r} \cdot \vec{s} = \sum_j x_j u_j = \underbrace{(x_1, x_2, \ldots, x_n)}_{\widehat{=}\,\vec{r}} \underbrace{\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}}_{\widehat{=}\, s^{\downarrow}} = \vec{r}\, s^{\downarrow}\,. \tag{1.2.9}$$

In view of the symmetry $\vec{r} \cdot \vec{s} = \vec{s} \cdot \vec{r}$, we thus have

$$\underbrace{\vec{r} \cdot \vec{s}}_{\substack{\text{inner product of}\\ \text{two row vectors}}} = \underbrace{\vec{r}\, s^{\downarrow} = \vec{s}\, r^{\downarrow}}_{\substack{\text{products of}\\ \text{the type}\\ \text{“row times}\\ \text{column”}}} = \underbrace{r^{\downarrow} \cdot s^{\downarrow}}_{\substack{\text{inner product of two}\\ \text{column vectors}}}\,. \tag{1.2.10}$$

The central identity here is, of course, consistent with the product rule for transposition, generally: $(AB)^{\mathrm{T}} = B^{\mathrm{T}} A^{\mathrm{T}}$, here:

$$(\vec{r}\, s^{\downarrow})^{\mathrm{T}} = s^{\downarrow\mathrm{T}} \vec{r}^{\,\mathrm{T}} = \vec{s}\, r^{\downarrow}\,, \quad \text{indeed}\,. \tag{1.2.11}$$

Upon combining

$$\vec{r} = \sum_j x_j \vec{e}_j \quad \text{and} \quad x_j = \vec{e}_j \cdot \vec{r} = \vec{r}\, e_j^{\downarrow} \tag{1.2.12}$$

into

$$\vec{r} = \sum_j \vec{r}\, e_j^{\downarrow} \vec{e}_j = \vec{r} \sum_j e_j^{\downarrow} \vec{e}_j \tag{1.2.13}$$

we meet an object of a new kind, the sum of products $e_j^{\downarrow}\vec{e}_j$ of “column times row” type. That is not a number but a *dyadic*, which would have a $n \times n$ matrix as its numerical representation. See, for example,

$$r^{\downarrow} \vec{s} \mathrel{\widehat{=}} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} (u_1, \ldots, u_n) = \begin{pmatrix} x_1 u_1 & x_1 u_2 & \cdots & x_1 u_n \\ x_2 u_1 & x_2 u_2 & \cdots & x_2 u_n \\ \vdots & \vdots & \ddots & \vdots \\ x_n u_1 & x_n u_2 & \cdots & x_n u_n \end{pmatrix}. \tag{1.2.14}$$

The particular dyadic that appears in (1.2.13) has the property that when it multiplies (on the left) the arbitrary vector $\vec{r}$, the outcome is this vector itself: it is the *unit dyadic*,

$$\vec{r}\,\overset{\downarrow\!\rightarrow}{1} = \vec{r} \quad \text{with} \quad \overset{\downarrow\!\rightarrow}{1} = \sum_j e_j^{\downarrow}\vec{e}_j \,. \tag{1.2.15}$$

The notation $\overset{\downarrow\!\rightarrow}{}$ reminds us that such a dyadic is like a column vector on the left, and like a row vector on the right.

The identification of the unit dyadic is consistent only if it also acts accordingly on the right. Indeed, it does,

$$\begin{aligned} \overset{\downarrow\!\rightarrow}{1}\,r^{\downarrow} = \sum_j e_j^{\downarrow}\vec{e}_j r^{\downarrow} &= \sum_j e_j^{\downarrow}\vec{e}_j \cdot \vec{r} \\ &= \sum_j e_j^{\downarrow} x_j = r^{\downarrow} \,. \end{aligned} \tag{1.2.16}$$

As this little calculation demonstrates, the statement

$$\sum_j e_j^{\downarrow}\vec{e}_j = \overset{\downarrow\!\rightarrow}{1} \tag{1.2.17}$$

expresses the *completeness* of the set of column vectors $e_j^{\downarrow}$, and also that of the set of row vectors $\vec{e}_j$, because we can expand any arbitrary vector $r^{\downarrow}$ as a linear combination of the $e_j^{\downarrow}$.

The statements of *orthonormality*,

$$\vec{e}_j e_k^{\downarrow} = \delta_{jk} \,, \tag{1.2.18}$$

and of completeness in (1.2.17) are two sides of the same coin. And, of course, there is nothing special here about the $\vec{e}_j$ set of vectors, the $\vec{f}_k$ basis vectors are also orthonormal,

$$\vec{f}_j f_k^{\downarrow} = \delta_{jk} \tag{1.2.19}$$

and complete,

$$\sum_k f_k^{\downarrow}\vec{f}_k = \overset{\downarrow\!\rightarrow}{1} \,. \tag{1.2.20}$$

In

$$r^{\downarrow} = \sum_j e_j^{\downarrow} x_j = \sum_k f_k^{\downarrow} y_k \tag{1.2.21}$$

we have two parameterizations of $r^{\downarrow}$. How does one express one set of coefficients in terms of the other, that is: How does one translate the x description into the y description and vice versa? That is easy! See,

$$\begin{aligned} x_j = \vec{e}_j r^{\downarrow} &= \vec{e}_j \sum_k f_k^{\downarrow} y_k \\ &= \sum_k \vec{e}_j f_k^{\downarrow} y_k = \sum_k (ef)_{jk} y_k \end{aligned} \tag{1.2.22}$$

with

$$(ef)_{jk} = \vec{e}_j f_k^{\downarrow} = \vec{e}_j \cdot \vec{f}_k \,; \tag{1.2.23}$$

and likewise

$$y_k = \sum_j (fe)_{kj} x_j \quad \text{with} \quad (fe)_{kj} = \vec{f}_k \cdot \vec{e}_j \,. \tag{1.2.24}$$

The two $n \times n$ transformation matrices composed of the matrix elements $(ef)_{jk}$ and $(fe)_{kj}$ are clearly transposes of each other. Furthermore, it follows from

$$\begin{aligned} x_j = \sum_k (ef)_{jk} y_k &= \sum_k (ef)_{jk} \sum_{j'} (fe)_{kj'} x_{j'} \\ &= \sum_{j'} \left(\sum_k (ef)_{jk} (fe)_{kj'} \right) x_{j'} \end{aligned} \tag{1.2.25}$$

that

$$\sum_k (ef)_{jk} (fe)_{kj'} = \delta_{jj'} \tag{1.2.26}$$

must hold. This is to say that the two transformation matrices are inverses of each other — hardly a surprise.

1-1 Use the definitions of $(ef)_{jk}$ and $(fe)_{kj'}$ to verify this relation directly.

1-2 What appears in $y_k = \sum_{k'} \boxed{?}_{kk'} y_{k'}$ as the result of converting y_k into x_j and then back to y_k? Verify here too that $\boxed{?}_{kk'} = \delta_{kk'}$.

Rather than converting one description into the other, we can ask how the two sets of basis vectors are related to each other. Since both sets are orthonormal and complete, the mapping

$$e_j^{\downarrow} \longrightarrow f_j^{\downarrow} = \overleftrightarrow{O}\, e_j^{\downarrow} \tag{1.2.27}$$

is a *rotation* in the n-dimensional space. Geometric intuition tells us that there must be a unique dyadic $\overleftrightarrow{O}$ that accomplishes this rotation. We find it by multiplying with $\vec{e}_j$ from the right

$$f_j^{\downarrow}\vec{e}_j = \overleftrightarrow{O}\, e_j^{\downarrow}\vec{e}_j\,, \tag{1.2.28}$$

followed by summing over j and exploiting the completeness of the $\vec{e}$ vectors,

$$\sum_j f_j^{\downarrow}\vec{e}_j = \sum_j \overleftrightarrow{O}\, e_j^{\downarrow}\vec{e}_j = \overleftrightarrow{O}\, \underbrace{\sum_j e_j^{\downarrow}\vec{e}_j}_{=\overleftrightarrow{1}}\,, \tag{1.2.29}$$

with the outcome

$$\overleftrightarrow{O} = \sum_j f_j^{\downarrow}\vec{e}_j\,. \tag{1.2.30}$$

As an exercise, we verify that it has the desired property:

$$\overleftrightarrow{O}\, e_j^{\downarrow} = \sum_k f_k^{\downarrow}\, \underbrace{\vec{e}_k e_j^{\downarrow}}_{=\delta_{kj}} = \sum_k f_k^{\downarrow}\delta_{kj} = f_j^{\downarrow}\,, \tag{1.2.31}$$

indeed. Further, we note that

$$\vec{f}_k\overleftrightarrow{O} = \sum_j \underbrace{\vec{f}_k f_j^{\downarrow}}_{=\delta_{kj}}\,\vec{e}_j = \sum_j \delta_{kj}\vec{e}_j = \vec{e}_k\,, \tag{1.2.32}$$

so that the same dyadic $\overleftrightarrow{O}$ also transforms the rows $\vec{f}_k$ into the $\vec{e}_k$s. Together with the transposed statements we thus have

$$\begin{aligned} f_j^{\downarrow} &= \overleftrightarrow{O}\, e_j^{\downarrow}\,, & \vec{f}_k\overleftrightarrow{O} &= \vec{e}_k\,, \\ \vec{f}_j &= \vec{e}_j\overleftrightarrow{O}^{\mathrm{T}}\,, & \overleftrightarrow{O}^{\mathrm{T}} f_k^{\downarrow} &= e_k^{\downarrow} \end{aligned} \tag{1.2.33}$$

with

$$\overleftrightarrow{O} = \sum_k f_k^{\downarrow}\vec{e}_k \quad \text{and} \quad \overleftrightarrow{O}^{\mathrm{T}} = \sum_j e_j^{\downarrow}\vec{f}_j\,. \tag{1.2.34}$$

Here, too, we can iterate the transformations, as in

$$\begin{aligned} f_j^{\downarrow} &= \overleftrightarrow{O}\, e_j^{\downarrow} = \overleftrightarrow{O}\,\overleftrightarrow{O}^{\mathrm{T}} f_j^{\downarrow} \\ \text{or}\quad \vec{e}_k &= \vec{f}_k \overleftrightarrow{O} = \vec{e}_k \overleftrightarrow{O}^{\mathrm{T}} \overleftrightarrow{O}\,, \end{aligned} \tag{1.2.35}$$

and conclude that

$$\overleftrightarrow{O}\,\overleftrightarrow{O}^{\mathrm{T}} = \overleftrightarrow{1} = \overleftrightarrow{O}^{\mathrm{T}} \overleftrightarrow{O}\,. \tag{1.2.36}$$

Dyadics with this property, namely: the transpose is the inverse,

$$\overleftrightarrow{O}^{\mathrm{T}} = \overleftrightarrow{O}^{-1}\,, \tag{1.2.37}$$

are called *orthogonal*, in analogy to the corresponding terminology for orthogonal matrices in linear algebra.

1-3 How are the transformation matrices $(ef)_{jk}$ and $(fe)_{kj}$ related to the orthogonal dyadic $\overleftrightarrow{O}$?

Actually, in linear algebra the most basic definition of an orthogonal transformation is that it leaves *all* inner products unchanged. That is, for any pair of vectors $\vec{r}, \vec{s}$ we should have

$$\big(\vec{r}\,\overleftrightarrow{O}\big)\cdot\big(\vec{s}\,\overleftrightarrow{O}\big) = \vec{r}\cdot\vec{s}\,, \tag{1.2.38}$$

and for any pair $r^{\downarrow}, s^{\downarrow}$ we should have

$$\big(\overleftrightarrow{O}\, r^{\downarrow}\big)\cdot\big(\overleftrightarrow{O}\, s^{\downarrow}\big) = r^{\downarrow}\cdot s^{\downarrow}\,. \tag{1.2.39}$$

Indeed, upon switching over to row-times-column products, we have

$$\begin{aligned} \vec{r}\, s^{\downarrow} &= \vec{r}\,\overleftrightarrow{O}\,\overleftrightarrow{O}^{\mathrm{T}} s^{\downarrow} \quad \text{from (1.2.38)} \\ \text{and}\quad \vec{r}\, s^{\downarrow} &= \vec{r}\,\overleftrightarrow{O}^{\mathrm{T}} \overleftrightarrow{O}\, s^{\downarrow} \quad \text{from (1.2.39),} \end{aligned} \tag{1.2.40}$$

and $\overleftrightarrow{O}\,\overleftrightarrow{O}^{\mathrm{T}} = \overleftrightarrow{O}^{\mathrm{T}} \overleftrightarrow{O} = \overleftrightarrow{1}$ are implied again.

1-4 One term in the sum of (1.2.17) is

$$\overset{\downarrow\rightarrow}{P}_j = e_j^{\downarrow}\vec{e}_j \,.$$

Show that $\overset{\downarrow\rightarrow}{P}_j^{\,2} = \overset{\downarrow\rightarrow}{P}_j$. What is, therefore, the geometrical significance of $\overset{\downarrow\rightarrow}{P}_j$? Repeat for

$$\overset{\downarrow\rightarrow}{P}_{jk} = \overset{\downarrow\rightarrow}{P}_j + \overset{\downarrow\rightarrow}{P}_k \quad \text{with} \quad j \neq k \,.$$

1.3 Dirac's kets and bras

We now return to the discussion of $\psi(x), \psi(p), \ldots$ as equivalent numerical descriptions of the same abstract entity, the state of affairs of the physical system under consideration. Following Paul A. M. Dirac, we symbolize the state by a so-called *ket*, for which we write $|\;\rangle$ if we mean just any state (as we do presently) and fill the gap with appropriate labels if we mean one of a specific set of states, such as $|1\rangle$, $|2\rangle$, $|3\rangle$, … or $|\alpha\rangle$, $|\beta\rangle$, …, whatever the convenient and fitting labels may be. Mathematically speaking, kets are vectors, elements of a complex vector space, which just says that we can add kets to get new ones, and we can multiply them with complex numbers to get other, related kets. More generally, any linear combination of kets is another ket.

It helps to think of kets as analogs of column-type vectors, and then

$$\begin{aligned} |\;\rangle &= \int \mathrm{d}x\, |x\rangle \psi(x) \\ &= \int \mathrm{d}p\, |p\rangle \psi(p) \end{aligned} \tag{1.3.1}$$

are analogs to the two decompositions of $r^{\downarrow}$ in (1.2.21). There are crucial differences, however. Then we were summing over discrete indices, now we are integrating over the continuous variables x and p that label the kets. Then we were dealing with real objects — the coordinates x_j and y_k are real numbers — now we have complex-valued wave functions, $\psi(x)$ and $\psi(p)$. But otherwise the analogy is rather close and well worth remembering.

A wave function $\psi(x)$ that is large only in a small x region describes an object that is very well localized in the sense that we can reliably predict that we shall find it in this small region. In the limit of ever smaller regions — eventually a single x value, a point — we would get $|\;\rangle \to |x\rangle$

in some sense and, therefore, $|x\rangle$ refers to the situation "object is at x, exactly". This, however, is no longer a real physical situation but rather the overidealized situation of that unphysical limit. As a consequence, ket $|x\rangle$ is not actually associated with a physically realizable state, it is a convenient mathematical fiction. The unphysical nature is perhaps most obvious when we recall that the perfect localization of the overidealized limit would require a control on the quantum object with infinite precision — and this is never available in an actual real-life experiment.

By the same token, ket $|p\rangle$ refers to the overidealized situation of infinitely sharp momentum, again a mathematical fiction, not a physical reality. Both $|x\rangle$ and $|p\rangle$ kets are extremely useful mathematical objects, but one must keep in mind that a physical ket $|\ \rangle$ *always* involves a range of x values and a range of p values. This range may be small, then we have a well-controlled position, or a well-controlled momentum, but it is invariably a finite range.

Kets $|\ \rangle$, $|x\rangle$, $|p\rangle$, ... are analogs of column-type vectors. They have their partners in the so-called *bras* $\langle\ |$, $\langle x|$, $\langle p|$, ..., which are analogs of row-type vectors. When dealing with the real vectors $r^{\downarrow}$, $\vec{r}$, we related the two kinds to each other by transposition, $r^{\downarrow} = \vec{r}^{\,\mathrm{T}}$. Now, however, the "coordinates" — that is the wave functions $\psi(x)$, $\psi(p)$ — are complex-valued. Therefore, mathematical consistency requires that we supplement transposition with complex conjugation, and thus have *hermitian conjugation*, or, as the physicists say, we

$$\text{"take the adjoint":}\quad |\ \rangle^{\dagger} = \langle\ |\,, \qquad \langle\ |^{\dagger} = |\ \rangle\,. \tag{1.3.2}$$

The built-in complex conjugation becomes visible as soon as we take the adjoint of a linear combination,

$$\begin{aligned}\bigl(|1\rangle\alpha + |2\rangle\beta\bigr)^{\dagger} &= \alpha^{*}\langle 1| + \beta^{*}\langle 2|\,,\\ \bigl(\gamma\langle 3| + \delta\langle 4|\bigr)^{\dagger} &= |3\rangle\gamma^{*} + |4\rangle\delta^{*}\,.\end{aligned} \tag{1.3.3}$$

In particular the adjoint statements to the decompositions of $|\ \rangle$ in (1.3.1) are

$$\langle\ | = \int \mathrm{d}x\, \psi(x)^{*}\langle x| = \int \mathrm{d}p\, \psi(p)^{*}\langle p|\,. \tag{1.3.4}$$

In further analogy with the column-type vectors of Section 1.2, the kets are also endowed with an inner product, so that the vector space of kets is an inner-product space or Hilbert space, the name honoring David Hilbert's contributions. The notation $|1\rangle\cdot|2\rangle$ of the inner product as a "dot product" is, however, not used at all. In the mathematical literature, inner products are commonly written as (,), so that the inner product of two kets would appear as $\big(|1\rangle, |2\rangle\big)$ — except that mathematicians are not fond of the ket and bra notation, Dirac's stroke of genius.

Instead, one follows the suggestion of (1.2.10) and understands the inner products of two kets, or two bras, as the analogs of row-times-column products. So the inner product of the ket

$$|1\rangle = \int \mathrm{d}x\, |x\rangle \psi_1(x) \tag{1.3.5}$$

with the ket

$$|2\rangle = \int \mathrm{d}x\, |x\rangle \psi_2(x) \tag{1.3.6}$$

is obtained by multiplying the bra

$$\langle 1| = \int \mathrm{d}x\, \psi_1(x)^* \langle x| \tag{1.3.7}$$

with the ket $|2\rangle$:

$$\begin{aligned}\langle 1|2\rangle &= \int \mathrm{d}x\, \psi_1(x)^* \langle x| \int \mathrm{d}x'\, |x'\rangle \psi_2(x') \\ &= \int \mathrm{d}x\, \psi_1(x)^* \int \mathrm{d}x'\, \langle x|x'\rangle \psi_2(x')\,,\end{aligned} \tag{1.3.8}$$

where the integration variable of (1.3.6) is changed to x' to avoid confusion. This is also the inner product of bras $\langle 1|$ and $\langle 2|$. As anticipated in (1.3.8), one writes only one vertical line in the bra-ket product $\langle 1|\,|2\rangle = \langle 1|2\rangle$ of bra $\langle 1|$ and ket $|2\rangle$ and speaks of a Dirac bracket or simply *bracket*.

In accordance with what the reader learned in whichever first course on quantum mechanics, we expect the inner product (1.3.8) to be given by

$$\langle 1|2\rangle = \int \mathrm{d}x\, \psi_1(x)^* \psi_2(x)\,, \tag{1.3.9}$$

so that we need $\langle x|x'\rangle$ such that

$$\int \mathrm{d}x'\, \langle x|x'\rangle \psi_2(x') = \psi_2(x) \tag{1.3.10}$$

for all $\psi_2(x)$. Thus, we infer

$$\langle x|x'\rangle = \delta(x - x') \tag{1.3.11}$$

which is to say that x kets and bras are pairwise orthogonal and normalized to the Dirac δ function. There is a longer discussion of the δ function in Section 4.1 of *Basic Matters*, and so we are content here with recalling the basic, defining property, namely,

$$\int \mathrm{d}x'\, \delta(x-x')f(x') = f(x) \tag{1.3.12}$$

for all the functions that are continuous near x.

The normalization (1.1.2) of the wave function,

$$\int \mathrm{d}x\, |\psi(x)|^2 = 1\,, \tag{1.3.13}$$

now appears as

$$\langle\;|\;\rangle = 1\,, \tag{1.3.14}$$

see:

$$\begin{aligned}\langle\;|\;\rangle &= \int \mathrm{d}x\,\psi(x)^*\langle x|\int \mathrm{d}x'\,|x'\rangle\psi(x') \\ &= \int \mathrm{d}x\,\psi(x)^* \underbrace{\int \mathrm{d}x'\,\underbrace{\langle x|x'\rangle}_{=\delta(x-x')}\psi(x')}_{=\psi(x)} \\ &= \int \mathrm{d}x\,|\psi(x)|^2\,. \end{aligned} \tag{1.3.15}$$

In other words, the physical kets $|\;\rangle$, and the physical bras $\langle\;|$, are of *unit length*.

We recall also the physical significance of the bracket $\langle 1|2\rangle$ in (1.3.9), after which it is named *probability amplitude*: Its squared modulus $|\langle 1|2\rangle|^2$ is the probability prob$(2\to 1)$ of finding the system in state 1, described by ket $|1\rangle$ and parameterized by the wave function $\psi_1(x)$, if the system is known to be in state 2, with ket $|2\rangle$ and wave function $\psi_2(x)$. We should not fail to note the symmetry possessed by prob$(2\to 1)$,

$$\mathrm{prob}(2\to 1) = |\langle 1|2\rangle|^2 = \mathrm{prob}(1\to 2)\,, \tag{1.3.16}$$

which is an immediate consequence of

$$\langle 1|2\rangle = \langle 2|1\rangle^*\,, \tag{1.3.17}$$

demonstrated by interchanging the labels, $1 \leftrightarrow 2$, in (1.3.9). The fundamental symmetry (1.3.16) is quite remarkable because it states that the probability of finding $|1\rangle$ when $|2\rangle$ is the case is always exactly equal to the probability of finding $|2\rangle$ when $|1\rangle$ is the case, although these probabilities can refer to two very different experimental situations.

There is a basic requirement of consistency in this context, namely that $\mathrm{prob}(2 \to 1) \leq 1$. Indeed, this is ensured by the well-known *Cauchy–Bunyakovsky–Schwarz inequality*, named after Augustin-Louis Cauchy, Viktor Y. Bunyakovsky, and K. Hermann A. Schwarz. This inequality is the subject matter of the following exercise.

1-5 For all bras $\langle a|$ and all kets $|b\rangle$, show that

$$\left|\langle a|b\rangle\right|^2 \leq \langle a|a\rangle\langle b|b\rangle \,,$$

and state under which condition the equal sign applies. Conclude that $\mathrm{prob}(2 \to 1) \leq 1$ because the physical bra $\langle 1|$ is normalized in accordance with (1.3.14), and so is the physical ket $|2\rangle$.

The orthonormality statement (1.3.11),

$$\langle x|x'\rangle = \delta(x - x') \,, \tag{1.3.18}$$

is the obvious analog of (1.2.18),

$$\vec{e}_j e_k^{\downarrow} = \delta_{jk} \,. \tag{1.3.19}$$

We expect that the analog of the completeness relation (1.2.17),

$$\sum_j e_j^{\downarrow}\vec{e}_j = \overset{\leftarrow\!\!\rightarrow}{1} \,, \tag{1.3.20}$$

reads

$$\int \mathrm{d}x \,|x\rangle\langle x| = 1 \,. \tag{1.3.21}$$

Strictly speaking, the symbol on the right is the identity operator — the operator analog of the unit dyadic $\overset{\leftarrow\!\!\rightarrow}{1}$ — but we will not be pedantic about it and write it just like the number 1. It will always be unambiguously clear from the context whether we mean the unit operator or the unit number. Likewise, the symbol 5, say, can mean the number 5 or 5 times the unit operator, depending on the context. For instance, in $5\langle x| = \langle x|5$ we have the number 5 on the left and 5 times the unit operator on the right. There will be no confusion arising from this convenience in notation.

But we must not forget to verify the *completeness relation* (1.3.21). "Verification" means here just the check that it is consistent with everything else we have so far. For example, is it true that $1|\ \rangle = |\ \rangle$? We check:

$$
\begin{aligned}
1|\ \rangle &= \underbrace{\left(\int \mathrm{d}x\,|x\rangle\langle x|\right)}_{=1} \underbrace{\int \mathrm{d}x'\,|x'\rangle\psi(x')}_{=|\ \rangle} \\
&= \int \mathrm{d}x\,|x\rangle \int \mathrm{d}x'\,\underbrace{\langle x|x'\rangle}_{=\delta(x-x')}\psi(x') \\
&= \int \mathrm{d}x\,|x\rangle\psi(x) = |\ \rangle, \quad \text{indeed}\,.
\end{aligned}
\tag{1.3.22}
$$

Similarly, we check that $\langle\ |1 = \langle\ |$. Another little calculation:

$$
\begin{aligned}
\langle 1|2\rangle &= \langle 1|1|2\rangle \\
&= \int \mathrm{d}x\,\psi_1(x)^*\langle x| \int \mathrm{d}x'\,|x'\rangle\langle x'| \int \mathrm{d}x''\,|x''\rangle\psi_2(x'') \\
&= \int \mathrm{d}x\,\psi_1(x)^* \underbrace{\int \mathrm{d}x'\,\underbrace{\langle x|x'\rangle}_{\delta(x-x')=} \underbrace{\int \mathrm{d}x''\,\underbrace{\langle x'|x''\rangle}_{=\delta(x'-x'')}\psi_2(x'')}_{=\psi_2(x')}}_{=\psi_2(x)} \\
&= \int \mathrm{d}x\,\psi_1(x)^*\psi_2(x)\,, \quad \text{all right as well.}
\end{aligned}
\tag{1.3.23}
$$

Finally, is $1^2 = 1$? Let us see,

$$
\begin{aligned}
1^2 &= \int \mathrm{d}x\,|x\rangle\langle x| \int \mathrm{d}x'\,|x'\rangle\langle x'| \\
&= \int \mathrm{d}x\,|x\rangle \int \mathrm{d}x'\,\underbrace{\langle x|x'\rangle}_{=\delta(x-x')}\langle x'| \\
&= \int \mathrm{d}x\,|x\rangle\langle x| = 1\,, \quad \text{indeed.}
\end{aligned}
\tag{1.3.24}
$$

In summary, we have for the position states

$$\begin{aligned} \text{adjoint relations:}\ & |x\rangle = \langle x|^\dagger\,, \quad \langle x| = |x\rangle^\dagger\,, \\ \text{orthonormality:}\ & \langle x|x'\rangle = \delta(x-x')\,, \\ \text{completeness:}\ & \int \mathrm{d}x\, |x\rangle\langle x| = 1\,. \end{aligned} \tag{1.3.25}$$

And, by the same token, the corresponding statements hold for the momentum states,

$$\begin{aligned} \text{adjoint relations:}\ & |p\rangle = \langle p|^\dagger\,, \quad \langle p| = |p\rangle^\dagger\,, \\ \text{orthonormality:}\ & \langle p|p'\rangle = \delta(p-p')\,, \\ \text{completeness:}\ & \int \mathrm{d}p\, |p\rangle\langle p| = 1\,, \end{aligned} \tag{1.3.26}$$

because we can repeat the whole line of reasoning with labels x consistently replaced by labels p.

1.4 *xp* transformation function

Since the two integrals in (1.3.1) are different parameterization for the same ket $|\ \rangle$, there must be well defined relations between the wave functions $\psi(x)$ and $\psi(p)$ and also between the kets $|x\rangle$ and $|p\rangle$. For the wave functions, the relations are the Fourier transformations of (1.1.3). We use them now to establish the corresponding statements that relate $|x\rangle$ and $|p\rangle$ to each other.

First note what is already implicit in (1.3.22), namely that

$$\begin{aligned} \langle x|\ \rangle &= \langle x| \int \mathrm{d}x'\, |x'\rangle\psi(x') \\ &= \int \mathrm{d}x'\, \underbrace{\langle x|x'\rangle}_{=\delta(x-x')} \psi(x') \\ &= \psi(x) \end{aligned} \tag{1.4.1}$$

or

$$\psi(x) = \langle x|\ \rangle \tag{1.4.2}$$

and (infer by analogy or repeat the argument)

$$\psi(p) = \langle p|\;\rangle \,. \tag{1.4.3}$$

Therefore we have

$$\begin{aligned}\langle x|\;\rangle = \psi(x) &= \int \mathrm{d}p\, \frac{\mathrm{e}^{\mathrm{i}xp/\hbar}}{\sqrt{2\pi\hbar}}\psi(p) \\ &= \int \mathrm{d}p\, \frac{\mathrm{e}^{\mathrm{i}xp/\hbar}}{\sqrt{2\pi\hbar}}\langle p|\;\rangle \,,\end{aligned} \tag{1.4.4}$$

and this must be true irrespective of the ket $|\;\rangle$ we are considering, so that

$$\langle x| = \int \mathrm{d}p\, \frac{\mathrm{e}^{\mathrm{i}xp/\hbar}}{\sqrt{2\pi\hbar}}\langle p| \tag{1.4.5}$$

follows. The adjoint statement reads

$$|x\rangle = \int \mathrm{d}p\, |p\rangle \frac{\mathrm{e}^{-\mathrm{i}px/\hbar}}{\sqrt{2\pi\hbar}} \,. \tag{1.4.6}$$

The inverse relations follow from

$$\begin{aligned}\langle p|\;\rangle = \psi(p) &= \int \mathrm{d}x\, \frac{\mathrm{e}^{-\mathrm{i}px/\hbar}}{\sqrt{2\pi\hbar}}\psi(x) \\ &= \int \mathrm{d}x\, \frac{\mathrm{e}^{-\mathrm{i}px/\hbar}}{\sqrt{2\pi\hbar}}\langle x|\;\rangle \,,\end{aligned} \tag{1.4.7}$$

which implies

$$\langle p| = \int \mathrm{d}x\, \frac{\mathrm{e}^{-\mathrm{i}px/\hbar}}{\sqrt{2\pi\hbar}}\langle x| \tag{1.4.8}$$

and

$$|p\rangle = \int \mathrm{d}x\, |x\rangle \frac{\mathrm{e}^{\mathrm{i}xp/\hbar}}{\sqrt{2\pi\hbar}} \,. \tag{1.4.9}$$

They are, of course, all variants of each other. The most basic statement, so far implicit, is that about $\langle x|p\rangle$, the *xp transformation function*:

$$\langle x|p\rangle = \underbrace{\langle x| \int \mathrm{d}x'\, |x'\rangle}_{=\int \mathrm{d}x'\, \delta(x-x')} \frac{\mathrm{e}^{\mathrm{i}x'p/\hbar}}{\sqrt{2\pi\hbar}} = \frac{\mathrm{e}^{\mathrm{i}xp/\hbar}}{\sqrt{2\pi\hbar}} \,, \tag{1.4.10}$$

that is

$$\langle x|p\rangle = \frac{\mathrm{e}^{\mathrm{i}xp/\hbar}}{\sqrt{2\pi\hbar}}\,, \tag{1.4.11}$$

which is the fundamental phase factor of Fourier transformation. It is worth memorizing this expression, as everything else follows from it, sometimes by using the adjoint relation

$$\langle p|x\rangle = \frac{\mathrm{e}^{-\mathrm{i}xp/\hbar}}{\sqrt{2\pi\hbar}}\,. \tag{1.4.12}$$

Again, we have the choice of repeating the argument, or we recognize it as a special case of the general statement (1.3.17).

As an illustration of the fundamental role of $\langle x|p\rangle$, we consider,

$$\begin{aligned}\psi(x) &= \langle x|\ \rangle = \langle x|1|\ \rangle \\ &= \langle x|\left(\int \mathrm{d}p\,|p\rangle\langle p|\right)|\ \rangle \\ &= \int \mathrm{d}p\,\langle x|p\rangle\langle p|\ \rangle \\ &= \int \mathrm{d}p\,\frac{\mathrm{e}^{\mathrm{i}xp/\hbar}}{\sqrt{2\pi\hbar}}\psi(p)\,,\end{aligned} \tag{1.4.13}$$

which takes us back to (1.1.3). Another application is

$$\begin{aligned}\delta(x-x') &= \langle x|x'\rangle = \langle x|1|x'\rangle \\ &= \langle x|\left(\int \mathrm{d}p\,|p\rangle\langle p|\right)|x'\rangle \\ &= \int \mathrm{d}p\,\langle x|p\rangle\langle p|x'\rangle \\ &= \int \mathrm{d}p\,\frac{\mathrm{e}^{\mathrm{i}xp/\hbar}}{\sqrt{2\pi\hbar}}\frac{\mathrm{e}^{-\mathrm{i}px'/\hbar}}{\sqrt{2\pi\hbar}} \\ &= \int \frac{\mathrm{d}p}{2\pi\hbar}\,\mathrm{e}^{\mathrm{i}(x-x')p/\hbar}\end{aligned} \tag{1.4.14}$$

which is the basic Fourier representation of the Dirac δ function. It appears in many forms, all of which are variants of

$$\int \mathrm{d}k\,\mathrm{e}^{\mathrm{i}kx} = 2\pi\delta(x)\,. \tag{1.4.15}$$

This, too, is an identity that is worth remembering.

The formulation of the position-momentum analog of the orthogonal transformation (1.2.27) requires some care because $|x\rangle \to |p\rangle$ is a mapping between objects of different metrical dimensions. Relations such as (1.3.9) or (1.3.13) tell us that the position wave function $\psi(x) = \langle x|\ \rangle$ has the metrical dimension of the reciprocal square root of a distance ($1/\sqrt{\mathrm{cm}}$, say), and likewise the momentum wave function $\psi(p) = \langle p|\ \rangle$ has the metrical dimension of the reciprocal square root of a momentum ($1/\sqrt{\mathrm{g\,cm/s}}$, for instance). And since the state ket $|\ \rangle$ is dimensionless, see (1.3.14), the bras $\langle x|$ and $\langle p|$ have these metrical dimensions as well, and so do the kets $|x\rangle$ and $|p\rangle$.

Therefore it is expedient to work with dimensionless quantities, for which purpose we introduce an arbitrary length scale a. Then, the analog of (1.2.27) reads

$$|x\rangle\sqrt{a} \to |p\rangle\sqrt{\hbar/a} = U|x\rangle\sqrt{a} \quad \text{for} \quad \frac{x}{a} = \frac{p}{\hbar/a}\,, \tag{1.4.16}$$

where we note that $\hbar/a$ is the corresponding momentum scale because Planck's constant has the metrical dimension of length $\times$ momentum. The operator U thus defined is given by

$$\begin{aligned} U = U1 = U\int \mathrm{d}x\,|x\rangle\langle x| &= \int \mathrm{d}x\,U|x\rangle\langle x| \\ &= \int \mathrm{d}x\,|p = x\hbar/a^2\rangle\sqrt{\hbar/a^2}\langle x| \end{aligned} \tag{1.4.17}$$

or, after substituting $x = ta$,

$$U = \int \mathrm{d}t\,|p = t\hbar/a\rangle\sqrt{\hbar}\langle x = ta|\,. \tag{1.4.18}$$

1-6 Verify that U brings about the transformation in (1.4.16). Then evaluate $\langle p|U$. Do you get what you expect?

The hermitian conjugation of (1.3.2) has the same product rule as transposition, see (1.2.11),

$$(AB)^\dagger = B^\dagger A^\dagger\,, \tag{1.4.19}$$

because the complex conjugation that distinguishes the two operations has no effect on the order of multiplication. In particular we have $\bigl(|p\rangle\langle x|\bigr)^\dagger =$

$|x\rangle\langle p|$, so that

$$U^\dagger = \int \mathrm{d}t \, |x = ta\rangle \sqrt{\hbar} \langle p = t\hbar/a| \tag{1.4.20}$$

is the adjoint of U. In analogy with the orthogonality statement (1.2.36), we expect

$$UU^\dagger = 1 = U^\dagger U \tag{1.4.21}$$

to hold. Let us verify the left identity:

$$\begin{aligned}
UU^\dagger &= \int \mathrm{d}t \, |p = t\hbar/a\rangle \sqrt{\hbar} \underbrace{\langle x = ta| \int \mathrm{d}t' \, |x = t'a\rangle}_{= \int \mathrm{d}t' \, \delta(ta - t'a)} \sqrt{\hbar} \langle p = t'\hbar/a| \\
&= \int \mathrm{d}t \, |p = t\hbar/a\rangle \hbar \underbrace{\int \mathrm{d}t' \, \delta(ta - t'a) \langle p = t'\hbar/a|}_{= (1/a)\langle p = t\hbar/a|} \\
&= \int \mathrm{d}p \, |p\rangle\langle p| = 1 \,.
\end{aligned} \tag{1.4.22}$$

It is all right indeed, and the right identity in (1.4.21) is demonstrated the same way. At an intermediate step the identity

$$\delta(ta - t'a) = \frac{1}{a} \delta(t - t') \qquad (a > 0) \tag{1.4.23}$$

is used, which is a special case of (5.1.109) in *Basic Matters*.

Operators with the property (1.4.21), that is: their adjoint is their inverse, are called *unitary operators*. They transform sets of kets or bras into equivalent sets, much like a rotation turns sets of vectors of column or row type into equivalent ones, and play a very important role in quantum mechanics.

1.5 Position operator, momentum operator, functions of them

We look for the object (electron, atom, ...) and find it at position x, with the probability of finding it inside a small vicinity around x given by $\mathrm{d}x \, |\psi(x)|^2$. This is then the probability distribution associated with the (random) variable x. Accordingly, the mean value of x is calculated as

$$\overline{x} = \int \mathrm{d}x \, |\psi(x)|^2 x \,, \tag{1.5.1}$$

and the mean value of x^2 as

$$\overline{x^2} = \int \mathrm{d}x \, |\psi(x)|^2 x^2 \,. \tag{1.5.2}$$

More generally we have

$$\overline{x^n} = \int \mathrm{d}x \, |\psi(x)|^2 x^n \tag{1.5.3}$$

for an arbitrary power, and

$$\overline{f(x)} = \int \mathrm{d}x \, |\psi(x)|^2 f(x) \tag{1.5.4}$$

for the mean value of an arbitrary function of the position variable x.

Recalling that $|\psi(x)|^2 = \psi(x)^*\psi(x) = \langle\ |x\rangle\langle x|\ \rangle$, we rewrite these expressions as

$$\begin{aligned}
\overline{x} &= \langle\ |\Bigl(\int \mathrm{d}x \, |x\rangle x \langle x|\Bigr)|\ \rangle \,, \\
\overline{x^2} &= \langle\ |\Bigl(\int \mathrm{d}x \, |x\rangle x^2 \langle x|\Bigr)|\ \rangle \,, \\
&\ \,\vdots \\
\overline{f(x)} &= \langle\ |\Bigl(\int \mathrm{d}x \, |x\rangle f(x) \langle x|\Bigr)|\ \rangle \,,
\end{aligned} \tag{1.5.5}$$

thereby isolating the specific state of the system — bra on the left, ket on the right — from the quantity that we are taking the average of,

$$\begin{aligned}
\overline{x} &\longrightarrow \int \mathrm{d}x \, |x\rangle x \langle x| \equiv X \,, \\
\overline{x^2} &\longrightarrow \int \mathrm{d}x \, |x\rangle x^2 \langle x| \equiv X^2 \,, \\
&\ \,\vdots \\
\overline{f(x)} &\longrightarrow \int \mathrm{d}x \, |x\rangle f(x) \langle x| \equiv f(X) \,.
\end{aligned} \tag{1.5.6}$$

The first line introduces the *position operator* X as the integral of $|x\rangle x\langle x|$, the second line introduces X^2, the square of X, as we can verify,

$$\begin{aligned}
XX &= \int \mathrm{d}x\, |x\rangle x \langle x| \int \mathrm{d}x'\, |x'\rangle x' \langle x'| \\
&= \int \mathrm{d}x\, |x\rangle x \underbrace{\int \mathrm{d}x'\, \underbrace{\langle x|x'\rangle}_{=\delta(x-x')} x' \langle x'|}_{=x\langle x|} \\
&= \int \mathrm{d}x\, |x\rangle x^2 \langle x| = X^2, \quad \text{indeed}\,.
\end{aligned} \tag{1.5.7}$$

And so forth, we have

$$X^n = \int \mathrm{d}x\, |x\rangle x^n \langle x| \tag{1.5.8}$$

for the powers of X. Then by linear combinations,

$$f(X) = \int \mathrm{d}x\, |x\rangle f(x) \langle x| \tag{1.5.9}$$

for all polynomial functions of x and, by approximation, finally for all reasonable functions of x. “Reasonable” means here that the numbers $f(x)$ have to be well defined function values for all real numbers x. Once we have gone through this argument we can just accept (1.5.9) as the definition of a function of position operator X. The integral on the right-hand side of (1.5.9) is an example for the *spectral decomposition* of an operator, here of $f(X)$.

1-7 Show that the eigenvalue equation

$$f(X)|x\rangle = |x\rangle f(x) \tag{1.5.10}$$

holds. What is $\langle x|f(X)$?

1-8 Consider $f(X)^\dagger = \int \mathrm{d}x\, |x\rangle f(x)^* \langle x|$ and compare $f(X)f(X)^\dagger$ with $f(X)^\dagger f(X)$. Which property must be possessed by $f(x)$ if $f(X)$ is its own adjoint, $f(X)^\dagger = f(X)$? Operators with this property are called *selfadjoint* or simply *hermitian* (Charles Hermite), whereby we ignore the subtle difference between the two terms in the mathematical literature. Conclude that all expectation values of a hermitian operator are real. This reality property can serve as an alternative definition of what is a hermitian operator.

Likewise, we have the *momentum operator* P,

$$P = \int \mathrm{d}p\,\big|p\big\rangle p\big\langle p\big|\,, \tag{1.5.11}$$

and can rely on the spectral decomposition

$$g(P) = \int \mathrm{d}p\,\big|p\big\rangle g(p)\big\langle p\big| \tag{1.5.12}$$

for all functions of P that derive from reasonable numerical functions $g(p)$. Once again, the reasoning is completely analogous and we need not repeat it.

1-9 Show the following fundamental relations:

$$\big\langle x\big|P = \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}\big\langle x\big|\,, \qquad X\big|p\big\rangle = \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial p}\big|p\big\rangle\,, \tag{1.5.13}$$

and their integral versions

$$\big\langle x\big|\,\mathrm{e}^{\mathrm{i}x'P/\hbar} = \big\langle x+x'\big|\,, \qquad \mathrm{e}^{\mathrm{i}p'X/\hbar}\big|p\big\rangle = \big|p+p'\big\rangle\,. \tag{1.5.14}$$

The latter involve the basic unitary operators associated with P and X; more about them in Section 1.8

Now, with the operator functions of X and P at hand, we return to (1.5.5) and note that

$$\underset{\substack{\text{average of}\\ \text{the numbers } x}}{\overline{x}} = \underbrace{\big\langle\ \big|X\big|\ \big\rangle}_{\substack{\text{position operator } X \text{ sandwiched}\\ \text{between state bra } \langle\ | \text{ and state ket } |\ \rangle}} \equiv \underbrace{\langle X\rangle}_{\substack{\text{"expectation value"}\\ \text{of position operator } X}} \tag{1.5.15}$$

which introduces a new notation, $\langle X\rangle$, that emphasizes the role played by the position operator X. One speaks of the "expectation value", a historical terminology that is, as so often, not fully logically but completely standard.

Similarly, we write $\big\langle f(X)\big\rangle$ for the expectation value of the operator-function $f(X)$ and $\langle P\rangle$, $\big\langle P^2\big\rangle$, ..., $\big\langle g(P)\big\rangle$, for the expectation values of P, P^2, ..., $g(P)$. We have introduced the latter, by analogy, in terms of integrals involving the momentum wave function $\psi(p)$, but we can, of course, also refer to the position wave function $\psi(x)$.

1-10 In particular, establish expressions (you know them from your first course on quantum mechanics) for the expectation values $\langle P\rangle$ and $\bigl\langle P^2\bigr\rangle$ in terms of $\psi(x)$.

1-11 Determine the normalization constant A for the position wave function

$$\psi(x)=\begin{cases} A\sin(2\pi x/L) & \text{for} \quad -L<x<L\,,\\ 0 & \text{for} \quad |x|>L\,,\end{cases}$$

and then calculate $\langle X\rangle$, $\bigl\langle X^2\bigr\rangle$, $\langle P\rangle$, and $\bigl\langle P^2\bigr\rangle$.

1.6 Traces and statistical operators

Given a ket $\bigl|1\bigr\rangle$ and a bra $\bigl\langle 2\bigr|$, we can multiply them in either order, thereby getting

$$\begin{aligned} &\text{the bracket } \bigl\langle 2\bigl|1\bigr\rangle\,, \ \text{a number}\,,\\ \text{or}\quad &\text{the “ket-bra” } \bigl|1\bigr\rangle\bigl\langle 2\bigr|\,, \ \text{an operator}\,.\end{aligned}$$

In the bracket $\bigl\langle 2\bigl|1\bigr\rangle$, the ingredients are no longer identifiable because very many pairs of a bra and a ket have the same number for the bracket. By contrast, given the ket-bra $\bigl|1\bigr\rangle\bigl\langle 2\bigr|$, one can identify the ingredient almost uniquely. Therefore, we cannot expect that there could be a mapping from the bracket to the ket-bra, but there is a mapping from the ket-bra to the bracket,

$$\bigl|1\bigr\rangle\bigl\langle 2\bigr| \to \bigl\langle 2\bigl|1\bigr\rangle\,.$$

It is called “taking the trace” and we write $\mathrm{tr}\{\cdots\}$ for it,

$$\mathrm{tr}\bigl\{\bigl|1\bigr\rangle\bigl\langle 2\bigr|\bigr\} = \bigl\langle 2\bigl|1\bigr\rangle\,; \tag{1.6.1}$$

read: the trace of $\bigl|1\bigr\rangle\bigl\langle 2\bigr|$ is $\bigl\langle 2\bigl|1\bigr\rangle$.

Before proceeding, let us look at the analog for column and row vectors:

$$r^{\downarrow}\vec{s} \longrightarrow \vec{s}\,r^{\downarrow} = \vec{s}\cdot\vec{r}\,, \tag{1.6.2}$$

or

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} (u_1, u_2, \ldots, u_n) = \begin{pmatrix} x_1u_1 & x_1u_2 & \cdots & x_1u_n \\ x_2u_1 & x_2u_2 & \cdots & x_2u_n \\ \vdots & \vdots & \ddots & \vdots \\ x_nu_1 & x_nu_2 & \cdots & x_nu_n \end{pmatrix}$$
$$\rightarrow x_1u_1 + x_2u_2 + \cdots + x_nu_n \,, \qquad (1.6.3)$$

the diagonal sum of the matrix for $r^{\downarrow}\vec{s}$. Clearly, if you only know the value of this sum, you cannot reconstruct the whole matrix, but given the matrix, you easily figure out the diagonal sum.

The linear structure for kets and bras is inherited by the trace. For example, consider

$$\bigl|1\bigr\rangle = \bigl|1_a\bigr\rangle\alpha + \bigl|1_b\bigr\rangle\beta \,, \qquad (1.6.4)$$

and compare the two ways of evaluating $\mathrm{tr}\bigl\{\bigl|1\bigr\rangle\bigl\langle 2\bigr|\bigr\}$,

$$\begin{aligned} \mathrm{tr}\bigl\{\bigl|1\bigr\rangle\bigl\langle 2\bigr|\bigr\} &= \bigl\langle 2\bigl|1\bigr\rangle = \bigl\langle 2\bigl|1_a\bigr\rangle\alpha + \bigl\langle 2\bigl|1_b\bigr\rangle\beta \\ &= \alpha\,\mathrm{tr}\bigl\{\bigl|1_a\bigr\rangle\bigl\langle 2\bigr|\bigr\} + \beta\,\mathrm{tr}\bigl\{\bigl|1_b\bigr\rangle\bigl\langle 2\bigr|\bigr\} \,, \\ \mathrm{tr}\bigl\{\bigl|1\bigr\rangle\bigl\langle 2\bigr|\bigr\} &= \mathrm{tr}\Bigl\{\Bigl(\bigl|1_a\bigr\rangle\alpha + \bigl|1_b\bigr\rangle\beta\Bigr)\bigl\langle 2\bigr|\Bigr\} \\ &= \mathrm{tr}\bigl\{\bigl|1_a\bigr\rangle\alpha\bigl\langle 2\bigr| + \bigl|1_b\bigr\rangle\beta\bigl\langle 2\bigr|\bigr\} \,. \end{aligned} \qquad (1.6.5)$$

This generalizes to

$$\mathrm{tr}\{\alpha A + \beta B\} = \alpha\,\mathrm{tr}\{A\} + \beta\,\mathrm{tr}\{B\} \qquad (1.6.6)$$

immediately, wherein A, B are operators and α, β are numbers.

We apply this to expectation values, such as

$$\langle A\rangle = \bigl\langle\ \bigl|A\bigr|\ \bigr\rangle \qquad (1.6.7)$$

where A is any operator, perhaps a function $f(X)$ of position operator X, or a function $g(P)$ of momentum operator P, or possibly something more complicated, like the symmetrized product $XP + PX$, say. Whatever the nature of operator A, we can read the above statement as

$$\langle A\rangle = \underbrace{\bigl\langle\ \bigl|A}_{\text{bra}}\underbrace{\bigl|\ \bigr\rangle}_{\text{ket}} = \mathrm{tr}\Bigl\{\underbrace{\bigl|\ \bigr\rangle}_{\text{ket}}\underbrace{\bigl\langle\ \bigl|A}_{\text{bra}}\Bigr\} \qquad (1.6.8)$$

or as

$$\langle A\rangle = \underbrace{\langle\ |}_{\textsf{bra}}\underbrace{A|\ \rangle}_{\textsf{ket}} = \mathrm{tr}\left\{\underbrace{A|\ \rangle}_{\textsf{ket}}\underbrace{\langle\ |}_{\textsf{bra}}\right\}, \tag{1.6.9}$$

which introduces $|\ \rangle\langle\ |$ as the mathematical object that refers solely to the state of affairs of the physical system,

$$\langle A\rangle = \mathrm{tr}\left\{\underbrace{|\ \rangle\langle\ |}_{\textsf{physical state}}\underset{\textsf{operator considered}}{A}\right\} = \mathrm{tr}\left\{\underset{\textsf{operator considered}}{A}\underbrace{|\ \rangle\langle\ |}_{\textsf{physical state}}\right\}. \tag{1.6.10}$$

What has been achieved here, is the complete separation of the mathematical entities that refer to the physical property (position, momentum, functions of them, such as energy), and to the state of affairs. There is one appropriate operator A for the physical property, irrespective of the particular state that is actually the case, and all that characterizes the actual situation is contained in the ket-bra

$$\rho \equiv |\ \rangle\langle\ |\,. \tag{1.6.11}$$

This is also an operator, but not one that describes a physical property, rather it is the *statistical operator* that summarizes all statistical aspects of the system as actually prepared. It is common to refer to the statistical operator as the *state operator* or simply the *state* of the physical system.

In particular, we extract probabilities from ρ as illustrated by the probability of finding the object between $x = a$ and $x = b$ ($a < b$). We recall that

$$\begin{aligned}\mathrm{prob}(a < x < b) &= \int_a^b \mathrm{d}x\, |\psi(x)|^2 \\ &= \int_{-\infty}^{\infty} \mathrm{d}x\, |\psi(x)|^2 \chi_{a,b}(x)\,,\end{aligned} \tag{1.6.12}$$

where

$$\chi_{a,b}(x) = \begin{cases} 1 & \text{if} \quad a < x < b\,, \\ 0 & \text{elsewhere,} \end{cases} \tag{1.6.13}$$

is the *mesa function* for the interval $a < x < b$. Thus,

$$\begin{aligned}\operatorname{prob}(a < x < b) &= \langle\ \big|\Big(\int \mathrm{d}x\, \big|x\big\rangle \chi_{a,b}(x) \big\langle x\big|\Big)\big|\ \rangle \\ &= \langle\ \big|\chi_{a,b}(X)\big|\ \rangle \\ &= \operatorname{tr}\big\{\chi_{a,b}(X)\big|\ \rangle\langle\ \big|\big\} \\ &= \operatorname{tr}\{\chi_{a,b}(X)\rho\}\ , \end{aligned} \tag{1.6.14}$$

where we recognize the expectation value of a function of position operator X, namely $\chi_{a,b}(X)$. Indeed, such a probability is an expectation value, and the argument is easily extended to other probabilities as well.

1-12 Find the operator A whose expectation value is the probability of finding the object in the range $x > 0$. Express A in terms of the *sign function*

$$\operatorname{sgn}(x) = \begin{cases} +1 & \text{for} \quad x > 0\,, \\ -1 & \text{for} \quad x < 0\,. \end{cases}$$

Repeat for the range $x < 0$, and then check that the two operators add up to the identity.

In (1.6.10) we found

$$\langle A\rangle = \operatorname{tr}\{A\rho\} = \operatorname{tr}\{\rho A\}\ , \quad \rho = \big|\ \big\rangle\big\langle\ \big|\,, \tag{1.6.15}$$

as if the order of multiplication did not matter. In fact, it does not. We demonstrate this by evaluating $\operatorname{tr}\{AB\}$ and $\operatorname{tr}\{BA\}$ for $A = \big|1\big\rangle\big\langle 2\big|$ and $B = \big|3\big\rangle\big\langle 4\big|$, which is all we need, because all operators are linear combinations of such ingredients, and we know already that the trace respects this linearity. Therefore, it suffices to consider these special operators. See, then,

$$\begin{aligned}\operatorname{tr}\{AB\} &= \operatorname{tr}\Big\{\underbrace{\big|1\big\rangle\big\langle 2\big|3\big\rangle}_{\text{ket}}\underbrace{\big\langle 4\big|}_{\text{bra}}\Big\} = \big\langle 4\big|1\big\rangle\big\langle 2\big|3\big\rangle\,, \\ \operatorname{tr}\{BA\} &= \operatorname{tr}\Big\{\underbrace{\big|3\big\rangle\big\langle 4\big|1\big\rangle}_{\text{ket}}\underbrace{\big\langle 2\big|}_{\text{bra}}\Big\} = \big\langle 2\big|3\big\rangle\big\langle 4\big|1\big\rangle\,. \end{aligned} \tag{1.6.16}$$

Indeed, they are the same because the numbers $\big\langle 4\big|1\big\rangle$ and $\big\langle 2\big|3\big\rangle$ can be multiplied in either order without changing the value of the product. Please note that $AB \neq BA$, as a rule, but their traces are the same.

Since the operators A, B themselves can be products, we have the more general rule

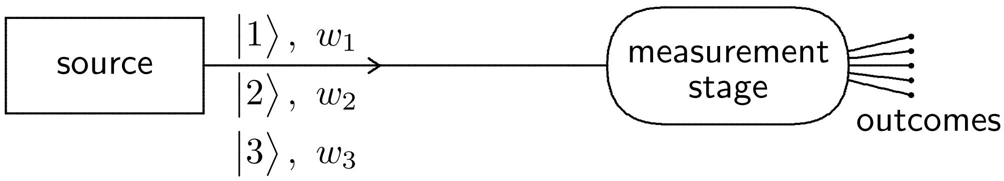
$$\mathrm{tr}\{ABC\} = \mathrm{tr}\{CAB\} = \mathrm{tr}\{BCA\} \tag{1.6.17}$$

for products of three factors and analogous statements about four, five, ... factors. All of them are summarized in the observation that

> the value of a trace does not change when the factors in a product are permuted cyclically.

This *cyclic property of the trace* is exploited very often.

So far, letter ρ was just an abbreviation for the ket-bra $|\ \rangle\langle\ |$, the product of the ket and the bra describing the actual state of affairs. Let us now move on and consider a more general situation:



We have a source that puts out the atoms either in state $|1\rangle$, or in state $|2\rangle$, or in state $|3\rangle$, whereby we have no clue what will be the case for the next atom except that we know that there are definite probabilities w_1, w_2, w_3 of occurrence for the three states, with $w_1 + w_2 + w_3 = 1$, of course, because nothing else can possibly happen.

At the measurement stage we perform a measurement of the physical property that is associated with operator A. Thus we have an expectation value that is given by

$$\langle A\rangle = w_1\langle 1|A|1\rangle + \underbrace{w_2}_{\text{probability of getting }|2\rangle}\underbrace{\langle 2|A|2\rangle}_{\text{expectation value if }|2\rangle\text{ is the case}} + w_3\langle 3|A|3\rangle \tag{1.6.18}$$

or

$$\begin{aligned}\langle A\rangle &= w_1\mathrm{tr}\{A|1\rangle\langle 1|\} + w_2\mathrm{tr}\{A|2\rangle\langle 2|\} + w_3\mathrm{tr}\{A|3\rangle\langle 3|\}\\ &= \mathrm{tr}\Big\{A\Big(|1\rangle w_1\langle 1| + |2\rangle w_2\langle 2| + |3\rangle w_3\langle 3|\Big)\Big\}\\ &= \mathrm{tr}\{A\rho\}\end{aligned} \tag{1.6.19}$$

with the statistical operator

$$\rho = \sum_{k=1}^{3} \big|k\big\rangle w_k \big\langle k\big| \,. \tag{1.6.20}$$

This ρ summarizes all that we know about the source: there are the states $\big|k\big\rangle$ and their statistical weights w_k.

There is nothing particular about the situation discussed, with three different states emitted by the source, there could be fewer or more. Accordingly, we have the more general case of

$$\rho = \sum_k \big|k\big\rangle w_k \big\langle k\big| \,, \quad w_k > 0 \,, \quad \sum_k w_k = 1 \,, \tag{1.6.21}$$

where the summation can have one or more terms, and it is understood that all kets and bras are normalized properly,

$$\big\langle k\big|k\big\rangle = 1 \quad \text{for all } k \,. \tag{1.6.22}$$

By measuring sufficiently many different physical properties, we can establish a body of experimental data that enables us to infer the statistical operator ρ with the desired precision. Then we know the statistical properties of the atoms emitted by the source, and this is *all* we can find out. We cannot, in particular, establish the ingredients $\big|k\big\rangle$ and their weights w_k, we can only know ρ. This is also the only really meaningful thing to know since it gives us all probabilities for the statistical predictions. Nothing else is needed. Nor is anything else available.

When we state that knowledge of ρ does not translate into knowledge of the ingredients from which it is composed in

$$\rho = \sum_k \big|k\big\rangle w_k \big\langle k\big| \,, \tag{1.6.23}$$

we mean of course that different right-hand sides can give the same ρ. To make this point it is quite enough to give one example. The simplest is

$$\rho = \frac{1}{2}\Big(\big|1\big\rangle\big\langle 1\big| + \big|2\big\rangle\big\langle 2\big|\Big) \quad \text{with} \quad \big\langle 1\big|2\big\rangle = 0 \,, \tag{1.6.24}$$

that is just two states mixed with equal weights of 50% for each. In view of their stated orthogonality, the kets

$$\big|\alpha\big\rangle = \frac{1}{\sqrt{2}}\Big(\big|1\big\rangle + \big|2\big\rangle\Big) \,, \quad \big|\beta\big\rangle = \frac{1}{\sqrt{2}}\Big(\big|1\big\rangle - \big|2\big\rangle\Big) \tag{1.6.25}$$

are also orthogonal and properly normalized. Then

$$\begin{aligned}
&\frac{1}{2}\Bigl(|\alpha\rangle\langle\alpha| + |\beta\rangle\langle\beta|\Bigr)\\
&= \frac{1}{2}\left(\frac{|1\rangle + |2\rangle}{\sqrt{2}}\,\frac{\langle 1| + \langle 2|}{\sqrt{2}} + \frac{|1\rangle - |2\rangle}{\sqrt{2}}\,\frac{\langle 1| - \langle 2|}{\sqrt{2}}\right)\\
&= \frac{1}{2}\Bigl(|1\rangle\langle 1| + |2\rangle\langle 2|\Bigr) = \rho
\end{aligned} \tag{1.6.26}$$

establishes

$$\rho = \frac{1}{2}\Bigl(|\alpha\rangle\langle\alpha| + |\beta\rangle\langle\beta|\Bigr) \tag{1.6.27}$$

which has different ingredients than the original ρ of (1.6.24). We speak of the two *blends* for one and the same *mixture* or *mixed state*. All that is relevant is the mixture ρ, not the particular ways in which one can blend it.

For

$$\rho = \frac{1}{2}\Bigl(|1\rangle\langle 2| + |2\rangle\langle 2|\Bigr) = \frac{1}{2}\Bigl(|\alpha\rangle\langle\alpha| + |\beta\rangle\langle\beta|\Bigr) \tag{1.6.28}$$

one can say that "it is *as if* we had 50% of $|1\rangle$ and 50% of $|2\rangle$" or one can say with equal justification that "it is *as if* we had 50% of $|\alpha\rangle$ and 50% of $|\beta\rangle$". But neither *as-if reality* is better than the other, both are on exactly the same footing, and there are many more as-if realities associated with this ρ.

1-13 Consider $|u\rangle = \frac{1}{\sqrt{2}}\Bigl(|1\rangle + \mathrm{i}|2\rangle\Bigr)$, $|v\rangle = \frac{1}{\sqrt{2}}\Bigl(|1\rangle - \mathrm{i}|2\rangle\Bigr)$ and evaluate $\frac{1}{2}\Bigl(|u\rangle\langle u| + |v\rangle\langle v|\Bigr)$. What do you conclude?

The basic probability is that of finding a particular state, $|0\rangle$, say. If state $|k\rangle$ is the case, this probability is $\bigl|\langle 0|k\rangle\bigr|^2$, as in (1.3.16), so more generally the probability is $\langle 0|\rho|0\rangle$. It must be nonnegative,

$$\langle 0|\rho|0\rangle \geq 0 \quad \text{for any choice of } |0\rangle\,. \tag{1.6.29}$$

In short: $\rho \geq 0$, which is a basic property of all statistical operators, their *positivity*. Other properties are that ρ is hermitian,

$$\rho = \rho^\dagger\,, \tag{1.6.30}$$

and normalized to unit trace,

$$\mathrm{tr}\{\rho\} = 1\,. \tag{1.6.31}$$

All these properties follow directly from the construction of ρ as a blend in (1.6.21).

We emphasize the physical significance of $\sum_k w_k = 1$. Suppose you perform a measurement that identifies the complete set of states $|a_n\rangle$, that is

$$\sum_n |a_n\rangle\langle a_n| = 1\,. \tag{1.6.32}$$

Then the probabilities of the various outcomes are $\langle a_n|\rho|a_n\rangle$ which are assuredly positive, and their sum must be 1:

$$\begin{aligned} 1 &= \sum_n \langle a_n|\rho|a_n\rangle = \sum_n \mathrm{tr}\{\rho|a_n\rangle\langle a_n|\} \\ &= \mathrm{tr}\Big\{\rho \underbrace{\sum_n |a_n\rangle\langle a_n|}_{=1}\Big\} = \mathrm{tr}\{\rho\}\,. \end{aligned} \tag{1.6.33}$$

That is: $\mathrm{tr}\{\rho\} = 1$ is just the statement that the probabilities of mutually exclusive events have unit sum.

The extreme situation of only one term is the one we started with, $\rho = |\ \rangle\langle\ |$. Then it is possible to almost identify the ingredients. "Almost" because of the phase arbitrariness,

$$\rho = |\ \rangle\langle\ | = \Big(|\ \rangle \mathrm{e}^{\mathrm{i}\varphi}\Big)\Big(\mathrm{e}^{-\mathrm{i}\varphi}\langle\ |\Big)\,, \tag{1.6.34}$$

according to which the pair

$$|\ \rangle \mathrm{e}^{\mathrm{i}\varphi}\,, \qquad \mathrm{e}^{-\mathrm{i}\varphi}\langle\ | \qquad (\varphi \text{ real}) \tag{1.6.35}$$

is as good as the pair $|\ \rangle$, $\langle\ |$. It is one of the advantages of using the statistical operator rather than $|\ \rangle$ and $\langle\ |$ that there is no phase arbitrariness in ρ. The statistical operator ρ is unique, its ingredients are not.

The situation of $\rho = |\ \rangle\langle\ |$ is also special because, for such a *pure* state, it is characteristically true that $\rho^2 = \rho$; see

$$\begin{aligned} \rho^2 = \Big(|\ \rangle\langle\ |\Big)\Big(|\ \rangle\langle\ |\Big) &= |\ \rangle\underbrace{\langle\ |\ \rangle}_{=1}\langle\ | \\ &= |\ \rangle\langle\ | = \rho\,. \end{aligned} \tag{1.6.36}$$

If there is more than one term in $\rho = \sum_k |k\rangle w_k \langle k|$ then $\rho^2 \neq \rho$, as is best illustrated by considering the trace

$$\begin{aligned} \operatorname{tr}\{\rho^2\} &= \sum_{j,k} \operatorname{tr}\{|j\rangle w_j \langle j|k\rangle w_k \langle k|\} \\ &= \sum_{j,k} w_j w_k \left|\langle j|k\rangle\right|^2 . \end{aligned} \tag{1.6.37}$$

Now, since $|j\rangle\langle j| \neq |k\rangle\langle k|$ if $j \neq k$ (we want *really* different ingredients), we have $\left|\langle j|k\rangle\right|^2 < 1$ for $j \neq k$, so that

$$\operatorname{tr}\{\rho^2\} < \sum_{j,k} w_j w_k = \underbrace{\left(\sum_j w_j\right)}_{=1} \underbrace{\left(\sum_k w_k\right)}_{=1} = 1 , \tag{1.6.38}$$

or $\operatorname{tr}\{\rho^2\} < 1$. Thus we have

$$\begin{aligned} &\operatorname{tr}\{\rho^2\} = 1 \quad \text{if} \quad \rho = |\ \rangle\langle\ | , \quad \rho^2 = \rho \\ \text{and} \quad &\operatorname{tr}\{\rho^2\} < 1 \quad \text{otherwise} . \end{aligned} \tag{1.6.39}$$

Therefore, the number $\operatorname{tr}\{\rho^2\}$ can serve as a crude measure of the purity of the state, it is maximal, $\operatorname{tr}\{\rho^2\} = 1$, for a pure state, $\rho = |\ \rangle\langle\ |$, and surely less than unity for all truly mixed states.

1.7 Algebraic completeness of operators X and P

We have the position operator X and the momentum operator P, functions $f(X)$, $g(P)$ of either one, and upon forming products and sums of such functions can introduce rather arbitrary functions of both X and P. And these general functions $f(X, P)$ comprise *all* possible operators for a degree of freedom of this sort. In other words: position X and momentum P are *algebraically complete*.

To demonstrate this we show that we can write any given operator A as a function of X and P, quite systematically. We begin with noting that $\langle x|A$ is a well defined bra and $A|p\rangle$ is a well defined ket, and that $\langle x|A|p\rangle$ is a uniquely specified set of numbers once A is stated. These numbers appear in

$$A = 1\,A\,1 = \int \mathrm{d}x\,|x\rangle\langle x|A \int \mathrm{d}p\,|p\rangle\langle p|$$
$$= \int \mathrm{d}x\,\mathrm{d}p\,|x\rangle\langle x|A|p\rangle\langle p|\,. \tag{1.7.1}$$

We divide and multiply by

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}\,\mathrm{e}^{\mathrm{i}xp/\hbar}\,, \tag{1.7.2}$$

which is never zero, to arrive at

$$A = \int \mathrm{d}x\,\mathrm{d}p\,|x\rangle\langle x|a(x,p)|p\rangle\langle p| \tag{1.7.3}$$

where

$$a(x,p) = \frac{\langle x|A|p\rangle}{\langle x|p\rangle} \tag{1.7.4}$$

is such that $A = 1$ is mapped onto $a(x,p) = 1$. Borrowing once again the terminology from classical mechanics, we call $a(x,p)$ a *phase-space function* of A.

1-14 Show that this mapping $A \to a(x,p)$ is linear. What is $a(x,p)$ for $A = f(X)$? For $A = g(P)$?

Further, consistent with the general rules

$$f(X) = \int \mathrm{d}x'\,|x'\rangle f(x')\langle x'|\,,$$
$$g(P) = \int \mathrm{d}p'\,|p'\rangle g(p')\langle p'| \tag{1.7.5}$$

we have

$$|x\rangle\langle x| = \int \mathrm{d}x'\,|x'\rangle\delta(x'-x)\langle x'| = \delta(X-x)\,,$$
$$|p\rangle\langle p| = \int \mathrm{d}p'\,|p'\rangle\delta(p'-p)\langle p'| = \delta(P-p) \tag{1.7.6}$$

and therefore

$$A = \int \mathrm{d}x\,\mathrm{d}p\;\delta(X-x)a(x,p)\delta(P-p)\,. \tag{1.7.7}$$

This equation already proves the case: we have expressed A as a function of X and P. But we can go one step further and evaluate the integrals over the δ functions, with the outcome

$$A = a(X,P)\Big|_{X,P\text{-ordered}} \equiv a(X;P) \tag{1.7.8}$$

where we must pay due attention to the structure of the previous expression (1.7.7). There all Xs stand to the left of all Ps, and this order must be preserved when we replace $x \to X$, $p \to P$ in $a(x,p)$.

We have thus achieved even more than what we really needed. Operator A is now expressed as an *ordered* function of X and P, for which the procedure gives a unique answer. Of course, we can interchange the roles of position and momentum in this argument and can equally well arrive at a unique P,X-ordered form, where all P operators are to the left of all X operators in products.

As a simple example, consider $A = PX$ for which

$$\begin{aligned}\langle x|A|p\rangle &= \langle x|PX|p\rangle \\ &= \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}\langle x|\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial p}|p\rangle \\ &= \left(\frac{\hbar}{\mathrm{i}}\right)^2\frac{\partial}{\partial x}\frac{\partial}{\partial p}\langle x|p\rangle \\ &= \left(\frac{\hbar}{\mathrm{i}}\right)^2\left(\frac{\mathrm{i}}{\hbar}+\left(\frac{\mathrm{i}}{\hbar}\right)^2 xp\right)\langle x|p\rangle\,, \end{aligned} \tag{1.7.9}$$

where the last step exploits the familiar explicit form of $\langle x|p\rangle$ in (1.7.2). Accordingly,

$$a(x,p) = \frac{\langle x|A|p\rangle}{\langle x|p\rangle} = xp + \frac{\hbar}{\mathrm{i}} \tag{1.7.10}$$

here, and we get the X,P-ordered form

$$A = PX = \left(xp + \frac{\hbar}{\mathrm{i}}\right)\Bigg|_{\substack{x\,\to\, X\\ p\,\to\, P\\ \text{ordered}}} = XP - \mathrm{i}\hbar\,. \tag{1.7.11}$$

The result is, of course, as expected inasmuch as we just get Werner Heisenberg's fundamental *commutation relation*

$$[X,P] = \mathrm{i}\hbar\,. \tag{1.7.12}$$

(See Section 3.2 in *Basic Matters* for the basic properties of commutators, in particular their linearity, expressed by the sum rule, and their product rule.) We recall the two extensions,

$$[f(X), P] = \mathrm{i}\hbar \frac{\partial f(X)}{\partial X}, \qquad [X, g(P)] = \mathrm{i}\hbar \frac{\partial g(P)}{\partial P}, \tag{1.7.13}$$

which are frequently used (see Section 5.1.4 in *Basic Matters*).

1-15 Show that, most generally,

$$[F(X,P), P] = \mathrm{i}\hbar \frac{\partial F(X,P)}{\partial X}, \qquad [X, F(X,P)] = \mathrm{i}\hbar \frac{\partial F(X,P)}{\partial P},$$

where $F(X,P)$ is *any* operator function of X and P. It is sufficient (why?) to consider the special case $F = |x\rangle\langle p|$.

1-16 Is there a difference between $\dfrac{\partial}{\partial X}\dfrac{\partial}{\partial P} f(X,P)$ and $\dfrac{\partial}{\partial P}\dfrac{\partial}{\partial X} f(X,P)$?

1-17 Find the X,P-ordered form of the commutator $\dfrac{1}{\mathrm{i}\hbar}\big[X^2, P^2\big]$.

A simple statistical operator for a pure state, $\rho = |\ \rangle\langle\ |$, is our next example. We have

$$\langle x|\rho|p\rangle = \langle x|\ \rangle\langle\ |p\rangle = \psi(x)\psi(p)^* \tag{1.7.14}$$

and then

$$\frac{\langle x|\rho|p\rangle}{\langle x|p\rangle} = \frac{\langle x|\ \rangle\langle\ |p\rangle}{\langle x|p\rangle} = \sqrt{2\pi\hbar}\,\psi(x)\,\mathrm{e}^{-\mathrm{i}xp/\hbar}\psi(p)^* \tag{1.7.15}$$

so that

$$\rho = \sqrt{2\pi\hbar}\,\psi(X)\,\mathrm{e}^{-\mathrm{i}X;\,P/\hbar}\psi(P)^\dagger\,. \tag{1.7.16}$$

Here

$$\begin{aligned}\mathrm{e}^{-\mathrm{i}X;\,P/\hbar} &= \mathrm{e}^{-\mathrm{i}XP/\hbar}\Big|_{X,P\text{-ordered}} \\ &= \sum_{k=0}^{\infty} \frac{1}{k!}\left(\frac{-\mathrm{i}}{\hbar}\right)^k X^k P^k \end{aligned} \tag{1.7.17}$$

is a basic *ordered exponential function*. Its adjoint is

$$\left(\mathrm{e}^{-\mathrm{i}X;\,P/\hbar}\right)^\dagger = \mathrm{e}^{\mathrm{i}P;\,X/\hbar} \tag{1.7.18}$$

as one verifies immediately, and since $\rho^\dagger = \rho$ we get

$$\rho = \sqrt{2\pi\hbar}\,\psi(P)\,\mathrm{e}^{\mathrm{i}P;\,X/\hbar}\psi(X)^\dagger \tag{1.7.19}$$

for the P, X-ordered version of ρ.

When the ordered form of an operator is at hand, it is particularly easy to evaluate its trace,

$$\begin{aligned}
\operatorname{tr}\{A\} &= \operatorname{tr}\left\{\int \mathrm{d}x\,|x\rangle\langle x|A\int \mathrm{d}p\,|p\rangle\langle p|\right\} \\
&= \int \mathrm{d}x\,\mathrm{d}p\,\langle x|A|p\rangle\langle p|x\rangle \\
&= \int \mathrm{d}x\,\mathrm{d}p\,\underbrace{\frac{\langle x|A|p\rangle}{\langle x|p\rangle}}_{a(x,p)=}\,\underbrace{\langle x|p\rangle\langle p|x\rangle}_{=1/(2\pi\hbar)}
\end{aligned} \tag{1.7.20}$$

so that

$$\operatorname{tr}\{A\} = \int \frac{\mathrm{d}x\,\mathrm{d}p}{2\pi\hbar}\,a(x,p)\,. \tag{1.7.21}$$

This has the appearance of a classical phase-space integral, counting one quantum state per phase-space area of $2\pi\hbar$, so to say.

Next, suppose operator A is given in its X, P-ordered form,

$$A = a(X;P)\,, \tag{1.7.22}$$

and the statistical operator is given as a P, X-ordered expression

$$\rho = r(P;X)\,. \tag{1.7.23}$$

Then we have

$$\begin{aligned}
\langle A\rangle &= \operatorname{tr}\{\rho A\} \\
&= \operatorname{tr}\left\{\int \mathrm{d}p\,|p\rangle\langle p|\rho\int \mathrm{d}x\,|x\rangle\langle x|A\right\} \\
&= \int \mathrm{d}x\,\mathrm{d}p\,\langle p|\rho|x\rangle\langle x|A|p\rangle\,,
\end{aligned} \tag{1.7.24}$$

for the expectation value of A, where we meet

$$\langle p|\rho|x\rangle = \langle p|r(\underset{\substack{\downarrow\\ p}}{P};\underset{\substack{\downarrow\\ x}}{X})|x\rangle = r(p,x)\langle p|x\rangle \tag{1.7.25}$$

and

$$\langle x|A|p\rangle = \langle x|a(\underset{\substack{\downarrow\\ x}}{X};\underset{\substack{\downarrow\\ p}}{P})|p\rangle = a(x,p)\langle x|p\rangle\,. \tag{1.7.26}$$

These just exploit the basic property of ordered operators, namely that X and P stand next to their respective eigenbras and eigenkets so that these operators can be equivalently replaced by their eigenvalues. Then,

$$\langle A\rangle = \int \mathrm{d}x\,\mathrm{d}p\; r(p,x)a(x,p)\,\underbrace{\langle p|x\rangle\langle x|p\rangle}_{=1/(2\pi\hbar)} \tag{1.7.27}$$

or

$$\langle A\rangle = \int \frac{\mathrm{d}x\,\mathrm{d}p}{2\pi\hbar}\, r(p,x)a(x,p)\,, \tag{1.7.28}$$

which looks even more like a classical phase-space integral of the product of a density $r(p,x)$ and a phase-space function $a(x,p)$, whereby $\dfrac{\mathrm{d}x\,\mathrm{d}p}{2\pi\hbar}$ suggests the injunction noted at (1.7.21), namely to count one quantum state per phase-space area of $2\pi\hbar$.

The seemingly classical appearance of (1.7.28) is striking, it is also profound, but we must keep in mind that we continue to talk about quantum mechanical traces and that the phase-space functions $r(p,x)$ and $a(x,p)$ are just particularly convenient numerical descriptions of the quantum mechanical operators ρ and A. We are *not* replacing quantum mechanics by some equivalent version of classical mechanics. There is no such thing.

1-18 Use the X,P-ordered and P,X-ordered forms of $\rho = \big|\ \big\rangle\big\langle\ \big|$ to evaluate $\mathrm{tr}\{\rho\}$ as a phase-space integral.

1.8 Weyl commutator, Baker–Campbell–Hausdorff relations

The unitary operators $\mathrm{e}^{\mathrm{i}p'X/\hbar}$ and $\mathrm{e}^{\mathrm{i}x'P/\hbar}$ of (1.5.14) do not commute, except for special values of x' and p', and we can find out what is the difference between applying them in either order by establishing the X,P-ordered version of the operator

$$A = \mathrm{e}^{\mathrm{i}x'P/\hbar}\,\mathrm{e}^{\mathrm{i}p'X/\hbar}\,, \tag{1.8.1}$$

which we thus define by its P,X-ordered version. We begin with

$$\begin{aligned}\langle x|A|p\rangle &= \underbrace{\langle x|\,\mathrm{e}^{\mathrm{i}x'P/\hbar}}_{\langle x+x'|=}\,\underbrace{\mathrm{e}^{\mathrm{i}p'X/\hbar}|p\rangle}_{=|p+p'\rangle} \\ &= \langle x+x'|p+p'\rangle = \frac{1}{\sqrt{2\pi\hbar}}\,\mathrm{e}^{\mathrm{i}(x+x')(p+p')/\hbar} \\ &= \underbrace{\frac{\mathrm{e}^{\mathrm{i}xp/\hbar}}{\sqrt{2\pi\hbar}}}_{=\langle x|p\rangle}\,\underbrace{\mathrm{e}^{\mathrm{i}xp'/\hbar}\,\mathrm{e}^{\mathrm{i}x'p/\hbar}\,\mathrm{e}^{\mathrm{i}x'p'/\hbar}}_{=a(x,p)}\,,\end{aligned} \tag{1.8.2}$$

and now the replacement $a(x,p) \to a(X;P)$ gives us

$$A = \mathrm{e}^{\mathrm{i}p'X/\hbar}\,\mathrm{e}^{\mathrm{i}x'P/\hbar}\,\mathrm{e}^{\mathrm{i}x'p'/\hbar}\,. \tag{1.8.3}$$

More explicitly, this says

$$\mathrm{e}^{\mathrm{i}p'X/\hbar}\,\mathrm{e}^{\mathrm{i}x'P/\hbar} = \mathrm{e}^{-\mathrm{i}x'p'/\hbar}\,\mathrm{e}^{\mathrm{i}x'P/\hbar}\,\mathrm{e}^{\mathrm{i}p'X/\hbar}\,, \tag{1.8.4}$$

which is Hermann K. H. Weyl's commutation relation for the basic unitary operators associated with X and P, the *Weyl commutator* for short.

How about combining the various exponentials into one? That requires some care because the arguments of the exponentials do not commute with each other. But nevertheless we can be systematic about it, and we could use several methods for this purpose. Let us do it with a sequence of unitary transformations.

As a preparation, we recall that

$$U^\dagger f(A)U = f(U^\dagger AU) \tag{1.8.5}$$

for any function of operator A (arbitrary) and any arbitrary operator U. One easily verifies that it is true for any power of A — for example

$$U^\dagger A^2U = U^\dagger AUU^\dagger AU = \left(U^\dagger AU\right)^2 \tag{1.8.6}$$

— and then it is true for all polynomials and finally for arbitrary functions.

1-19 If $|a\rangle$ is an eigenket of A, $A|a\rangle = |a\rangle a$, then $f(A)|a\rangle = |a\rangle f(a)$. Why? Conclude that $U^\dagger|a\rangle$ is an eigenket of $f(U^\dagger AU)$.

We also note that the differentiation rules in (1.7.13) imply

$$\begin{aligned}\Bigl[X,\,\mathrm{e}^{\mathrm{i}g(P)}\Bigr] &= -\hbar g'(P)\,\mathrm{e}^{\mathrm{i}g(P)}\,,\\ \Bigl[\mathrm{e}^{\mathrm{i}f(X)},P\Bigr] &= -\hbar f'(X)\,\mathrm{e}^{\mathrm{i}f(X)}\,,\end{aligned}\tag{1.8.7}$$

which are equivalent to

$$\begin{aligned}\mathrm{e}^{-\mathrm{i}g(P)}X\,\mathrm{e}^{\mathrm{i}g(P)} &= X-\hbar g'(P)\,,\\ \mathrm{e}^{\mathrm{i}f(X)}P\,\mathrm{e}^{-\mathrm{i}f(X)} &= P-\hbar f'(X)\,,\end{aligned}\tag{1.8.8}$$

with primes indicating differentiation with respect to the argument.

In a first step we write

$$\mathrm{e}^{\mathrm{i}(p'X+x'P)/\hbar}=\mathrm{e}^{\mathrm{i}p'\left(X+\frac{x'}{p'}P\right)/\hbar}\quad\text{for}\quad p'\neq 0\tag{1.8.9}$$

and note that

$$X+\frac{x'}{p'}P=\mathrm{e}^{\frac{\mathrm{i}}{2\hbar}\frac{x'}{p'}P^2}X\,\mathrm{e}^{-\frac{\mathrm{i}}{2\hbar}\frac{x'}{p'}P^2}\tag{1.8.10}$$

and therefore

$$\begin{aligned}\mathrm{e}^{\mathrm{i}(p'X+x'P)/\hbar} &= \exp\Bigl(\mathrm{i}p'\,\mathrm{e}^{\frac{\mathrm{i}}{2\hbar}\frac{x'}{p'}P^2}X\,\mathrm{e}^{-\frac{\mathrm{i}}{2\hbar}\frac{x'}{p'}P^2}/\hbar\Bigr)\\ &= \mathrm{e}^{\frac{\mathrm{i}}{2\hbar}\frac{x'}{p'}P^2}\,\mathrm{e}^{\mathrm{i}p'X/\hbar}\,\mathrm{e}^{-\frac{\mathrm{i}}{2\hbar}\frac{x'}{p'}P^2}\,.\end{aligned}\tag{1.8.11}$$

Now we remember that we want to have a factor $\mathrm{e}^{\mathrm{i}p'X/\hbar}$ on the right eventually, so we write

$$\mathrm{e}^{\mathrm{i}(p'X+x'P)/\hbar}=\mathrm{e}^{\frac{\mathrm{i}}{2\hbar}\frac{x'}{p'}P^2}\,\mathrm{e}^{\mathrm{i}p'X/\hbar}\,\mathrm{e}^{-\frac{\mathrm{i}}{2\hbar}\frac{x'}{p'}P^2}\,\mathrm{e}^{-\mathrm{i}p'X/\hbar}\,\mathrm{e}^{\mathrm{i}p'X/\hbar}\tag{1.8.12}$$

and observe that

$$\mathrm{e}^{\mathrm{i}p'X/\hbar}\,\mathrm{e}^{-\frac{\mathrm{i}}{2\hbar}\frac{x'}{p'}P^2}\,\mathrm{e}^{-\mathrm{i}p'X/\hbar}=\exp\biggl(-\frac{\mathrm{i}}{2\hbar}\frac{x'}{p'}\Bigl(\mathrm{e}^{\mathrm{i}p'X/\hbar}P\,\mathrm{e}^{-\mathrm{i}p'X/\hbar}\Bigr)^2\biggr)\tag{1.8.13}$$

wherein

$$\mathrm{e}^{\mathrm{i}p'X/\hbar}P\,\mathrm{e}^{-\mathrm{i}p'X/\hbar}=P-p'\,,\tag{1.8.14}$$

implying

$$\mathrm{e}^{\mathrm{i}(p'X+x'P)/\hbar}=\mathrm{e}^{\frac{\mathrm{i}}{2\hbar}\frac{x'}{p'}P^2}\,\mathrm{e}^{-\frac{\mathrm{i}}{2\hbar}\frac{x'}{p'}(P-p')^2}\,\mathrm{e}^{\mathrm{i}p'X/\hbar}.\tag{1.8.15}$$

The first and second exponentials on the right are functions of P only and so there is no problem in combining them into one,

$$\begin{aligned}
\mathrm{e}^{\frac{\mathrm{i}}{2\hbar}\frac{x'}{p'}P^2}\,\mathrm{e}^{-\frac{\mathrm{i}}{2\hbar}\frac{x'}{p'}(P-p')^2} &= \mathrm{e}^{\frac{\mathrm{i}}{2\hbar}\frac{x'}{p'}\left[P^2-(P^2-p')^2\right]} \\
&= \mathrm{e}^{\frac{\mathrm{i}}{2\hbar}\frac{x'}{p'}(2P-p')p'} \\
&= \mathrm{e}^{\mathrm{i}x'P/\hbar}\,\mathrm{e}^{-\frac{\mathrm{i}}{2}x'p'/\hbar}\,. \qquad (1.8.16)
\end{aligned}$$

Accordingly,

$$\begin{aligned}
\mathrm{e}^{\mathrm{i}(p'X+x'P)/\hbar} &= \mathrm{e}^{\mathrm{i}x'P/\hbar}\,\mathrm{e}^{\mathrm{i}p'X/\hbar}\,\mathrm{e}^{-\frac{\mathrm{i}}{2}x'p'/\hbar} \\
&= \mathrm{e}^{\mathrm{i}p'X/\hbar}\,\mathrm{e}^{\mathrm{i}x'P/\hbar}\,\mathrm{e}^{\frac{\mathrm{i}}{2}x'p'/\hbar} \qquad (1.8.17)
\end{aligned}$$

where the second equality is that of the right-hand sides in (1.8.1) and (1.8.2). These are examples of the famous *Baker–Campbell–Hausdorff relations* among exponential functions of operators, named after Henry F. Baker, John E. Campbell, and Felix Hausdorff.

1-20 What is $\langle x|\,\mathrm{e}^{\mathrm{i}(p'X+x'P)/\hbar}$? What is $\mathrm{e}^{\mathrm{i}(p'X+x'P)/\hbar}|p\rangle$? Introduce the operator

$$R = \int \frac{\mathrm{d}x'\,\mathrm{d}p'}{2\pi\hbar}\,\mathrm{e}^{\mathrm{i}(p'X+x'P)/\hbar}$$

and find $\langle x|R$ and $R|p\rangle$.

1-21 Show that

$$\operatorname{tr}\left\{\mathrm{e}^{\mathrm{i}(p'X+x'P)/\hbar}\,\mathrm{e}^{-\mathrm{i}(p''X+x''P)/\hbar}\right\} = 2\pi\hbar\,\delta(x'-x'')\,\delta(p'-p'')\,.$$

1-22 For any (reasonable) operator function $A(X,P)$ we can define its *characteristic function*

$$a(x,p) = \operatorname{tr}\left\{A(X,P)\,\mathrm{e}^{\mathrm{i}(xP+pX)/\hbar}\right\}\,.$$

This is a mapping of operator $A(X,P)$ on the phase-space function $a(x,p)$. [The present $a(x,p)$ is *not* the same as the one in (1.7.4).] Show that the inverse map is given by

$$A(X,P) = \int \frac{\mathrm{d}x\,\mathrm{d}p}{2\pi\hbar}\,\mathrm{e}^{\mathrm{i}(xP+pX)/\hbar}a(x,p)\,.$$

Hint: It is enough to show this for all ket-bras of the form $|x'\rangle\langle p'|$. Why?

1-23 Find $a(x,p)$ for $A(X,P)=X^n$, $n=0,1,2,\ldots$.