Handout #5: The statistical algorithm refined, the uncertainty principle, Schrödinger's Equation, states of multi-particle systems

I. Projection operators

First, a word about *operators*: In general, an operator **A** on a vector space is just a function that takes vectors as input and spits out vectors as output. Almost all the operators we will ever care about are *linear operators*; so let us define them:

An operator **A** is *linear* iff, for all scalars a,b, and vectors ϕ,ψ, $\mathbf{A}(a\phi + b\psi) = a\mathbf{A}(\phi) + b\mathbf{A}(\psi)$.

Now we can lead up to a definition of *projection operators*, which are a particularly useful species of linear operator. We will begin with a basic fact about vector spaces (more exactly: those equipped with an inner product—hence those for which a notion of "orthogonality" is well-defined): given any two vectors Φ and Ψ, there is a unique choice of scalar c and vector $\Psi^{\perp}$ such that

(i) $\Phi = c\Psi + \Psi^{\perp}$, and

(ii) $\Psi^{\perp}$ is orthogonal to Ψ.

In other words, there is a unique way to write Φ as a linear combination of Ψ with a vector orthogonal to Ψ.

Now we can say what a projection operator *onto a vector* is: The projection operator onto the vector Ψ—written "$\mathbf{P}_{\Psi}$"—is simply the linear operator obeying the following equation:

$\mathbf{P}_{\Psi}\Phi = c\Psi$,

where $\Phi = c\Psi + \Psi^{\perp}$, as above. That is, $\mathbf{P}_{\Psi}\Phi$ returns the component of Φ that is parallel to the vector Ψ. Finally, it is very easy to show that if Ψ' is parallel to Ψ, then $\mathbf{P}_{\Psi'} = \mathbf{P}_{\Psi}$. Given this fact, it would be a little more accurate—but a little less convenient—to write $\mathbf{P}_{\Psi}$ as $\mathbf{P}_S$, where S is understood to be the (one-dimensional) *subspace* spanned by Ψ. Never mind.

Now, here is another basic fact about vector spaces (again more exactly: those equipped with an inner product): given any vector Φ and subspace S (of whatever dimensionality), there is a unique choice of vectors Ψ and $\Psi^{\perp}$ such that

(i) $\Phi = \Psi + \Psi^{\perp}$,

(ii) Ψ lies in the subspace S, and

(iii) $\Psi^{\perp}$ is orthogonal to Ψ.

In other words, there is a unique way to write Φ as a linear combination of a vector in S with a vector orthogonal to S.

Now we can say what a projection operator onto a *subspace* is: The projection operator onto the subspace S—written "$\mathbf{P}_S$"—is simply the linear operator obeying the following equation:

$\mathbf{P}_S\Phi = \Psi$,

where $\Phi = \Psi + \Psi^{\perp}$, as above. That is, $\mathbf{P}_S\Phi$ returns the component of Φ that lies in the subspace S.

All of this is about to be put to use.

II. An even more efficient statement of the statistical algorithm

Let's start with the rendition of the statistical algorithm presented in Handout #4, but stated this time more formally. Bowing to convention, we will use the more common expression "measurement of an observable" in place of the much less loaded "experiment":

First rendition: If the state of a system is represented by the normalized vector ψ, then the probability that a measurement of the observable A (represented by orthonormal basis $\{\phi_i\}$, where each vector ϕ_i is associated with the outcome o_i) performed on the system will yield as outcome the value α is $\Sigma|a_j|^2$, where the index j ranges over all those values such that $o_j = \alpha$, and where $\psi = \Sigma a_i \phi_i$.

In other words, we start by writing ψ as a linear combination of the elements in the basis $\{\phi_i\}$. Suppose we have done so; i.e., we have found all the coefficients a_i such that

$\psi = \Sigma a_i \phi_i$.

Then we collect together all those coefficient a_j such that the basis element ϕ_j is associated with the outcome α that we are interested in. Finally, we square these coefficients, and add the results: that gives us the probability that the outcome of the measurement will be α.

Here is another way to proceed. First, we notice that the basis elements ϕ_j that are associated with the outcome α collectively pick out a *subspace*: namely, the subspace consisting of all possible linear combinations of them. Let us call this subspace "S_α". And let $\mathbf{P}_\alpha$ be the projection operator onto this subspace. Then it is fairly easy to show that

$\mathbf{P}_\alpha \psi = \Sigma a_j \phi_j$,

where, again, the index j ranges over all those values such that $o_j = \alpha$. It is also easy to show that the *length* of the vector $\mathbf{P}_\alpha \psi$ is

$\langle \mathbf{P}_\alpha \psi \mid \mathbf{P}_\alpha \psi \rangle = \langle \psi \mid \mathbf{P}_\alpha \psi \rangle = \Sigma |a_j|^2$.

Putting all this together, we get what will turn out to be a more convenient statement of the statistical algorithm:

Second rendition: If the state of a system is represented by the normalized vector ψ, then the probability that a measurement of the observable A (represented by orthonormal basis $\{\phi_i\}$, where each vector ϕ_i is associated with the outcome o_i) performed on the system will yield as outcome the value α is $<\psi \mid \mathbf{P}_\alpha\psi>$, where $\mathbf{P}_\alpha$ is the projection operator onto the subspace S_α spanned by all those basis vectors ϕ_j such that $o_j = \alpha$.

Let us finally provide a rendition of the statistical algorithm that makes explicit use of the representation of observables by Hermitian operators. First, some definitions:

Linear operator **A** is *Hermitian* iff, for every ϕ,ψ, $<\phi \mid \mathbf{A}\psi> = <\mathbf{A}\phi \mid \psi>$.

Vector ϕ is an *eigenvector* of operator **A** iff there is some scalar c such that $\mathbf{A}\phi = c\phi$; the scalar c is called its *eigenvalue*. We will say that **A** *admits of a complete set of eigenvectors* iff there is some orthonormal basis $\{\phi_i\}$ consisting entirely of eigenvectors of **A**.

Suppose that **A** is a Hermitian operator that has eigenvectors. (Not every Hermitian operator does; we're mostly going to ignore those that don't.) Suppose, in fact, that the set of its eigenvectors is rich enough that *any* vector can be expressed as a linear combination of eigenvectors of **A**. Then there are several points to note about these eigenvectors.

The first, rather minor point is that for any non-zero eigenvector ϕ of **A**, the associated eigenvalue c must be real. That is because

$<\phi \mid \mathbf{A}\phi> = <\mathbf{A}\phi \mid \phi>$ since **A** is Hermitian; therefore

$<\phi \mid c\phi> = <c\phi \mid \phi>$ since ϕ is an eigenvector of **A** with eigenvalue c; therefore

$c<\phi \mid \phi> = c^*<\phi \mid \phi>$ by the definition of the inner product; therefore

$c = c^*$ since $<\phi \mid \phi> \neq 0$.

All too often, you see authors making a big deal out of this result, for the following reason: Since, at the end of the day, we will be using eigenvalues of Hermitian operators to represent measurement outcomes, it is supposed to be a **Very Good Thing** that those eigenvalues are real—since, allegedly, it would be **Completely Mysterious** how a complex number could represent a measurement outcome. This is a piece of utter foolishness. You might pause to figure out why.

Next, suppose that ϕ and ψ are eigenvectors of **A** with the same eigenvalue c. Then for any scalars a and b,

$\mathbf{A}(a\phi + b\psi) = a\mathbf{A}(\phi) + b\mathbf{A}(\psi)$ since A is linear;

$= ac\phi + bc\psi$

$= c(a\phi + b\psi)$.

Thus, any linear combination of ϕ and ψ is an eigenvector of **A** with eigenvalue c. What this means is that the set of eigenvectors of **A** with eigenvalue c is in fact a *subspace*.

Finally, suppose that ϕ and ψ are eigenvectors of **A** with *different* eigenvalues c and d, respectively. Then

$\langle\phi \mid \mathbf{A}\psi\rangle = \langle\phi \mid d\psi\rangle = d\langle\phi \mid \psi\rangle$; but also

$\langle\phi \mid \mathbf{A}\psi\rangle = \langle\mathbf{A}\phi \mid \psi\rangle = \langle c\phi \mid \psi\rangle = c\langle\phi \mid \psi\rangle$. So

$c\langle\phi \mid \psi\rangle = d\langle\phi \mid \psi\rangle$.

But since $c \neq d$, this equation can only hold if $\langle\phi \mid \psi\rangle = 0$. Thus we conclude that eigenvectors of **A** with different eigenvalues must be orthogonal.

What all this means is that we can think of **A** as picking out a set of subspaces, where (i) each subspace is associated with a unique eigenvalue of **A**; (ii) any two subspaces are orthogonal to

each other; and (iii) these subspaces span the whole vector space. Making use of this observation, we arrive at the

Third rendition: If the state of a system is represented by the normalized vector ψ, then the probability that a measurement of the observable A (represented by Hermitian operator **A**) performed on the system will yield as outcome the value α is $<\psi \mid \mathbf{P}_\alpha\psi>$, where $\mathbf{P}_\alpha$ is the projection operator onto the subspace of eigenvectors of **A** with eigenvalue α.

III. Incompatible observables, and the Uncertainty Principle

Suppose we have a system in state ϕ, and we perform a measurement of observable A (represented by Hermitian operator **A**) on it. One quantity we can calculate is the *expected value* of the outcome; we get this by weighting each possible outcome by its probability, and summing the results (we are assuming here that **A** admits of a complete set of eigenvectors). Let us denote this quantity "$<\mathbf{A}>_\phi$". It's a relatively straightforward consequence of the statistical algorithm that $<\mathbf{A}>_\phi = <\phi \mid \mathbf{A}\phi>$.

Suppose that the observed outcome of the measurement is α; then we will say that the *deviation* of the outcome from the expected value is equal to $(\alpha - <\mathbf{A}>_\phi)^2$ (we take the square so that the deviation is always a positive number). We can then calculate the expected value of the deviation, sometimes denoted $<(\Delta\mathbf{A})^2>_\phi$; this turns out to be equal to $<\phi \mid \mathbf{A}^2\phi> - <\phi \mid \mathbf{A}\phi>^2$. Intuitively, the quantity $<(\Delta\mathbf{A})^2>_\phi$—sometimes called the *uncertainty* of A (with respect to state ϕ)—measures how widely distributed the outcomes of the measurement of A are likely to be.

Heisenberg's Uncertainty Principle states that for any two observables A and B, and any state ϕ, the product of the uncertainties of A and B (with respect to ϕ) is greater than or equal to a certain non-negative bound, which is a function of **A**, **B**, and ϕ. Precisely:

$$<(\Delta\mathbf{A})^2>_\phi <(\Delta\mathbf{B})^2>_\phi \geq \tfrac{1}{4} \; |<\phi \mid (\mathbf{AB} - \mathbf{BA})\phi>|^2.$$

If this bound is positive, then, intuitively, the principle says that either a measurement of A must have outcomes that are to some extent uncertain, or a measurement of B must.

Two observables A and B are said to be *incompatible* iff the Hermitian operators **A** and **B** which represent them do not commute; i.e., iff $\mathbf{AB} \neq \mathbf{BA}$; otherwise they are said to be compatible. The intuitive punchline of the Uncertainty Principle is, then, this: If two observables are incompatible, then there will be states ϕ such that—no matter how carefully we do our measurements—at least one of a measurement of A and of B must have outcomes that are to some extent uncertain. Note well what the Uncertainty Principle does *not* say: it does not say, for example, that it is impossible to simultaneously measure incompatible observables; nor does it say that the act of measurement invariably perturbs the system being measured, or that there is some irreducible limit to how small we can make such perturbation. It doesn't even come close to saying such things.

IV. Unitary operators

So far, we have a mathematical framework that allows us to represent the states of systems *at a moment*; we don't yet have a picture of how these states change over time. Developing this picture will require introducing a new class of linear operators, called *unitary* operators. The official definition is this: An operator **U** is unitary iff it is linear, has an inverse, and for every ϕ,

$\langle \mathbf{U}\phi \mid \mathbf{U}\phi \rangle = \langle \phi \mid \phi \rangle$. We will work with an unofficial definition—provably equivalent to the official one—which we arrive at by considering the rotation operators in $\mathcal{R}^2$.

Recall (from Hughes, section 1.2) that a rotation operator in $\mathcal{R}^2$ is a linear operator that rotates any vector through a fixed angle, leaving the length of the vector unchanged. Notice that if **R** is a rotation, and $\{\phi_1, \phi_2\}$ is an orthonormal basis, then $\{\mathbf{R}\phi_1, \mathbf{R}\phi_2\}$ is also an orthonormal basis. We generalize this fact to arrive at our unofficial definition: Given some orthonormal basis $\{\phi_i\}$ for vector space $\mathcal{V}$, an operator **U** is unitary iff it is linear, and the set of vectors $\{\mathbf{U}\phi_i\}$ is also an orthonormal basis for $\mathcal{V}$. Thus, the unitary operators on $\mathcal{V}$ are just the analogues of the rotations on $\mathcal{R}^2$.

It follows that **U** has an inverse. For define the operator $\mathbf{U}^{-1}$ this way: $\mathbf{U}^{-1}$ is linear, and for each vector $\mathbf{U}\phi_i$, $\mathbf{U}^{-1}(\mathbf{U}\phi_i) = \phi_i$. (This definition is legitimate, since $\{\mathbf{U}\phi_i\}$ is an orthonormal basis. Note the in order to exactly specify a linear operator, it is only necessary to specify how it operates on each member of some basis; linearity does the rest.) Then for any ψ,

$$\begin{aligned}
\mathbf{U}^{-1}(\mathbf{U}\psi) &= \mathbf{U}^{-1}(\mathbf{U}\Sigma_i\langle\phi_i \mid \psi\rangle\phi_i) \\
&= \Sigma_i \langle\phi_i \mid \psi\rangle\mathbf{U}^{-1}(\mathbf{U}\phi_i) \\
&= \Sigma_i \langle\phi_i \mid \psi\rangle\phi_i \\
&= \psi.
\end{aligned}$$

So $\mathbf{U}^{-1}$ is indeed the inverse of **U**.

It also follows that, for any ψ and γ, $\langle\psi \mid \gamma\rangle = \langle\mathbf{U}\psi \mid \mathbf{U}\gamma\rangle$. To see this, note first that it follows immediately from the definition of **U** in the special case where $\psi = \phi_j$ and $\gamma = \phi_k$. To prove the general case, we notice that

$$\langle \psi \mid \gamma \rangle = \langle (\Sigma_i \langle \phi_i \mid \psi \rangle \phi_i) \mid (\Sigma_j \langle \phi_j \mid \gamma \rangle \phi_j) \rangle$$

$$= \Sigma_{ij} \langle \psi \mid \phi_i \rangle \langle \phi_j \mid \gamma \rangle \langle \phi_i \mid \phi_j \rangle$$

$$= \Sigma_i \langle \psi \mid \phi_i \rangle \langle \phi_i \mid \gamma \rangle$$

and that

$$\langle \mathbf{U}\psi \mid \mathbf{U}\gamma \rangle = \langle \mathbf{U}(\Sigma_i \langle \phi_i \mid \psi \rangle \phi_i) \mid \mathbf{U}(\Sigma_j \langle \phi_j \mid \gamma \rangle \phi_j) \rangle$$

$$= \Sigma_{ij} \langle \psi \mid \phi_i \rangle \langle \phi_j \mid \gamma \rangle \langle \mathbf{U}\phi_i \mid \mathbf{U}\phi_j \rangle$$

$$= \Sigma_i \langle \psi \mid \phi_i \rangle \langle \phi_i \mid \gamma \rangle.$$

Hence in particular, $\langle \psi \mid \psi \rangle = \langle \mathbf{U}\psi \mid \mathbf{U}\psi \rangle$. So if an operator is unitary in our unofficial sense, then it is also unitary in the official sense.

The converse implication also holds. The key step in the proof is to show that, if **U** is linear, has an inverse, and for every ϕ, $\langle \mathbf{U}\phi \mid \mathbf{U}\phi \rangle = \langle \phi \mid \phi \rangle$, then for every pair of orthogonal vectors ϕ and ψ, $\langle \mathbf{U}\phi \mid \mathbf{U}\psi \rangle = \langle \phi \mid \psi \rangle = 0$. To show this, observe that

$$\langle \mathbf{U}(\phi + \psi \mid \mathbf{U}(\phi + \psi \rangle = \langle \phi + \psi \mid \phi + \psi \rangle = \langle \phi \mid \phi \rangle + \langle \psi \mid \psi \rangle$$

and that

$$\langle \mathbf{U}(\phi + \psi \mid \mathbf{U}(\phi + \psi \rangle = \langle \mathbf{U}\phi \mid \mathbf{U}\phi \rangle + \langle \mathbf{U}\phi \mid \mathbf{U}\psi \rangle + \langle \mathbf{U}\psi \mid \mathbf{U}\phi \rangle + \langle \mathbf{U}\psi \mid \mathbf{U}\psi \rangle$$

$$= \langle \phi \mid \phi \rangle + \langle \psi \mid \psi \rangle + \langle \mathbf{U}\phi \mid \mathbf{U}\psi \rangle + \langle \mathbf{U}\psi \mid \mathbf{U}\phi \rangle.$$

It follows that $\langle \mathbf{U}\phi \mid \mathbf{U}\psi \rangle + \langle \mathbf{U}\psi \mid \mathbf{U}\phi \rangle = 0$.

Similarly,

$$\langle \mathbf{U}(\phi + i\psi \mid \mathbf{U}(\phi + i\psi \rangle = \langle \phi + i\psi \mid \phi + i\psi \rangle = \langle \phi \mid \phi \rangle + \langle \psi \mid \psi \rangle$$

and

$$\langle \mathbf{U}(\phi + i\psi \mid \mathbf{U}(\phi + i\psi \rangle = \langle \mathbf{U}\phi \mid \mathbf{U}\phi \rangle + i\langle \mathbf{U}\phi \mid \mathbf{U}\psi \rangle - i\langle \mathbf{U}\psi \mid \mathbf{U}\phi \rangle + \langle \mathbf{U}\psi \mid \mathbf{U}\psi \rangle$$

$$= \langle \phi \mid \phi \rangle + \langle \psi \mid \psi \rangle + i\langle \mathbf{U}\phi \mid \mathbf{U}\psi \rangle - i\langle \mathbf{U}\psi \mid \mathbf{U}\phi \rangle.$$

It follows that $\langle \mathbf{U}\phi \mid \mathbf{U}\psi \rangle - \langle \mathbf{U}\psi \mid \mathbf{U}\phi \rangle = 0$.

Hence, $< \mathbf{U}\phi | \mathbf{U}\psi > = 0$.

So the unofficial definition is in fact equivalent to the official one: unitary operators are exactly those linear operators that map orthonormal bases into orthonormal bases.

One bit of unfinished business, before we move on to a description of the dynamics of states. The unofficial definition of unitary operators was made relative to a particular choice of orthonormal basis $\{\phi_i\}$. In fact, this choice is arbitrary, in the sense that if **U** maps a particular orthonormal basis $\{\phi_i\}$ into another orthonormal basis $\{\mathbf{U}\phi_i\}$, then for any *other* orthonormal basis $\{\psi_i\}$, $\{\mathbf{U}\psi_i\}$ is *also* an orthonormal basis. To see this, note first that $\{\mathbf{U}\psi_i\}$ is obviously an orthonormal *set*; this follows from the fact that $<\mathbf{U}\psi_j \mid \mathbf{U}\psi_k> = <\psi_j \mid \psi_k>$. Now suppose that $\{\mathbf{U}\psi_i\}$ is not a *basis*. Then there must be some vector γ which is orthogonal to all of the $\mathbf{U}\psi_i$. But then it follows that the vector $\mathbf{U}^{-1}\gamma$ is orthogonal to all of the ψ_i, which contradicts the fact that $\{\psi_i\}$ is a basis.

V. Dynamics of states

Suppose that **A** is an Hermitian operator on $\mathcal{V}$ that has a complete set of eigenvectors $\{\phi_i\}$; in fact, let $\{\phi_i\}$ be an orthonormal basis for $\mathcal{V}$. (There will always be such Hermitian operators, when $\mathcal{V}$ is a Hilbert space—even if $\mathcal{V}$ is infinite-dimensional. Proving this is left as an exercise for the ambitious.) Then, letting t be an arbitrary real number, we define the operator $e^{-it\mathbf{A}}$ this way:

$$e^{-it\mathbf{A}} = \Sigma_{j=0,\infty} (-it\mathbf{A})^n/n!.$$

(You may recognize in this sum the usual Taylor-series expansion of e^x.) So defined, $e^{-it\mathbf{A}}$ is linear; further, if ϕ is an eigenvector of **A** with eigenvalue a, we have

$$e^{-it\mathbf{A}}(\phi) = \Sigma_{j=0,\infty} (-it\mathbf{A})^n/n! \, (\phi)$$

$$= \Sigma_{j=0,\infty} [(-it)^n/n!] \, \mathbf{A}^n(\phi)$$

$$= \Sigma_{j=0,\infty} [(-it)^n/n!] \, a^n(\phi)$$

$$= \Sigma_{j=0,\infty} (-ita)^n/n! \, (\phi)$$

$$= e^{-ita}\phi.$$

Let $\mathbf{U}_A(t) = e^{-it\mathbf{A}}$. Then, making use of the above result, we can verify that

$\mathbf{U}_A(0)\phi = \phi$;

$\mathbf{U}_A(t_1 + t_2)\phi = \mathbf{U}_A(t_1)\mathbf{U}_A(t_2)\phi$.

These equations hold for *any* eigenvector of **A**. Hence—since **A** has a complete set of eigenvectors—they hold *in general*. So we in fact have (letting **I** denote the identity operator)

(1) $\mathbf{U}_A(0) = \mathbf{I}$;

(2) $\mathbf{U}_A(t_1 + t_2) = \mathbf{U}_A(t_1)\mathbf{U}_A(t_2)$.

Notice now that for any t, $\langle \mathbf{U}_A(t)\phi \mid \mathbf{U}_A(t)\phi \rangle = \langle e^{-ita}\phi \mid e^{-ita}\phi \rangle = e^{ita}e^{-ita}\langle \phi \mid \phi \rangle = \langle \phi \mid \phi \rangle$. This holds for any eigenvector of **A**; hence in particular it holds for each member of the orthonormal basis $\{\phi_i\}$. It follows (as you may care to verify) that for *any* vector ψ, $\langle \mathbf{U}_A(t)\psi \mid \mathbf{U}_A(t)\psi \rangle = \langle \psi \mid \psi \rangle$. Further, (1) and (2) guarantee that $\mathbf{U}_A(t)$ has an inverse—namely, $\mathbf{U}_A(-t)$. So for each real number t, $\mathbf{U}_A(t)$ is a unitary operator.

Now, suppose that S is some physical system whose possible physical states are represented on Hilbert space $\mathcal{H}_S$. There is a particular Hermitian operator **H** on $\mathcal{H}_S$—called the *Hamiltonian operator* for S—which is typically taken to represent the *energy* of S. Further, if S is a *closed*

system—which we can take to mean, roughly, "isolated from its surroundings"—then the evolution of its physical state is as follows:

(3) Let ψ_t be the state of S at time t. Then the state of S at any time t' is $\psi_{t'} = \mathbf{U}_H(t' - t)\psi_t$ (provided S remains closed throughout the relevant time interval).

Notice the importance of (1) and (2): Clearly, we must have $\mathbf{U}_H(0)\psi_t = \psi_t$; (1) guarantees this. Further, it must be consistent to hold that $\psi_{t_2} = \mathbf{U}_H(t_2 - t_1)\psi_{t_1}$, $\psi_{t_1} = \mathbf{U}_H(t_1 - t_0)\psi_{t_0}$, and $\psi_{t_2} = \mathbf{U}_H(t_2 - t_0)\psi_{t_0}$; (2) guarantees that it is.

(3) is, in essence, a statement of *Schrödinger's Equation*, which is the fundamental equation of motion for quantum mechanics. Notice that the evolution it describes is fully deterministic: provided the system S remains closed, its state at any one time determines its state at all other times. If you smell trouble at this point, your nose is in good working order.

VI. Tensor products

Suppose a system S consists of n particles. Then we know that the possible states of S are represented by means of some Hilbert space $\mathcal{H}_S$, and the possible states of each particle i are represented by means of some Hilbert spaces $\mathcal{H}_i$. But questions remain: How is $\mathcal{H}_S$ related to each of the $\mathcal{H}_i$? How are states of S related to states of the individual particles that make it up? In particular, is the state of S fixed, once we know the states of the particles? Conversely, are the states of the particles fixed, once we know the state of S?

We'll begin with the first question: The answer is that $\mathcal{H}_S$ is the *tensor product* of the $\mathcal{H}_i$. To explain what this means, we'll consider just the case of a two-particle system; the generalization to many-particle systems will be obvious. Note that we are not interested in giving a rigorous

definition of "tensor product"; for that you can consult any decent linear algebra text. What follows aims at a more user-friendly, working account of the notion.

We write $\mathcal{H}_S = \mathcal{H}_1 \otimes \mathcal{H}_2$; $\mathcal{H}_S$ is then a vector space with the following features:

(i) For any $\phi \in \mathcal{H}_1$ and $\psi \in \mathcal{H}_2$, there is a vector $\phi \otimes \psi \in \mathcal{H}_S$;

(ii) If $\{\phi_i\}$ is an orthonormal basis for $\mathcal{H}_1$ and $\{\psi_j\}$ is an orthonormal basis for $\mathcal{H}_2$, then the set of vectors $\{\phi_i \otimes \psi_j\}$ is an orthonormal basis for $\mathcal{H}_S$;

(iii) The inner product on $\mathcal{H}_S$ is specified in terms of the inner products on $\mathcal{H}_1$ and $\mathcal{H}_2$ by the following stipulation: $\langle \phi_1 \otimes \psi_1 \mid \phi_2 \otimes \psi_2 \rangle = \langle \phi_1 \mid \phi_2 \rangle \langle \psi_1 \mid \psi_2 \rangle$;

(iv) If **A** and **B** are linear operators on $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively, then we can define their *product* $\mathbf{A} \otimes \mathbf{B}$ to be that linear operator on $\mathcal{H}_S$ such that, for any vector $\phi \otimes \psi$, $(\mathbf{A} \otimes \mathbf{B})(\phi \otimes \psi) = (\mathbf{A}\phi) \otimes (\mathbf{B}\psi)$.

Let's unpack this a bit. First, observe that if ϕ_1 and ϕ_2 are orthogonal, then regardless of whether ψ_1 and ψ_2 are orthogonal, the vectors $\phi_1 \otimes \psi_1$ and $\phi_2 \otimes \psi_2$ will be orthogonal, given (iii). Likewise, if ψ_1 and ψ_2 are orthogonal, then regardless of whether ϕ_1 and ϕ_2 are orthogonal, the vectors $\phi_1 \otimes \psi_1$ and $\phi_2 \otimes \psi_2$ will be orthogonal. Finally, if ϕ and ψ are both normalized (i.e., $\langle \phi \mid \phi \rangle = 1 = \langle \psi \mid \psi \rangle$, then of course so is $\phi \otimes \psi$. All of which means that *part* of what is stipulated in (ii) is redundant, given (i) and (iii): namely, the set of vectors $\{\phi_i \otimes \psi_j\}$ is guarantee to be an orthonormal *set* of vectors. What is *added* by (ii) is the stipulation that this set is a *basis* for the vector space $\mathcal{H}_S$.

Next, it can be helpful to note that in many ways, the operation "$\otimes$" can be treated like *multiplication*. In particular, it's distributive; i.e., for arbitrary scalars a,b and arbitrary vectors $\gamma_1, \gamma_2 \in \mathcal{H}_1$ and $\nu_1, \nu_2 \in \mathcal{H}_2$, both of the following hold:

(1) $(a\gamma_1 + b\gamma_2)\otimes\nu_1 = a(\gamma_1\otimes\nu_1) + b(\gamma_2\otimes\nu_1)$;

(2) $\gamma_1\otimes(a\nu_1 + b\nu_2) = a(\gamma_1\otimes\nu_1) + b(\gamma_1\otimes\nu_2)$.

It will be a useful little exercise to actually *prove* that (1) and (2) hold. Here's how to do so. Let $\{\phi_i\otimes\psi_j\}$ be an orthonormal basis for $\mathcal{H}_1\otimes\mathcal{H}_2$. Then we can show that these equalities hold by showing that, in each case, the inner products of the two vectors with an arbitrary element of $\{\phi_i\otimes\psi_j\}$ are the same. For, quite generally, if Φ and Ψ are two vectors in some vector space **V**, then $\Phi = \Psi$ iff, for every element Γ of some orthonormal basis for **V**, $\langle\Psi \mid \Gamma\rangle = \langle\Phi \mid \Gamma\rangle$.

By the definition of the inner product for $\mathcal{H}_1\otimes\mathcal{H}_2$,

$\langle(a\gamma_1 + b\gamma_2)\otimes\nu_1 \mid \phi_i\otimes\psi_j\rangle$

$= \langle(a\gamma_1 + b\gamma_2) \mid \phi_i\rangle\langle\nu_1 \mid \psi_j\rangle$

$= \{a^*\langle\gamma_1 \mid \phi_i\rangle + b^*\langle\gamma_2 \mid \phi_i\rangle\}\langle\nu_1 \mid \psi_j\rangle$

$= a^*\langle\gamma_1 \mid \phi_i\rangle\langle\nu_1 \mid \psi_j\rangle + b^*\langle\gamma_2 \mid \phi_i\rangle\langle\nu_1 \mid \psi_j\rangle$

$= a^*\langle\gamma_1\otimes\nu_1 \mid \phi_i\otimes\psi_j\rangle + b^*\langle\gamma_2\otimes\nu_2 \mid \phi_i\otimes\psi_j\rangle$

$= \langle a(\gamma_1\otimes\nu_1) \mid \phi_i\otimes\psi_j\rangle + \langle b(\gamma_2\otimes\nu_2) \mid \phi_i\otimes\psi_j\rangle$

$= \langle a(\gamma_1\otimes\nu_1) + b(\gamma_2\otimes\nu_1) \mid \phi_i\otimes\psi_j\rangle$.

The proof of (2) is a fairly obvious variation on this proof.

Next, we should observe that there was something a little fishy about (ii), which said that if $\{\phi_i\}$ is an orthonormal basis for $\mathcal{H}_1$ and $\{\psi_j\}$ is an orthonormal basis for $\mathcal{H}_2$, then the set of vectors $\{\phi_i\otimes\psi_j\}$ is an orthonormal basis for $\mathcal{H}_S$. For suppose we had chosen a *different* orthonormal basis $\{\gamma_i\}$ for $\mathcal{H}_1$ and $\{\nu_j\}$ for $\mathcal{H}_2$; what guarantees that the resulting set $\{\gamma_i\otimes\nu_j\}$ would *also* have constituted an orthonormal basis for $\mathcal{H}_S$? Happily, we can now say precisely what guarantees this: each vector in $\{\phi_i\otimes\psi_j\}$ can be written as a linear combination of the

vectors in $\{\gamma_i \otimes \nu_j\}$, and likewise each vector in $\{\gamma_i \otimes \nu_j\}$ can be written as a linear combination of the vectors in $\{\phi_i \otimes \psi_j\}$. Suppose, for example, that $\phi_1 = \Sigma a_i \gamma_i$ and $\psi_1 = \Sigma b_j \nu_j$. Then

$$\begin{aligned} \phi_1 \otimes \psi_1 &= (\Sigma a_i \gamma_i) \otimes (\Sigma b_j \nu_j) \\ &= \Sigma\Sigma a_i b_j \gamma_i \otimes \nu_j. \end{aligned}$$

The last point to emphasize is one that will turn out to be **VERY IMPORTANT:** Not every vector in $\mathcal{H}_S$ can be written in the form $\phi \otimes \psi$, where $\phi \in \mathcal{H}_1$ and $\psi \in \mathcal{H}_2$. In fact, if γ_1 is not parallel to γ_2, and ν_1 is not parallel to ν_2, then $\gamma_1 \otimes \nu_1 + \gamma_2 \otimes \nu_2$ cannot be written in this simple product form (proving this is left as an exercise). However, (ii) guarantees that every vector in $\mathcal{H}_S$ can at least be written as a linear combination of such vectors. The next section explains why this point is so important.

VII. Limitations of the representation of states by vectors

The fact that not every vector in $\mathcal{H}_1 \otimes \mathcal{H}_2$ can be written in the form $\phi \otimes \psi$ has the noteworthy consequence that the following four theses are jointly inconsistent:

1. For any system S, every possible state of S is represented by some normalized vector in the associated Hilbert space $\mathcal{H}_S$.

2. For any system S, every normalized vector in the associated Hilbert space $\mathcal{H}_S$ represents some possible state of S.

3. Non-parallel vectors cannot represent the same state.

4. If the state of particle 1 is represented by $\phi \in \mathcal{H}_1$, and the state of particle 2 is represented by $\psi \in \mathcal{H}_2$, then the state of the two-particle system consisting of particles 1 and 2 is represented by $\phi \otimes \psi \in \mathcal{H}_1 \otimes \mathcal{H}_2$.

In a moment, we'll consider which of these four claims ought to be abandoned; but let us first see why they are inconsistent. We need only focus on the two-particle system S consisting of particles 1 and 2.

Let $\phi_1,\phi_2 \in \mathcal{H}_1$ be non-parallel, normalized vectors; likewise, let $\psi_1,\psi_2 \in \mathcal{H}_2$ be non-parallel, normalized vectors. Choose a scalar c such that $c\phi_1\otimes\psi_1 + c\phi_2\otimes\psi_2$ is normalized. Then claim 2 tells us that this vector represents a possible state of S. According to claim 1, every possible state of particle 1 is represented by some normalized vector in $\mathcal{H}_1$, and similarly for particle 2. So consider a possible situation in which the state of S is represented by $c\phi_1\otimes\psi_1 + c\phi_2\otimes\psi_2$: let γ represent the state of particle 1 in that situation, and ν represent the state of particle 2. By claim 4, it follows that $\gamma\otimes\nu$ represents the state of S, in this situation. We know that $\gamma\otimes\nu \neq c\phi_1\otimes\psi_1 + c\phi_2\otimes\psi_2$. But so far, we have no inconsistency, since it is in general possible for distinct vectors to represent the same state. However, not only are the vectors $\gamma\otimes\nu$ and $c\phi_1\otimes\psi_1 + c\phi_2\otimes\psi_2$ non-identical, they are also non-parallel. (Why?) It therefore follows from claim 3 that they do not represent the same state—a contradiction.

So one of the four claims has to go. Which one?

To begin, observe that claim 3 derives it support from the statistical algorithm: If there are sufficiently many observables that we can measure, then there will always be *some* measurement for which non-parallel vectors give different outcome probabilities. But in that case, they must represent different states. In the case at hand, it suffices to produce *one* state of the form $c\phi_1\otimes\psi_1 + c\phi_2\otimes\psi_2$ (with ϕ_1 and ϕ_2 non-parallel, etc.) which can be "separated" in this way from any state of the form $\gamma\otimes\nu$. Physics provides examples aplenty: for instance, the singlet spin state

$\frac{1}{\sqrt{2}}(z^+\Box z^- - z^-\Box z^+)$ has this feature. (We've seen this state before: it's the one that gives rise to the perfectly correlated spin-experiments.) So it looks like claim 3 is not the culprit.

Consider now claim 2. It's not enough merely to deny this claim; one must also state that *no* vector of the form $c\phi_1\otimes\psi_1 + c\phi_2\otimes\psi_2$ represents a possible state of the two-particle system. Again, the theoretical treatment of a host of actual systems violates this requirement. So it looks like claim 2 is not the culprit.

Claim 4 in fact follows from a fundamental principle (to be discussed below) about how the state of a composite system determines the states of its component parts. That principle is not wholly uncontroversial—although as far as I can tell, it's about 97-98% uncontroversial. Suffice it to say that it forms a part of every standard presentation of ordinary, non-relativistic quantum mechanics.

That leaves claim 1 as the one that ought to be abandoned. Here I'll add two comments: First, it was already suspect on experimental grounds. For example, we *seem* to be able to prepare spin-$\frac{1}{2}$ particles in states such that they have a 50% chance of going up, no matter the direction of spin we choose; no vector will reproduce those probabilities, in accordance with the statistical algorithm. (I say "seem" because if the given spin-measurement statistics are the *only* information we have, we can always insist that the measured particles aren't really prepared to be in the same state, but rather that the preparation procedure randomly puts them in, say, either z^+ or z^-; more on this later.) Second, as we'll see in the next section, abandoning it is not that big a deal; there is a natural generalization of the representation of states which has the usual vector representation as a special case. So we don't *lose* anything by giving up claim 1.

VIII. A non-redundant representation of physical states

Thus far, our representation of physical states has been highly redundant. For two vectors represent the *same* state iff they give rise to exactly the same probabilities. (Recall that all we mean by "state of a system"—so far!—is a compilation of the experimental probabilities associated with the system. Given *that* understanding of "state of a system", the claim just stated follows trivially. But even those who invest the quantum mechanical representations of states with more physical significance endorse this claim, at least as long as they aren't being misled by arguments such as van Fraassen's (see Assignment #4).) Thus if ϕ is normalized, it represents the same state as $c\phi$, for any c such that $|c| = 1$, since $\langle c\phi \mid \mathbf{P}c\phi\rangle = c^*c\langle\phi \mid \mathbf{P}\phi\rangle = \langle\phi \mid \mathbf{P}\phi\rangle$, for any projection operator **P**. At the same time, there is rather a lot of waste, since every non-normalized vector represents nothing whatsoever. Let us see whether we can rectify this aesthetic oversight. Along the way, we'll solve the problem introduced in the last section: How can we represent the states of particles 1 and 2, when the two-particle system has the state $c\phi_1\otimes\psi_1 + c\phi_2\otimes\psi_2$?

For certain linear operators **A**—called "trace-class" operators (see Hughes, section 5.1)—the following sum, called the *trace* of **A**, is well-defined and real-valued ($\{\phi_i\}$ is an orthonormal basis for our Hilbert space $\mathcal{H}_S$):

$\mathrm{Tr}(\mathbf{A}) = \Sigma_i\langle\phi_i \mid \mathbf{A}\phi_i\rangle$.

It's easily shown that the value of $\mathrm{Tr}(\mathbf{A})$ does not depend on the choice of orthonormal basis $\{\phi_i\}$. For let $\{\psi_j\}$ be some other orthonormal basis. Since $\{\psi_j\}$ and $\{\phi_i\}$ are both orthonormal bases, for each i, $\phi_i = \Sigma_j\langle\psi_j \mid \phi_i\rangle\psi_j$; likewise, for each j, $\psi_j = \Sigma_i \langle\phi_i \mid \psi_j\rangle\phi_i$. Therefore,

$\Sigma_i \langle\phi_i \mid \mathbf{A}\phi_i\rangle = \Sigma_i < \Sigma_j \langle\psi_j \mid \phi_i\rangle\psi_j \mid \mathbf{A}\phi_i\rangle$

$= \Sigma_{ij} \langle\phi_i \mid \psi_j\rangle\langle\psi_j \mid \mathbf{A}\phi_i\rangle$

$= \Sigma_{ij} \langle\psi_j \mid \mathbf{A}\langle\phi_i \mid \psi_j\rangle\phi_i\rangle$

$= \Sigma_j \langle \psi_j \mid \mathbf{A}(\Sigma_i \langle \phi_i \mid \psi_j \rangle \phi_i) \rangle$

$= \Sigma_j \langle \psi_j \mid \mathbf{A}\psi_j \rangle$.

Next, if **A** and **B** are trace-class operators, then for any scalars a and b, Tr(a**A** + b**B**) = aTr(**A**) + bTr(**B**). (This is more or less obvious.) Finally, we will make use of (but not prove) the fact that if **A** is a trace-class operator and **B** is a bounded linear operator (i.e., $|\mathbf{B}\phi| \leq b|\phi|$, for all ϕ and some real number b), then **AB** is a trace-class operator, and Tr(**AB**) = Tr(**BA**) (see Hughes, pp. 137-8).

Projection operators are bounded linear operators; projection operators onto finite-dimension subspaces are trace-class operators. So let $\mathbf{P}_\alpha$ project onto the subspace of eigenvectors of Hermitian operator **A** with eigenvalue α, and consider the quantity $\mathrm{Tr}(\mathbf{P}_\alpha\mathbf{P}_\phi)$; this is well-defined, since from the last paragraph we know that $\mathbf{P}_\alpha\mathbf{P}_\phi$ is a trace-class operator. To evaluate this, choose an orthonormal basis $\{\phi_i\}$ where $\phi = \phi_k$. Then

$\mathrm{Tr}(\mathbf{P}_\alpha\mathbf{P}_\phi) = \Sigma_i \langle \phi_i \mid \mathbf{P}_\alpha\mathbf{P}_\phi\phi_i \rangle = \langle \phi \mid \mathbf{P}_\alpha\phi \rangle$.

This is familiar: it is just the probability that a measurement of the observable represented by **A** yields outcome α. So, instead of representing states (redundantly) by normalized vectors, we could equally well represent them (non-redundantly) by projection operators onto one-dimensional subspaces, replacing the old algorithm for calculating probabilities of measurement outcomes with this new one:

If a measurement of the observable A (represented by Hermitian operator **A**) is performed on a system S in state $\mathbf{P}_\phi$, then the probability that the outcome α is obtained is equal to $\mathrm{Tr}(\mathbf{P}_\alpha\mathbf{P}_\phi)$, where $\mathbf{P}_\alpha$ is the projection operator onto the subspace of eigenvectors of **A** with eigenvalue α.

In fact, our final, fully general representation of states will not be as one-dimensional projection operators, but as *density operators*. We define a density operator to be any weighted

sum $\Sigma_i\ c_i\mathbf{P}_i$, where the c_i are non-negative real numbers such that $\Sigma_i\ c_i = 1$, and the $\mathbf{P}_i$ are projection operators. Thus, if $\mathcal{H}_S$ is the Hilbert space for system S, then every density operator **D** on $\mathcal{H}_S$ represents a possible physical state of S. We refine the algorithm for calculating outcome probabilities as follows:

If a measurement of the observable A (represented by Hermitian operator **A**) is performed on a system S in state **D**, then the probability that the outcome α is obtained is equal to $\mathrm{Tr}(\mathbf{P}_\alpha\mathbf{D})$, where $\mathbf{P}_\alpha$ is the projection operator onto the subspace of eigenvectors of **A** with eigenvalue α.

If we assume that sufficiently many Hermitian operators represents measurable observable quantities, then the representation of states by density operators will be non-redundant: if $\mathbf{D}_1 \neq \mathbf{D}_2$, that is, there will be some measurement for which $\mathbf{D}_1$ and $\mathbf{D}_2$ yield different outcome probabilities, from which it follows that they must represent different states (I'll omit the proof).

One important bit of terminology: If there is some ψ such that $\mathbf{D} = \mathbf{P}_\psi$, then we say that **D** is a *pure* state; if *not*, we say that **D** is a *mixed* state. So far, pure states are the only states we've been dealing with; so it might seem at this point that there's no need for mixed states. Next, we'll see that there is more than aesthetics at work in choosing this representation of states: mixed states are in fact indispensable.

Let **A** be an Hermitian operator on $\mathcal{H}_1$ representing some observable pertaining to particle 1 alone (e.g., the particle's spin in some direction). Let **I** be the identity operator on $\mathcal{H}_2$. Then $\mathbf{A}\otimes\mathbf{I}$ is an Hermitian operator on $\mathcal{H}_S = \mathcal{H}_1\otimes\mathcal{H}_2$. We now make the fundamental assumption that **A** and $\mathbf{A}\otimes\mathbf{I}$ *represent the same observable*, in the sense that *either* can be used to calculate outcome probabilities for a measurement of this observable. More precisely, suppose that $\mathbf{D}_S$ is a density

operator on $\mathcal{H}_S$ that represents the state of S, and $\mathbf{D}_1$ is a density operator on $\mathcal{H}_1$ that represents the state of particle 1. Then we require that the following hold:

For all Hermitian operators $\mathbf{A}$ on $\mathcal{H}_1$, $\mathrm{Tr}(\mathbf{AD}_1) = \mathrm{Tr}((\mathbf{A}\otimes\mathbf{I})\mathbf{D}_S)$.

This is sometimes called the *Reduction of States* principle (restricted, here, to the particular case of a two-particle system).

Recall the questions raised above: (i) How is $\mathcal{H}_S$ related to each of the $\mathcal{H}_i$? (ii) How are states of S related to states of the individual particles that make it up? (iii) In particular, is the state of S fixed, once we know the states of the particles? (iv) Conversely, are the states of the particles fixed, once we know the state of S? We can now answer each of these.

(i) In the general case of an n-particle system S, $\mathcal{H}_S$ is the tensor product of the $\mathcal{H}_i$:

$\mathcal{H}^S = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \ldots \otimes \mathcal{H}_n$.

(ii) The states of S are related to the states of the particles that compose it via the Reduction of States principle.

(iii) The state of S is *not* fixed, once we know the states of the individual particles.

To see this, consider again the case of the two-particle system. Choose orthogonal, normalized vectors $\phi_1,\phi_2 \in \mathcal{H}_1$ and $\psi_1,\psi_2 \in \mathcal{H}_2$. Let $\Psi = \frac{1}{\sqrt{2}}(\phi_1\square\psi_1 + \phi_2\square\psi_2)$. Let $\mathbf{D}_1 = \frac{1}{2}(\mathbf{P}_{\phi_1} + \mathbf{P}_{\phi_2})$; $\mathbf{D}_2 = \frac{1}{2}(\mathbf{P}_{\psi_1} + \mathbf{P}_{\psi_2})$. Then for any Hermitian operator $\mathbf{A}$ on $\mathcal{H}_1$ and $\mathbf{B}$ on $\mathcal{H}_2$,

$\mathrm{Tr}(\mathbf{AD}_1) = \frac{1}{2}(\langle\phi_1 \mid \mathbf{A}\phi_1\rangle + \langle\phi_2 \mid \mathbf{A}\phi_2\rangle)$; and

$\mathrm{Tr}(\mathbf{BD}_2) = \frac{1}{2}(\langle\psi_1 \mid \mathbf{B}\psi_1\rangle + \langle\psi_2 \mid \mathbf{B}\psi_2\rangle)$.

Now suppose that $\mathbf{D}_S = \mathbf{P}_\Psi$. Then, as is easily verified,

$\mathrm{Tr}((\mathbf{A}\otimes\mathbf{I})\mathbf{D}_S) = \langle\Psi \mid (\mathbf{A}\otimes\mathbf{I})\Psi\rangle = \frac{1}{2}(\langle\phi_1 \mid \mathbf{A}\phi_1\rangle + \langle\phi_2 \mid \mathbf{A}\phi_2\rangle)$; and

$\mathrm{Tr}((\mathbf{I}\otimes\mathbf{B})\mathbf{D}_S) = \langle\Psi \mid (\mathbf{I}\otimes\mathbf{B})\Psi\rangle = \frac{1}{2}(\langle\psi_1 \mid \mathbf{B}\psi_1\rangle + \langle\psi_2 \mid \mathbf{B}\psi_2\rangle)$.

Invoking the Reduction principle, we conclude that if the state of S is represented by $\mathbf{D}_S = \mathbf{P}_\Psi$, then the state of particle 1 is $\mathbf{D}_1$, and the state of particle 2 is $\mathbf{D}_2$.

Suppose on the other hand that $\mathbf{D}_S = \frac{1}{2}(\mathbf{P}_{\phi_1\Box\psi_1} + \mathbf{P}_{\phi_2\Box\psi_2})$. Then, again, it is easy to verify that $\mathrm{Tr}((\mathbf{A}\otimes\mathbf{I})\mathbf{D}_S) = \frac{1}{2}(\langle\phi_1 \mid \mathbf{A}\phi_1\rangle + \langle\phi_2 \mid \mathbf{A}\phi_2\rangle)$, etc. Again invoking the Reduction principle, we conclude that if the state of S is represented by $\mathbf{D}_S = \frac{1}{2}(\mathbf{P}_{\phi_1\Box\psi_1} + \mathbf{P}_{\phi_2\Box\psi_2})$, then the state of particle 1 is $\mathbf{D}_1$, and the state of particle 2 is $\mathbf{D}_2$. But these two possible states of S are *not* equivalent, since there are Hermitian operators on $\mathcal{H}_S$ for which they give different probabilities. For example, if $\mathbf{D}_S = \mathbf{P}_\Psi$, then $\mathrm{Tr}(\mathbf{P}_\Psi\mathbf{D}_S) = 1$; but if $\mathbf{D}_S = \frac{1}{2}(\mathbf{P}_{\phi_1\Box\psi_1} + \mathbf{P}_{\phi_2\Box\psi_2})$, then $\mathrm{Tr}(\mathbf{P}_\Psi\mathbf{D}_S) = \frac{1}{2}$. So if we knew only that particle 1 had state $\mathbf{D}_1$ and particle 2 state $\mathbf{D}_2$, we wouldn't yet know the state of the composite system S.

(iv) However, the state of S *does* determine the states of the particles which compose it. Again, consider the two-particle system. Suppose the state of S is $\mathbf{D}_S$, and suppose that, consistent with the Reduction principle, the state of particle 1 could be either **D** or **D**'. That is, for every Hermitian operator **A** on $\mathcal{H}_1$, $\mathrm{Tr}((\mathbf{A}\otimes\mathbf{I})\mathbf{D}_S) = \mathrm{Tr}(\mathbf{AD})$ *and* $\mathrm{Tr}((\mathbf{A}\otimes\mathbf{I})\mathbf{D}_S) = \mathrm{Tr}(\mathbf{AD}')$. Then for every such **A**, $\mathrm{Tr}(\mathbf{AD}) = \mathrm{Tr}(\mathbf{AD}')$. Hence **D** and **D**' must represent the same state; further, since the representation in terms of density operators is *nonredundant*, **D** = **D**'.