Recall the (pure) states of a quantum system are described by vectors in Hilbert space, $|\phi\rangle$, and corresponding dual vectors, $\langle \phi |$. The inner product $\langle \phi | \psi \rangle$ satisfies
\[ \langle \phi | \psi \rangle = \langle \psi | \phi \rangle^* . \quad (4.1) \]
A state vector is a unit vector, a vector of unit norm,
\[ \langle \phi | \phi \rangle = 1. \quad (4.2) \]
Physically, $\langle \phi | \psi \rangle$ represents a probability amplitude: That is, $|\langle \phi | \psi \rangle|^2$ is the probability of finding the system in the state $|\phi\rangle$ given that it was previously determined to be in (prepared in) the state $|\psi\rangle$. Suppose we introduce an orthonormal basis set $\{ |j\rangle \}$:
\[ \langle j | k \rangle = \delta_{jk} . \quad (4.3) \]
Such a set is complete. If $| \rangle$ is any vector,
\[ | \rangle = \sum_j c_j | j \rangle, \quad c_j = \langle j | \rangle. \quad (4.4) \]
That is,
\[ | \rangle = \sum_j | j \rangle \langle j | \rangle, \quad (4.5) \]
or, since this is true for any state,
\[ \sum_j | j \rangle \langle j | = 1. \quad (4.6) \]
The probability of finding the system in the state $|j\rangle$ is
\[ P_j = |c_j|^2 = |\langle j | \rangle|^2 = \langle j | \langle j | \rangle, \quad (4.7) \]
or if we define a “density operator” for the pure state $| \rangle$ by
\[ \rho = | \rangle \langle |, \quad (4.8) \]
the probability is
\[ P_j = \langle j | \rho | j \rangle. \]  
(4.9)

Note that \( \rho \) has the following properties:
\[ \text{Tr} \rho = \langle | | \rangle = 1, \]  
(4.10)
\[ \rho^2 = | \rangle \langle | = | \rangle \langle | = \rho. \]  
(4.11)

Physical properties or “observables” are represented by linear operators, \( A \), which carries a vector \( | \phi \rangle \) to another vector \( A|\phi\rangle \). The dual of this vector is \( \langle \phi | A^\dagger \), where \( A^\dagger \) is the adjoint or Hermitian conjugate of \( A \). A physical property is represented by a self-adjoint operator: \( A = A^\dagger \). Eigenvectors of \( A \) represent states in which \( A \) assumes a definite value, its eigenvalue:
\[ A|a'\rangle = a'|a'\rangle, \quad \langle a'|A = (a')^\ast (a'|; \]  
(4.12)
where we’ve used the eigenvalue to label the vector. Note that for a self-adjoint operator
\[ \langle a''|A|a'\rangle = a'\langle a''|a'\rangle = \langle a'|A|a''\rangle^\ast = (a'')^\ast \langle a'|a''\rangle, \]  
(4.13)
from which conclude
1. If \( a'' = a', \) \( a' = (a')^\ast \), that is, the eigenvalues must all be real.
2. If \( a' \neq a'' \), \( \langle a''|a'\rangle = 0 \), which means that eigenvectors belonging to distinct eigenvalues are orthogonal.

The expectation value of \( A \) in a state \( | \rangle \) is
\[ \langle A \rangle = \langle A | \rangle = \sum_j \langle j | \rangle \langle j | A | \rangle = \sum_j \langle j | A \rangle \langle j | \rangle = \sum_j \langle j | A \rho | j \rangle, \]  
(4.14)
or
\[ \langle A \rangle = \text{Tr} A \rho = \text{Tr} \rho A. \]  
(4.15)

Up to this point, we have been thinking of pure states. More generally, quantum systems will be found in mixed states. That is, we know only the probability \( p_n \) of finding the system in some pure state \( |n\rangle \), described by the density operator
\[ \rho = \sum_n |n\rangle p_n \langle n|, \]  
(4.16)
where \( \{ |n\rangle \} \) is some complete set of states, and
\[ \sum_n p_n = 1, \quad 1 \geq p_n \geq 0. \]  
(4.17)

Now
\[ \text{Tr} \rho = \sum_n \langle n | n \rangle p_n = 1, \]  
(4.18)
as before, but now
\[ \rho^2 = \sum_{n,n'} |n\rangle p_n \langle n'| p_{n'} \langle n' | = \sum_n |n\rangle p_n^2 \langle n|, \] (4.19)
so
\[ \text{Tr} \rho^2 = \sum_n p_n^2 \leq 1, \] (4.20)
where the trace of \( \rho^2 \) is one only if \( p_n = 1 \) for some \( n \) and zero otherwise. Thus,
\[ \text{Tr} \rho^2 = 1 \] (4.21)
characterizes a pure state.

Now let’s consider the expectation value of some physical observable \( A \) in a mixed state:
\[ \langle A \rangle = \sum_n p_n \langle n| A| n \rangle = \text{Tr} \sum_n |n\rangle p_n \langle n| A = \text{Tr} \rho A. \] (4.22)
The trace is independent of basis. Indeed, such a new basis is related to the old by a unitary transformation,
\[ |j\rangle = U |j\rangle, \quad \langle j| = \langle j| U^\dagger, \] (4.23)
where \( U^\dagger = U^{-1} \). Therefore, in the new basis,
\[ \langle A \rangle = \sum_j \langle j| \rho A| j \rangle = \sum_j \langle j| U^{-1} \rho A U | j \rangle \]
\[ = \text{Tr} (U^{-1} \rho A U) = \text{Tr} \rho A U U^{-1} \]
\[ = \text{Tr} \rho A = \langle A \rangle. \] (4.24)
Recall, classically, we had Liouville’s theorem for the phase-space distribution:
\[ \frac{d\rho}{dt} = 0 = \frac{\partial \rho}{\partial t} + \{ \rho, H \}, \] (4.25)
in terms of the Poisson bracket. What happens for the quantum density operator? Recall, in the Schrödinger picture, the time evolution of states is given by the Schrödinger equation,
\[ i\hbar \frac{\partial}{\partial t} |t\rangle = H |t\rangle. \] (4.26)
The density operator
\[ \rho(t) = \sum_n p_n |n,t \rangle \langle n,t| \] (4.27)
then satisfies
\[ i\hbar \frac{\partial}{\partial t} \rho(t) = \sum_n p_n H |n,t \rangle \langle n,t| - \sum_n p_n |n,t \rangle \langle n,t| H \]
\[ = [H, \rho], \] (4.28)
in terms of the commutator,
\[ [A, B] \equiv AB - BA. \]  (4.29)

Therefore the quantum-mechanical Liouville equation is
\[ \frac{\partial}{\partial t} \rho(t) + \frac{1}{i\hbar} [\rho, H] = 0. \]  (4.30)

Note that upon comparison with Eq. (4.25) we have the correspondence principle relation:
\[ \frac{1}{i\hbar}[\cdot, \cdot] \rightarrow \{\cdot, \cdot\}. \]  (4.31)

We will usually be interested in stationary distributions, for which
\[ \frac{\partial}{\partial t} \rho(t) = 0 \Rightarrow [\rho, H] = 0. \]  (4.32)

Suppose the observable \( A = A(q_i, p_i) \) is not explicitly time dependent. Then, for a stationary distribution
\[ 0 = \frac{d}{dt} \langle A \rangle = \text{Tr} \left( A \frac{\partial}{\partial t} \rho \right) = -\frac{1}{i\hbar} \text{Tr} \left( [\rho, H] A \right) \]  (4.33)
in the Schrödinger picture. This must be true for all \( A \), in particular for \( A = |j\rangle\langle k| \), so we conclude that \( [\rho, H] = 0 \).

This result will also be true in the Heisenberg picture, where states are time independent, but operators move. The two pictures are related by a unitary transformation,
\[ A_H(t) = U^{-1}(t) A_S U(t), \quad |\rangle_H = U^{-1}(t)|\rangle_S, \]  (4.34)
where the unitary time evolutionary operator satisfies
\[ i\hbar \frac{d}{dt} U(t) = H_S U(t) = U(t) H_H. \]  (4.35)

The Heisenberg picture Hamiltonian is time independent,
\[ i\hbar \frac{d}{dt} H_H = [H_H, H_H] = 0, \]  (4.36)
whereas the other observables satisfy
\[ i\hbar \frac{d}{dt} A_H = [A_H, H_H]. \]  (4.37)

The density operator depends on time in the Schrödinger picture, so it satisfies in the Heisenberg picture
\[ i\hbar \frac{d}{dt} \rho_H = [\rho_H, H_H] + U^{-1} \left( i\hbar \frac{\partial}{\partial t} \rho_S \right) U \]
\[ = [\rho_H, H_H] - U^{-1} [\rho_S, H_S] U = [\rho_H, H_H] - [\rho_H, H_H] = 0, \]  (4.38)
as must be.

In the Heisenberg picture

\[
0 = \frac{d}{dt} \langle A \rangle = \text{Tr} \rho \frac{d}{dt} A \\
= \frac{1}{\ihbar} \text{Tr} \rho [A, H] = \frac{1}{\ihbar} \text{Tr} [H, \rho] A,
\]

so again we conclude that

\[ [\rho, H] = 0. \tag{4.40} \]

This says, once again, that \( \rho \) is a constant of the motion, and the only reasonable choice, if translational or rotational invariance is not present, is

\[ \rho = \rho(H). \tag{4.41} \]

Thus in the basis in which the energy is diagonal,

\[ \rho(H) = \sum_{i,k} p(E_i |E_i, k\rangle \langle E_i, k|, \tag{4.42} \]

where \( p(E_i) = \rho(E_i) \) and \( k \) labels quantum numbers other than the energy.