

A Few Notes on Density Operators, Expectation Values and Matrix Shapes

David Meyer

dmm@{1-4-5.net,uoregon.edu}

Last update: January 18, 2017

1 Introduction

These notes started life as an experiment in drawing matrices and their shapes (see Section 4). However, it has evolved into a more ad-hoc collection of notes covering a few topics in quantum mechanics. So its a WIP. We start with a review of Orthonormality, Completeness, and Projection...

2 Orthonormality, Completeness, and Projection

As we saw above, unitary matrices are matrices which satisfy

$$\mathbf{U}^{-1} = \mathbf{U}^\dagger \tag{1}$$

Unitary matrices are ubiquitous and important in quantum mechanics, in particular because they have the following unique and useful properties: Orthonormality, Completeness, and Projection [3]. We'll briefly look at each of these below¹.

2.1 Orthonormality

We can rewrite Equation 1 as

$$\mathbf{U}^\dagger \mathbf{U} = \mathbf{I} \tag{2}$$

where \mathbf{I} is the *identity* matrix. What Equation 2 is really telling us is that the columns of the matrix \mathbf{U} form a set of orthonormal vectors.

¹I will use the notation $(x_1, \dots, x_n)^T$ and $[x_1, \dots, x_n]^T$ interchangeably in the following discussion.

Note that we can interpret a matrix as a row vector where the entries are the columns \mathbf{v}_i of \mathbf{U} . That is

$$\mathbf{U} = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \dots \quad \mathbf{v}_N]$$

Similarly, \mathbf{U}^{-1} can be written as a column vector where the entries are the row vectors \mathbf{v}_i^\dagger :

$$\mathbf{U}^{-1} = \mathbf{U}^\dagger = \begin{bmatrix} \mathbf{v}_1^\dagger \\ \mathbf{v}_2^\dagger \\ \vdots \\ \mathbf{v}_N^\dagger \end{bmatrix}$$

Now we can see that

$$\begin{aligned} \mathbf{U}^\dagger \mathbf{U} &= \begin{bmatrix} \mathbf{v}_1^\dagger \\ \mathbf{v}_2^\dagger \\ \vdots \\ \mathbf{v}_N^\dagger \end{bmatrix} [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \dots \quad \mathbf{v}_N] \\ &= \begin{bmatrix} \mathbf{v}_1^\dagger \cdot \mathbf{v}_1 & \mathbf{v}_1^\dagger \cdot \mathbf{v}_2 & \mathbf{v}_1^\dagger \cdot \mathbf{v}_3 & \dots & \mathbf{v}_1^\dagger \cdot \mathbf{v}_N \\ \mathbf{v}_2^\dagger \cdot \mathbf{v}_1 & \mathbf{v}_2^\dagger \cdot \mathbf{v}_2 & \mathbf{v}_2^\dagger \cdot \mathbf{v}_3 & \dots & \mathbf{v}_2^\dagger \cdot \mathbf{v}_N \\ \mathbf{v}_3^\dagger \cdot \mathbf{v}_1 & \mathbf{v}_3^\dagger \cdot \mathbf{v}_2 & \mathbf{v}_3^\dagger \cdot \mathbf{v}_3 & \dots & \mathbf{v}_3^\dagger \cdot \mathbf{v}_N \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{v}_N^\dagger \cdot \mathbf{v}_1 & \mathbf{v}_N^\dagger \cdot \mathbf{v}_2 & \mathbf{v}_N^\dagger \cdot \mathbf{v}_3 & \dots & \mathbf{v}_N^\dagger \cdot \mathbf{v}_N \end{bmatrix} \\ &= \mathbf{I} \end{aligned}$$

or in Dirac notation [2]

$$\begin{aligned}
\mathbf{U}^\dagger \mathbf{U} &= \begin{bmatrix} \mathbf{v}_1^\dagger \\ \mathbf{v}_2^\dagger \\ \vdots \\ \mathbf{v}_N^\dagger \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_N \end{bmatrix} \\
&= \begin{bmatrix} \langle v_1 | \\ \langle v_2 | \\ \vdots \\ \langle v_N | \end{bmatrix} \begin{bmatrix} |v_1\rangle & |v_2\rangle & \dots & |v_N\rangle \end{bmatrix} \\
&= \begin{bmatrix} \langle v_1 | v_1 \rangle & \langle v_1 | v_2 \rangle & \langle v_1 | v_3 \rangle & \dots & \langle v_1 | v_N \rangle \\ \langle v_2 | v_1 \rangle & \langle v_2 | v_2 \rangle & \langle v_2 | v_3 \rangle & \dots & \langle v_2 | v_N \rangle \\ \langle v_3 | v_1 \rangle & \langle v_3 | v_2 \rangle & \langle v_3 | v_3 \rangle & \dots & \langle v_3 | v_N \rangle \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \langle v_N | v_1 \rangle & \langle v_N | v_2 \rangle & \langle v_N | v_3 \rangle & \dots & \langle v_N | v_N \rangle \end{bmatrix} \\
&= \mathbf{I}
\end{aligned}$$

What we can notice² here is that since $(\mathbf{U}^\dagger \mathbf{U})_{ij} = (\mathbf{U}^{-1} \mathbf{U})_{ij} = \delta_{ij}$, the columns of \mathbf{U} can be written as the inner product $\langle v_i | v_j \rangle = \delta_{ij}$. Said another way, the vectors v_i form an orthonormal set. In particular, if $\mathbf{V} = \{v_j\}$ is an orthonormal set, then for $v_i, v_j \in \mathbf{V}$, the inner product $\langle v_i | v_j \rangle = \delta_{ij}$. See Section 4 for a brief discussion on matrix shapes.

2.2 Completeness

From $\mathbf{U}^\dagger \mathbf{U} = \mathbf{I}$ we saw that we could derive orthonormality. But we also expect that $\mathbf{U} \mathbf{U}^\dagger = \mathbf{I}$. It turns out that we can get something interesting by observing this. In particular

$$\mathbf{U} \mathbf{U}^\dagger = \begin{bmatrix} |v_1\rangle & |v_2\rangle & |v_3\rangle & \dots & |v_N\rangle \end{bmatrix} \begin{bmatrix} \langle v_1 | \\ \langle v_2 | \\ \langle v_3 | \\ \vdots \\ \langle v_N | \end{bmatrix}$$

² δ_{ij} is the Kronecker Delta function [4], $\delta_{ij} = \begin{cases} 1 & \text{when } i = j \\ 0 & \text{when } i \neq j \end{cases}$

If we multiply this out we find that

$$|v_1\rangle \langle v_1| + |v_2\rangle \langle v_2| + \cdots + |v_N\rangle \langle v_N| = \sum_{i=1}^N |v_i\rangle \langle v_i| = \mathbf{I} \quad (3)$$

Equation 3 is known as the *completeness* relation.

Completeness turns out to be useful and is a sort of a "dual" of orthonormality. While orthonormality is kind of an "inner product" ($\mathbf{U}^\dagger \mathbf{U}$), completeness is like an outer product in that $\mathbf{U} \mathbf{U}^\dagger$ is a sum over i of $|v_i\rangle \langle v_i|$, although the shapes might be seen as reversed (see Section 4 on shapes).

2.3 Projection

To get an idea of what projection is all about, consider the expansion of a vector into components in a basis:

$$|w\rangle = \sum_{i=1}^N w_i |v_i\rangle \quad (4)$$

Now, if the set of vectors basis vectors $\{v_i\}$ are orthonormal, then we know that

$$w_i = \langle v_i | w \rangle$$

and substituting back into Equation 4 we get

$$|w\rangle = \sum_{i=1}^N \langle v_i | w \rangle |v_i\rangle$$

Interestingly, there is another way to derive this result: use the completeness relation, which is simply a fancy but useful way to write \mathbf{I} :

$$|w\rangle = \mathbf{I} \cdot |w\rangle = \left(\sum_{i=1}^N |v_i\rangle \langle v_i| \right) |w\rangle = \sum_{i=1}^N |v_i\rangle \langle v_i | w \rangle$$

In words, we were able to use the completeness relation to project a vector onto its components in a particular basis.

For example, we know that for vectors $|\alpha\rangle$ and $|\beta\rangle$, we can take the inner product between them by using their components in a basis $\{v_i\}$:

$$\langle \alpha | \beta \rangle = \sum_{i=1}^N a_i^* b_i$$

where $a_i = \langle v_i | \alpha \rangle$ and $b_i = \langle v_i | \beta \rangle$. Interestingly, we can again derive this using the completeness relation:

$$\begin{aligned} \langle \alpha | \beta \rangle &= \langle \alpha | \mathbf{I} | \beta \rangle & \# \langle \alpha | \beta \rangle &= \langle \alpha | | \beta \rangle = \langle \alpha | \mathbf{I} | \beta \rangle \\ &= \langle \alpha | \left(\sum_{i=1}^N |v_i\rangle \langle v_i| \right) | \beta \rangle & \# \sum_{i=1}^N |v_i\rangle \langle v_i| &= \mathbf{I} \text{ (Equation 3)} \\ &= \sum_{i=1}^N \langle \alpha | v_i \rangle \langle v_i | \beta \rangle & \# &\text{rearrange} \\ &= \sum_{i=1}^N \langle v_i | \alpha \rangle^* \langle v_i | \beta \rangle & \# \langle a | b \rangle &= \langle b | a \rangle^* \text{ so } \langle \alpha | v_i \rangle = \langle v_i | \alpha \rangle^* \\ &= \sum_{i=1}^N a_i^* b_i & \# a_i^* &= \langle v_i | \alpha \rangle^* \text{ and } b_i = \langle v_i | \beta \rangle \end{aligned}$$

3 Expectation Values

Consider an observable \mathbf{A} in the pure state $|\psi\rangle$. The expectation value $\langle A \rangle_\psi$ is given by

$$\langle A \rangle_\psi = \langle \psi | A | \psi \rangle \quad (5)$$

where $\dim(|\psi\rangle) = 1 \times n$, $\dim(A) = n \times n$, and $\dim(|\psi\rangle) = n \times 1$.

So why is $\langle A \rangle_\psi$ an expectation? Well, first, if A is an observable for a system with a discrete set of values $\{a_1, a_2, \dots, a_N\}$, then this observable is represented by a Hermitean operator \hat{A} that has these discrete values as its eigenvalues, and associated eigenstates $\{|a_n\rangle\}$, for $n = 1, 2, 3, \dots$ satisfying the eigenvalue equation $\hat{A} |a_n\rangle = a_n |a_n\rangle$. I drop the "hat" in most of the below.

First, observe that $\langle a_n | A = a_n \langle a_n |$. Why?

$$\begin{aligned}
A |a_n\rangle &= a_n |a_n\rangle && \# \text{ eigenvalue equation for } A \ (A\mathbf{v} = \lambda\mathbf{v}) \\
\implies (A |a_n\rangle)^\dagger &= (a_n |a_n\rangle)^\dagger && \# \text{ conjugate transpose both sides} \\
\implies |a_n\rangle^\dagger A^\dagger &= |a_n\rangle^\dagger a_n^\dagger && \# (AB)^\dagger = B^\dagger A^\dagger \\
\implies |a_n\rangle^\dagger A^\dagger &= a_n^\dagger |a_n\rangle^\dagger && \# \text{ rearrange } (a_n^\dagger \text{ is a scalar}) \\
\implies |a_n\rangle^\dagger A &= a_n^\dagger |a_n\rangle^\dagger && \# A \text{ is Hermitean so } A = A^\dagger \\
\implies |a_n\rangle^\dagger A &= a_n^* |a_n\rangle^\dagger && \# a_n^\dagger = a_n^* \ (a_n \text{ is a scalar}) \\
\implies \langle a_n | A &= a_n^* \langle a_n | && \# |a_n\rangle^\dagger = \langle a_n | \\
\implies \langle a_n | A &= a_n \langle a_n | && \# a_n^* = a_n
\end{aligned} \tag{6}$$

But why does $a_n^* = a_n$ (last line of (6))? Well, consider

$$\begin{aligned}
AX &= \lambda X && \# \text{ eigenvalue equation} \\
\implies X^\dagger A^\dagger &= X^\dagger \lambda^\dagger && \# (AB)^\dagger = B^\dagger A^\dagger \\
\implies X^\dagger A^\dagger &= \lambda^\dagger X^\dagger && \# \text{ rearrange } (\lambda^\dagger \text{ is a scalar}) \\
\implies X^\dagger A^\dagger &= \lambda^* X^\dagger && \# \lambda^\dagger = \lambda^* \ (\lambda \text{ is a scalar}) \\
\implies X^\dagger A &= \lambda^* X^\dagger && \# A^\dagger = A \text{ since } A \text{ is Hermitean} \\
\implies X^\dagger A &= X^\dagger \lambda^* && \# \text{ rearrange} \\
\implies X^\dagger AX &= X^\dagger \lambda^* X && \# \text{ multiply both sides by } X
\end{aligned} \tag{7}$$

Now notice that if we multiply both sides of the original eigenvalue equation ($AX = \lambda X$) by X^\dagger we get $X^\dagger AX = X^\dagger \lambda X$. We know from (7) that $X^\dagger AX = X^\dagger \lambda^* X$ and therefore that $X^\dagger \lambda^* X = X^\dagger \lambda X$. This implies that $\lambda^* = \lambda$, so $\lambda \in \mathbb{R}$. Similarly $a_n^* = a_n$ so $a_n \in \mathbb{R}$.

Another way to look at this is to assume the computational basis³ and then

³The approach taken in (6) doesn't seem to require this assumption.

$$\begin{aligned}
\langle a_n | A &= a_n \langle n | A & \# \langle a_n | &= a_n [0 \quad \dots \quad 1 \quad \dots 0] = a_n \langle n | \\
&= a_n \langle n | A^\dagger & \# A &\text{ is Hermitian so } A = A^\dagger \\
&= a_n \langle n | \begin{bmatrix} \langle a_1 | \\ \vdots \\ \langle a_n | \\ \vdots \\ \langle a_N | \end{bmatrix} & \# A^\dagger &= \begin{bmatrix} \langle a_1 | \\ \vdots \\ \langle a_n | \\ \vdots \\ \langle a_N | \end{bmatrix} \\
&= a_n [0 \quad \dots \quad 1 \quad \dots 0] \begin{bmatrix} \langle a_1 | \\ \vdots \\ \langle a_n | \\ \vdots \\ \langle a_N | \end{bmatrix} & \# \langle n | &= [0 \quad \dots \quad 1 \quad \dots 0] \\
&= a_n \langle a_n | & \# \langle n | &\text{ selects the } n^{th} \text{ element of } A^\dagger, \langle a_n |
\end{aligned}$$

In any event, now we have $\langle a_n | A = a_n \langle a_n |$. So we can observe that

$$\begin{aligned}
\langle A \rangle_\psi &= \langle \psi | A | \psi \rangle & \# \text{ definition of } \langle A \rangle_\psi \text{ for } \textit{pure} \text{ state } |\psi\rangle \\
&= \langle \psi | I A | \psi \rangle & \# I \cdot A = A \\
&= \langle \psi | \left(\sum_{n=1}^N |a_n\rangle \langle a_n| \right) A | \psi \rangle & \# \sum_{n=1}^N |a_n\rangle \langle a_n| = \mathbf{I} \text{ (Equation 3)} \\
&= \sum_{n=1}^N \langle \psi | a_n \rangle \langle a_n | A | \psi \rangle & \# \text{ rearrange} \\
&= \sum_{n=1}^N \langle \psi | a_n \rangle a_n \langle a_n | \psi \rangle & \# \langle a_n | A = a_n \langle a_n | \text{ (see above)} \\
&= \sum_{n=1}^N \langle \psi | a_n \rangle \langle a_n | \psi \rangle a_n & \# \text{ rearrange} \\
&= \sum_{n=1}^N |\langle \psi | a_n \rangle|^2 a_n & \# |\langle \psi | a_n \rangle|^2 = \langle \psi | a_n \rangle \langle \psi | a_n \rangle^* = \langle \psi | a_n \rangle \langle a_n | \psi \rangle \\
&= \sum_{n=1}^N p(a_n) a_n & \# |\langle \psi | a_n \rangle|^2 = p(a_n), \text{ the probability of observing eigenvalue } a_n \\
&= \sum_{n=1}^N \frac{N_n}{N} a_n & \# N_n \text{ is the number of times } a_n \text{ has been measured} \\
&= \mathbb{E}[A] & \# \mathbb{E}[X] = \sum_{n=1}^N p(X_n) X_n
\end{aligned}$$

So the expectation value for the result of a measurement represented by a self-adjoint operator A , $\langle A \rangle_\psi$, is the weighted average of all possible outcomes under A , that is, $\mathbb{E}[A]$.

4 Shapes

One way to visualize $\langle A \rangle_\psi$ is

$$\begin{aligned} \langle A \rangle_\psi &\rightarrow \underbrace{[\dots\dots\dots]}_{1 \times n} \underbrace{\begin{bmatrix} \dots & \dots & \dots & \dots \\ \vdots & \ddots & \vdots & \\ \dots & \dots & \dots & \dots \end{bmatrix}}_{n \times n} \underbrace{\begin{bmatrix} \dots \\ \vdots \\ \dots \end{bmatrix}}_{n \times 1} \\ &\rightarrow c \end{aligned}$$

where $c \in \mathbb{C}$.

The *density operator* ρ for pure state $|\psi\rangle$ is given by $\rho = |\psi\rangle \langle\psi|$. The shape of ρ is

$$\rho \rightarrow \underbrace{\begin{bmatrix} \dots \\ \vdots \\ \dots \end{bmatrix}}_{n \times 1} \underbrace{[\dots\dots\dots]}_{1 \times n} \rightarrow \underbrace{\begin{bmatrix} \dots & \dots & \dots & \dots \\ \vdots & \ddots & \vdots & \\ \dots & \dots & \dots & \dots \end{bmatrix}}_{n \times n}$$

The shape of the inner product of two $n \times 1$ column vectors $\langle \mathbf{u}, \mathbf{v} \rangle = \langle u|v \rangle = \mathbf{u}^T \mathbf{v}$ is

$$\mathbf{u}^T \mathbf{v} \rightarrow \underbrace{[\dots\dots\dots]}_{1 \times n} \underbrace{\begin{bmatrix} \dots \\ \vdots \\ \dots \end{bmatrix}}_{n \times 1} \rightarrow c$$

where $c \in \mathbb{C}$. The shape of the outer product $\mathbf{u} \otimes \mathbf{v} = |u\rangle \langle v| = \mathbf{u} \mathbf{v}^T$ is

$$\mathbf{u} \mathbf{v}^T \rightarrow \underbrace{\begin{bmatrix} \dots \\ \vdots \\ \dots \end{bmatrix}}_{n \times 1} \underbrace{[\dots\dots\dots]}_{1 \times n} \rightarrow \underbrace{\begin{bmatrix} \dots & \dots & \dots & \dots \\ \vdots & \ddots & \vdots & \\ \dots & \dots & \dots & \dots \end{bmatrix}}_{n \times n}$$

5 The Density ρ and the Trace of an Operator D

So ρ is an $n \times n$ linear operator with $\text{Tr}(\rho) = \text{Tr}(|\psi\rangle\langle\psi|) = \langle\psi|\psi\rangle$. In addition, $\text{Tr}(|\psi_i\rangle\langle\psi_i|) = \langle\psi_i|\psi_i\rangle = \delta_{ii} = 1$, and if $\{|\psi_i\rangle\}$ is an orthonormal basis then $\text{Tr}(|\psi_i\rangle\langle\psi_j|) = \langle\psi_i|\psi_j\rangle = \delta_{ij}$.

The density matrix [1] ρ has the following important properties:

Projection:	$\rho^2 = \rho$
Hermiticity:	$\rho^\dagger = \rho$
Normalization:	$\text{Tr}(\rho) = 1$
Positivity:	$\rho \geq 1$

The *trace* of an operator D , $\text{Tr}(D)$, is defined to be $\text{Tr}(D) = \sum_{i=1}^n \langle n| D |n\rangle$. Now, suppose $D = |\psi\rangle\langle\phi|$. Then we can see that $\text{Tr}(D) = \text{Tr}(|\psi\rangle\langle\phi|) = \langle\phi|\psi\rangle$ as follows:

$$\begin{aligned}
 \text{Tr}(D) &= \sum_{n=1}^N \langle n| D |n\rangle && \# \text{ definition of } \text{Tr}(D) \\
 &= \sum_{n=1}^N \langle n| (|\psi\rangle\langle\phi|) |n\rangle && \# D = |\psi\rangle\langle\phi| \\
 &= \sum_{n=1}^N \langle n| |\psi\rangle\langle\phi| |n\rangle && \# \text{ drop parens} \\
 &= \sum_{n=1}^N \langle n|\psi\rangle\langle\phi|n\rangle && \# \langle a|b\rangle = \langle a| |b\rangle \\
 &= \sum_{n=1}^N \langle\phi|n\rangle\langle n|\psi\rangle && \# \text{ rearrange} \\
 &= \langle\phi| \left(\sum_{n=1}^N |n\rangle\langle n| \right) |\psi\rangle && \# \text{ neither } \phi \text{ nor } \psi \text{ depend on } n \\
 &= \langle\phi| I |\psi\rangle && \# \sum_{n=1}^N |n\rangle\langle n| = \mathbf{I} \text{ (Equation 3)} \\
 &= \langle\phi| |\psi\rangle && \# \langle\phi| I = \langle\phi| \text{ and } I |\psi\rangle = |\psi\rangle \\
 &= \langle\phi|\psi\rangle && \# \langle\phi|\psi\rangle = \langle\phi| |\psi\rangle
 \end{aligned}$$

So the trace of the outer product $|\psi\rangle\langle\phi|$, $\text{Tr}(|\psi\rangle\langle\phi|)$, is the inner product $\langle\phi|\psi\rangle$.

A simple theorem relates the expectation value of an observable A in a state represented by a density matrix ρ to the trace of A :

$$\langle A \rangle_\rho = \text{Tr}(\rho A) \tag{8}$$

The proof of Equation 8 is also pretty simple:

$$\begin{aligned}
\text{Tr}(\rho A) &= \text{Tr}(|\psi\rangle \langle\psi| A) && \# \rho \equiv |\psi\rangle \langle\psi| \\
&= \sum_{n=1}^N \langle n| |\psi\rangle \langle\psi| A |n\rangle && \# \text{definition of } \text{Tr}(\cdot) \\
&= \sum_{n=1}^N \langle n|\psi\rangle \langle\psi| A |n\rangle && \# \langle n|\psi\rangle = \langle n| |\psi\rangle \\
&= \sum_{n=1}^N \langle\psi| A |n\rangle \langle n|\psi\rangle && \# \text{rearrange} \\
&= \langle\psi| A \left(\sum_{n=1}^N |n\rangle \langle n| \right) |\psi\rangle && \# \text{neither } A \text{ nor } \psi \text{ depend on } n \\
&= \langle\psi| A \cdot I |\psi\rangle && \# \sum_{n=1}^N |n\rangle \langle n| = \mathbf{I} \text{ (Equation 3)} \\
&= \langle\psi| A |\psi\rangle && \# \mathbf{A} \cdot \mathbf{I} = \mathbf{A} \\
&= \langle A \rangle_\psi && \# \langle A \rangle_\psi = \langle\psi| A |\psi\rangle \text{ (Equation 5)}
\end{aligned}$$

6 A More General View of the Density Operator

Consider an ensemble of identical quantum systems. The system has probability w_i to be in quantum state $|\psi_i\rangle$. Here $\langle\psi_i|\psi_i\rangle = 1$, but the states $|\psi_i\rangle$ aren't necessarily orthogonal to one another. That means that out of all the examples in the ensemble, a fraction w_i are in state $|\psi_i\rangle$, with $w_i > 0$ and $\sum_i w_i = 1$.

The expectation value for the result of a measurement represented by a self-adjoint operator A is

$$\langle A \rangle_\psi = \sum_i w_i \langle\psi_i| A |\psi_i\rangle \quad (9)$$

We can write the expectation value in a different way using a basis $|K\rangle$ as

$$\begin{aligned}
\langle A \rangle_\psi &= \sum_i w_i \langle \psi_i | A | \psi_i \rangle && \# \text{ definition of } \langle A \rangle_\psi, \text{ Equation 9} \\
&= \sum_i w_i \langle \psi_i | I A I | \psi_i \rangle && \# \mathbf{A} = \mathbf{I} \cdot \mathbf{A} \cdot \mathbf{I} \\
&= \sum_i w_i \langle \psi_i | \left(\sum_J |J\rangle \langle J| \right) A \left(\sum_K |K\rangle \langle K| \right) | \psi_i \rangle && \# \sum_J |J\rangle \langle J| = \mathbf{I}, \sum_K |K\rangle \langle K| = \mathbf{I} \\
&= \sum_i w_i \sum_{J,K} \langle \psi_i | J \rangle \langle J | A | K \rangle \langle K | \psi_i \rangle && \# \text{ rearrange} \\
&= \sum_i w_i \sum_{J,K} \langle K | \psi_i \rangle \langle \psi_i | J \rangle \langle J | A | K \rangle && \# \text{ rearrange} \\
&= \sum_{J,K} \sum_i w_i \langle K | \psi_i \rangle \langle \psi_i | J \rangle \langle J | A | K \rangle && \# \text{ none of } A, J, \text{ or } K \text{ depend on } i \\
&= \sum_{J,K} \langle K | \left(\sum_i w_i | \psi_i \rangle \langle \psi_i | \right) | J \rangle \langle J | A | K \rangle && \# \text{ rearrange} \\
&= \sum_{J,K} \langle K | \rho | J \rangle \langle J | A | K \rangle && \# \rho \equiv \sum_i w_i | \psi_i \rangle \langle \psi_i | \\
&= \sum_K \langle K | \rho I A | K \rangle && \# \sum_J |J\rangle \langle J| = \mathbf{I} \\
&= \sum_K \langle K | \rho A | K \rangle && \# \mathbf{I} \cdot \mathbf{A} = \mathbf{A} \\
&= \text{Tr}(\rho A) && \# \text{Tr}(D) = \sum_n \langle n | D | n \rangle
\end{aligned}$$

6.1 Properties of the Density Operator

As mentioned above, there are several important properties of the density operator ρ . The first of these is that $\text{Tr}(\rho) = 1$. This follows from w_i has $w_i > 0$ and $\sum_i w_i = 1$.

Next, ρ is self-adjoint: $\rho^\dagger = \rho$. Because it is self-adjoint, ρ has eigenvectors $|J\rangle$ with eigenvalues λ_J and the eigenvectors form a basis for vector space. Thus ρ has a standard spectral representation

$$\rho = \sum_J \lambda_J |J\rangle \langle J|$$

We can express λ_J as $\lambda_J = \langle J | \rho | J \rangle$. Then

$$\begin{aligned}
\lambda_J &= \langle J | \rho | J \rangle && \# \\
&= \langle J | \left(\sum_i w_i | \psi_i \rangle \langle \psi_i | \right) | J \rangle && \# \rho = \sum_i w_i | \psi_i \rangle \langle \psi_i | \\
&= \sum_i w_i \langle J | \psi_i \rangle \langle \psi_i | J \rangle && \# \text{ rearrange} \\
&= \sum_i w_i \langle J | \psi_i \rangle \langle J | \psi_i \rangle^* && \# \langle J | \psi_i \rangle^* = \langle \psi_i | J \rangle \\
&= \sum_i w_i | \langle J | \psi_i \rangle |^2 && \# \langle J | \psi_i \rangle \langle J | \psi_i \rangle^* = | \langle J | \psi_i \rangle |^2
\end{aligned}$$

Since $w_i > 0$ and $|\langle J|\psi_i\rangle|^2 > 0$, each eigenvalue must be non-negative, that is, $\lambda_J \geq 0$. In addition, the trace of ρ is the sum of its eigenvalues, so $\sum_J \lambda_J = 1$. Since each eigenvalue is non-negative, $\lambda_J \leq 1$.

Another way to see why $|\langle a_n|\psi\rangle|^2 = p(a_n)$:

$$\begin{aligned} |\psi\rangle &= I|\psi\rangle & \# \mathbf{I} \cdot \mathbf{X} &= \mathbf{X} \\ &= \sum_n |a_n\rangle \langle a_n| |\psi\rangle & \# \sum_n |a_n\rangle \langle a_n| &= I \\ &= \sum_n |a_n\rangle \langle a_n|\psi\rangle & \# \langle a_n| |\psi\rangle &= \langle a_n|\psi\rangle \end{aligned}$$

So $\langle a_n|\psi\rangle$ is the amplitude of $|a_n\rangle$, making $|\langle a_n|\psi\rangle|^2 = p(a_n)$.

7 Acknowledgements

References

- [1] Frank Porter. Physics 125c Course Notes: Density Matrix Formalism. <http://www.cithec.caltech.edu/~fcp/physics/quantumMechanics/densityMatrix/densityMatrix.pdf>, 2011. [Online; accessed 20-Dec-2018].
- [2] F. Gieres. Mathematical surprises and Dirac’s formalism in quantum mechanics. *Reports on Progress in Physics*, 63:1893–1931, December 2000.
- [3] J. D. Cresser. Probability, expectation value and uncertainty. <http://physics.mq.edu.au/~jcresser/Phys301/Chapters/Chapter14.pdf>, 2007. [Online; accessed 20-Dec-2018].
- [4] Wikipedia contributors. Kronecker delta — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Kronecker_delta&oldid=862709627, 2018. [Online; accessed 11-December-2018].