# ... 2 ...

# The Mathematical Formalism and the Standard Way of Thinking about It

There is an algorithm (and the name of that algorithm, of course, is quantum mechanics) for predicting the behaviors of physical systems, which correctly predicts all of the unfathomable-looking behaviors of the electron in the story in Chapter 1, and there is a standard way of interpreting that algorithm (that is, a way attempting to fathom those behaviors, a way of attempting to confront the fact of superposition) which can more or less be traced back to some sayings of Niels Bohr.[1] This chapter will describe that algorithm and rehearse that standard way of talking about it, and then it will apply them both, in some detail, to that story.

## Mathematical Preliminaries

Let me say a few things, to begin with, about the particular mathematical language in which it is most convenient to write the algorithm down.

Let's start with something about vectors. A good way to think about vectors is to think about arrows. A vector is a mathematical object, an abstract object, which (like an arrow) is characterized by

1. The story of the evolution of this standard way of thinking is a very long and complicated one, and it will be completely ignored here. The far more obscure question of what Bohr himself really thought about these issues will be ignored too. What will matter for us is the legacy which Bohr and his followers have left, by whatever route, and whatever they themselves may have originally thought, to modern physics. That legacy, as it stands now, can be characterized fairly clearly. The name of that legacy is the Copenhagen interpretation of quantum mechanics.

a direction (the direction in which the arrow is pointing) and a magnitude (the length of the arrow).

Think of a coordinate system with a specified origin point. Every distinct geometrical point in the space mapped out by such a coordinate system can be associated with some particular (and distinct) vector, as follows: the vector associated with any given point (in that given coordinate system) is the one whose tip lies at the given point and whose tail lies at the origin. The length of that vector is the distance between those two points, and the direction of that vector is the direction from the origin to the given point (see figure 2.1).

The infinite collection of vectors associated with *all* the points in such a space is referred to as a *vector space.*

Spaces of points can be characterized by (among other things) their dimensionality, and spaces of vectors can too. The dimension of a given vector space is just the dimension of the associated space of points. That latter dimension, of course, is equal to the number of magnitudes, the number of coordinates, that need to be specified in order to pick out (given a coordinate system) some particular geometrical point.

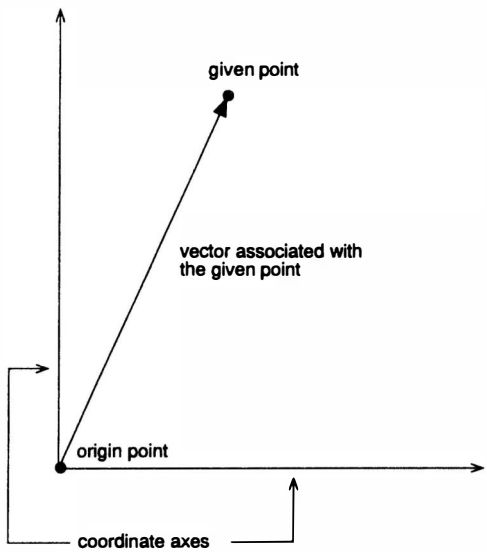



*Figure 2.1*

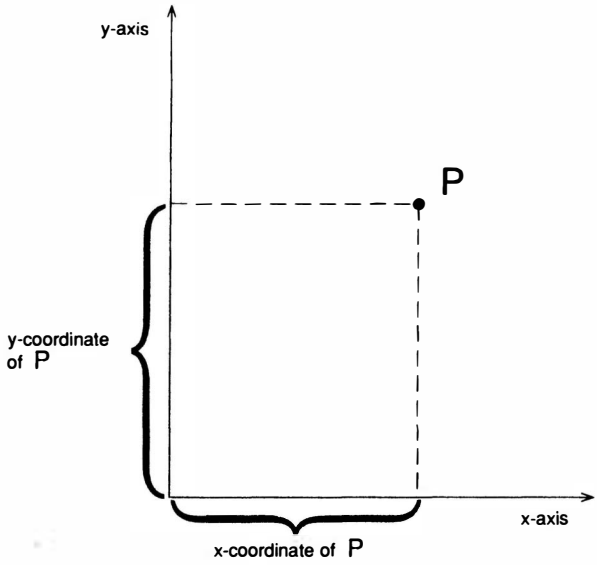



*Figure 2.2*

Figure 2.2, for example, shows a two-dimensional space, a *plane* of points, wherein (given the indicated coordinate system) two coordinates need to be specified (the $x$-coordinate and the $y$-coordinate) in order to pick out a point. The reader can convince herself that a line of points forms a one-dimensional space, and that the space we move around in has three dimensions. Spaces of points with more dimensions than that are hard to visualize, but the formal handling (that is: the mathematical handling) of such spaces is not a problem.

Let's introduce a notation for vectors: the symbols $|\ \rangle$ around some expression will henceforth indicate that that expression is the name of the vector; so that, for example, $|A\rangle$ will denote the vector called $A$. That's the notation most commonly used in the literature of quantum mechanics.

Vectors can be added to one another. Here's how: To add $|A\rangle$ to $|B\rangle$, move the tail of $|B\rangle$ to the tip of $|A\rangle$ (without altering the length or the direction of either in the process). The sum of $|A\rangle$ and $|B\rangle$ (which is written $|A\rangle + |B\rangle$) is defined to be that vector ($|C\rangle$) whose tail now coincides with the tail of $|A\rangle$ and whose tip now coincides

with the tip of $|B\rangle$ (see figure 2.3). The sum of any two vectors in any particular vector space is always another vector in that same space (that, indeed, is part of the definition of a vector space). Think, for example, of the spaces discussed above.

That fact is going to be important. Vectors, in quantum mechanics, are going to represent physical states of affairs. The *addition* of vectors will turn out to have something to do with the *superposition* of physical states of affairs. The fact that two vectors can be added together to form a third will turn out to accommodate, within the algorithm, the fact that certain physical states of affairs, states like being white, are superpositions of certain other states of affairs, states like being hard and being soft; but of all this more later.

Vectors can be multiplied too. There are two ways to multiply them. First of all, they can be multiplied by numbers. The vector $5|A\rangle$, say, is defined to be that vector whose direction is the same as the direction of $|A\rangle$ and whose length is 5 times the length of $|A\rangle$. $5|A\rangle = |A\rangle + |A\rangle + |A\rangle + |A\rangle + |A\rangle$. Of course, if $|A\rangle$ is an element of a certain vector space, any number *times* $|A\rangle$ will be an element of that space too.

The other way to multiply vectors is to multiply them by other vectors. The multiplication of a vector by another vector yields a number (not a vector!). $|A\rangle$ times $|B\rangle$ (which is written $\langle A|B\rangle$) is defined to be the following number: the length of $|A\rangle$ times the length of $|B\rangle$ times the cosine of the angle, $\theta$, between $|A\rangle$ and $|B\rangle$.

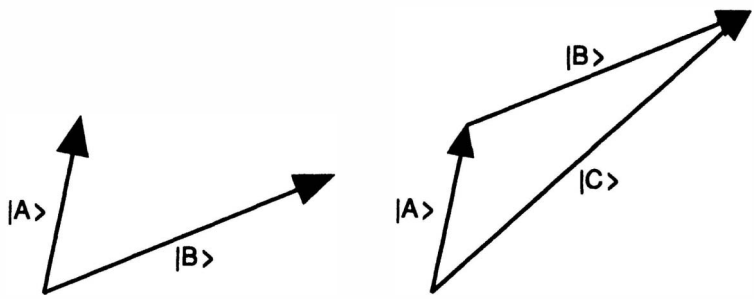



*Figure 2.3*

The length of $|A\rangle$ (also called the *norm* of $|A\rangle$, which is written $|A|$) is obviously equal to the square root of the number $\langle A|A\rangle$, since the cosine of 0° (0° is the angle between $|A\rangle$ and itself) is equal to 1.

So, vectors plus vectors are vectors, and vectors times numbers are vectors, and vectors times vectors are numbers.

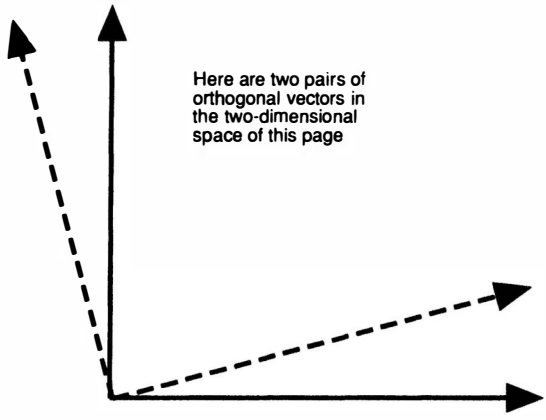
Here's a slightly more sophisticated way of defining a vector space: a vector space is a collection of vectors such that the sum of any two vectors in the collection is also a vector in the collection, and such that any vector in the collection times any (real) number is also a vector in the collection. Such collections (by the way) clearly have to be infinite. Think, again, of the examples of spaces described above.

If $|A| \neq 0$ and $|B| \neq 0$ and yet $\langle A|B\rangle = 0$ (that is: if the angle between $|A\rangle$ and $|B\rangle$ is 90°, since cos 90° = 0), then $|A\rangle$ and $|B\rangle$ are said to be orthogonal to one another. *Orthogonal* just means perpendicular.

Here's another definition of dimension: The dimension of a vector space is equal (by definition) to the maximum number (call that number $N$) of vectors $|A_1\rangle$, $|A_2\rangle$, . . . $|A_N\rangle$ which can be chosen in the space such that for all values of $i$ and $j$ from 1 through $N$ such that $i \neq j$, $\langle A_i|A_j\rangle = 0$. That is, the dimension of a space is equal to the number of mutually perpendicular directions in which vectors within that space can point.

Given a space, there are generally lots of ways to pick out those directions. Pick a vector, at random, from an $N$-dimensional space. It will always be possible to find a set of $N - 1$ other vectors in that space which are all orthogonal to that original vector and to one another. In most cases, given that original vector, there will still be many such orthogonal sets (or, rather, an infinity of such sets) to choose from. Figure 2.4 shows some examples.

Think of an $N$-dimensional space. Think of any collection of $N$ mutually orthogonal vectors in that space, and suppose that the norm, the length, of each of those vectors happens to be 1. Such a set of vectors is said to form an orthonormal *basis* of that $N$-dimensional space. *Ortho* is for orthogonal, *normal* is for norm-1, and here's why sets of vectors like that are called *bases* of their spaces: Suppose that the set $|A_1\rangle$, $|A_2\rangle$, . . . $|A_N\rangle$ forms a basis of a




All four of these vectors, moreover, are orthogonal to any vector pointing directly out of the page

*Figure 2.4*

certain $N$-dimensional vector space; it turns out that any vector whatever in that space (call it $|B\rangle$) can be expressed as the following sort of sum:

$$|B\rangle = b_1|A_1\rangle + b_2|A_2\rangle + \ldots b_N|A_N\rangle \tag{2.1}$$

where the $b_i$ are all simply numbers—more particularly, simply the following numbers:

$$b_i = \langle B|A_i\rangle \tag{2.2}$$

So any vector in a vector space can be "built up" (as in (2.1)) out of the elements of any basis of that space. All that is illustrated, for a two-dimensional space, in figure 2.5.

Bases end up amounting to precisely the same thing as coordinate systems: given a coordinate system for an $N$-dimensional point space, $N$ numbers (the coordinate values) will suffice to pick out a point; given a basis of an $N$-dimensional vector space, $N$ numbers (the $b_i$ of equation (2.1)) will suffice to pick out a vector. Vectors which are of norm 1 and which point along the perpendicular coordinate axes of an $N$-dimensional point space will constitute an

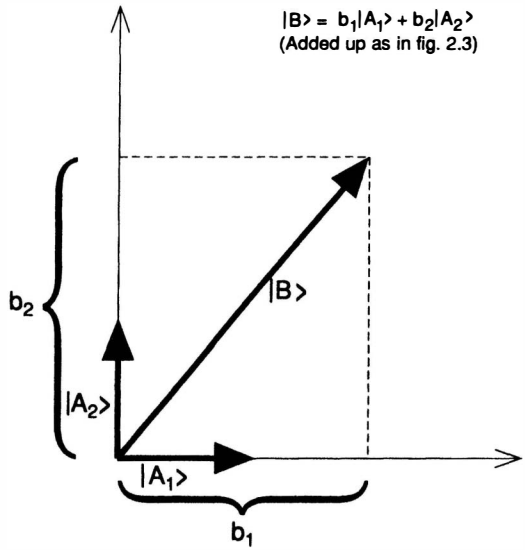



*Figure 2.5*

orthonormal basis of the associated $N$-dimensional vector space, and vice versa.

For any space of more than a single dimension, there will be an infinity of equivalently good orthornormal bases to choose from. Any vector in that space will be writable, a la (2.1), in terms of any of those bases, but of course, for a given vector $|B\rangle$, the numbers $b_i$ in (2.1) (which, by the way are called expansion coefficients) will differ from basis to basis. Figure 2.6 shows how that works.

Now, it happens to be the case that for any three vectors $|A\rangle$, $|B\rangle$, and $|C\rangle$, the product $|A\rangle$ times the vector $(|B\rangle + |C\rangle)$ is equal to the product $|A\rangle$ times $|B\rangle$ plus the product $|A\rangle$ times $|C\rangle$:

$$\langle A\|B\rangle + |C\rangle\rangle = \langle A|B\rangle + \langle A|C\rangle, \tag{2.3}$$

and that can be shown to entail, for any two vectors $|M\rangle$ and $|Q\rangle$, that

$$|M\rangle + |Q\rangle = (m_1 + q_1)|A_1\rangle + (m_2 + q_2)|A_2\rangle + \ldots + (m_N + q_N)|A_N\rangle \tag{2.4a}$$

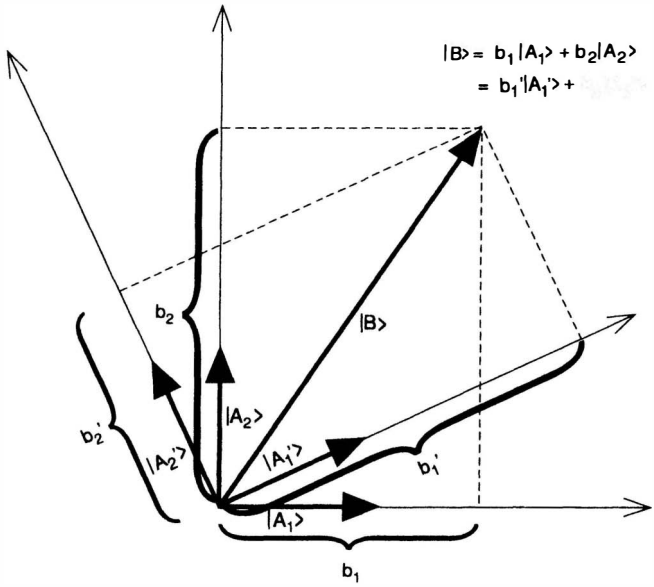



*Figure 2.6*

and that

(2.4b) $\langle M|Q\rangle = m_1q_1 + m_2q_2 + \ldots + m_Nq_N$

wherein the $m_i$ and $q_i$ are the expansion coefficients of $|M\rangle$ and $|Q\rangle$, respectively, in any particular basis $|A_i\rangle$. The numbers $q_i$ and $m_i$ will, of course, depend on the choice of basis, but note that the sum of their products in (2.4b) (which is equal to $\langle M|Q\rangle$, which depends only on which vectors $|M\rangle$ and $|Q\rangle$ happen to *be,* and not on which basis we happen to map them out in) will not. That sum, rather, will be *invariant* under changes of basis.

Suppose that we have agreed to settle on some particular basis for some particular vector space. Once that's done, all that will be required for us to pick out some particular vector ($|Q\rangle$, say) will be to specify the numbers (the expansion coefficients) $q_i$ of $|Q\rangle$ for that particular basis. Those $N$ numbers then (once the basis is chosen)

can serve to represent the vector. Those numbers are usually written down in a column; for example:

(2.5)
$$|Q\rangle = \begin{bmatrix} 1 \\ 5 \\ -3/2 \end{bmatrix}$$
$$= \text{the three–dimensional vector for which} \begin{cases} \langle Q|A_1\rangle = 1 \\ \langle Q|A_2\rangle = 5 \\ \langle Q|A_3\rangle = -3/2 \end{cases}$$

(see equation (2.2)), where the $|A_i\rangle$ are the chosen basis vectors. It follows from (2.4b) that the norm (the length) of any vector $|Q\rangle$ will be equal to the square root of the sum of the squares of its expansion coefficients. That number, too, must obviously be invariant under changes of basis.

That's all that will concern us about vectors. The other sorts of mathematical objects which we shall need to know something about are *operators*.

Operators are mechanisms for making new vectors out of old ones. An operator on a vector space, more particularly, is some definite prescription for taking every vector in that space into some other vector; it is a mapping (for those readers who know the mathematical meaning of that word) of a vector space into itself.

Let's introduce a notation. Suppose that the operator called $O$ is applied to the vector $|B\rangle$ (that is: suppose that the prescription called $O$ is carried out on the vector $|B\rangle$). The result of that operation, of that procedure, is written:

(2.6) $O|B\rangle$

Then what was just said about operators can be expressed like this:

(2.7) $O|B\rangle = |B'\rangle$ for any vector $|B\rangle$ in the vector space on which $O$ is an operator.

where $|B'\rangle$ is some vector in the same space as $|B\rangle$.

Here are some examples. One example is the "unit" operator

(that's the prescription which instructs us to multiply every vector in the space by the number 1, to transform every vector into itself). The unit operator is the one for which

(2.8) $$O_U|B\rangle = |B\rangle = |B'\rangle$$

Another example is the operator "multiply every vector by the number 5." Another example is the operator "rotate every vector clockwise by 90° about some particular vector $|C\rangle$" (see figure 2.7). Another example is the operator "map every vector in the space into some particular vector $|A\rangle$."

The particular sorts of operators which will play a vital role in the quantum-mechanical algorithm are *linear operators*. Linear operators are, by definition, operators which have the following properties:

(2.9a) $$O(|A\rangle + |B\rangle) = O|A\rangle + O|B\rangle$$

and

(2.9b) $$O(c|A\rangle) = c(O|A\rangle)$$

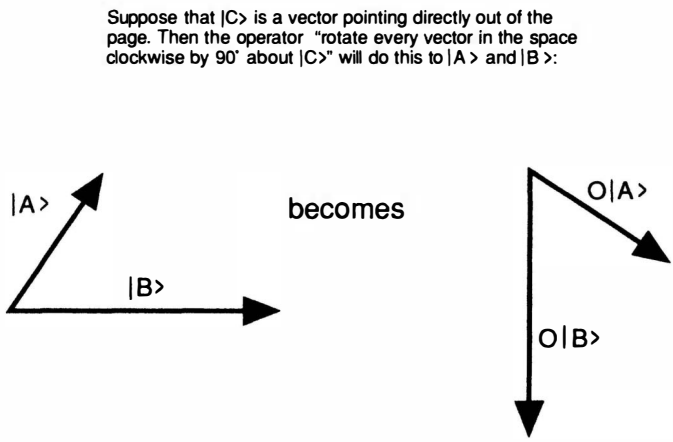



*Figure 2.7*

for any vectors $|A\rangle$ and $|B\rangle$ and any number $c$. Here's what (2.9a) says: take that vector which is the sum of two other vectors $|A\rangle$ and $|B\rangle$ (such sums, remember, are always vectors), and operate on that sum with any linear operator. The resultant (new) vector will be that vector which is the sum of the new vector produced by operating on $|A\rangle$ with $O$ and the new vector produced by operating on $|B\rangle$ with $O$. What (2.9b) says is that the vector produced by operating on $c$ times $|A\rangle$ with $O$ is the same as $c$ times the vector produced by operating on $|A\rangle$ itself with $O$, for any number $c$.[2]

Now, the two conditions in (2.9) pick out a very particular sort of operator. They are by no means properties of operators in general. Let me leave it as an exercise for the reader to show, for example, that of the four operators just now described, the first three are linear and the last one isn't.

Linear operators are very conveniently representable by arrays of numbers. We learned it was possible, remember, to represent any $N$-dimensional *vector*, given a choice of basis, by $N$ *numbers* (a la equation (2.5)); and it similarly turns out to be possible to represent any linear operator (the linearity is crucial here) on an $N$-dimensional vector space by $N^2$ numbers. Those $N^2$ numbers are traditionally arranged not in a column (as in equation (2.5), for vectors), but in a *matrix*, as (for a two-dimensional operator, say) follows:

$$O = \begin{bmatrix} O_{11} & O_{12} \\ O_{21} & O_{22} \end{bmatrix} \tag{2.10}$$

The numbers $O_{ij}$ in (2.10) are defined to be

$$O_{ij} = \langle A_i | O | A_j \rangle \tag{2.11}$$

That is: the number $O_{ij}$ is the vector $O|A_j\rangle$ multiplied by the vector $|A_i\rangle$ (such products of vectors, remember, are always *numbers*),

2. The two parts of (2.9) aren't completely independent of one another, by the way. Note, for example, that in the event that $c$ is an integer, (2.9b) is entailed by (2.9a).

where the $|A_N\rangle$ are the chosen basis vectors of the space. There's a rule for multiplying operator matrices by vector columns, which is:

$$\begin{bmatrix} O_{11} & O_{12} \\ O_{21} & O_{22} \end{bmatrix} \times \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} (O_{11}b_1 + O_{12}b_2) \\ (O_{21}b_1 + O_{22}b_2) \end{bmatrix} \tag{2.12}$$

Note that the right-hand side of (2.12) is a vector column; so this rule stipulates that the product of an operator matrix and a vector column is a new vector column.

Here's why all this notation is useful: it turns out (we won't prove it here) that any linear operator whatever can be uniquely specified (given a basis choice) by specifying the $N^2$ $O_{ij}$ of equations (2.10) and (2.11) (just as any *vector* can be uniquely specified by specifying the $N$ $b_i$ of equations (2.1) and (2.2) and (2.5)); and it turns out that for any linear operator $O$, we can calculate $O$'s effect on any vector $|B\rangle$ simply by multiplying the $O$-matrix by the $|B\rangle$-column (given, as always, a basis choice) as in (2.12). That is, for any linear operator $O$ and any vector $|B\rangle$:

$$\begin{aligned} |B'\rangle &= \begin{bmatrix} O_{11} & O_{12} \\ O_{21} & O_{22} \end{bmatrix} \times \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} (O_{11}b_1 + O_{12}b_2) \\ (O_{21}b_1 + O_{22}b_2) \end{bmatrix} \\ &= \underbrace{(O_{11}b_1 + O_{12}b_2)}_{\text{these are } \textit{numbers}}|A_1\rangle + \underbrace{(O_{21}b_1 + O_{22}b_2)}_{\text{these are } \textit{numbers}}|A_2\rangle = |B'\rangle \end{aligned} \tag{2.13}$$

where $|A_i\rangle$ are the chosen basis vectors.[3] (The next-to-last equality follows from equations (2.1) and (2.2) and (2.5)).

3. Perhaps it's worth saying all that out in words: In order to calculate the effect of any linear operator $O$ on any vector $|B\rangle$, first choose a basis, then calculate the $|B\rangle$ column vector in that basis by means of formula (2.2); then calculate the operator matrix in that basis by means of formula (2.11); then multiply that column vector by that operator matrix by means of formula (2.12); and the result of that multiplication will be the column vector, in that same basis, of the new vector $|B'\rangle$ (that is, the vector obtained by operating with $O$ on $|B\rangle$).

One more definition will be useful. If it happens to be the case for some particular operator $O$ and some particular vector $|B\rangle$ that

$$O|B\rangle = @|B\rangle \tag{2.14}$$

where @ is some number—that is, if the new vector generated by operating on $|B\rangle$ with $O$ happens to be a vector pointing in the same direction as $|B\rangle$—then $|B\rangle$ is said to be an *eigenvector* of $O$, with *eigenvalue* @ (where @ is the length of that new vector relative to the length of $|B\rangle$).

Certain vectors will in general be eigenvectors of some operators and not of some others; certain operators will in general have some vectors, and not others, as eigenvectors, and other operators will have *other* vectors as eigenvectors. The operator-eigenvector relation, however, depends only on the vector and the operator in question, and not at all on the basis in which we choose to write those objects down. In other words, if the eigenvector-operator relation obtains between the vector column and the operator matrix of a certain vector and a certain operator in a certain particular basis, then it can be shown that the same relation, with the same eigenvalue, will obtain between the vector column and the operator matrix in any basis whatever of that space.

Here are some examples: all vectors are eigenvectors of the unit operator, and all have eigenvalue 1; and similarly (but with eigenvalue 5) for the operator "multiply every vector by 5." All vectors of the form $@|C\rangle$, where @ is any number, are eigenvectors of the operator "rotate every vector about $|C\rangle$ by 90°"; all those vectors have eigenvalue 1, and there are no other eigenvectors of that operator. The four-dimensional space operator (written down in some particular basis)

$$O = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 3/2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -7 \end{bmatrix} \tag{2.15}$$

has eigenvectors (written in that same basis)

(2.16) $$|A\rangle = \begin{bmatrix}1\\0\\0\\0\end{bmatrix} \quad |B\rangle = \begin{bmatrix}0\\1\\0\\0\end{bmatrix} \quad |C\rangle = \begin{bmatrix}0\\0\\1\\0\end{bmatrix} \quad |D\rangle = \begin{bmatrix}0\\0\\0\\1\end{bmatrix}$$

with eigenvalues 5, 3/2, 2, and −7, respectively. Any number *times* $|A\rangle$ or $|B\rangle$ or $|C\rangle$ or $|D\rangle$ will be an eigenvector of $O$ too, with the same eigenvalue; but vectors like $|A\rangle + |B\rangle$ won't be eigenvectors of $O$.

## Quantum Mechanics

Now we're in a position to write out the algorithm. It pretty much all boils down to five principles.

*(A) Physical States.* Physical situations, physical states of affairs, are represented in this algorithm by vectors. They're called *state vectors.* Here's how that works: Every physical system (that is: every physical object, and every collection of such objects), to begin with is associated in the algorithm with some particular vector space; and the various possible physical states of any such system are stipulated by or correspond to vectors, and more particularly to vectors of length 1, in that system's associated space; and every such vector is taken to pick out some particular such state; and the states picked out by all those vectors are taken to comprise all of the possible physical situations of that system (the correspondence isn't precisely one-to-one, however: we shall soon discover, for example, that for any vector $|A\rangle$ of length 1, $-|A\rangle$ must necessarily pick out the same physical state as $|A\rangle$ does).

This will turn out to be a very apt way to represent states, since (as I mentioned before) the possibility of "superposing" two states to form another gets reflected in the algorithm by the possibility of adding (or subtracting) two vectors to form another.

*(B) Measurable Properties.* Measurable properties of physical systems (such properties are referred to as *observables,* in the quan-

tum-mechanical literature) are represented in the algorithm by linear operators on the vector spaces associated with those systems. There's a rule that connects those operators (and their properties) and those vectors (and their physical states), which runs as follows: If the vector associated with some particular physical state happens to be an eigenvector, with eigenvalue (say) $a$, of an operator associated with some particular measurable property of the system in question (in such circumstances, the state is said to be an "eigenstate" of the property in question), then that state has the value $a$ of that particular measurable property.

Let's try all that out. Let's construct a vector space in which the state of being hard and the state of being soft can be represented. Suppose we let the following two two-dimensional column vectors stand for hardness and softness:

$$|\text{hard}\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad |\text{soft}\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{2.17}$$

Notice that if we adopt (2.17), $\langle\text{hard}|\text{soft}\rangle = 0$ (see equations (2.4) and (2.5)). As a matter of fact, the two vectors in (2.17) constitute a basis of the two-dimensional space which they inhabit. That particular basis, by the way, is precisely the one in which the vector columns in (2.17) have been written down (that is: the relevant basis vectors $|A_1\rangle$ and $|A_2\rangle$ of equation (2.5) are, in the case of (2.17), precisely $|\text{hard}\rangle$ and $|\text{soft}\rangle$).

What operator should represent the hardness property? Let's try this:

$$\text{hardness operator} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{2.18}$$

where we stipulate that "hardness = +1" means "hard"
and that "hardness = −1" means "soft"

So far all this works out right: $|\text{hard}\rangle$ and $|\text{soft}\rangle$ of equation (2.17) are, indeed, as the reader can now easily confirm, eigenvectors of

the hardness operator of equation (2.18), with the appropriate eigenvalues.

Let's push this example further. Remember that it seemed to us in Chapter 1 that the "black" and "white" states must both be superpositions of both of the "hard" and "soft" states; and remember (from the present chapter) that the superposition of physical states is supposed to correspond somehow to the addition or subtraction of their respective state vectors; and remember that the sum or the difference of any two vectors in any particular vector space is necessarily yet another vector in that same space. All that suggests that the states of being white and being black ought to be representable by vectors in this space too, and that there ought to be a color operator on this space. Let's try this one, written down in the basis of equation (2.17):

(2.19)
$$|\text{black}\rangle = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \qquad |\text{white}\rangle = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

$$\text{color operator} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \begin{array}{l} \text{"color} = +1\text{" means "black"} \\ \text{"color} = -1\text{" means "white"} \end{array}$$

That works out right too: The reader can show that the various stipulations of (2.19) are all consistent with one another, as are the stipulations of (2.17) and (2.18). Furthermore $\langle\text{black}|\text{white}\rangle = 0$ too; and $|\text{black}\rangle$ and $|\text{white}\rangle$ constitute another basis of this space.

Now, it follows from (2.4a) that:

(2.20)
$$\text{if} \quad |A\rangle = \begin{bmatrix} a \\ b \end{bmatrix} \quad \text{and} \quad |B\rangle = \begin{bmatrix} c \\ d \end{bmatrix}$$

in some particular basis, then

$$|A\rangle + |B\rangle = \begin{bmatrix} (a + c) \\ (b + d) \end{bmatrix}$$

in that same basis (and the same applies, of course, to vector columns of any dimension).

Notice, then, how beautifully (2.17) and (2.18) and (2.19) reflect the principles of superposition and incompatibility. First of all, it follows from (2.17) and (2.19) that:

(2.21) $|\text{black}\rangle = 1/\sqrt{2}|\text{hard}\rangle + 1/\sqrt{2}|\text{soft}\rangle$

$|\text{white}\rangle = 1/\sqrt{2}|\text{hard}\rangle - 1/\sqrt{2}|\text{soft}\rangle$

$|\text{hard}\rangle = 1/\sqrt{2}|\text{black}\rangle + 1/\sqrt{2}|\text{white}\rangle$

$|\text{soft}\rangle = 1/\sqrt{2}|\text{black}\rangle - 1/\sqrt{2}|\text{white}\rangle$

So sums and differences of vectors, in the algorithm, *do* denote *superpositions* of physical states; and (just as we concluded in the last chapter) states of definite color *are* superpositions of different hardness states, and states of definite hardness *are* superpositions of different color states.

Moreover, look how well the forms of the hardness and color operators confirm all this: It's easy to verify that the "black" and "white" vectors aren't eigenvectors of the hardness operator, and that the "hard" and "soft" vectors aren't eigenvectors of the color operator. The hardness and color operators are (just as they ought to be) incompatible with one another, in the sense that states of definite hardness (that is: states whose vectors are eigenvectors of the hardness operator) apparently have no assignable color value (since those vectors aren't eigenvectors of the color operator) and vice versa.

So it turns out that the descriptions of color and of hardness and of all the relations between them can be subsumed within a single, two-dimensional vector space. That space is referred to within the quantum-mechanical literature as the *spin* space, and color and hardness are referred to as *spin* properties.

Let's get back to the enumeration of the five principles.

*(C) Dynamics.* Given the state of any physical system at any "initial" time (given, that is, the vector which represents the state of that system at that time), and given the forces and constraints to

which that system is subject, there is a prescription whereby the state of that system at any *later* time (that is, the vector at any later time) can, in principle, be calculated. There is, in other words, a *dynamics* of the state vector; there are deterministic laws about how the state vector of any given system, subject to given forces and constraints, changes with time. Those laws are generally cast in the form of an equation of motion, and the name of that equation, for nonrelativistic systems, is the *Schrödinger* equation.

Since every state vector must, by definition, be a vector of length one, the changes in state vectors dictated by the dynamical laws are exclusively changes of direction, and never of length.

Here's an important property of the quantum-mechanical dynamical laws: Suppose that a certain system, subject to certain specified forces and constraints and whose state vector at time $t_1$ is $|A\rangle$, evolves, in accordance with the laws, into the state $|A'\rangle$ at time $t_2$; and suppose that that same system, subject to those same forces and constraints, if its state vector at $t_1$ is, rather, $|B\rangle$, evolves, in accordance with those laws, into the state $|B'\rangle$ at time $t_2$. Then, the laws dictate that if that same system, subject to those same forces and constraints, were, rather, in the state $\alpha|A\rangle + \beta|B\rangle$ at time $t_1$, then its state at time $t_2$ will be $\alpha|A'\rangle + \beta|B'\rangle$ (where $|A\rangle$ and $|B\rangle$ can be any state vectors at all). This property of the laws will concern us a good deal later on. The name of this property is *linearity* (and note that there is indeed a resemblance between "linearity" as applied to dynamical laws, here, and "linearity" as applied to operators, as in the two equations in (2.9)).

*(D) The Connection with Experiment.* So far, almost nothing in these principles has touched upon the results of measurements. All we have is a stipulation in (B) that the physical state whose state vector is an eigenvector, with eigenvalue $a$, of the operator associated with some particular measurable property will have the value $a$ for that property; and presumably it follows that a measurement of that property, carried out on a system which happens to *be* in that state, will produce the result $a$. But much more needs to be said about the results of measurements than that! What if we measure a certain property of a certain physical system at a moment

when (as must happen in the vast majority of cases) the state vector of that system does not happen to be an eigenvector of that property operator? What if, say, we measure the color of a hard electron, an electron in a superposition of being white and being black? What happens then? Principle (B) is of no help here. A new principle shall have to be introduced to settle the question, which runs as follows:

Suppose we have before us a system whose state vector is $|a\rangle$, and we carry out a measurement of the value of property $B$ on that system, where the eigenvectors of the property operator for $B$ are $|B = b_i\rangle$, with eigenvalues $b_i$ (i.e., $B|B = b_i\rangle = b_i|B = b_i\rangle$ for all $i$). According to quantum mechanics, the outcome of such a measurement is a matter of *probability;* and (more particularly) quantum mechanics stipulates that the probability that the outcome of this measurement will be $B = b_i$ is equal to:

$$(2.22) \qquad (\langle a|B = b_i\rangle)^2$$

Note that (as must be the case for probability) the number denoted by the above formula will always be less than or equal to 1; and note that in the special case of eigenvectors covered by principle (B), (2.22) yields (as it should) the probability 1. And note that it follows from (2.17) and (2.19) and (2.22) that the probability that a black electron will be found by a hardness measurement to be, say, soft, is (precisely as we have learned to expect) ½.

And this is where it emerges that the correspondence between states and vectors of length 1 isn't precisely one-to-one. First of all, it follows from equation (2.3) that (for any vectors $|a\rangle$ and $|b\rangle$ and any number @) $\langle a|@|b\rangle = @\langle a|b\rangle$. Now, since the probability (2.22) depends only on the *square* of the product of the vectors involved, and since $(1x)^2 = (-1x)^2$, it follows that the probability of any result of any measurement carried out on a system in the state $|a\rangle$ will be identical to the probability of that same result of that same measurement carried out on a system in the state $-|a\rangle$. Vectors $|a\rangle$ and $-|a\rangle$, then, have precisely the same observable consequences; which is to say (as is customary in the quantum-mechanical litera-

ture) that the vectors $|a\rangle$ and $-|a\rangle$ represent precisely the same physical state.

*(E) Collapse.* Measurements (as I remarked in Chapter 1) are always, in principle, repeatable. Once a measurement is carried out and a result is obtained, the state of the measured system must be such as to guarantee that if that measurement is repeated, the same result will be obtained.[4]

Consider what that entails about the state vector of the measured system. Something happens to that state vector when the measurement occurs. If, say, a measurement of an observable called $O$ is carried out on a system called $S$, and if the outcome of that measurement is $O = @$, then, whatever the state vector of $S$ was just prior to the measurement of $O$, *the state vector of $S$ just after that measurement must necessarily be an eigenvector of $O$* with eigenvalue @. The effect of measuring an observable must necessarily be to *change* the state vector of the measured system, to "collapse" it, to make it "jump" from whatever it may have been just prior to the measurement into some eigenvector of the measured observable operator. Which particular such eigenvector it gets changed into is of course determined by the outcome of the measurement; and note that that outcome, in accordance with principle (D), is a matter of probability. It's at this point, then, and at no point other than this one, that an element of pure chance enters into the evolution of the state vector.

Those are the principles of quantum mechanics. They are the most precise mechanism for predicting the outcomes of experiments on physical systems ever devised. No exceptions to them have ever been discovered. Nobody expects any.

Suppose that we should like to predict the behavior of some particular physical system by means of this algorithm. How, exactly,

4. Supposing, of course, that there has been no "tampering" in the interim; and supposing that not enough time has elapsed for the natural dynamics of the measured system itself to bring about changes in the value of the measured observable.

do we go about that? The first thing to do is to identify the vector space associated with that system: the space wherein all the possible physical states of that system can be represented. Given a precise physical description of the system, there are systematic techniques for doing that. Then the operators associated with the various measurable properties of that system need to be identified. There are techniques for doing that too. With that done, the specific correspondences between individual physical states and individual vectors can be mapped out (the vector which corresponds to the state wherein a certain measurable property has a certain value, for example, will be the one which is an eigenvector, with that eigenvalue, of the operator associated with that property). Then the present state vector of the system can be ascertained by means of measurements, and then (given the various forces and constraints to which the system will be subject) the state vector of any future time can be calculated by means of the prescription of principle (C), and then the probabilities of particular outcomes of a measurement carried out at some such future time can be calculated by means of principle (D), and the *effect* of such a measurement on the state vector can be taken into account by means of principle (E). And then principle (C) can be applied yet again, to that *new* state vector (the state vector which emerges from the measurement) to calculate the state vector of this system yet farther in the future, up to the moment when the next measurement occurs, whereupon principles (D) and (E) can be reapplied, and so on.

Notice, by the way, that principle (E) stipulates that under certain particular circumstances (namely, when a measurement occurs) the state vector evolves in a certain particular way (it "collapses" onto an eigenvector of the measured observable operator). Notice, too, that principle (C) is supposed to be a completely *general* account of how the state vector evolves under *any* circumstances. If that's all so, a question of consistency necessarily arises: it seems like (E) ought to be just a special case of (C), that (E) ought to be deducible from (C). But it isn't easy to see how that could be so, since the changes in the state vector stipulated by (E) are probabilistic, whereas those stipulated by (C) are, invariably, deterministic. This

is going to require some worrying about, but let's not start that just yet; that worrying will commence in earnest in Chapter 4.

As I mentioned before, there is a standard way of talking, which students of physics are traditionally required to master along with this algorithm, about what superpositions are. That line, that way of dealing with the apparent contradiction of Chapter 1, boils down to this: the right way to think about superpositions of, say, being black and being white is to think of them as situations wherein color predicates cannot be applied, situations wherein color talk is unintelligible. Talking and inquiring about the color of an electron in such circumstances is (on this view) like talking or inquiring about, say, whether or not the number 5 is still a bachelor. On this view, then, the contradictions of Chapter 1 go away. On this view, it just isn't so that hard electrons are not black and not white and not both and not neither, since color talk of any kind, about hard electrons, simply has no meaning at all. And that's the way things are, on this view, for all sorts of superposition: superpositions are situations wherein the superposed predicates just don't apply.

Of course, once an electron has been *measured* to be white or black, then it *is* white or black (then, in other words, color predicates surely do apply). Measuring the color of a hard electron, then, isn't a matter of ascertaining what the color of that hard electron is; rather, it is a matter of first changing the state of the measured electron into one to which the color predicate applies, and to which the hardness predicate cannot apply (this is the "collapse" of principle (E)), and *then* of ascertaining the color of that newly created, color-applicable state. Measurements in quantum mechanics (and particularly within this interpretation of quantum mechanics) are very active processes. They aren't processes of merely learning something; they are invariably processes which drastically change the measured system.

That's what's at the heart of the standard view. The rest (of which I shall have much more to say later on) is details.

* * *

Here (before we move on to particular cases) are a few more general technicalities.

First, the vector spaces which are made use of in quantum mechanics are *complex* vector spaces. A complex vector space is one in which it's permissible to multiply vectors not merely by real numbers but by *complex* (i.e., real or imaginary or both) numbers in order to produce new vectors. In complex vector spaces, the expansion coefficients of vectors in given bases (the $b_i$ of equation (2.1)) may be complex numbers too. That will necessitate a few refinements of what's been introduced thus far.

In complex vector spaces, the formula for the product of two vectors, written in terms of their expansion coefficients in some particular basis (that is, the formula (2.4b)), needs to be changed, very slightly (what, precisely, it gets changed into need not concern us here), in order to guarantee that the norm of any vector (that is, its length: $\sqrt{\langle A|A\rangle}$) remains, under all circumstances, a positive real number. Formula (2.22) for probabilities needs to be altered very slightly too, since, in complex spaces, $\langle A|B\rangle$ and, hence, (2.22) may be complex numbers (and yet probabilities must necessarily be real, positive numbers between 0 and 1). The solution is to change (2.22) to

$$|\langle a|B = b_i\rangle|^2 \tag{2.23}$$

where the vertical bars denote absolute value (or "distance from zero," which is invariably a real, positive number). Equation (2.23) stipulates that the probability that a measurement of $B$ on a system in the state $|a\rangle$ will produce the outcome $B = b_i$ is equal to the square of the distance from 0 of the complex number $\langle a|B = b_i\rangle$; and that probability, so defined, will invariably be a real and positive number. Formula (2.22), by the way, will entail not only that $|A\rangle$ and $-|A\rangle$ represent the same physical state (we've already seen that to be the case), but, more generally, that $|A\rangle$ and $@|A\rangle$ represent the same state, where @ may be any one of the infinity of complex numbers of absolute value 1.

The elements of the operator matrices of linear operators on

complex vector spaces (that is, the numbers $O_{ij}$ of (2.10) and (2.11)) can be complex numbers too. Nonetheless, it may happen to some such operators that all of their eigenvectors are associated only with real eigenvalues (albeit, perhaps, not all of their matrix elements $O_{ij}$, and perhaps even none of them, are real). Linear operators like that are called *Hermitian* operators; and it's clear from principle (B) (since, of course, the values of physically measurable quantities are always real numbers) that the operators associated with measurable properties must necessarily be Hermitian operators.

Here are some facts about Hermitian operators:

(1) If two vectors are both eigenvectors of the same Hermitian operator, and if the eigenvalues associated with those two eigenvectors are two different (real) numbers, then the two vectors in question are necessarily orthogonal to each other.

That pretty much had to be so, if this algorithm is going to work out right; otherwise, measurements wouldn't be repeatable. The different eigenvalues of a property operator, after all, correspond to different values of that property; and (if measurements of a property are to be repeatable) having a certain value of a certain property must entail that subsequent measurements of that property will certainly not find any other value of it;[5] and that (given principle (D)) will require that state vectors connected with different values of the same measurable property ($|\text{black}\rangle$ and $|\text{white}\rangle$, say, or $|\text{hard}\rangle$ and $|\text{soft}\rangle$) be orthogonal to one another.

(2) Any Hermitian operator on an $N$-dimensional space will always have at least one set of $N$ mutually orthogonal eigenvectors. Which is to say: it will always be possible to form a basis of the space out of the eigenvectors of any Hermitian operator; different bases, of course, for different operators. Consider, for example, the hardness operator of equation (2.18) and the color operator of equation (2.19).

(3) The reader ought to be able to persuade herself, now, of the following: if a Hermitian operator on an $N$-dimensional space

5. Supposing, once again, that no tampering, and no dynamical evolution, has gone on in the meantime.

happens to have $N$ different eigenvalues, then there is a *unique* vector in the space (or, rather, unique modulo multiplication by numbers) associated with each different one of those eigenvalues; and of course the set of all eigenvectors of length 1 of that operator will form a unique basis of that space (or, rather, unique modulo multiplication by numbers of absolute value 1). Operators like that are called *complete* or *nondegenerate* operators.

(4) Any Hermitian operator on a given space will invariably be associated with some measurable property of the physical system connected with that space (this is just a somewhat more informative version of the first part of principle (B)).

(5) Any vector whatever in a given space will invariably be an eigenvector of some complete Hermitian operator on that space. That, combined with fact (4) and principle (B), will entail that any quantum state whatever of a given physical system will invariably be associated with some definite value of some measurable property of that system.

All this turns out to entail (among other things) that every quantum-mechanical system necessarily has an infinity of mutually incompatible measurable properties. Think (just to have something concrete to talk about) of the space of possible spin states of an electron. There are, to begin with, a continuous infinity of different such states (since there are a continuous infinity of vectors of length 1 in a two-dimensional space); moreover, given any one of those states, there are clearly a continuous infinity of different possible states which are not orthogonal to it. And, by facts (3) and (5) above, every state in this space is necessarily the only eigenstate associated with a certain particular eigenvalue of a certain particular complete operator, and, by fact (1), none of the continuous infinity of states which aren't orthogonal to the state in question can possibly be eigenstates of the same complete operator. What's more, the complete operators of which those other states *are* eigenstates clearly can't even be *compatible* with the operator in question. And so (since all this applies to every state in the space) there must necessarily be a continuous infinity of mutually incompatible complete measurable properties, of which color and hardness are only two.

It will be useful, for what comes later, to give two more of those properties names. The vectors

$$1/2|\text{black}\rangle + \sqrt{3}/2|\text{white}\rangle \quad \text{and} \quad \sqrt{3}/2|\text{black}\rangle - 1/2|\text{white}\rangle$$

are both of length 1 and are orthogonal to one another (and aren't orthogonal to any of the eigenvectors of color or hardness), and so it follows that there must be a complete observable of which they are both eigenstates, with different eigenvalues (which can always be set at $+1$ and $-1$, respectively). Let's call that observable "gleb." And the vectors

$$1/2|\text{black}\rangle - \sqrt{3}/2|\text{white}\rangle \quad \text{and} \quad \sqrt{3}/2|\text{black}\rangle + 1/2|\text{white}\rangle$$

are both of length 1 and are orthogonal to one another (and aren't orthogonal to any of the eigenvectors of color or hardness or gleb), and so it follows that there must be a complete observable of which *they* are both eigenstates, with different eigenvalues (which can always be set at $+1$ and $-1$, respectively). Let's call that observable "scrad." Of course, the eigenstates of gleb and scrad (just like those of color and hardness) both form different bases of the spin space.

Finally, there are rules (never mind what those rules are, precisely) for adding and subtracting matrices to or from one another, and for multiplying them by one another. The *commutator* of two matrices $A$ and $B$, which is denoted by the symbol $[A,B]$, is defined to be the object $AB - BA$ (the rules for multiplying matrices by one another entail that the order of multiplication counts: $AB$ isn't necessarily the same as $BA$).

Now, it can be shown that in the event that $[A,B] = 0$ (that is, in the event that $AB$ *is* equal to $BA$), $A$ and $B$ *share* at least one set of eigenvectors which form a basis of the space. A little reflection will confirm that the operator matrices of incompatible observables can't possibly share any such complete basis of eigenvectors (since such eigenvectors would correspond to definite value states of both observables at the same time). It must be the case, then, that *the commutators of incompatible observable matrices are nonzero.* So

the property of commutativity (that is, the condition $[A,B] = 0$) turns out to be a convenient mathematical test for compatibility. Moreover, in cases of incompatible observables, the commutator of the two observables in question turns out to be extremely useful for assessing the *degree* of their incompatibility.[6]

## Coordinate Space

Let's begin to apply all this. Let's see, in some detail, how to set up a quantum-mechanical representation, and a quantum-mechanical dynamics, of some simple physical system. Forget about color and hardness for the moment. Think of a familiar sort of particle, one with only the familiar sorts of physical properties: position and velocity and momentum and energy and things like that.

Here's a way to get started: We know, from hundreds of years of experience, that the behaviors of relatively big particles, with relatively big masses (particles you can see, like rocks and baseballs and planets) are very well described by the classical mechanics of Newton. That entails something about the quantum theory of particles: whatever that theory ends up predicting about the strange, tiny particles of Chapter 1, it ought to predict that everyday particles, subject to everyday circumstances, will behave in the

6. Perhaps the notion of there being various different degrees of incompatibility requires some elucidation. Here's what the idea is (or here's what it is, at any rate, in the simplest case, when the observables involved are both complete):

Consider two complete and incompatible observables (call them $A$ and $B$) of some physical system. If, when any particular eigenstate of $A$ obtains, the outcome of a measurement of $B$ can be predicted (by means of formula (2.23)) with something approaching certainty (that is: if, for each eigenvector of $A$, there is some particular eigenvector of $B$ such that the product of those two vectors is something approaching one), then $A$ and $B$ are said to be only very slightly incompatible. But if (at the other extreme), when any particular eigenstate of $A$ obtains, the probabilities of the various possible outcomes of a measurement of $B$ are all the same (that is: if knowing the value of $A$ gives us no information whatever about the outcome of an upcoming measurement of $B$), *then* $A$ and $B$ are said to be *maximally* incompatible.

So (for example) color and hardness (which are maximally incompatible observables) are a good deal more incompatible with one another than color and scrad are.