Chapter 14

The density matrix

Prerequisites: Chapters 2 – 5. Preferably also Chapters 6 and 7

The density operator, or as it is more commonly called, the density matrix, is a key tool in quantum mechanics that allows us to connect the world of quantum mechanics with the world of statistical mechanics, and we introduce it briefly here. This is an important connection. Just as we needed statistical ideas in complicated classical systems (e.g., large collections of atoms or molecules), we will need the same ideas in complicated quantum mechanical systems.

We will also examine one rather elementary but important application in optics; we will use the density matrix to turn the rather extreme and unrealistic idea of infinitely sharp "δ-function" optical transitions that emerged from the simple perturbation theory model of Chapter 7 into more physical absorption lines with finite width, just as we actually see in atomic and molecular spectra, for example.

14.1 Pure and mixed states

So far in thinking about the state of a system, the only randomness we have considered is that involved in quantum-mechanical measurement. We have presumed that otherwise everything else was definite. Suppose, for example, we were thinking about the state of polarization of a photon. In the way we have so far considered states, we could write a general state of polarization as

$$|\psi\rangle = a_H |H\rangle + a_V |V\rangle \tag{14.1}$$

where $|H\rangle$ means a horizontally polarized photon state and $|V\rangle$ means a vertically polarized one. There is apparently randomness on quantum mechanical measurement; if we performed a measurement, using, e.g., a polarizing beamsplitter oriented to separate horizontal and vertical polarizations to different outputs with different detectors, we expect a probability of $|a_H|^2$ of measuring the photon in the horizontal polarization and $|a_V|^2$ of measuring it in the vertical polarization.

Since we must have $|a_H|^2 + |a_V|^2 = 1$ by normalization, we could also choose to write

$$a_H = \cos\theta \quad a_V = \exp(i\delta)\sin\theta \tag{14.2}$$

The fact that $\sin^2\theta + \cos^2\theta = 1$ ensures that this way of writing the state is properly normalized. For the case where $\delta = 0$, we are describing ordinary linear polarization of a light field, and θ has the simple physical meaning of the angle of the optical electric vector relative to the horizontal axis. It is also possible with light that there can be some phase delay between the horizontal and vertical components of the optical electric field, and that phase difference

can be expressed through the $\exp(i\delta)$ term. When $\delta \neq 0$, the field is in general "elliptically polarized", which is the most general possible state of polarization of a propagating photon, and the special cases of $\delta = \pm\pi/2$ with $\theta = 45^\circ$ give the two different kinds of circularly polarized photons (right and left circularly polarized)[1].

An important point about such a state is that we can always build a polarizing filter that will pass a photon of any specific polarization, 100% of the time. Even if the photon was in the most general possible elliptically polarized state, we could arrange to delay only the horizontal polarization by a compensating amount $-\delta$ to make the photon linearly polarized[2], and then orient a linear polarizer at an angle θ so that the photon was always passed through. Such polarization compensators and filters are routine optical components. When we can make such a polarization filter so that we will get 100% transmission of the photons, we say that the photons are in a "pure" state (here, Eq. (14.1)). The states of photons or anything else we have considered so far have all been pure states in this sense; we can at least imagine that an appropriate filter could be made to pass any particles that are in any one such specific quantum mechanical state with 100% efficiency.[3]

But such "pure" states are by no means the only ones we will encounter. Suppose that we have a beam of light that is a mixture from two different, and quite independent, lasers, lasers "1" and "2", giving laser beams of possibly even different colors. Presume that laser 1 contributes a fraction P_1 of the total number of photons, and laser 2 contributes a fraction P_2. Then the probability that a given photon is from laser 1 is P_1 and similarly there is probability P_2 it is from laser 2. We presume also that these two lasers give photons of two possibly different polarization states, $|\psi_1\rangle$ and $|\psi_2\rangle$ respectively, in the beam.

Something is quite measurably different about this "mixed" state; no setting of our polarizing filter will in general pass 100% of the photons. If we set the polarization filter to pass all the photons in state $|\psi_1\rangle$, it will in general not pass all the photons in state $|\psi_2\rangle$, and *vice versa*. If we set the polarizing filter in any other state, it will not pass 100% of either set of photons. This measurable difference tells us that we cannot simply write this mixed state as some linear combination of the two different polarization states in the fashion we have used up to now for linear combinations of quantum mechanical states. If we were able to do that, e.g., in some linear combination of the form $b_1|\psi_1\rangle + b_2|\psi_2\rangle$, we would be able to construct a polarizing filter that would pass 100% of the photons in this state (which is a pure state), and we know here that no such polarization filter is possible.[4]

---

[1] "Right circular polarization" means that, if we were to look back towards the source of a plane wave, i.e., looking against the direction of propagation, the electric field vector would be rotating, with constant length, in a clockwise direction, and this corresponds to $\delta = \pi/2$. "Left circular polarization" similarly corresponds to $\delta = -\pi/2$ and anti-clockwise rotation.

[2] A birefringent material, which has different refractive indices on two axes at right angles to one another, can create such a relative phase delay if it is oriented in the correct direction and is of the right thickness.

[3] If the reader prefers to think in electron spin states rather than photon polarization states, we can construct filters out of Stern-Gerlach-like magnet setups, with beam blocks to block the undesired spin components. Spin states can also be described in the same form as in Eqs. (14.1) and (14.2), using spin-up and spin-down components along a given axis instead of the horizontal and vertical components of the photon case. The rest of the mathematics proceeds essentially identically.

[4] It is true that, if the photons from the two different lasers were actually both in the same mode, i.e., exactly the same frequency and spatial form, and of defined relative phase, we could construct such a

The difference between pure and mixed states can have important consequences for other measurable quantities. Suppose, for example, that, for some particle with mass we have a potential well, such as the simple "infinite" one-dimensional potential well we considered earlier, with infinitely high potential barriers around a layer of some thickness, L_z, in the z direction. If we put the particle in a pure state that is an equal linear superposition of the lowest two states of this well, $|\psi\rangle = (1/\sqrt{2})(|\psi_1\rangle + |\psi_2\rangle)$, the position of this particle, as given formally by the expectation value $\langle z \rangle$ of the $\hat{z}$ position operator, will oscillate back and forwards in time in the well because of the different time-evolution factors $\exp(-iE_1t/\hbar)$ and $\exp(-iE_2t/\hbar)$ for the two energy eigenstates (with energies E_1 and E_2 respectively). Suppose instead we take an ensemble of identical potential wells, and randomly prepare half of them with the particle in the lowest state and half of them with the particle in the second state. Statistically, since we do not know which wells are which, at least before performing any measurements, each of these wells is in a mixed state, with 50% probability of being in either the first or second state. Now we evaluate the expectation value $\langle z \rangle$ of the $\hat{z}$ position operator for each potential well. In each well in this ensemble, $\langle z \rangle$ evaluates to the position of the center of the well since both these wavefunctions are equally balanced about the center (the lowest being symmetric and the second being antisymmetric). The average, $\overline{\langle z \rangle}$, of all of the expectation values from each of the different wells (what we can call an "ensemble average") is also zero, and there is no oscillation in time. Hence again the mixed state and the pure state have quite different properties. Again it would not be correct simply to write the mixed state as a linear combination of the form $b_1|\psi_1\rangle + b_2|\psi_2\rangle$.

We could also consider a slightly more complicated version of the ensemble of potential wells, one in which each well is skewed, as would be the case if we applied electric field perpendicular to the wells for the case of a charged particle like an electron in the well. Then the expectation values $\langle z \rangle$ of the position would be different for the first and second states of the well (and neither one would in general be in the center of the well), with some specific values $\langle z \rangle = z_1$ for the first state and $\langle z \rangle = z_2$ for the second state. In the case of the pure state, with a combination of these two states in each well, we would still expect oscillation. We could now instead presume a mixed state with possibly different probabilities P_1 and P_2 respectively that we had prepared a given well in the first or second state. In this mixed state, we would still have no oscillation, and our ensemble average value of the measured position would now be

$$\overline{\langle z \rangle} = P_1 z_1 + P_2 z_2 \equiv \sum_{j=1}^{2} P_j \langle \psi_j | \hat{z} | \psi_j \rangle \qquad (14.3)$$

More generally for a mixed state, we expect that the ensemble average expectation value for some operator $\hat{A}$ corresponding to an observable quantity can be written

$$\overline{\langle A \rangle} = \sum_j P_j \langle \psi_j | \hat{A} | \psi_j \rangle \qquad (14.4)$$

for some set of different quantum mechanical state preparations $|\psi_j\rangle$ made with respective probabilities P_j.

Note, incidentally, that there is no requirement in mixed states that the different $|\psi_j\rangle$ are orthogonal. We could be considering several different polarization states that are quite close to

pure state by combining them, but that would violate our presumption above that the lasers are independent.

one another in angle. For example, there might be some fluctuation in time in the precise output polarization of some laser, perhaps because some mirror in the laser cavity is subject to vibrations. Then we would have to consider a mixed state of many different possible similar but not identical polarizations.

The question now is whether we can find a convenient and powerful algebra for representing such mixed states and their properties, so that we can get to results such as $\overline{\langle A \rangle}$ by a more elegant method than simply adding up probabilities and measured values as in Eq. (14.3). Ideally also, that method would give the correct results even if the mixed state became a pure one, i.e., we simply had one of the $P_j = 1$ and all the others zero, giving us one unified approach for pure and mixed states. We have already concluded that the linear superposition form $b_1 |\psi_1\rangle + b_2 |\psi_2\rangle$ will not work for representing mixed states, so we may need to go beyond such a simple addition. The answer to this question is to introduce the density operator.

Problem

14.1.1 Suppose that we are measuring the value of the spin magnetic moment of electrons. We take the spin magnetic dipole moment operator to be $\hat{\boldsymbol{\mu}}_e = g\mu_B \hat{\boldsymbol{\sigma}}$. We will compare the average value we measure in two different states, both of which are equal mixtures of x and y spin character for the electrons.

(i) Consider the pure spin state $|s_p\rangle = (1/\sqrt{3})(|s_x\rangle + |s_y\rangle)$. Here $|s_x\rangle$ and $|s_y\rangle$ are respectively spin states oriented along the $+x$ and $+y$ directions. [Hint: see Eq. (12.27) and the associated discussion of the Bloch sphere to see how to write out these two states.]

> (a) Show that this pure state is normalized.
> (b) Find the expected value (which will be a vector) of the spin magnetic dipole moment (i.e., the average result on measuring the magnetic dipole on multiple successive electrons all prepared in this state).
> (c) What is the magnitude of this average dipole moment (i.e., the length of the vector)?

(ii) Consider now the mixed spin state, with equal probabilities of the electrons being in the pure state $|s_x\rangle$ and the pure state $|s_y\rangle$.

> (a) What is the resulting average expected value (again a vector) of the spin magnetic dipole moment when measuring an ensemble of successive electrons prepared this way?
> (b) What is the magnitude of this ensemble average value?

(iii) What differences are there between the measured magnetic dipole moments in the two different cases of pure and mixed states?

14.2 Density operator

However we are going to represent the mixed state, it must obviously contain the probabilities P_j and the pure states $|\psi_j\rangle$, but it must not simply be a linear combination of the states. The structure we propose instead is the density operator

$$\rho = \sum_j P_j |\psi_j\rangle \langle \psi_j| \tag{14.5}$$

We see that this is an operator because it contains the outer products of state vectors (i.e., $|\psi_j\rangle\langle\psi_j|$). Though the mathematics of operating with this operator takes on the same algebraic rules as any other operators we have been using, we have deliberately left the "hat" off the top of this operator to emphasize that its physical meaning and use are quite different from other operators we have considered. ρ is *not* an operator representing some physical observable. Rather, ρ is representing the *state* (in general, a mixed state) of the system.

If ρ is a useful way of representing the mixed state, it must allow us to calculate quantities like the ensemble average measured value $\overline{\langle A \rangle}$ for any physical observable with corresponding operator $\hat{A}$. In fact, if we can evaluate $\overline{\langle A \rangle}$ for any physically observable quantity, then ρ will be the most complete way we can have of describing this mixed quantum mechanical state because it will tell us the value we will get of any measurable quantity, to within our underlying statistical uncertainties.

14.3 Density matrix and ensemble average values

To see how to use the density operator, first let us write it out in terms of some complete orthonormal basis set, $|\phi_m\rangle$. First we expand each of the pure states $|\psi_j\rangle$ in this set, obtaining

$$|\psi_j\rangle = \sum_u c_u^{(j)} |\phi_u\rangle \tag{14.6}$$

where the superscript (j) means these are the coefficients for the expansion of the specific state $|\psi_j\rangle$. Then we use Eq. (14.6) and its adjoint in Eq. (14.5) to obtain

$$\rho = \sum_j P_j \left(\sum_u c_u^{(j)} |\phi_u\rangle \right) \left(\sum_v \left(c_v^{(j)} \right)^* \langle \phi_v | \right)$$
$$= \sum_{u,v} \left(\sum_j P_j c_u^{(j)} \left(c_v^{(j)} \right)^* \right) |\phi_u\rangle\langle\phi_v| \tag{14.7}$$

Written this way, the matrix representation of ρ is now clear. We have for a matrix element in this basis

$$\rho_{uv} \equiv \langle \phi_u | \rho | \phi_v \rangle = \sum_j P_j c_u^{(j)} \left(c_v^{(j)} \right)^* \equiv \overline{c_u c_v^*} \tag{14.8}$$

where we have also introduced and defined the idea of the ensemble average of the coefficient product $\overline{c_u c_v^*}$. Given the form Eq. (14.8), we now more typically talk of ρ as the density matrix (rather than the density operator), with matrix elements given as in Eq. (14.8). The density matrix is just a way of representing the density operator, but since we essentially always represent the density operator this way, the two terms are in practice used interchangeably, with "density matrix" being the more common usage.

There are several important properties of the density matrix we can deduce from Eq. (14.8).

(i) The density matrix is Hermitian, i.e., explicitly

$$\rho_{vu} \equiv \sum_j P_j c_v^{(j)} \left(c_u^{(j)} \right)^* = \left(\sum_j P_j c_u^{(j)} \left(c_v^{(j)} \right)^* \right)^* = \rho_{uv}^* \tag{14.9}$$

Because the density matrix is Hermitian, so also is the density operator since the density matrix is just a representation of the density operator.

(ii) The diagonal elements ρ_{mm} give us the probabilities of finding the system in a specific one of the basis states $|\phi_m\rangle$. $c_m^{(j)}(c_m^{(j)})^* \equiv |c_m^{(j)}|^2$ is the probability for any specific pure state j that we will find the system in (basis) state m. Hence adding these up with probabilities P_j gives us the overall probability of finding the system in state m in the mixed state. (The meaning of the off-diagonal elements is more subtle, but quite important. They are a measure of what we

could loosely think of as the "coherence" between different states in the system, and we will return to discuss this below.)

(iii) The sum of the diagonal elements of the density matrix is unity, i.e., remembering that we can formally write the sum of the diagonal elements of some matrix or operator as the trace (*Tr*) of the matrix or operator,

$$Tr(\rho) = \sum_m \rho_{mm} = \sum_m \sum_j P_j \left| c_m^{(j)} \right|^2 = \sum_j P_j \sum_m \left| c_m^{(j)} \right|^2 = \sum_j P_j = 1 \qquad (14.10)$$

because (a) the state $\left| \psi_j \right\rangle$ is normalized (so $\sum_m | c_m^{(j)} |^2 = 1$), and (b) the sum of all the probabilities P_j of the various pure states $\left| \psi_j \right\rangle$ in the mixed state must be 1.

Now we come to a key trick with the density matrix. Let us consider an operator $\hat{A}$ corresponding to some physical observable, and specifically consider the product $\rho \hat{A}$, i.e.,

$$\rho \hat{A} = \sum_{u,v} \left(\sum_j P_j c_u^{(j)} \left(c_v^{(j)} \right)^* \right) \left| \phi_u \right\rangle \left\langle \phi_v \right| \hat{A} \qquad (14.11)$$

We can therefore write some diagonal element of the resulting matrix as

$$
\begin{aligned}
\left\langle \phi_q \left| \rho \hat{A} \right| \phi_q \right\rangle &= \sum_{u,v} \left(\sum_j P_j c_u^{(j)} \left(c_v^{(j)} \right)^* \right) \left\langle \phi_q \middle| \phi_u \right\rangle \left\langle \phi_v \left| \hat{A} \right| \phi_q \right\rangle \\
&= \sum_{u,v} \left(\sum_j P_j c_u^{(j)} \left(c_v^{(j)} \right)^* \right) \delta_{qu} \left\langle \phi_v \left| \hat{A} \right| \phi_q \right\rangle \\
&= \sum_v \sum_j P_j c_q^{(j)} \left(c_v^{(j)} \right)^* \left\langle \phi_v \left| \hat{A} \right| \phi_q \right\rangle
\end{aligned}
\qquad (14.12)
$$

Then the sum of the all of these diagonal elements is

$$
\begin{aligned}
\sum_q \left\langle \phi_q \left| \rho \hat{A} \right| \phi_q \right\rangle &= \sum_j P_j \left(\sum_v \left(c_v^{(j)} \right)^* \left\langle \phi_v \right| \right) \hat{A} \left(\sum_q c_q^{(j)} \left| \phi_q \right\rangle \right) \\
&= \sum_j P_j \left\langle \psi_j \left| \hat{A} \right| \psi_j \right\rangle
\end{aligned}
\qquad (14.13)
$$

Note that the result here is exactly the result for the ensemble average value $\overline{\langle A \rangle}$ of the expectation value of the operator $\hat{A}$ for this specific mixed state as given in Eq. (14.4) above. Hence we have a key result of density matrix theory

$$\overline{\langle A \rangle} = Tr\left(\rho \hat{A} \right) \qquad (14.14)$$

We have therefore found that the density matrix, through the use of the relation (14.14), describes any measurable ensemble average property of a mixed state. Hence the density matrix does indeed give as full a description as possible of a mixed state.

Note that this result, Eq. (14.14), is completely independent of the basis used to calculate the trace – the basis $\left| \phi_m \right\rangle$ could be any set that is complete for the problem of interest. (This invariance of the trace with respect to any complete orthonormal basis is a general property of traces of operators.) Note also that, if we have the system in a pure state $\left| \psi \right\rangle$, in which case $P = 1$ for that state, and there is zero probability for any other state, then we recover the usual result for the expectation value, i.e., $Tr\left(\rho \hat{A} \right) = \left\langle \psi \left| \hat{A} \right| \psi \right\rangle = \langle A \rangle$, so the density matrix description gives the correct answers for pure or mixed states.

Problems

14.3.1 Suppose we have a set of photons in a mixed state, with probabilities $P_1 = 0.2$ and $P_2 = 0.8$ respectively of being in the two different pure states

$$|\psi_1\rangle = |\psi_H\rangle \text{ and } |\psi_2\rangle = \frac{3}{5}|\psi_H\rangle + \frac{4i}{5}|\psi_V\rangle$$

where $|\psi_H\rangle$ and $|\psi_V\rangle$ are the normalized and orthogonal basis states representing horizontal and vertical polarization respectively. ($|\psi_1\rangle$ therefore is a horizontally polarized state, and $|\psi_2\rangle$ is an elliptically polarized state.) Write the density matrix for this state, in the $|\psi_H\rangle$ and $|\psi_V\rangle$ basis, with $\langle\psi_H|\rho|\psi_H\rangle$ as the top left element.

14.3.2 Consider the mixed spin state, with equal probabilities of the electrons being in the pure state $|s_x\rangle$ and the pure state $|s_y\rangle$. Here $|s_x\rangle$ and $|s_y\rangle$ are respectively spin states oriented along the $+x$ and $+y$ directions. (See Problem 14.1.1)

(i) Evaluate the density operator ρ on the z spin basis (i.e., $|\uparrow\rangle$ and $|\downarrow\rangle$)

(ii) Now write this density operator as a density matrix, with the term in $|\uparrow\rangle\langle\uparrow|$ in the top left element.

(iii) Taking the spin magnetic dipole moment operator to be $\hat{\boldsymbol{\mu}}_e = g\mu_B\hat{\boldsymbol{\sigma}}$, evaluate $\hat{\boldsymbol{\mu}}_e$ as a matrix on the same z spin basis (i.e., $|\uparrow\rangle$ and $|\downarrow\rangle$), with the element $\langle\uparrow|\hat{\boldsymbol{\mu}}_e|\uparrow\rangle$ in the top left corner.

(iv) Using the expression of the form $\overline{\langle A\rangle} = Tr(\rho\hat{A})$, evaluate the ensemble average expectation value for the spin magnetic dipole moment in this mixed state. [Hint: the answer should be the same as that for Problem 14.1.1 (ii)(a).]

14.4 Time-evolution of the density matrix

When we want to understand how a quantum mechanical system in some pure state $|\psi_j\rangle$ evolves, we can use the Schrödinger equation

$$\hat{H}|\psi_j\rangle = i\hbar\frac{\partial}{\partial t}|\psi_j\rangle \tag{14.15}$$

How can we describe the evolution of a mixed state? In principle, we can do this by considering each pure state in the mixture, and appropriately averaging the result, but there is a more elegant approach, which is directly to calculate the time-evolution of the density matrix. The density matrix, after all, contains all the available information about the mixed state. To see how to construct the appropriate equation for the density matrix, we start with the Schrödinger equation (14.15), and substitute using the expansion, Eq. (14.6), in some basis set $|\phi_n\rangle$, to obtain

$$i\hbar\sum_n \frac{\partial c_n^{(j)}(t)}{\partial t}|\phi_n\rangle = \sum_n c_n^{(j)}(t)\hat{H}|\phi_n\rangle \tag{14.16}$$

where we have put all of the time dependence of the state into the coefficients $c_n^{(j)}(t)$. Now operating from the left of Eq. (14.16) with $\langle\phi_m|$, we have

$$i\hbar\frac{\partial c_m^{(j)}(t)}{\partial t} = \sum_n c_n^{(j)}(t)H_{mn} \tag{14.17}$$

where $H_{mn} = \langle \phi_m | \hat{H} | \phi_n \rangle$ is a matrix element of the Hamiltonian. Also, we can take the complex conjugate of both sides of Eq. (14.17). Noting that $\hat{H}$ is Hermitian, i.e., $H_{mn}^* = H_{nm}$, we have

$$-i\hbar \frac{\partial \left(c_m^{(j)}(t) \right)^*}{\partial t} = \sum_n \left(c_n^{(j)}(t) \right)^* H_{nm} \tag{14.18}$$

or, trivially, changing indices

$$-i\hbar \frac{\partial \left(c_n^{(j)}(t) \right)^*}{\partial t} = \sum_s \left(c_s^{(j)}(t) \right)^* H_{sn} \tag{14.19}$$

But, from (14.8),

$$\frac{\partial \rho_{mn}}{\partial t} = \sum_j P_j \left(c_m^{(j)} \frac{\partial \left(c_n^{(j)} \right)^*}{\partial t} + \left(c_n^{(j)} \right)^* \frac{\partial c_m^{(j)}}{\partial t} \right) \tag{14.20}$$

So, substituting from Eqs. (14.17) and (14.19) (and changing the summation index in Eq. (14.17) from n to q), we have

$$\frac{\partial \rho_{mn}}{\partial t} = \sum_j P_j \left(\frac{i}{\hbar} c_m^{(j)} \sum_q \left(c_q^{(j)} \right)^* H_{qn} - \frac{i}{\hbar} \left(c_n^{(j)} \right)^* \sum_s c_s^{(j)} H_{ms} \right)$$
$$= \frac{i}{\hbar} \left(\sum_q \left(\sum_j P_j c_m^{(j)} \left(c_q^{(j)} \right)^* \right) H_{qn} - \sum_s H_{ms} \left(\sum_j P_j c_s^{(j)} \left(c_n^{(j)} \right)^* \right) \right) \tag{14.21}$$

So, using the definition Eq. (14.8) of the density matrix elements, we have

$$\frac{\partial \rho_{mn}}{\partial t} = \frac{i}{\hbar} \left(\sum_q \rho_{mq} H_{qn} - \sum_s H_{ms} \rho_{sn} \right)$$
$$= \frac{i}{\hbar} \left(\left(\rho \hat{H} \right)_{mn} - \left(\hat{H} \rho \right)_{mn} \right) \tag{14.22}$$
$$= \frac{i}{\hbar} \left[\rho, \hat{H} \right]_{mn}$$

or equivalently,

$$\frac{\partial \rho}{\partial t} = \frac{i}{\hbar} \left[\rho, \hat{H} \right] \tag{14.23}$$

Eq. (14.23) is therefore the equation that governs the time-evolution of the density matrix, and hence of the mixed state. This is a very useful equation in applications of the density matrix. It is sometimes known as the Liouville equation because it has the same form as the Liouville equation of classical mechanics.[5]

---

[5] The Liouville equation of classical mechanics is the equation of motion for the phase space probability distribution.

14.5 Interaction of light with a two-level "atomic" system

One simple and useful example of the use of the density matrix is the interaction of light with an electron system with two energy levels. We presume these levels have energies E_1 and E_2, with corresponding energy eigenfunctions $|\psi_1\rangle$ and $|\psi_2\rangle$. We will also presume for simplicity that the system is much smaller than an optical wavelength, so an incident optical field E will simply be uniform across the system, and we will take E to be polarized in the z direction. We will take a simple "electric dipole" interaction between the light and the electron, so that the change of energy of the electron as it is displaced by an amount z is $e\mathsf{E}z$. Hence we can take the perturbing Hamiltonian, in a semiclassical approximation for simplicity, to be

$$\hat{H}_p = e\mathsf{E}\,z = -\mathsf{E}\,\hat{\mu} \tag{14.24}$$

where $\hat{\mu}$ is the operator corresponding to the electric dipole, with matrix elements

$$\mu_{mn} = -e\langle\psi_m|z|\psi_n\rangle \tag{14.25}$$

so that

$$\left(\hat{H}_p\right)_{mn} \equiv H_{pmn} = -\mathsf{E}\mu_{mn} \tag{14.26}$$

For simplicity also, we will choose the states $|\psi_1\rangle$ and $|\psi_2\rangle$ both to have definite parity, so that

$$\mu_{11} = \mu_{22} = 0 \text{ and hence } H_{p11} = H_{p22} = 0 \tag{14.27}$$

and we are free to choose the relative phase of the two wavefunctions such that μ_{12} is real so that we have

$$\mu_{12} = \mu_{21} \equiv \mu_d \tag{14.28}$$

Hence the dipole operator of this system can be written as

$$\hat{\mu} = \begin{bmatrix} 0 & \mu_d \\ \mu_d & 0 \end{bmatrix} \tag{14.29}$$

and the perturbing Hamiltonian is

$$\hat{H}_p = \begin{bmatrix} 0 & -\mathsf{E}\mu_d \\ -\mathsf{E}\mu_d & 0 \end{bmatrix} \tag{14.30}$$

The unperturbed Hamiltonian $\hat{H}_o$ is just a 2 x 2 diagonal matrix on this basis, with E_1 and E_2 as the diagonal elements, so the total Hamiltonian is

$$\hat{H} = \hat{H}_o + \hat{H}_p = \begin{bmatrix} E_1 & -\mathsf{E}\mu_d \\ -\mathsf{E}\mu_d & E_2 \end{bmatrix} \tag{14.31}$$

The density matrix is also a 2 x 2 matrix because there are only two basis states under consideration here, and in general we can write it as

$$\rho = \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix} \tag{14.32}$$

The dipole induced in this system is important for two different reasons. First, as above, we see that it is closely related to the perturbing Hamiltonian. Second, it represents the response of the system to the electric field. Formally, the polarization of a system in response to electric field is the dipole moment per unit volume, and the relation between polarization and electric field allows us to define the electric susceptibilities or dielectric constants that we typically use to describe the optical properties of materials. So we particularly want to know the expectation value or ensemble average value of the dipole. We have not yet defined what the state of this system is, but we can still use (14.14) to write

$$\overline{\langle \mu \rangle} = Tr\left(\rho \hat{\mu} \right) \tag{14.33}$$

Using Eqs. (14.29) and (14.32), we have

$$\rho \hat{\mu} = \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix} \begin{bmatrix} 0 & \mu_d \\ \mu_d & 0 \end{bmatrix} = \begin{bmatrix} \rho_{12}\mu_d & \rho_{11}\mu_d \\ \rho_{22}\mu_d & \rho_{21}\mu_d \end{bmatrix} \tag{14.34}$$

Hence

$$\overline{\langle \mu \rangle} = \mu_d \left(\rho_{12} + \rho_{21} \right) \tag{14.35}$$

Now let us try to evaluate the behavior of the density matrix in time so we can deduce the behavior of $\overline{\langle \mu \rangle}$. We have, from Eq. (14.23) with the definitions of ρ from Eq. (14.32) and $\hat{H}$ from Eq. (14.31),[6]

$$
\begin{aligned}
\frac{d\rho}{dt} &= \frac{i}{\hbar}\left(\rho \hat{H} - \hat{H}\rho \right) \\
&= \frac{i}{\hbar}\left(\begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix} \begin{bmatrix} E_1 & -\mathsf{E}\mu_d \\ -\mathsf{E}\mu_d & E_2 \end{bmatrix} - \begin{bmatrix} E_1 & -\mathsf{E}\mu_d \\ -\mathsf{E}\mu_d & E_2 \end{bmatrix} \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix} \right) \\
&= \frac{i}{\hbar}\begin{bmatrix} -\mathsf{E}\mu_d\left(\rho_{12} - \rho_{21} \right) & -\mathsf{E}\mu_d\left(\rho_{11} - \rho_{22} \right) + \left(E_2 - E_1 \right)\rho_{12} \\ -\mathsf{E}\mu_d\left(\rho_{22} - \rho_{11} \right) + \left(E_1 - E_2 \right)\rho_{21} & -\mathsf{E}\mu_d\left(\rho_{21} - \rho_{12} \right) \end{bmatrix}
\end{aligned} \tag{14.36}
$$

There are two specific useful equations we can deduce from this. The first concerns $d\rho_{12}/dt$ or $d\rho_{21}/dt$. We only need consider one of these since ρ_{12} and ρ_{21} are complex conjugates of one another because of the Hermiticity of the density matrix. We have, taking the "2 − 1" element of both sides,

$$
\begin{aligned}
\frac{d\rho_{21}}{dt} &= \frac{i}{\hbar}\left(\left(\rho_{11} - \rho_{22} \right)\mathsf{E}\mu_d - \left(E_2 - E_1 \right)\rho_{21} \right) \\
&= -i\omega_{21}\rho_{21} + i\frac{\mu_d}{\hbar}\mathsf{E}\left(\rho_{11} - \rho_{22} \right)
\end{aligned} \tag{14.37}
$$

where $\hbar\omega_{21} = E_2 - E_1$. The second equation we will write relates to the change in $\rho_{11} - \rho_{22}$, which is essentially the fractional population difference between the lower and upper states.

$$\frac{d}{dt}\left(\rho_{11} - \rho_{22} \right) = 2i\frac{\mu_d}{\hbar}\mathsf{E}\left(\rho_{21} - \rho_{21}^* \right) \tag{14.38}$$

---

[6] We have changed from a partial derivative in time to a total derivative because there are no other variables in this problem.

where we have used the Hermiticity of ρ (which tells us that $\rho_{12} = \rho_{21}^*$).

This analysis and these equations (14.37) and (14.38) hint at various behaviors we might expect as we shine light on this "atom". For example, Eq. (14.38) shows that the relative fractional population of the upper and lower states of this atom is likely to be changing in time because of the presence of field E.

The analysis so far has made no approximations. In particular, this is not a perturbation theory analysis. Solving the pair of coupled equations (14.37) and (14.38) would cover any possible behavior of this idealized system. Of course, there is nothing in these equations so far that was not in the original time-dependent Schrödinger equation; solving that separately for each of the possible pure starting states $|\psi_j\rangle$ of interest, and then averaging the resulting expectation values for some quantity of interest, such as the dipole moment, would give us the same results as we would get from our density matrix analysis so far.

A key benefit of the density matrix is, however, that we can model additional random processes that lie outside the idealized problem, and about which we may know little. Suppose we consider the fractional population difference $\rho_{11} - \rho_{22}$ between the "lower" and "upper" states. We might know that, in thermal equilibrium, and with no optical electric field present, this difference settles down to some particular value, $(\rho_{11} - \rho_{22})_o$. Suppose then that, perhaps due to optical absorption from the lower to the upper state, we have some specific different fractional population difference $\rho_{11} - \rho_{22}$. Experience tells us that such systems often settle back to $(\rho_{11} - \rho_{22})_o$ with an exponential decay with a time constant T_1. This settling might result from random collisions of an atom, e.g., with the walls of the box containing the atom, or possibly with other atoms, or possibly just due to spontaneous emission. In this model, these processes can change the probabilities P_j of the various pure states, and hence we expect this density matrix picture could be appropriate. If we believe all that,[7] then we could hypothesize that we could add an appropriate term to Eq. (14.38), e.g.,

$$\frac{d}{dt}(\rho_{11} - \rho_{22}) = 2i\frac{\mu_d}{\hbar}\mathsf{E}(\rho_{21} - \rho_{21}^*) - \frac{(\rho_{11} - \rho_{22}) - (\rho_{11} - \rho_{22})_o}{T_1} \qquad (14.39)$$

We can see that, if there is no driving optical field E, Eq. (14.39) will indeed lead to an exponential settling of the fractional population difference to its equilibrium value $(\rho_{11} - \rho_{22})_o$ with a time constant T_1.[8]

In practice, though, just changing Eq. (14.38) into Eq. (14.39) is not sufficient for modeling such systems. We have to consider a similar process also for the off-diagonal elements of the density matrix, as in Eq. (14.37). To understand this, we need to understand the meaning of the off-diagonal elements.

Within any given pure state j, the product $c_u^{(j)}\left(c_v^{(j)}\right)^*$ is something that is in general oscillating. There is a time dependence $\exp(-iE_u t/\hbar)$ built into $c_u^{(j)}$ and associated with basis state u. Similarly $\left(c_v^{(j)}\right)^*$ has a time dependence $\exp(iE_v t/\hbar)$, so the product has an oscillation of the form $\exp(-i(E_u - E_v)t/\hbar)$. In our two-level system, even if we manage to prepare the system in one pure state with such an oscillation in this product, as time evolves we can imagine that our simple system can get scattered into another pure state k with some probability, possibly

---

[7] Or if we are just too lazy to construct a better model, such as correctly including spontaneous emission from first principles.

[8] This "T_1" relaxation of populations is sometimes called a "longitudinal relaxation".

even one in which the fractional populations ρ_{11} and ρ_{22} are unchanged, but in which the phases of the coefficients $c_1^{(k)}$ and $c_2^{(k)}$ are different. At any given time, therefore, we may have an ensemble of different possibilities for the quantum mechanical state, and in general these different possibilities will have different phases of oscillation. If we have sufficiently many such random phases that are sufficiently different in our mixed state, then the ensemble average of a product $c_u c_v^*$ for different u and v, i.e., $\overline{c_u c_v^*}$, will average out to zero. But this ensemble average is simply the off-diagonal density matrix element ρ_{uv}, as defined in Eq. (14.8).

Hence, we can see that these off-diagonal elements contain information about the coherence of the populations in different states. The collisions or other processes that could scatter into states with different phases for the expansion coefficients can be called "dephasing" processes. The simplest model we can construct is to postulate that such dephasing processes acting alone on the off-diagonal density matrix elements would give rise to an exponential settling of any off-diagonal element to zero, with some time constant T_2.[9] Hence we can postulate adding a term $-\rho_{21}/T_2$ to Eq. (14.37) to obtain

$$\frac{d\rho_{21}}{dt} = -i\omega_{21}\rho_{21} + i\frac{\mu_d}{\hbar}\mathsf{E}\left(\rho_{11} - \rho_{22}\right) - \frac{\rho_{21}}{T_2} \tag{14.40}$$

The dephasing rate or loss of coherence (the "T_2" process) is always comparable to or faster than the population relaxation (the "T_1" process)[10] because any process that relaxes the population by such a decay also destroys the coherence.[11] In this equation, in the absence of an optical field E, ρ_{21} would execute an oscillation at approximately frequency ω_{21}, decaying to zero amplitude approximately exponentially with a time constant T_2.

The equations (14.39) and (14.40) now constitute our model for our two-level system, including collision, relaxation, and dephasing processes that lie outside the Hamiltonian we know about for the system. We are particularly interested in solving for the behavior of the system for the case of an oscillating electric field, for example, of the form

$$\mathsf{E}(t) = \mathsf{E}_o \cos\omega t = \frac{\mathsf{E}_o}{2}\left(\exp(i\omega t) + \exp(-i\omega t)\right) \tag{14.41}$$

To solve the equations in this case, we make a substitution and one final simplifying approximation. As is often the case, it is useful to take the underlying oscillation out of the variables of interest. We therefore choose to define a new "slowly varying" quantity

$$\beta_{21}(t) = \rho_{21}(t)\exp(i\omega t) \tag{14.42}$$

and substitute using this to obtain, instead of Eqs. (14.39) and (14.40),

$$\frac{d}{dt}\left(\rho_{11} - \rho_{22}\right) = i\frac{\mu_d}{\hbar}\mathsf{E}_o\left(\beta_{21} - \beta_{21}^*\right) - \frac{\left(\rho_{11} - \rho_{22}\right) - \left(\rho_{11} - \rho_{22}\right)_o}{T_1} \tag{14.43}$$

---

[9] This "T_2" relaxation of populations is sometimes called a "transverse relaxation".

[10] Old professors often become incoherent long before they have stopped talking.

[11] More precisely, we can write $(1/T_2) = (1/2T_1) + (1/T_2')$ where T_2' describes dephasing processes that are in addition to the population relaxation. See, e.g., M. Fox, *Quantum Optics* (Oxford, 2006), p. 181

$$\frac{d\beta_{21}}{dt} = i\left(\omega - \omega_{21}\right)\beta_{21} + i\frac{\mu_d}{2\hbar}E_o\left(\rho_{11} - \rho_{22}\right) - \frac{\beta_{21}}{T_2} \tag{14.44}$$

where we have also made the approximation of dropping all terms proportional to $\exp(\pm 2i\omega t)$. Such terms will average out to zero over timescales of cycles, and hence will make relatively little contribution to the resulting values of $\rho_{11} - \rho_{22}$ and β_{21} that are obtained formally by integrating these equations in practice.[12] These equations (14.43) and (14.44) are often known as the optical Bloch equations.[13] In terms of β_{21}, the ensemble average of the dipole moment is now, from Eq. (14.35),

$$\begin{aligned}\overline{\langle\mu\rangle} &= \mu_d\left(\beta_{12}\exp\left(i\omega t\right) + \beta_{21}\exp\left(-i\omega t\right)\right) \\ &= 2\mu_d\left[\operatorname{Re}\left(\beta_{21}\right)\cos\omega t + \operatorname{Im}\left(\beta_{21}\right)\sin\omega t\right]\end{aligned} \tag{14.45}$$

where we have used the fact that $\beta_{21} = \beta_{12}^*$, which follows from the definition, Eq. (14.42) and the fact that the density matrix itself is Hermitian.

Now let us solve the equations (14.43) and (14.44) for one important and useful case. We want to know what happens in the "steady state" – i.e., for a monochromatic field and when the system has settled down. In that case, we expect first that the fractional population difference, $\rho_{11} - \rho_{22}$, will no longer be changing, so $d(\rho_{11} - \rho_{22})/dt = 0$. We also expect that any coherent responses we have from the system, as seen for example in the off-diagonal density matrix element ρ_{21}, will have settled down to following the appropriate driving field terms. Since we already took out the variation of the form $\exp(-i\omega t)$ in setting up the slowly-varying element β_{21}, we therefore expect that $d\beta_{21}/dt = 0$ in the steady state.

Therefore, setting the left-hand sides of both (14.43) and (14.44) to zero, we can solve these equations. Adding Eq. (14.43) and its complex conjugate gives us an expression for the real part of β_{21}, and similarly subtracting them gives the imaginary part. These can then be substituted into Eq. (14.44), leading to solution for $\rho_{11} - \rho_{22}$ and hence for all the desired variables. Hence we have

$$\rho_{11} - \rho_{22} = \left(\rho_{11} - \rho_{22}\right)_o \frac{1 + \left(\omega - \omega_{21}\right)^2 T_2^2}{1 + \left(\omega - \omega_{21}\right)^2 T_2^2 + 4\Omega^2 T_2 T_1} \tag{14.46}$$

$$\operatorname{Im}\left(\beta_{21}\right) = \frac{\Omega T_2\left(\rho_{11} - \rho_{22}\right)_o}{1 + \left(\omega - \omega_{21}\right)^2 T_2^2 + 4\Omega^2 T_2 T_1} \tag{14.47}$$

$$\operatorname{Re}\left(\beta_{21}\right) = \frac{\left(\omega_{21} - \omega\right)\Omega T_2^2\left(\rho_{11} - \rho_{22}\right)_o}{1 + \left(\omega - \omega_{21}\right)^2 T_2^2 + 4\Omega^2 T_2 T_1} \tag{14.48}$$

where $\Omega = \mu_d E_o / 2\hbar$.

To understand what these equations are telling us, we now presume that we have some large number N of such systems ("atoms") per unit volume. The population difference between the number in the lower state and the number in the higher state (per unit volume) is therefore

---

[12] This approximation is known as the "rotating wave" approximation.

[13] The actual Bloch equations in their simplest form were derived to describe another "two-level" system, namely spin-1/2 nuclei in a magnetic field. The optical Bloch equations here are mathematically closely analogous. See, e.g., M. Fox, *Quantum Optics* (Oxford, 2006) for a recent introductory discussion.

$\Delta N = N\left(\rho_{11} - \rho_{22}\right)$, and the population difference in the absence of the optical field is $\Delta N_o = N\left(\rho_{11} - \rho_{22}\right)_o$. Then instead of Eq. (14.46) we can write

$$\Delta N = \Delta N_o \frac{1 + \left(\omega - \omega_{21}\right)^2 T_2^2}{1 + \left(\omega - \omega_{21}\right)^2 T_2^2 + 4\Omega^2 T_2 T_1} \tag{14.49}$$

In general in electromagnetism, the (static) polarization P is defined as

$$\mathsf{P} = \varepsilon_o \chi \mathsf{E} \tag{14.50}$$

where χ is the susceptibility. When we have an oscillating field, the response of the medium, and hence the polarization, can be out of phase with the electric field, and then it is convenient to generalize the idea of susceptibility. We can formally think of it as a complex quantity with real and imaginary parts χ' and χ'' respectively, or equivalently we can explicitly write the response to a real field $\mathsf{E}_o \cos \omega t$ as

$$\mathsf{P} = \varepsilon_o \mathsf{E}_o \left(\chi' \cos \omega t + \chi'' \sin \omega t \right) \tag{14.51}$$

It is also generally true in electromagnetism that the polarization is the dipole moment per unit volume. Hence here we can also write

$$\mathsf{P} = N\overline{\langle \mu \rangle} \tag{14.52}$$

Hence, putting Eqs. (14.45), (14.47), (14.48), (14.51), and (14.52) together, we can write

$$\chi'(\omega) = \frac{\mu_d^2 T_2 \Delta N_o}{\varepsilon_o \hbar} \frac{\left(\omega_{21} - \omega\right) T_2}{1 + \left(\omega - \omega_{21}\right)^2 T_2^2 + 4\Omega^2 T_2 T_1} \tag{14.53}$$

$$\chi''(\omega) = \frac{\mu_d^2 T_2 \Delta N_o}{\varepsilon_o \hbar} \frac{1}{1 + \left(\omega - \omega_{21}\right)^2 T_2^2 + 4\Omega^2 T_2 T_1} \tag{14.54}$$

In electromagnetism, the in-phase component of the polarization (and hence χ' - the real part of χ) is responsible for refractive index, and the quadrature (i.e., 90 degrees shifted) component (and hence χ'' - the imaginary part of χ) is responsible for optical absorption. These behaviors can be checked by formally solving the wave equation, as derived from Maxwell's equations, with these susceptibilities present. In this particular case, therefore, what we have found is an expression for the refractive index and absorption of an atomic absorption line.

If we consider the case where the electric field amplitude is small, then $\Omega \simeq 0$, and we have the normal "linear" refraction variation

$$\chi'(\omega) = \frac{\mu_d^2 T_2 \Delta N_o}{\varepsilon_o \hbar} \frac{\left(\omega_{21} - \omega\right) T_2}{1 + \left(\omega - \omega_{21}\right)^2 T_2^2} \tag{14.55}$$

and a Lorentzian absorption line

$$\chi''(\omega) = \frac{\mu_d^2 T_2 \Delta N_o}{\varepsilon_o \hbar} \frac{1}{1 + \left(\omega - \omega_{21}\right)^2 T_2^2} \tag{14.56}$$

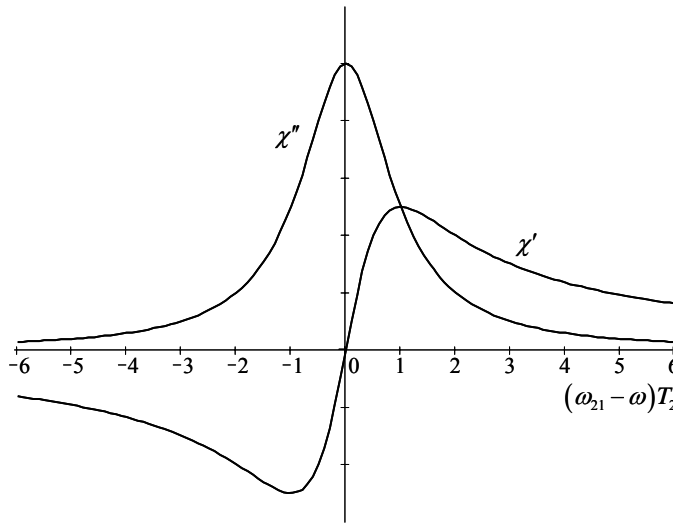associated with an "atomic" or "two-level" transition. These functions are sketched in Fig. 14.1.

Fig. 14.1. Sketch of the Lorentzian line shape for χ'' and the corresponding line shape for χ', as in Eqs. (14.56) and (14.55) respectively.

Note that, finally, we have got rid of the δ-function behavior of absorbing transitions that we got from simple time-dependent perturbation theory. Now we have an absorbing line with a much more reasonable shape, with a width. The parameter of the width is $1/T_2$, i.e., the more dephasing "collisions" there are with the "atom", the wider the absorption line. Incidentally, any process that leads to the recovery of the atom from its excited state back to its lower state will also cause a decay of the off-diagonal elements, and so even if there are no additional dephasing processes, the absorbing line will still have a width. When the only recovery process is spontaneous emission, the resulting line-width is called the natural line-width. We can view the line-width of the absorbing transition as being consistent with the energy-time uncertainty principle – if the state (or the coherence of the state) only persists for a finite time, then the state cannot have a well-defined energy, hence the line-width.

The approach we have taken so far is not perturbation theory. Within its simplifications and the approximations of simple times T_1 and T_2 to describe the relaxation of the system, it is exact for all fields. In particular, the results we have here model absorption saturation for this system. If we keep trying to absorb more and more photons into an ensemble of these "atoms", there will be fewer and fewer atoms in their lower states (and also more and more in their upper states), which removes these atoms from absorption. (It also allows them to show stimulated emission from the upper to the lower state, which is a process that is exactly the opposite of absorption.) Hence the absorption should progressively disappear as we go to higher intensities. This process is built in to the model we have here. The quantity Ω^2 is proportional to the electric field squared, which in turn is proportional to the intensity I of the light field. Hence we can write $4\Omega^2 T_2 T_1 \equiv I/I_S$ where I_S is called the saturation intensity. Hence, for example, on resonance ($\omega_{21} = \omega$), we have

$$\chi''(\omega) \propto \frac{1}{1 + I/I_S} \tag{14.57}$$

This equation is extensively used to describe the process of "absorption saturation" that is often encountered in absorbing systems illuminated with the high intensities available from lasers.

Problem

14.5.1   Suppose we have been driving an ensemble of "two-level" atoms with an energy separation between the two levels of $\hbar\omega_{21}$ and a (real) dipole matrix element μ_d between these two levels. The atoms have been illuminated with an electric field of the form $\mathsf{E}_o \cos \omega t$ for a sufficiently long time that the system has reached a steady state. We work in the approximation where we have characteristic times T_1 and T_2 to describe the population recovery time and the dephasing time, respectively. For simplicity also, we presume we are at a low temperature so that the equilibrium condition (in the absence of optical electric field) is that all the atoms are in their lower state.

(i)   Write down expressions for the values of the fractional difference $\rho_{11} - \rho_{22}$ between the lower and upper state populations of the atoms, and for the ensemble average $\langle \mu \rangle$ of the polarization of these atoms at time $t = 0$.

(ii)   At time $t = 0$, we suddenly cut off this driving field (i.e., we suddenly set $\mathsf{E}_o = 0$). Find expressions for the subsequent behavior of $\rho_{11} - \rho_{22}$ and $\langle \mu \rangle$. [Note: you may neglect the additional electric field that results from the polarization, for simplicity assuming that the total electric field is always zero for $t \geq 0$. You may use expressions based on the "rotating wave" approximation for this part.]

(iii) Presuming that $T_1 \gg T_2 \gg 1/\omega_{21}$ (e.g., $T_1 = 4T_2 = 20 \times 2\pi / \omega_{21}$), and for simplicity presuming we are operating far from resonance, i.e., $|\omega - \omega_{21}| \gg 1/T_2$, sketch the resulting behavior as a function of time of both the fraction of the population of atoms that are in their upper state, and the ensemble average of the polarization (in arbitrary units for both these quantities), indicating all the characteristic times on your sketch. [Note: it is easier here *not* to use the slowly-varying or rotating wave expressions for the off-diagonal elements.]

(iv) It might be that radiative recombination is the only process by which the atoms recover to their ground state (and hence that is the only process that contributes to T_1). You may find in part (ii) above that the ensemble average of the overall magnitude of the polarization decays faster than the population recovers to its ground state. Explain why that rapid polarization decay can still be consistent with the longer overall (radiative) recovery of the population.

14.6 Density matrix and perturbation theory

The discussion above of a two level system showed an exact solution of a simple problem, including both the linear response and also the non-linear response of absorption saturation. Just as in solutions of Schrödinger's equation, for more complicated systems, exact solutions are usually not possible. Again, just as for Schrödinger's time-dependent equation, we can use perturbation theory, but now with Eq. (14.22) or Eq. (14.23) for the time evolution of the density matrix instead of Schrödinger's equation. One common approach is to generalize the kinds of relaxation time approximations we introduced above, writing instead of Eq. (14.22)

$$\frac{\partial \rho_{mn}}{\partial t} = \frac{i}{\hbar}\left[\rho, \hat{H} \right]_{mn} - \gamma_{mn}\left(\rho_{mn} - \rho_{mno} \right) \qquad (14.58)$$

Here ρ_{mno} is the value to which density matrix element ρ_{mn} settles in the absence of excitation, and γ_{mn} is the "relaxation rate" of element ρ_{mn}. For our two-level system problem above, if we had chosen $\gamma_{11} = \gamma_{22} = 1/T_1$ and $\gamma_{21} = \gamma_{12} = 1/T_2$, then Eq. (14.58) would lead to Eqs. (14.39) and (14.40) that we had solved essentially exactly for that system.

In the common perturbation theory approach, one starts with the equations of the form (14.58) for each of the density matrix elements, and constructs a perturbation theory just as we had

done in Chapter 7 above when starting with the time-dependent Schrödinger equation. In fact, this density matrix version is the one commonly used for calculating non-linear optical coefficients, rather than the simple Schrödinger equation version of Chapter 7, because the inclusion of relaxation avoids all of the singularities that otherwise occur when the transition energy and the photon energy coincide. For reasons of space, we will not repeat this version of perturbation theory here, referring the reader to the literature on non-linear optics.[14]

Further reading

K. Blum, *Density Matrix Theory and Applications* (Plenum, New York, 1996) gives a comprehensive and accessible discussion of the density matrix and all its many applications.

M. Fox, *Quantum Optics* (Oxford, 2006) discusses the density matrix and its application to two-level systems in a clear introductory account.

A. Yariv, *Quantum Electronics (3rd Edition)* (Wiley, New York, 1989) gives an introductory account of the density matrix applied to a two-level system, in an approach similar to that we have taken here.

14.7 Summary of concepts

Pure state

A pure state in practice is one that can be written as a simple linear combination of quantum mechanical states, e.g., as in the form $b_1|\psi_1\rangle + b_2|\psi_2\rangle$. (It is not necessary that $|\psi_1\rangle$ and $|\psi_2\rangle$ are orthogonal.)

Ensemble average

If we have a system that can occupy any of a set of different states with probabilities P_n, then we can imagine constructing a set or "ensemble" of N replicas of the system, with P_1N of them in state 1, P_2N in state 2, and so on. An ensemble average is the average of some quantity when measured in each of the members of this ensemble.

Mixed state

A mixed state is one in which we only know probabilities P_j that the system is in one of the different pure states $|\psi_j\rangle$. (It is not necessary that the $|\psi_j\rangle$ are orthogonal.)

Ensemble average measured value of a quantity

For a system in a quantum-mechanical mixed state as above, the ensemble average value of a quantity that is represented by an operator $\hat{A}$ is

$$\overline{\langle A \rangle} = \sum_j P_j \langle \psi_j | \hat{A} | \psi_j \rangle \tag{14.4}$$

Density operator

The density operator is a way of representing a system that is in a mixed state, and is given by

---

[14] See in particular the excellent treatment in R. W. Boyd, *Nonlinear Optics* (Academic, New York, 1992), pp 116 – 148.

$$\rho = \sum_j P_j \left| \psi_j \right\rangle \left\langle \psi_j \right| \tag{14.5}$$

As a special case, with one particular $P_j = 1$, it can also represent a pure state, so the density operator and its associated formalism can be used for both pure and mixed states.

Note that, though it is an operator, in quantum mechanics it is not like other operators that represent variables, such as position, momentum or energy; instead, it represents the state of a system. To emphasize this difference, we represent it without the "hat" used on other operators.

Density matrix

The density operator is almost always represented by a matrix, in which case the matrix elements are

$$\rho_{uv} \equiv \left\langle \phi_u \left| \rho \right| \phi_v \right\rangle = \sum_j P_j c_u^{(j)} \left(c_v^{(j)} \right)^* \equiv \overline{c_u c_v^*} \tag{14.8}$$

where the c's are the expansion coefficients of the pure states on some complete orthonormal basis $\left| \phi_u \right\rangle$

$$\left| \psi_j \right\rangle = \sum_u c_u^{(j)} \left| \phi_u \right\rangle \tag{14.6}$$

and the "bar" notation on $\overline{c_u c_v^*}$ denotes an ensemble average.

The density matrix is Hermitian (and hence so also is the density operator), and

$$Tr(\rho) = 1 \tag{14.10}$$

Ensemble average from the density matrix

For a quantity represented by the quantum mechanical operator $\hat{A}$, the ensemble average measured value of that quantity when the system is in a state represented by density matrix ρ is given by

$$\overline{\langle A \rangle} = Tr\left(\rho \hat{A} \right) \tag{14.14}$$

which applies for both pure and mixed states.

Equation of motion for the density matrix

Just as the time-dependent Schrödinger equation can give the time-evolution of a system in a pure state represented by a wavefunction ψ, so also the equation

$$\frac{\partial \rho}{\partial t} = \frac{i}{\hbar} \left[\rho, \hat{H} \right] \tag{14.23}$$

can give the time-evolution of a system in a mixed (or pure) state represented by the density operator ρ. This same equation can also be written explicitly for each element of the corresponding density matrix ρ as

$$\frac{\partial \rho_{mn}}{\partial t} = \frac{i}{\hbar} \left[\rho, \hat{H} \right]_{mn} \tag{14.22}$$

Relaxation time approximation

When the system interacts in a random fashion with other elements that are not included in the Hamiltonian $\hat{H}$, the overall behavior can often be modeled by phenomenologically adding relaxation rates to obtain, instead of (14.22)

$$\frac{\partial \rho_{mn}}{\partial t} = \frac{i}{\hbar}\left[\rho, \hat{H}\right]_{mn} - \gamma_{mn}\left(\rho_{mn} - \rho_{mno}\right) \tag{14.58}$$

where ρ_{mno} is the equilibrium value of the density matrix element ρ_{mn} in the absence of excitation.

Optical Bloch equations

A specific pair of equations that follow from the relaxation time approximation above, with an additional approximation of the neglect of terms oscillating at 2ω (the "rotating wave" approximation), are the optical Bloch equations

$$\frac{d}{dt}\left(\rho_{11} - \rho_{22}\right) = i\frac{\mu_d}{\hbar}E_o\left(\beta_{21} - \beta_{21}^*\right) - \frac{\left(\rho_{11} - \rho_{22}\right) - \left(\rho_{11} - \rho_{22}\right)_o}{T_1} \tag{14.43}$$

$$\frac{d\beta_{21}}{dt} = i\left(\omega - \omega_{21}\right)\beta_{21} + i\frac{\mu_d}{2\hbar}E_o\left(\rho_{11} - \rho_{22}\right) - \frac{\beta_{21}}{T_2} \tag{14.44}$$

which describe an optically excited two-level system. Here $\beta_{21}(t) = \rho_{21}(t)\exp(i\omega t)$, and the system, with dipole matrix element μ_d, is being excited with an electric field of the form $E_o \cos \omega t$. These equations lead to the classic "Lorentzian" line shape of atomic transitions, and also explain absorption saturation in a simple model.