

TWO STATE SYSTEMS

© R. L. Jaffe 2003

1 Introduction

We have spent the first part of 8.05 setting up the kinematic and dynamical framework of quantum physics. We have developed a structure that is much more general than wave mechanics. Now is a good time to work through examples of operator quantum mechanics to solidify our understanding. Fortunately many beautiful and relatively modern applications of quantum mechanics involve the time development of systems where only *two quantum states* are important. A two-dimensional Hilbert space is so simple that all of the properties of these systems can be displayed analytically without approximation. Such systems are not only of pedagogical interest. Many interesting modern developments in quantum theory and applications to the real world involve systems where the restriction to two quantum states is natural. In these notes we introduce the general description of two state systems using the ammonia molecule as an example.

Applications to nuclear magnetic resonance, the ammonia maser, neutrino oscillations, and the physics of kaons, will soon follow.

2 General formalism

Suppose we can isolate and focus on two states of a quantum system to the exclusion of all others. It might be that the system only has two states. The classic example is the intrinsic spin of an electron or proton, a purely quantum mechanical label that takes on only two values: $+$ or $-$. Or it might be that all other degrees of freedom of the system are “frozen out”, requiring much more energy to excite. An example is the configuration of the ammonia molecule, NH_3 , where the nitrogen sits either above ($+$) or below ($-$) the plane formed by the three hydrogens. Excitations that rearrange the molecule in more significant ways or ionize it require much more energy than is required to flip the nitrogen from $+$ to $-$.

Whatever the reason, we shall assume that our Hilbert space is spanned by two basis states, $|+\rangle$ and $|-\rangle$, that we assume to be orthonormal, $\langle\pm|\pm\rangle = 1$, $\langle\pm|\mp\rangle = 0$ and complete, $\sum_{i=\pm} |i\rangle\langle i| = I$, where I is the identity operator.

2.1 Hamiltonian

In *isolation* this system is described by a *constant* Hamiltonian with eigenvalues E_{\pm} . For simplicity we assume \mathcal{H} is diagonal in the $|\pm\rangle$ basis. If it weren't we would find its eigenvectors, change to that basis and rename those states $+$ and $-$. We can choose the zero of energy so that the eigenvalues of \mathcal{H} are $\pm\epsilon/2$,

$$\mathcal{H}|\pm\rangle = \pm\frac{\epsilon}{2}|\pm\rangle. \quad (1)$$

From this we can construct the *matrix elements* of \mathcal{H} in this basis,

$$\begin{aligned} \langle\pm|\mathcal{H}|\pm\rangle &= \pm\frac{\epsilon}{2} \\ \langle\pm|\mathcal{H}|\mp\rangle &= 0. \end{aligned} \quad (2)$$

2.2 General Observable

Any quantity of physical interest must be associated with a hermitian operator, \mathcal{A} . To find the matrix representation of \mathcal{A} in the \pm basis we use completeness,

$$\begin{aligned} \mathcal{A} &= \left[\sum_{i=\pm} |i\rangle\langle i| \right] \mathcal{A} \left[\sum_{j=\pm} |j\rangle\langle j| \right] \\ &= \sum_{i,j=\pm} \langle i|\mathcal{A}|j\rangle \times |i\rangle\langle j| \end{aligned} \quad (3)$$

The operator $|i\rangle\langle j|$ transforms state j into state i . If we represent the state $|+\rangle$ by $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|-\rangle$ by $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ then $|i\rangle\langle j|$ is represented by the matrix with zeros everywhere except a 1 in the i th row and j th column. Let us define $\langle i|\mathcal{A}|j\rangle = a_{ij}$. Then we conclude that the operator \mathcal{A} is associated with the matrix

$$\mathcal{A} \Rightarrow \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad (4)$$

in the \pm basis. For example, \mathcal{H} is associated with the matrix

$$\mathcal{H} \Rightarrow \begin{pmatrix} \frac{\epsilon}{2} & 0 \\ 0 & -\frac{\epsilon}{2} \end{pmatrix}. \quad (5)$$

With respect to a given basis an operator is represented by a (2×2) matrix. When we want to emphasize the difference between the operator and its matrix representation we will use the \Rightarrow symbol as in (5). Usually, however, we will be less precise and relate operators to their matrix representations with equal signs.

Returning to the most general physical observable, hermiticity restricts the $\{a_{ij}\}$ by $a_{ij}^* = a_{ji}$. Without loss of generality we can define $a_{++} \equiv a_0 + a_3$, $a_{--} \equiv a_0 - a_3$, $a_{+-} \equiv a_1 - ia_2$, and $a_{-+} \equiv a_1 + ia_2$, where a_0 , a_1 , a_2 and a_3 are four real constants:

$$\mathcal{A} \Rightarrow \begin{pmatrix} a_0 + a_3 & a_1 - ia_2 \\ a_1 + ia_2 & a_0 - a_3 \end{pmatrix}, \quad (6)$$

So the operator \mathcal{A} has been replaced by a set of four c -numbers — a considerable simplification.

2.3 Pauli Matrices

When Pauli first encountered matrices of this form in the study of spin, he defined a set of standard 2×2 hermitian matrices now known as the *Pauli matrices*,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (7)$$

The Pauli matrices together with the identity matrix $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ form a complete set of 2×2 hermitian matrices. [They are a maximal linearly independent set.] They are *normalized* so that $\text{Tr } \sigma_k^2 = 2$, where Tr denotes the matrix space.

Using the Pauli matrices, \mathcal{H} can be written

$$\mathcal{H} \Rightarrow \frac{\epsilon}{2} \sigma_3 \quad (8)$$

and our arbitrary observable can be written

$$\mathcal{A} \Rightarrow a_0 I + \vec{a} \cdot \vec{\sigma}, \quad (9)$$

i.e.

$$\langle i | \mathcal{A} | j \rangle = a_0 \delta_{ij} + \sum_{k=1}^3 a_k \sigma_{ij}^k \quad (10)$$

Note that we can use a vector notation for the sum over the three σ matrices. This looks like a mere convenience, an accident of the fact that there are three non-trivial 2×2 Hermitian matrices, allowing us to make use of a familiar notation. Later, we will encounter examples where $\vec{\sigma}$ and \vec{a} *do* refer to three-dimensional vectors, and when we study the general behavior of two state systems more thoroughly we will see that the analogy goes still deeper.

The Pauli matrices have many simple properties that make it very convenient to use them in the manipulation of operators. First, their commutators and anticommutators are very simple:

$$\begin{aligned} [\sigma_j, \sigma_k] &= 2i\varepsilon_{jkl}\sigma_l \\ \{\sigma_j, \sigma_k\} &= 2\delta_{jk} \end{aligned} \quad (11)$$

where we use the summation convention that repeated indices are assumed to be summed over the labels 1, 2, and 3.¹ Also, $\{A, B\} = AB + BA$ is the *anticommutator*. The derivation of these relations is left to the problems. From (11) it follows that

$$\vec{a} \cdot \vec{\sigma} \vec{b} \cdot \vec{\sigma} = \vec{a} \cdot \vec{b} + i\sigma \cdot \vec{a} \times \vec{b}, \quad (12)$$

and

$$\text{Tr} \vec{a} \cdot \vec{\sigma} \vec{b} \cdot \vec{\sigma} \vec{c} \cdot \vec{\sigma} = 2i\vec{a} \cdot \vec{b} \times \vec{c}. \quad (13)$$

Sometimes it is useful to introduce non-hermitian combinations of the Pauli matrices. If we define $\sigma_{\pm} \equiv \frac{1}{\sqrt{2}}(\sigma_1 \pm i\sigma_2)$, then

$$\vec{a} \cdot \vec{\sigma} = a_+ \sigma_- + a_- \sigma_+ + a_3 \sigma_3 \quad (14)$$

¹ ε_{jkl} is the *Levi-Civita* symbol defined by $\varepsilon_{123} = 1$ and by the fact that it changes sign when any pair of its indices are exchanged. Thus $\varepsilon_{213} = \varepsilon_{321} = -1$ and $\varepsilon = 0$ if any two indices are the same.

where $a_{\pm} = \frac{1}{\sqrt{2}}(a_1 \pm ia_2)$. The commutation relations of the set $\{\sigma_{\pm}, \sigma_3\}$ are easily derived,

$$\begin{aligned} [\sigma_3, \sigma_{\pm}] &= \pm 2\sigma_{\pm} \\ [\sigma_+, \sigma_-] &= 2\sigma_3. \end{aligned} \quad (15)$$

Unitary matrices constructed as exponentials of the Pauli matrices have particularly simple properties. If we define $\mathcal{M}(\vec{a})$ by

$$\begin{aligned} \mathcal{M}(\vec{a}) &\equiv \exp i\vec{a} \cdot \vec{\sigma} \\ &\equiv 1 + i\vec{a} \cdot \vec{\sigma} + \frac{i^2}{2!}(\vec{a} \cdot \vec{\sigma})^2 + \frac{i^3}{3!}(\vec{a} \cdot \vec{\sigma})^3 + \dots \end{aligned} \quad (16)$$

then it is easy to show that

$$\mathcal{M}(\vec{a}) = I \cos a + i\vec{\sigma} \cdot \hat{a} \sin a, \quad (17)$$

where $a \equiv |\vec{a}|$ and $\hat{a} \equiv \vec{a}/a$. We leave the derivations of these results to the problems and leave any further ones to the point that we need them.

2.4 Time Dependence of an Isolated Two State System

The time development of an isolated system (one in which the Hamiltonian does not depend explicitly on the time) is generated by the unitary operator,

$$\mathcal{U}(t) = e^{-i\mathcal{H}t}. \quad (18)$$

Suppose $|\psi\rangle$ ($= |\psi\rangle_H = |\psi(0)\rangle_S$) is an arbitrary state defined as a superposition of the fixed basis vectors $|\pm\rangle$ at time $t = 0$. Then in accordance with our discussion of Heisenberg and Schroedinger pictures, we define,

$$\begin{aligned} |\psi\rangle_H &= a_+|+\rangle + a_-|-\rangle \\ |\psi(t)\rangle_S &= \mathcal{U}(t)|\psi(0)\rangle_S \\ &= a_+(t)|+\rangle + a_-(t)|-\rangle \end{aligned} \quad (19)$$

a_+ and a_- define the state at time $t = 0$. In the Schroedinger picture they evolve with time by the action of $\mathcal{U}(t)$ — $a_{\pm}(t) = e^{\mp i\epsilon t/2} a_{\pm}$.

It is particularly easy to follow the time dependence of observables in the Heisenberg picture. Let \mathcal{A} be an arbitrary observable defined by the

c -number constants a_0, \dots, a_3 , *a la* (9). The general relation between an operator in Schroedinger picture, \mathcal{A}_S and Heisenberg picture, $\mathcal{A}_H(t)$ is given by

$$\mathcal{A}_H(t) = e^{i\mathcal{H}t/\hbar} \mathcal{A}_S e^{-i\mathcal{H}t/\hbar} \quad (20)$$

or in this case,

$$\begin{aligned} &= e^{\frac{i\epsilon t}{2\hbar}\sigma_3} (a_0 I + \vec{a} \cdot \vec{\sigma}) e^{-\frac{i\epsilon t}{2\hbar}\sigma_3} \\ &= a_0 I + a_3 \sigma_3 + e^{i\lambda\sigma_3} (a_1 \sigma_1 + a_2 \sigma_2) e^{-i\lambda\sigma_3}, \end{aligned} \quad (21)$$

where $\lambda = \epsilon t/2\hbar$. It is possible to carry out the operator algebra exactly in this case. We define,

$$\Sigma_k(\lambda) = e^{i\lambda\sigma_3} \sigma_k e^{-i\lambda\sigma_3}, \quad (22)$$

so

$$\begin{aligned} \frac{d\Sigma_k}{d\lambda} &= i e^{i\lambda\sigma_3} [\sigma_3, \sigma_k] e^{-i\lambda\sigma_3} \\ &= -2\varepsilon_{3kj} \Sigma_j. \end{aligned} \quad (23)$$

Specifically, $d\Sigma_1/d\lambda = -2\Sigma_2$ and $d\Sigma_2/d\lambda = 2\Sigma_1$. Using the initial values $\Sigma_k(0) = \sigma_k$, it is easy to integrate these coupled first order (matrix) equations,

$$\begin{aligned} \Sigma_1(\lambda) &= \sigma_1 \cos 2\lambda - \sigma_2 \sin 2\lambda \\ \Sigma_2(\lambda) &= \sigma_2 \cos 2\lambda + \sigma_1 \sin 2\lambda \end{aligned} \quad (24)$$

Thus we have found the time dependence of the most general operator:

$$\mathcal{A}_H(t) = a_0 I + a_3 \sigma_3 + a_1 \left(\sigma_1 \cos \frac{\epsilon t}{\hbar} - \sigma_2 \sin \frac{\epsilon t}{\hbar} \right) + a_2 \left(\sigma_2 \cos \frac{\epsilon t}{\hbar} + \sigma_1 \sin \frac{\epsilon t}{\hbar} \right) \quad (25)$$

Let us analyse this result. The piece of an operator proportional to the identity, a_0 , labels both states the same way and is of no interest to us. It may be regarded as an overall property of the system that does not vary in the problem at hand. The piece proportional to σ_3 commutes with \mathcal{H} , so it labels the two energy eigenstates differently and distinguishes between them. The pieces of \mathcal{A} proportional to σ_1 and σ_2 vary sinusoidally in time with a frequency determined by the energy difference between the two eigenstates.

We are now ready to look at a few simple examples of two state systems.

3 Ammonia Molecule I

The presentation here follows the discussion in Feynman's *Lectures on Physics*, Volume III, §9. Ammonia (NH_3) forms a tetrahedral structure where the nitrogen atom sits above or below the plane formed by the three hydrogen atoms. When all other degrees of freedom (electronic, vibrational, rotational) are in their respective ground states, the NH_3 molecule still has two distinct configurations: $|+\rangle$ — when the N is *above* the plane of the H_3 , and $|-\rangle$ — when it is *below*. Since the two configurations are related by mirror reflection in the plane of the hydrogens, it is not surprising that the expectation value of the energy is the same in the two states: $\langle +|\mathcal{H}|+\rangle = \langle -|\mathcal{H}|-\rangle \equiv E_0$. However, atomic interactions mix the two states.² Without any information about the interaction we can nonetheless summarize it by a single parameter, Δ , the magnitude of the off-diagonal matrix element of \mathcal{H} , $\langle +|\mathcal{H}|-\rangle = \langle -|\mathcal{H}|+\rangle \equiv -\Delta$. We can choose the phase of the states so that Δ is real.³

In Nature Δ is small. It is much smaller than the energy required to excite electronic, vibrational or rotational degrees of freedom (which require optical, infrared and far-infrared photons respectively). Δ corresponds to excitations in the microwave part of the electromagnetic spectrum, a frequency of $\Delta/\hbar \approx 24,000$ Megacycles, corresponding to a wavelength of ≈ 1.25 cm.

In the $\{\pm\}$ basis, \mathcal{H} is represented by the matrix,

$$\mathcal{H} \Rightarrow \begin{pmatrix} E_0 & -\Delta \\ -\Delta & E_0 \end{pmatrix} \quad (26)$$

Clearly $|+\rangle$ and $|-\rangle$ are not energy eigenstates. It is easy to see that $|1, 2\rangle \equiv \frac{1}{\sqrt{2}}(|+\rangle \mp |-\rangle)$ are eigenstates of (26) with eigenvalues $E_0 \pm \Delta$ respectively. So in the new basis, \mathcal{H} has the form described in the previous section,

$$\mathcal{H} = E_0 I + \sigma_3 \Delta. \quad (27)$$

As an example of an observable, let us choose the probability to observe the system in the state $|+\rangle$ minus the probability to find it in the state $|-\rangle$

²A nice example of the quantum mechanical phenomenon of *tunnelling*: the nitrogen atom sits stably on either side of the hydrogen plane, but has a small amplitude to tunnel through the plane where it would be classically forbidden.

³Note, we have chosen the phase of the states so that $\langle +|\mathcal{H}|-\rangle$ is negative. This is merely a convenience. You should convince yourself that we have the freedom to do this.

divided by the sum, which is unity. Our first job is to find the operator corresponding to this observation. Let us call the operator \mathcal{A} . We know its expectation value in *any* state. If the state is $|\psi\rangle = a_+|+\rangle + a_-|-\rangle$ then the expectation value is $\langle\psi|\mathcal{A}|\psi\rangle = |a_+|^2 - |a_-|^2$. An operator that does the job is $\mathcal{A} = [|+\rangle\langle+|] - [|-\rangle\langle-|]$. Since this choice works for *any* state, it is unique. In matrix terms

$$\mathcal{A} \Rightarrow \sigma_3, \quad (28)$$

in the $\{\pm\}$ basis.

We are now faced with a simple example of a common problem: the Hamiltonian for this system is simple (represented by a diagonal matrix — or *diagonal* for short) is one basis, while the operator of interest is diagonal in another. We have to choose one basis or the other. Let's choose to work in the $\{1, 2\}$ basis. We use m, n for $\{1, 2\}$ state labels and i, j for $\{\pm\}$ state labels. We need $\langle m|\mathcal{A}|n\rangle$, where m, n are either 1 or 2. Inserting complete sets $I = \sum_{i=\pm} |i\rangle\langle i|$, we get

$$\langle m|\mathcal{A}|n\rangle = \sum_{i=\pm} \sum_{j=\pm} \langle m|i\rangle\langle i|\mathcal{A}|j\rangle\langle j|n\rangle \quad (29)$$

The matrix $\langle j|n\rangle$ defines the unitary transformation from the \pm basis to the $\{1, 2\}$ basis. Since the basis states are related by

$$\begin{aligned} |1\rangle &= \frac{1}{\sqrt{2}}(|+\rangle - |-\rangle) \\ |2\rangle &= \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle), \end{aligned} \quad (30)$$

we can read off the matrix elements of the matrix $\mathcal{T}_{jn} = \langle j|n\rangle$

$$\mathcal{T} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \quad (31)$$

Substituting this definition of $\langle j|n\rangle$ in terms of \mathcal{T} back into (29) we obtain the matrix representation of \mathcal{A} in the $\{1, 2\}$ basis,

$$\mathcal{A} \Rightarrow \mathcal{T}^\dagger \sigma_3 \mathcal{T} = \sigma_1. \quad (32)$$

To find the time dependence of \mathcal{A} we substitute into (24),⁴

$$\mathcal{A}(t) = \sigma_1 \cos \frac{2\Delta t}{\hbar} - \sigma_2 \sin \frac{2\Delta t}{\hbar} \quad (33)$$

To find out the time dependence of the asymmetry in any state we take the expectation value of the Heisenberg operator, (33), in the (time independent) Heisenberg state. \mathcal{A} vanishes in an energy eigenstate, $\langle 1, 2 | \mathcal{A} | 1, 2 \rangle = 0$ as it must. So the energy eigenstates populate the up and down configurations equally. If, however, we prepare a state in the up configuration at $t = 0$, $|+\rangle$, then the asymmetry oscillates with time, $\langle + | \mathcal{A}(t) | + \rangle = \cos \frac{2\Delta t}{\hbar}$. One virtue of the Heisenberg picture is that the entire range of time evolution of operator matrix elements is manifest in the Heisenberg operator. It is easy, for example, to construct an initial state that selects out the $\sin \frac{2\Delta t}{\hbar}$ term in (33). We will return to the ammonia molecule after exploring several relatively simple and interesting two-state problems.

⁴I have gone through this example of changing an operator representation from one basis to another in lots of detail to provide an illustration of the general methods discussed in lecture.