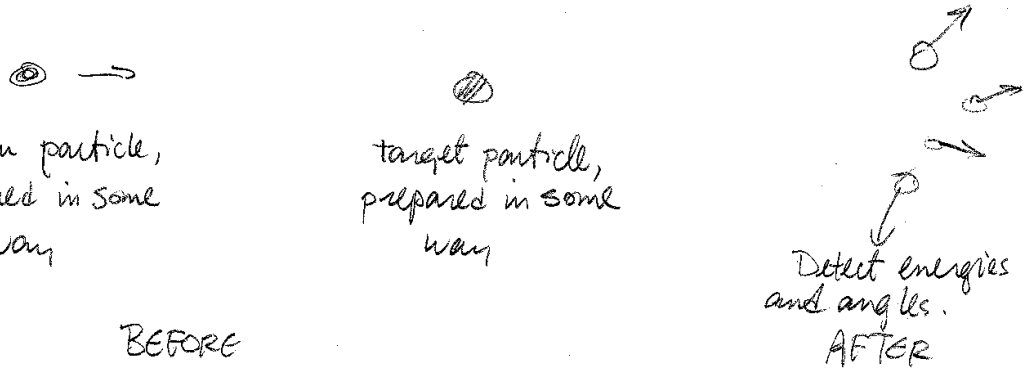
Chap. 2 Formulation of Quantum Mechanics

We are now armed with a certain amount of mathematical machinery or at least vocabulary. Let's now start as close to the "beginning" of physics as we can, and see how we would describe it mathematically.

Postulate 1 To each dynamical variable [i.e. physical observable] there corresponds a linear operator, and the possible values of the dynamical variable are the eigenvalues of the operator.

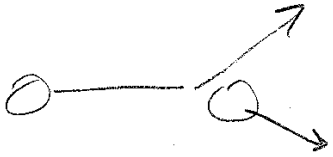
This is reasonably basic. As we know, the physical observables have discrete and continuous values (bound states, free translation) which can correspond to eigenvalues. This postulate doesn't gain us much until we have a prescription for assigning the physical observables.

To this end, let's think about measurements. Ballentine uses the simple scattering experiment as a prototypical experiment



Let's start on a crude level. We throw rocks at rocks and observe the results. Each time we get a different result. So to make the system more controlled, we try to "prepare" the rocks identically.

At first, we just say we pick them up from the same stream. Then, we say that we round them to the same size and shape. For low velocity rocks, sometimes they bounce in similar directions [perhaps even the same direction to the accuracy of our measurements].



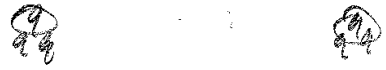
Sometimes they break. Since we get some situations which a reproducibly similar (but not exactly the same) we try to prepare our systems better. We try to make them the same to some accuracy determined by the experiment.

Making them the "same" is a relative term. It drives us to atomic and subatomic systems:

show video.

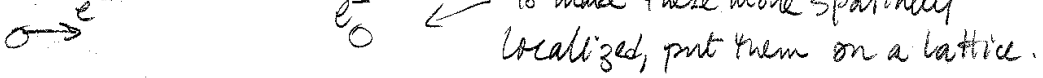
Video shows that there is uncontrolled structural motion in the nuclei (which is a quantum effect, Fermi motion)

Let's try protons



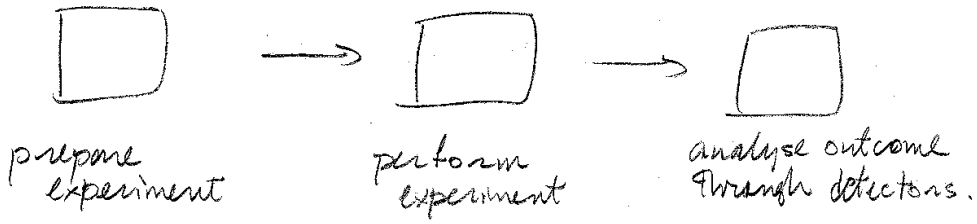
But protons have a characteristic size $\langle r^2 \rangle^{1/2} \sim 0.8 \times 10^{-15}$ m. Maybe their structure has Fermi motion as well due to their quark constituents. Hmmm. Electrons appear to be points $\langle r^2 \rangle^{1/2} < 10^{-17}$ for all of our experimental concerns.

So, we prepare our electrons and scatter them. We may find similar results.



but we never find event-to-event identical results.

Hmm. The point is that we are driven to a statistical description of the experiment, rather than a deterministic description.



It doesn't matter whether there is an intrinsic indeterminism in nature, or whether measurement processes destroy or change the system they measure, the point is that the system is statistical.

We talk, then, of –

state preparation procedure: a repeatable process that yields a well-defined probability distribution for all observables.

state: a set of systems with a specified probability distribution for each observable.
A state can also be identified with an ensemble of similarly prepared systems.

This leads us to:

Postulate 2: To each state there corresponds a unique state operator $\hat{\rho}$ [also called the density matrix]. The average value of a dynamical variable $\mathcal{R}$, represented by the linear operator $\hat{R}$ in the virtual ensemble of events that may result from a state preparation procedure is

$$\langle \mathcal{R} \rangle = \frac{\mathrm{Tr}(\hat{\rho}\hat{R})}{\mathrm{Tr}\,\hat{\rho}}$$

Virtual meaning a set of similarly prepared systems which may not coexist in time.

The trace operation for matrices is $\mathrm{Tr}\,A = \sum_i a_{ii}$

For operators, it is $\mathrm{Tr}(\hat{A}) = \sum_i \langle \phi_i | \hat{A} | \phi_i \rangle$ where $\{\phi_i\}$ is an orthonormal basis set.

What is $\hat{\rho}$? First, it is convenient to normalize $\hat{\rho}$ to unit trace so the denominator in $\langle \mathcal{R} \rangle$ can be dropped:

$$\mathrm{Tr}\,\hat{\rho} = 1 \qquad (1)$$

To find out more about $\hat{\rho}$, let's construct a projection operator $\hat{P}_u$ from kets $|u\rangle$ with unit norm.

$$\hat{P}_u \equiv |u\rangle\langle u|.$$

Now, the eigenvalues of $\hat{P}_u$ must be 0 or 1:

if $\hat{P}_u$ operates on $|u\rangle$ $\quad \hat{P}_u|u\rangle = |u\rangle\langle u|u\rangle = |u\rangle$

if $\hat{P}_u$ operates on $|v\rangle$ and $\langle v|u\rangle = 0$ $\quad \Rightarrow \hat{P}_u|v\rangle = 0$

if $\hat{P}_u$ operates on any other vector, it's not an eigenvalue equation ($\hat{P}_u|\phi\rangle = \text{const}\,|u\rangle$, not $\text{const}\times|\phi\rangle$)

$$= \sum_i \langle\varphi_i|\hat{\rho}|u\rangle\langle u|\varphi_i\rangle = \sum_i \underbrace{\langle u|\varphi_i\rangle}_{=\langle u|}\langle\varphi_i|\hat{\rho}|u\rangle$$

So, $\mathrm{Tr}(\hat{\rho}\hat{P}_u) = \langle u|\hat{\rho}|u\rangle$ ← must be real if eigenvalues are real.

(II) If this is true $\forall\ |u\rangle$, then $\hat{\rho} = \hat{\rho}^\dagger$ (from Theorem 1).

[This is just the demand that the prop. of finding the system in state $|u\rangle$ be real].

Lastly, the average of a variable [$\hat{P}_u$] that takes on ~~non-~~only non-negative values must also be non-negative: $\langle u|\hat{\rho}|u\rangle \geq 0$.

If this is true for all $|u\rangle$, then $\hat{\rho}$ is a non-negative operator.

(III)

Unfortunately, we cannot prove II $\not\Rightarrow$ III since we cannot prove that all projection operators correspond to observables. [i.e. the constraints apply only if true for all $|u\rangle$, hence all projection ops.].

So we need another postulate:

Postulate 2a To each state there corresponds a unique state operator $\hat{\rho}$ which must be Hermitian, non-negative, and of unit trace.

Now, if all dynamical observables are real, so that all $\langle \hat{R} \rangle$ are real, then we can place a constraint on $\hat{R}$. Let's now assume that $\hat{\rho}$ can be represented as

$$\hat{\rho} = |\Psi\rangle\langle\Psi| \qquad \text{where } |\Psi\rangle \text{ is some vector of unit norm.}$$

Obviously, this $\hat{\rho}$ satisfies the 3 conditions in Postulate 2a. (although we cannot say that $|\Psi\rangle$ is a physical state). Then

$$\mathrm{Tr}(\hat{\rho}\hat{R}) = \mathrm{Tr}\left(|\Psi\rangle\langle\Psi|\hat{R}\right) = \langle\Psi|\hat{R}|\Psi\rangle$$

If this be real for all $|\Psi\rangle$, then $\hat{R}^\dagger = \hat{R}$ by Theorem 1. So, if we demand that all physical observables are real, then

Postulate 1a To each dynamical observable there is a Hermitian operator whose eigenvalues are the possible values of the dynamical variable.

In fact, it is the Hermiticity of the operators which is important, rather than the reality of the observables. One could always change some phases and have operators whose mathematical values are "complex" while being physically "real". There are even cases where the eigenvalues are real but the average is complex because the eigenvector set is not complete. Hence, we need Post. 1a.

General States and Pure States

From the above, we have that $\hat{\rho}$ must satisfy

$$\text{Tr}\,\hat{\rho} = 1 \quad (A) \qquad \hat{\rho} = \hat{\rho}^{\dagger} \quad (B) \qquad \langle u|\hat{\rho}|u\rangle \geq 0 \quad \forall |u\rangle \quad (C)$$

Because of (B), then $\hat{\rho}$ has a spectral representation

$$\hat{\rho} = \sum_n \rho_n |\varphi_n\rangle\langle\varphi_n| \qquad \leftarrow \text{assume this is discrete.}$$

$$(A) \rightarrow \text{Tr}\left(\sum_n \rho_n |\varphi_n\rangle\langle\varphi_n|\right) = \sum_n \rho_n \langle\varphi_n|\varphi_n\rangle = \sum_n \rho_n = 1. \quad (D)$$

$$(B) \rightarrow \rho_n = \rho_n^* \quad (E)$$

$$(C) \rightarrow \rho_n \geq 0 \quad (F)$$

Some of these are invertible (E)+(F)→(C). Further, if we combine (D) + (F) we get

$$0 \leq \rho_n \leq 1 \quad (G)$$

(the latter inequality since a term in a sum cannot exceed the sum if all terms are positive).

The set of all state operators $\hat{\rho}^i$ [not states $|u\rangle$] form a convex set. This means that if two or more operators $\hat{\rho}^i$ satisfy (A) - (C), then so does any combination

$$\hat{\rho} = \sum_i a_i \hat{\rho}^i$$

provided $0 \leq a_i \leq 1$ and $\sum_i a_i = 1$.

In the above proofs, we used a representation

$$\hat{\rho} = |\Psi\rangle\langle\Psi| \qquad (D1)$$

definition 1.

as a means of linking reality of observables with hermiticity of operators. We were encouraged to try this representation because of the spectral theorem. In fact, this form for $\hat{\rho}$ is referred to as a pure state operator (where the motivation for the name is obvious) and $|\psi\rangle$ is referred to as a state vector.

For a pure state: $\langle R\rangle = Tr\left(|\psi\rangle\langle\psi|\hat{R}\right) = \langle\psi|\hat{R}|\psi\rangle$.

Clearly, the value of $\langle R\rangle$ is independent of the form of $|\psi\rangle$ up to $e^{i\alpha}|\psi\rangle$, where α is a phase angle.

A. Equivalent definitions of a pure state operator $\hat{\rho}$ are

$$\hat{\rho}^2 = \hat{\rho} \qquad (D2)$$

and

$$Tr(\hat{\rho}^2) = 1 \qquad (D3)$$

The equivalence of (D2) to (D1) goes as:

First, it is necessary, since (D1) looks like a projection operator, which is of the form (D2)

Second, it is sufficient since: consider the eigenvalues of $\hat{\rho}$: ρ_n

using the spectral theorem expansion $\hat{\rho} = \sum_n \rho_n |\varphi_n\rangle\langle\varphi_n|$ ①

then $\hat{\rho}^2 = \sum_i \sum_j \rho_i \rho_j |\varphi_i\rangle\langle\varphi_i|\varphi_j\rangle\langle\varphi_j| = \sum \rho_n^2 |\varphi_n\rangle\langle\varphi_n|$ ②

$\therefore$ ① + ② $\Rightarrow \rho_n^2 = \rho_n$

$\Rightarrow \rho_n = 0$ or 1 only.

But ① $\Rightarrow \sum_n \rho_n = 1$ $\therefore$ only 1 of $\rho_n = 1$, the rest are 0.

Hence $\hat{\rho} = |\Psi_n\rangle\langle\varphi_n|$.

Theorem: A pure state cannot be expressed as a nontrivial combination of other states, but a nonpure state can always be so expressed.

Proof: (Ballentine pg. 39)

The second part of the Theorem is trivial because of the spectral representation. The first part we prove by assuming the contrary:

Suppose

$$\hat{\rho} = \sum_i a_i \hat{\rho}^i \qquad 0 \le a_i \le 1 \;,\; \sum_i a_i = 1$$

Now

$$\mathrm{Tr}\,(\hat{\rho}^2) = \sum_i \sum_j a_i a_j \,\mathrm{Tr}\,(\hat{\rho}^i \hat{\rho}^j)$$

But each ρ^i has a spectral representation

$$\hat{\rho}^i = \sum_n \rho_n^i \, |\varphi_n^i\rangle\langle\varphi_n^i|$$

$$\Rightarrow \mathrm{Tr}\,(\hat{\rho}^i \hat{\rho}^j) = \sum_n \sum_m \rho_n^i \rho_m^j \,\mathrm{Tr}\,\{ |\varphi_n^i\rangle\langle\varphi_n^i|\varphi_m^j\rangle\langle\varphi_m^j| \}$$

$$= \sum_n \sum_m \rho_n^i \rho_m^j \, |\langle\varphi_n^i|\varphi_m^j\rangle|^2$$

But this is $\le |\langle\varphi_n^i|\varphi_n^i\rangle| \, |\langle\varphi_m^j|\varphi_m^j\rangle|$

$$\le 1$$

$$\therefore \mathrm{Tr}\,(\hat{\rho}^i \hat{\rho}^j) \le 1$$

But condition (D3) says that $\mathrm{Tr}(\hat{\rho}^2) = 1$, not ≤ 1.

So the only way the $=$ sign holds is if there is only one n and m contributing to the sum. That is, $\hat{\rho}^i$ and $\hat{\rho}^j$ are pure state operators. The equality sign also implies that $\rho^i = \rho^j \;\forall\, i,j$

In other words, all the ρ's have to be the same, so there is no non-trivial combination of ρ's.

QED.

If we indeed do have a non-pure state, then its representation in terms of pure state operators is not unique. For example, suppose we write

$$\hat{\rho} = \frac{1}{2}|x\rangle\langle x| + \frac{1}{2}|y\rangle\langle y|$$

where $|x\rangle$ and $|y\rangle$ are pure states. Now, suppose we have two other states $|u\rangle$ and $|v\rangle$ which can be constructed from $|x\rangle$ and $|y\rangle$ via

$$|x\rangle = \sqrt{a}\,|u\rangle + \sqrt{1-a}\,|v\rangle$$
$$|y\rangle = \sqrt{a}\,|u\rangle - \sqrt{1-a}\,|v\rangle$$

Then solving for $|u\rangle$, $|v\rangle$ and substituting, we find

$$\hat{\rho} = a|u\rangle\langle u| + (1-a)|v\rangle\langle v|.$$

In other words, we can form other representations of $\hat{\rho}$ which are not simply phase changes of $|\psi\rangle$ (which leave $\hat{\rho}$ invariant).

Probability Distributions (pg. 41)

The expression $\langle R\rangle = \mathrm{Tr}(\hat{\rho}\hat{R})$ allows us to calculate average values of R. Can we actually get the probability distribution if we know $\hat{\rho}$? To extract the distribution, we need a result for how a function of R behaves.

Call $F(\hat{R})$ an operator which is a function of $\hat{R}$. Let $F(R)$ be the corresponding observable. Going back to the spectral function result $f(\hat{A}) = \sum_i f(a_i)|q_i\rangle\langle q_i|$, then we expect that if R has the value r, then $F(R)$ has the value $F(r)$. Suppose, then, that $g(r)dr$ is the probability that the observable lies between r and $r+dr$. Then

$$\langle F(R)\rangle = \int_{-\infty}^{\infty} F(r)\,g(r)\,dr \quad (1) \qquad = \mathrm{Tr}\left(\hat{\rho}F(\hat{R})\right) \quad (2)$$

By suitable choice of $F(\hat{R})$, we can extract $g(r)$. We do this for two situations: Discrete spectrum and continuous spectrum.

Discrete:

Let $\hat{R}$ be self-adjoint, since we are interested in such operators by Postulate 1a. We can represent $\hat{R}$ by its eigenvalues r_n through

$$\hat{R} = \sum_n r_n |r_n\rangle\langle r_n|$$

Suppose now we consider the θ-function $F(R) = \theta(r-R)$

scalar $\boxed{F(\hat{R}) = \theta(r-\hat{R})}$

[recall $\theta(x) = 0 \quad x<0$, $\theta(x) = 1 \quad x>1$]

This gives $$\langle \theta(r-R)\rangle = \int_{-\infty}^{r} g(r')\,dr' \qquad (3)$$

$$= \text{Prob}(R < r \,|\, \rho)$$

This reads: given ρ, the probability that R is less than r.

But we also have from (2): $\langle \theta(r-R)\rangle = \text{Tr}\{\hat{\rho}\,\theta(r-\hat{R})\}$

$$= \text{Tr}\Big\{\hat{\rho}\sum_n \theta(r-r_n)|r_n\rangle\langle r_n|\Big\}$$

$$= \sum_n \langle r_n|\hat{\rho}|r_n\rangle\,\theta(r-r_n) \qquad (4)$$

(3) + (4)

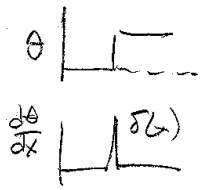
$$\Rightarrow \int_{-\infty}^{r} g(r')dr' = \sum_n \langle r_n|\hat{\rho}|r_n\rangle\,\theta(r-r_n). \qquad (5)$$

Since $g(r)$ is a probability density, then we also have

$$g(r) = \frac{\partial}{\partial r}\,\text{Prob}(R < r \,|\, \rho) \qquad (6)$$

⑤→⑥ $\Rightarrow \quad g(r) = \frac{\partial}{\partial r}\left[\sum_n \langle r_n|\hat{\rho}|r_n\rangle\, \Theta(r-r_n)\right]$

$= \sum_n \langle r_n|\hat{\rho}|r_n\rangle\, \delta(r-r_n) \quad ⑦$



Note that $g(r) = 0$ if r is not an eigenvalue. This, of course, is just what we expect for a discrete spectrum.

The probability that $\mathcal{R} = r$, in a discrete spectrum is then

$$\text{Prob}(\mathcal{R}=r) = \lim_{\epsilon\to 0}\left\{\text{Prob}(R_\rho < r+\epsilon\,|\,\rho) - \text{Prob}(R_\rho < r-\epsilon\,|\,\rho)\right\}$$

$$= \sum_n \langle r_n|\hat{\rho}|r_n\rangle \delta_{r,r_n}$$

Let's recast this a little using the projection operator $P(r) = \sum_n |r_n\rangle\langle r_n|\delta_{r,r_n}$

$$\Rightarrow \text{Prob}(\mathcal{R}=r) = \text{Tr}\left(\hat{\rho}\,P(r)\right)$$

Finally, if $\hat{\rho}$ is a pure state $\hat{\rho} = |\psi\rangle\langle\psi|$ and r_n is <u>non-degenerate</u> then for a specific eigenvalue

$$\text{Prob}(\mathcal{R}=r_n) = \text{Tr}\left(|\psi\rangle\langle\psi|r_n\rangle\langle r_n|\right) = |\langle r_n|\psi\rangle|^2.$$

[See notes on page 2-15 for discussion of non-diagonal $\hat{\rho}$].

Suppose that for some state, the value of $\mathcal{R}$ is equal to one of the discrete eigenvalues r_0. If the probability for this to happen is unity, then from

$$\text{Prob}(\mathcal{R}=r_0) = \sum_n \langle r_n|\hat{\rho}|r_n\rangle \delta_{r_0,r_n} = \langle r_0|\hat{\rho}|r_0\rangle$$

$= 1$ by assumption

But, by the $\text{Tr}(\hat{\rho}^2) \le 1$ condition, which applies to any state, pure or otherwise, we also have

$$\sum_m \sum_n \langle r_n|\hat{\rho}|r_m\rangle\langle r_m|\hat{\rho}|r_n\rangle \le 1$$

The sum condition $\sum_m \sum_n = 1$ is fulfilled by just one term $\langle r_0|\hat{\rho}|r_0\rangle = 1$. So all other "off-diagonal" terms $\langle r_m|\hat{\rho}|r_n\rangle$ <u>must vanish</u> <u>for</u> this <u>case</u>

This can only be true if $\hat{\rho} = |r_0\rangle\langle r_0|$, a pure state.
So what we find is

$\text{Prob}(R = r_0 | \rho) = 1 \Rightarrow \hat{\rho}$ is a pure state $|r_0\rangle\langle r_0|$

where r_0 is a discrete eigenvalue

$\hat{\rho} \cdot |r_0\rangle$ is an <u>eigenstate</u> corresponding to r_0.

[This is just the usual starting point for the eigenvalue problem].

<u>Continuous spectrum</u>

We finish off by running through the same arguments as above, but for the continuous spectrum. We let $\hat{Q}$ be a self-adjoint operator having a continuous spectrum, with spectral representation:

$$\hat{Q} = \int q' |q'\rangle\langle q'| \, dq'$$

The eigenvectors are un-normalizable in general, satisfying

$$\langle q' | q'' \rangle = \delta(q' - q'')$$

We let $g(q)\,dq'$ be the probability density for Q lying between q and $q + dq$, and using the θ-function for $F(\hat{Q})$ as before, find

$$\langle \theta(q - Q) \rangle = \int_{-\infty}^{q} g(q')\,dq' = \text{Prob}(Q < q | \rho)$$

We go through the same steps as before to get

$$g(q) = \langle q | \hat{\rho} | q \rangle,$$

and for a pure state $|\psi\rangle$,

$$g(q) = |\langle q | \psi \rangle|^2.$$

filter function; state function

Recap: If ρ = pure state, then $\text{Prob}(R = r_n) = |\langle r_n | \psi \rangle|^2$. (doesn't say this is = 1).

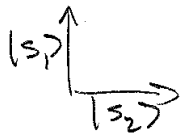
If $\text{Prob}(R = r_n) = 1$, then $\rho = |r_n\rangle\langle r_n|$ = pure state.

One last discussion about pure and mixed states. Suppose we have a state which is polarized:

PURE 

state each element of state.

Then we can choose a basis vector $|\psi\rangle$ which is parallel to the spin direction, so that $\hat{\rho} = |\psi\rangle$. Of course, for a degenerate system one might have



in which case we can rotate one of $|s_1\rangle$ & $|s_2\rangle$ to point along the spin direction of the state.

Suppose we had a mixed state like:

State  =

less than unit length

elements.

Even though we can rotate a set of basis vectors in the sum

$$\frac{1}{5}|\uparrow\rangle\langle\uparrow| + \frac{1}{5}|\uparrow\rangle\langle\uparrow| + \frac{1}{5}\cdots$$

to look like $\rho = \frac{1}{4}|\uparrow\rangle\langle\uparrow|$ ← These still aren't equivalent.

it isn't a pure state since $\mathrm{Tr}(\hat{\rho}^2) \neq 1$.

① Thermodynamic systems with $e^{-\beta E_n}|E_n\rangle\langle E_n|$ are mixed states. ② The spectral representation $\sum \rho_n |r_n\rangle\langle r_n|$ was for a specific observable. A measurement must yield one value of the observable. If all values of all measurements are the same, then the system is a pure state.

A different emphasis can be placed on the density matrix by considering the idea of coherence. If we expand a state vector

$$|\psi\rangle = \sum_i a_i |q_i\rangle$$

Then there is a fixed phase relation between the components which make up $|\psi\rangle$, and the state is coherent. Such a state will not, in general, result in a diagonal form for $\hat{\rho} = |\psi\rangle\langle\psi|$. But we can always find a matrix $\hat{R}$ to diagonalize $\hat{\rho}$ and change $|\psi\rangle \rightarrow |\psi'\rangle = \hat{R}|\psi\rangle$

$$\hat{\rho}' = |\psi'\rangle\langle\psi'| = \hat{R}|\psi\rangle\langle\psi|\hat{R}^{-1} = \hat{R}\hat{\rho}\hat{R}^{-1}$$ where $\hat{\rho}'$ is diagonal.

Note that the eigenstates of $\hat{\rho}'$ are not necessarily the eigenstates of the set $|q_i\rangle$. This form for $\hat{\rho}'$ is $\begin{pmatrix} a_0 & & \\ & a_1 & \\ & & \ddots \end{pmatrix}$.

An incoherent state is of the form

$$\hat{\rho} = \sum_i a_i |q_i\rangle\langle q_i| = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}.$$

This "matrix" is already of diagonal form, but has more than 1 non-zero diagonal element.

Pure (Coherent)

$\hat{\rho}$ may have off-diagonal elements, but can be made to look like $\begin{pmatrix} a & 0 & 0 & \\ 0 & 1-a & 0 & \\ 0 & 0 & 0 & \\ & & & \ddots \end{pmatrix}$

Mixed (Incoherent)

$\hat{\rho}$ may be diagonal, but not $\begin{pmatrix} 0 & 0 & 0 & \\ 0 & 1 & 0 & \\ 0 & 0 & 0 & \\ & & & \ddots \end{pmatrix}$