Time dependent perturbation theory[1]

D. E. Soper[2]
University of Oregon
11 May 2012

I offer here some background for Chapter 5 of J. J. Sakurai, *Modern Quantum Mechanics*.

1 The problem

Let the hamiltonian for a system of interest have the form

$$H(t) = H_0 + V(t) \ . \tag{1}$$

Here H_0 is time-independent. We assume that we know the eigenvectors and eigenvalues of H_0. The interaction hamiltonian V can be time independent or time dependent.

We wish to solve the Schrödinger equation

$$i\,\frac{d}{dt}\,\big|\psi(t)\big\rangle = (H_0 + V(t))\big|\psi(t)\big\rangle \ , \tag{2}$$

working order by order in powers of V.

2 The interaction picture

In the first section of these notes, we have been working in the usual Schrödinger picture. Now let us switch to the "interaction picture" defined by

$$\big|\psi_{\rm I}(t)\big\rangle = e^{iH_0 t}\big|\psi_{\rm S}(t)\big\rangle \ . \tag{3}$$

One easily derives (see the derivation below for $U_I(t,t_0)$) that

$$i\,\frac{d}{dt}\,\big|\psi_{\rm I}(t)\big\rangle = V_{\rm I}(t)\big|\psi_{\rm I}(t)\big\rangle \ , \tag{4}$$

[2]soper@uoregon.edu

where

$$V_{\rm I}(t) = e^{iH_0 t} V_{\rm S}(t) e^{-iH_0 t} \quad . \tag{5}$$

Note that $V_{\rm I}(t)$ is time dependent even if $V_{\rm S}(t)$ is time independent (unless V commutes with H_0). Note also that $V_{\rm I}(t)$ at one time generally does not commute with $V_{\rm I}(t)$ at another time. Finally, note that if, for some time interval, $V_{\rm S}(t)$ vanishes, then $\big|\psi_{\rm I}(t)\big\rangle$ does not evolve during that time interval. That is, essentially, the point of working in the interaction picture.

To express the evolution in the interaction picture using a starting time t_0 at which initial conditions are specified, it is convenient to write

$$\begin{aligned} \big|\psi_{\rm I}(t)\big\rangle &= e^{iH_0 t}\big|\psi_{\rm S}(t)\big\rangle \\ &= e^{iH_0 t} U_{\rm S}(t,t_0)\big|\psi_{\rm S}(t_0)\big\rangle \\ &= e^{iH_0 t} U_{\rm S}(t,t_0) e^{-iH_0 t}\big|\psi_{\rm I}(t_0)\big\rangle \quad . \end{aligned} \tag{6}$$

Here $U_{\rm S}(t,t_0)$ is the Schrödinger picture evolution operator, equal to

$$\exp(-iH(t-t_0))$$

if the hamiltonian is time independent. When the hamiltonian depends on time, $U_{\rm S}(t,t_0)$ is more complicated, but all that we need is the differential equation that it obeys,

$$i\frac{d}{dt}\, U_{\rm S}(t,t_0) = H(t)\, U_{\rm S}(t,t_0) \quad . \tag{7}$$

Thus we have

$$\big|\psi_{\rm I}(t)\big\rangle = U_{\rm I}(t,t_0)\big|\psi_{\rm I}(t_0)\big\rangle \quad , \tag{8}$$

where

$$U_{\rm I}(t,t_0) = e^{iH_0 t} U_{\rm S}(t,t_0) e^{-iH_0 t_0} \quad . \tag{9}$$

We have

$$\begin{aligned} i\,\frac{d}{dt}\, U_{\rm I}(t,t_0) &= e^{iH_0 t}[-H_0 + H_{\rm S}(t)] U_{\rm S}(t,t_0) e^{-iH_0 t_0} \\ &= e^{iH_0 t} V_{\rm S}(t) U_{\rm S}(t,t_0) e^{-iH_0 t_0} \\ &= [e^{iH_0 t} V_{\rm S}(t) e^{-iH_0 t}]\, e^{iH_0 t} U_{\rm S}(t,t_0) e^{-iH_0 t_0} \quad . \end{aligned} \tag{10}$$

Thus

$$i\,\frac{d}{dt}\, U_{\rm I}(t,t_0) = V_{\rm I}(t)\, U_{\rm I}(t,t_0) \tag{11}$$

with the initial condition

$$U_{\rm I}(t_0,t_0) = 1 \quad . \tag{12}$$

3 Perturbative solution

We can write an integral equation for $U_{\rm I}(t,t_0)$,

$$U_{\rm I}(t,t_0) = 1 - i\int_{t_0}^{t} d\tau\ V_{\rm I}(\tau)\, U_{\rm I}(\tau,t_0)\ \ . \tag{13}$$

The reader should check that if $U_{\rm I}$ obeys this integral equation then it obeys both the differential equation and the boundary condition.

We can think of this as follows. The operator $U_{\rm I}(t,t_0)$ tells us that the system evolves from time t_0 to time t. The system can evolve in two ways. First, it may be that nothing happens: the "1" term. Second, it may be that the system evolves up to some time τ, then an interaction happens at time τ, then nothing happens from time τ to time t. We integrate over possible times τ.

It is easy to solve this by iteration starting at $U_{\rm I}(t,t_0) = 1$ so as to generate a solution that is a power series in V. We write

$$U_I(t,t_0) = \sum_{n=0}^{\infty} U_I^{(n)}(t,t_0)\ \ , \tag{14}$$

where $U_I^{(n)}(t,t_0)$ is proportional to exactly n powers of $V_{\rm I}$. Then

$$U_I^{(0)}(t,t_0) = 1\ \ . \tag{15}$$

The integral equation gives

$$U_{\rm I}^{(n+1)}(t,t_0) = -i\int_{t_0}^{t} d\tau\ V_{\rm I}(\tau)\, U_{\rm I}^{(n)}(t,t_0)\ \ . \tag{16}$$

Substituting $U_I^{(0)} = 1$, we get

$$U_{\rm I}^{(1)}(t,t_0) = -i\int_{t_0}^{t} d\tau_1\ V_{\rm I}(\tau_1)\ \ . \tag{17}$$

Then we get

$$U_{\rm I}^{(2)}(t,t_0) = (-i)^2\int_{t_0}^{t} d\tau_2 \int_{t_0}^{\tau_2} d\tau_1\ V_{\rm I}(\tau_2)\, V_{\rm I}(\tau_1)\ \ . \tag{18}$$

Continuing in this manner, we get

$$U_{\rm I}^{(n)}(t,t_0) = (-i)^n \int_{t_0}^{t} d\tau_n \cdots \int_{t_0}^{\tau_3} d\tau_2 \int_{t_0}^{\tau_2} d\tau_1 \; V_{\rm I}(\tau_n) \cdots V_{\rm I}(\tau_2)\, V_{\rm I}(\tau_1) \quad . \tag{19}$$

Another way to write the integration is

$$\begin{aligned} U_{\rm I}^{(n)}(t,t_0) = (-i)^n \int_{t_0}^{t} d\tau_n \cdots \int_{t_0}^{t} d\tau_2 \int_{t_0}^{t} d\tau_1 \; \theta(\tau_1 < \tau_2 \cdots < \tau_{n-1} < \tau_n) \\ \times V_{\rm I}(\tau_n) \cdots V_{\rm I}(\tau_2)\, V_{\rm I}(\tau_1) \quad . \end{aligned} \tag{20}$$

That is, we integrate over the times τ_j in the interval from t_0 to t subject to the restriction that

$$\tau_1 < \tau_2 \cdots < \tau_{n-1} < \tau_n \; . \tag{21}$$

4 Shorthand notation

There is a convenient shorthand for writing our result. We could write

$$\begin{aligned} U_{\rm I}^{(n)}(t,t_0) = (-i)^n \int_{t_0}^{t} d\xi_n \cdots \int_{t_0}^{t} d\xi_1 \\ \times \frac{1}{n!} \sum_{\pi \in S_n} \theta(\xi_{\pi(1)} < \xi_{\pi(2)} \cdots < \xi_{\pi(n-1)} < \xi_{\pi(n)}) \\ \times V_{\rm I}(\xi_{\pi(n)}) \cdots V_{\rm I}(\xi_{\pi(2)})\, V_{\rm I}(\xi_{\pi(1)}) \quad . \end{aligned} \tag{22}$$

This may seem perverse. I have just made up new variable names ξ_j and matched them up to the τ_j in according to the $n!$ permutations of n objects. The $n!$ terms each gives the same integral. It is important that I always put the operators $V(\tau)$ that have the later times to the left in the product of operators.

I can write this as

$$\begin{aligned} U_{\rm I}^{(n)}(t,t_0) = (-i)^n \int_{t_0}^{t} d\xi_n \cdots \int_{t_0}^{t} d\xi_1 \\ \times \frac{1}{n!} \sum_{\pi \in S_n} \theta(\xi_{\pi(1)} < \xi_{\pi(2)} \cdots < \xi_{\pi(n-1)} < \xi_{\pi(n)}) \\ \times T\, V_{\rm I}(\xi_n) \cdots V_{\rm I}(\xi_2)\, V_{\rm I}(\xi_1) \quad . \end{aligned} \tag{23}$$

Here the notation "T" instructs me to put the operators $V(\tau)$ that have the later times to the left in the product of operators. Then I can use

$$\sum_{\pi\in S_n} \theta(\xi_{\pi(1)} < \xi_{\pi(1)} \cdots < \xi_{\pi(n-1)} < \xi_{\pi(n)}) = 1 \quad . \tag{24}$$

This gives

$$U_{\rm I}^{(n)}(t,t_0) = (-i)^n \frac{1}{n!} \int_{t_0}^{t} d\xi_n \cdots \int_{t_0}^{t} d\xi_1 \; T\, V_{\rm I}(\xi_n) \cdots V_{\rm I}(\xi_2)\, V_{\rm I}(\xi_1) \quad . \tag{25}$$

With the "T" notation, the whole series can be written as

$$U_{\rm I}(t,t_0) = T \exp\left(-i \int_{t_0}^{t} d\tau \; V_{\rm I}(\tau) \right) \quad . \tag{26}$$

This is a nice, compact expression. What it *means* is that one is supposed to expand the exponential and time-order the operators. That is, what it means is Eq. (20).

5 Two level system

As an example, let us consider a two level system subject to a perturbing interaction that links the two levels and oscillates with frequency ω,

$$V_{\rm S}(t) = e^{-i\omega t} \begin{pmatrix} 0 & \gamma \\ \gamma & 0 \end{pmatrix} \quad . \tag{27}$$

Then the perturbing potential in the interaction picture is

$$V_{\rm I}(t) = e^{-i\omega t} \begin{pmatrix} e^{i\omega_1 t} & 0 \\ 0 & e^{i\omega_2 t} \end{pmatrix} \begin{pmatrix} 0 & \gamma \\ \gamma & 0 \end{pmatrix} \begin{pmatrix} e^{-i\omega_1 t} & 0 \\ 0 & e^{-i\omega_2 t} \end{pmatrix} \quad . \tag{28}$$

where ω_2 is the unperturbed energy of the upper level, ω_1 is the unperturbed energy of the lower level. When we multiply the matrices, we obtain

$$V_{\rm I}(t) = \begin{pmatrix} 0 & \gamma e^{-2i\tilde{\omega}_2 t} \\ \gamma e^{-2i\tilde{\omega}_1 t} & 0 \end{pmatrix} \quad , \tag{29}$$

where

$$\begin{aligned} \tilde{\omega}_2 &= \frac{1}{2}\left[\omega - (\omega_2 - \omega_1)\right] \quad , \\ \tilde{\omega}_1 &= \frac{1}{2}\left[\omega - (\omega_1 - \omega_2)\right] \end{aligned} \tag{30}$$

We take the initial time to be $t_0 = 0$.

With this perturbation, the evolution operator evaluated to first order is

$$\begin{aligned} U_{\rm I}(t,0) &= 1 - i\int_0^t d\tau\ V_I(\tau) + \cdots \\ &= 1 - i\gamma \int_0^t d\tau\ \begin{pmatrix} 0 & e^{-2i\tilde{\omega}_2 t} \\ e^{-2i\tilde{\omega}_1 t} & 0 \end{pmatrix} + \cdots \\ &= 1 - i\gamma \begin{pmatrix} 0 & \frac{i}{2\tilde{\omega}_2}[e^{-2i\tilde{\omega}_2 t} - 1] \\ \frac{i}{2\tilde{\omega}_1}[e^{2i\tilde{\omega}_1 t} - 1] & 0 \end{pmatrix} + \cdots \\ &= 1 - i\gamma \begin{pmatrix} 0 & \frac{i}{2\tilde{\omega}_2} e^{-i\tilde{\omega}_2 t}[e^{-i\tilde{\omega}_2 t} - e^{i\tilde{\omega}_2 t}] \\ \frac{i}{2\tilde{\omega}_1} e^{-i\tilde{\omega}_1 t}[e^{-i\tilde{\omega}_1 t} - e^{i\tilde{\omega}_1 t}] & 0 \end{pmatrix} + \cdots \\ &= 1 - i\gamma \begin{pmatrix} 0 & \frac{1}{\tilde{\omega}_2} e^{-i\tilde{\omega}_2 t}\sin(\tilde{\omega}_2 t) \\ \frac{1}{\tilde{\omega}_1} e^{-i\tilde{\omega}_1 t}\sin(\tilde{\omega}_1 t) & 0 \end{pmatrix} + \cdots \ . \end{aligned} \tag{31}$$

If the initial condition is

$$\psi_{\rm I}(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{32}$$

then at later times $\psi_{\rm I}(t) = U_I(t,0)\psi_{\rm I}(0)$ is given by

$$\psi_{\rm I}(t) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - i\gamma\, \frac{\sin(\tilde{\omega}_2 t)}{\tilde{\omega}_2} e^{-i\tilde{\omega}_2 t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \cdots \ . \tag{33}$$

This appears to be singular for $\tilde{\omega}_2 \to 0$, but note that the factor $\sin(\tilde{\omega}_2 t)$ vanishes for $\tilde{\omega}_2 \to 0$ so that the singularity is cancelled.

The probability to find the system in the excited state after time t is

$$|C_2(t,\tilde{\omega}_2)|^2 \equiv |\langle 2|\psi_I(t)\rangle|^2 = \gamma^2\, \frac{\sin^2(\tilde{\omega}_2 t)}{\tilde{\omega}_2^2} \ . \tag{34}$$

This solution applies when γ is small. If we do not have $\gamma^2/\tilde{\omega}_2 \ll 1$ then first order perturbation theory is not accurate and we need the exact solution for the two level system with a harmonic perturbation, which is analyzed in Sakurai.

6 Behavior integrated over energy

Let us consider a slightly generalized problem. Suppose that we have a perturbation

$$V_S(t) = e^{-i\omega t} V_0 \tag{35}$$

and suppose that the perturbation can take the system from a starting state $\big|1\big\rangle$ with unperturbed energy ω_1 at time $t_0 = 0$ to a range of states $\big|j\big\rangle$ that have unperturbed energies ω_j. Then the first order transition amplitude to get from state $\big|1\big\rangle$ to state $\big|j\big\rangle$ is

$$\begin{aligned}
\big\langle j\big|U_I^{(1)}(t,0)\big|1\big\rangle &= -i\int_0^t d\tau\ \big\langle j\big|V_I(\tau)\big|1\big\rangle \\
&= -i\int_0^t d\tau\ \big\langle j\big|e^{iH_0\tau}V_0e^{-i\omega\tau}e^{-iH_0\tau}\big|1\big\rangle \\
&= -i\int_0^t d\tau\ e^{-i(\omega+\omega_1-\omega_j)\tau}\big\langle j\big|V_0\big|1\big\rangle
\end{aligned} \tag{36}$$

Let us define

$$\begin{aligned}
\tilde{\omega}_j &= \frac{1}{2}\left[\omega - (\omega_j - \omega_1)\right] \\
\gamma_j &= \big\langle j\big|V_0\big|1\big\rangle
\end{aligned} \tag{37}$$

Then

$$\begin{aligned}
\big\langle j\big|U_I^{(1)}(t,0)\big|1\big\rangle &= -i\gamma_j\int_0^t d\tau\ e^{-2i\tilde{\omega}_j\tau} \\
&= \frac{\gamma_j}{2\tilde{\omega}_j}e^{-i\tilde{\omega}_jt}[e^{-i\tilde{\omega}_jt} - e^{i\tilde{\omega}_jt}] \\
&= -i\,\frac{\gamma_j}{\tilde{\omega}_j}e^{-i\tilde{\omega}_jt}\sin(\tilde{\omega}_jt)
\end{aligned} \tag{38}$$

Thus the probability to get to state $\big|j\big\rangle$, evaluated in lowest order perturbation theory, is

$$\big|\big\langle j\big|\psi_I(t)\big\rangle\big|^2 = |\gamma_j|^2\,\frac{\sin^2(\tilde{\omega}_jt)}{\tilde{\omega}_j^2}\quad . \tag{39}$$

The result in this form may be just what we want. However, in many cases, this is not what we want. First, we want t to be large. Second, the function $\big|\big\langle j\big|\psi_I(t)\big\rangle\big|^2$ will often appear inside an integral over $\tilde{\omega}_j$. For instance, there may be a continuous range of excited states $\big|j\big\rangle$. If we want the probability to go to any of these states, we need to sum over j, which amounts, for the case of a continuous spectrum, to integrating over ω_j. That is, we want to integrate over $\tilde{\omega}_j$.

In order to see what we get in a general way, suppose that we are interested in an integral

$$I[h,t] \equiv \int d\tilde{\omega}_j \; \frac{\sin^2(\tilde{\omega}_j t)}{\tilde{\omega}_j^2} \; h(\tilde{\omega}_j) \quad . \tag{40}$$

Here $h(\tilde{\omega}_j)$ represents whatever function of $\tilde{\omega}_j$ appears in our problem. It includes a factor of $|\gamma_j|^2$. Our analysis will apply to any function as long as it is smooth and well behaved at large $\tilde{\omega}_j$. We want to know the limiting behavior of $I[h,t]$ for large t.

This is an easy problem. We note that $\sin^2(\tilde{\omega}_j t)/\tilde{\omega}_j^2$ is very sharply peaked at $\tilde{\omega}_j = 0$, with a width $\delta\tilde{\omega}_j \sim 1/t$. For this reason, only values of $\tilde{\omega}_j$ near zero matter in the integral. This allows us to approximate $h(\tilde{\omega}_j)$ by $h(0)$, giving

$$I[h,t] \sim h(0) \int d\tilde{\omega}_j \; \frac{\sin^2(\tilde{\omega}_j t)}{\tilde{\omega}_j^2} \quad . \tag{41}$$

All that we have to do now is to approximate the integral for $t \to \infty$. Change variables to $x = \tilde{\omega}_j t$, giving

$$I[h,t] \sim h(0)\, t \int dx \; \frac{\sin^2(x)}{x^2} \quad . \tag{42}$$

The integral is finite and equals π. Thus

$$I[h,t] \sim h(0)\, \pi\, t \quad . \tag{43}$$

That is

$$\frac{\sin^2(\tilde{\omega}_j t)}{\tilde{\omega}_j^2} \sim \pi\, t\, \delta(\tilde{\omega}_j) \qquad \text{for } t \to \infty \quad . \tag{44}$$

The presence of the delta function indicates that this large t limiting behavior applies inside an integration.

This result has a physical interpretation. Recall that the probability to get to state $\big|j\big\rangle$, evaluated in lowest order perturbation theory, is

$$|\big\langle j\big|\psi_I(t)\big\rangle|^2 = |\gamma_j|^2 \, \frac{\sin^2(\tilde{\omega}_j t)}{\tilde{\omega}_j^2} \quad . \tag{45}$$

We have now seen that for large t this becomes

$$|\big\langle j\big|\psi_I(t)\big\rangle|^2 \sim |\gamma_j|^2 \, \pi\, t\, \delta(\tilde{\omega}_j) \quad . \tag{46}$$

Therefore the probability per unit time to get to state $\big|j\big\rangle$ is

$$\frac{1}{t}\,|\big\langle j\big|\psi_I(t)\big\rangle|^2 \sim |\gamma_j|^2\,2\pi\,\delta(\omega-(\omega_j-\omega_1))\ \ . \tag{47}$$

Here I have used the definition of $\tilde{\omega}$ and moved a factor 1/2 outside of the delta function.

7 Behavior integrated over energy II

The analysis of the previous section has followed, more or less, that in Sakurai. Here is another way to think about it.

Suppose that the matrix element of the perturbing potential between state $\big|1\big\rangle$ and state $\big|j\big\rangle$ is

$$\big\langle j\big|V_I(t)\big|1\big\rangle = \gamma_j e^{-i(\omega-\omega_j+\omega_1)t}\,g(t)\ \ . \tag{48}$$

What we had before was that $g(t)=1$ and that we specified the initial condition that the system started in state $\big|1\big\rangle$ at $t=0$. Now, let's suppose that the system started in state $\big|1\big\rangle$ at $t_0=-\infty$. However, $V_I(t)$ is turned off for times far in the past. Then the interaction is turned on and stays on for a long time. Finally the interaction is turned off. We describe turning the potential on and off with the function $g(t)$. We want $g(t)=0$ for times far in the past and for times far in the future, while $g(t)=1$ for intermediate times. We can imagine that the turning on is smooth. We could also take $g(t)=\theta(0<t<T)$, which would give the same results as in the previous section.

If $g(t)=\theta(0<t<T)$, then T is the time that the perturbation was on. We can generalize this. We note that g is proportional to, say, the amplitude of the electromagnetic field that perturbs the system. The intensity of the field is $|g(t)|^2$. We define the effective time that the field is on (weighted by the intensity) by

$$T=\int dt\ |g(t)|^2\ \ . \tag{49}$$

Now we are ready for perturbation theory. The amplitude to find the

system in state $\big|j\big\rangle$ at a time t far in the future is

$$\begin{aligned}\big\langle j\big|U_I^{(1)}(\infty,-\infty)\big|1\big\rangle &= -i\int dt\ \big\langle j\big|V_I(t)\big|1\big\rangle \\ &= -i\gamma_j\int dt\ e^{-i(\omega-\omega_j+\omega_1)t}\,g(t) \\ &= -i\gamma_j\,\tilde{g}(-\omega+\omega_j-\omega_1)\ ,\end{aligned} \tag{50}$$

where $\tilde{g}(\omega')$ is the Fourier transform of $g(t)$. The probability to find the system in state $\big|j\big\rangle$ at a time t far in the future is

$$|\big\langle j\big|U_I^{(1)}(\infty,-\infty)\big|1\big\rangle|^2 = |\gamma_j|^2\,|\tilde{g}(-\omega+\omega_j-\omega_1)|^2\ . \tag{51}$$

Now $g(t)$ is very flat as a function of time – almost 1 for a long time. Therefore $\tilde{g}(\omega')$ as a function of frequency is very sharply peaked near $\omega'=0$ – almost proportional to $\delta(\omega')$. That is, for any function $h(\omega_j)$ we will have

$$\int d\omega_j|\tilde{g}(-\omega+\omega_j-\omega_1)|^2h(\omega_j) \sim h(\omega+\omega_1)\int d\omega_j|\tilde{g}(-\omega+\omega_j-\omega_1)|^2\ . \tag{52}$$

That is,

$$|\big\langle j\big|U_I^{(1)}(\infty,-\infty)\big|0\big\rangle|^2 \sim |\gamma_j|^2\,2\pi\delta(-\omega+\omega_j-\omega_1)\times I\ , \tag{53}$$

where

$$I=\int\frac{d\omega'}{2\pi}\ |\tilde{g}(\omega')|^2 \tag{54}$$

We can use the general theorem

$$\int\frac{d\omega'}{2\pi}\ |\tilde{g}(\omega')|^2 = \int dt\ |g(t)|^2 \tag{55}$$

together with our definition Eq. (49) to obtain

$$I=T\ . \tag{56}$$

Thus

$$|\big\langle j\big|U_I^{(1)}(\infty,-\infty)\big|1\big\rangle|^2 \sim |\gamma_j|^2\,2\pi\delta(-\omega+\omega_j-\omega_1)\times T\ , \tag{57}$$

We conclude that the probability to find the system in state $\big|j\big\rangle$ at a time far in the future, per unit time that the potential was on, is

$$\frac{1}{T}\,|\big\langle j\big|U_I^{(1)}(\infty,-\infty)\big|1\big\rangle|^2 \sim |\gamma_j|^2\,2\pi\delta(-\omega+\omega_j-\omega_1)\ . \tag{58}$$

This assumes that the time T is very large and that we will integrate over one or more of the frequencies that appear in the delta function. There has been an approximation: the delta function is not really a delta function but rather a sharply peaked function with a width $\delta\omega'$ of order $1/T$; the function is normalized so that the area under the peak is 1.

8 Cross section

Eq. (58) gives the probability per unit time for the system to get to a state $\big|j\big\rangle$ if it starts in state $\big|1\big\rangle$. There can also be a range $\mathcal{R}$ of states, in which case we want to sum over j in the range. The probability per unit time to find the system in a state $j \in \mathcal{R}$ is then

$$\frac{1}{T}\,P(\mathcal{R}) = \sum_j \theta(j \in \mathcal{R}) |\gamma_j|^2\, 2\pi\delta(-\omega + \omega_j - \omega_1) \ . \tag{59}$$

In applications, the sum here may be an integral.

In the following section, we consider an atom placed in a beam of photons. We will represent the beam of photons by a classical electromagnetic field, and then in Sec. 10 we will see how to do the same thing with a real quantum description of photons. With either description, we need the key concept of cross section. The cross section to find $j \in \mathcal{R}$ is defined to be

$$\sigma(\mathcal{R}) = \frac{1}{\mathcal{F}T}\,P(\mathcal{R}) \ . \tag{60}$$

Here $\mathcal{F}$ is the "flux" of incoming photons: the number of photons per unit time per unit area. Note that then $\sigma(\mathcal{R})$ has the dimensions of area. In a simple classical picture in which the atom has area A and every photon that hits the atom excites it to a state $j \in \mathcal{R}$, the cross section is precisely A.

In first order perturbation theory, the cross section is

$$\sigma(\mathcal{R}) = \sum_j \theta(j \in \mathcal{R}) |\gamma'_j|^2\, 2\pi\delta(-\omega + \omega_j - \omega_1) \ . \tag{61}$$

Here

$$|\gamma'_j|^2 = \frac{|\gamma_j|^2}{\mathcal{F}} = \frac{|\big\langle j\big|V\big|1\big\rangle|^2}{\mathcal{F}} \ . \tag{62}$$

As we will see, V is proportional to the strength of the classical electromagnetic field strength that we use in the calculation and $\mathcal{F}$ is proportional to the square of the field strength, so $|\gamma'_j|^2$ is independent of the field strength.

In some cases, the photon beam can consist of a mixture of photons with different frequencies, so that the number of photons crossing the position of the atom per unit time and per unit area in the frequency range $d\omega$ is $d\mathcal{F} = \rho(\omega)d\omega$. This would be the case, for instance, for a beam of sunlight impinging on an atom. In this case, the probability per unit time to find the system in a state $j \in \mathcal{R}$ is

$$\frac{1}{T}\,P(\mathcal{R}) = \int d\omega\ \rho(\omega)\sigma(\mathcal{R},\omega)\ \ , \tag{63}$$

where $\sigma(\mathcal{R},\omega)$ is the cross section for a monochromatic beam. Thus in first order perturbation theory,

$$\frac{1}{T}\,P(\mathcal{R}) = \int d\omega\ \rho(\omega)\sum_j \theta(j\in\mathcal{R})|\gamma'_j|^2\,2\pi\delta(-\omega+\omega_j-\omega_1)\ \ . \tag{64}$$

In this case, even if there is only one, discrete, final state $\langle j|$, the factor $\delta(-\omega+\omega_j-\omega_1)$ is absorbed by the integration over the frequencies of the initial photon.

9 Absorbing and emitting photons

In quantum mechanics, one describes the electromagnetic field using the vector potential $\vec{A}(\vec{x},t)$ and the scalar potential $\phi(\vec{x},t)$. It is useful to choose the Coulomb gauge

$$\vec{\nabla}\cdot\vec{A} = 0\ \ . \tag{65}$$

In that gauge, a beam of light is described by

$$\vec{A}(\vec{x},t) = 2A_0\,\vec{\varepsilon}\cos(\vec{k}\cdot\vec{x}-\omega t)\ \ . \tag{66}$$

That is

$$\vec{A}(\vec{x},t) = A_0\,\vec{\varepsilon}\exp(-i(\vec{k}\cdot\vec{x}-\omega t)) + A_0\,\vec{\varepsilon}\exp(+i(\vec{k}\cdot\vec{x}-\omega t))\ \ . \tag{67}$$

Here $\vec{\varepsilon}$ gives the polarization direction of the light; $\vec{\varepsilon}$ is normalized to $\vec{\varepsilon}^2 = 1$ and $\vec{k}\cdot\vec{\varepsilon} = 0$. The wave is moving in the $\vec{k}$ direction. The angular frequency

of the radiation is $\omega = |\vec{k}|$. Note that I have taken units with $c = 1$. The scalar potential associated with the radiation is zero. However, in an atom there will be a potential $\phi(\vec{x})$ associated with the electric field of the nucleus or the nucleus plus other electrons.

The hamiltonian for an electron of charge $-e$ interacting with the electromagnetic field is[3]

$$\begin{aligned} H &= \frac{1}{2m}\,(\vec{p} + e\vec{A}(\vec{x},t))^2 - e\phi(\vec{x},t) \\ &= \frac{1}{2m}\,\vec{p}^{\,2} + \frac{e}{2m}\,(\vec{p}\cdot\vec{A} + \vec{A}\cdot\vec{p}) + \frac{e^2}{2m}\,\vec{A}^{\,2} - e\phi \\ &= \frac{1}{2m}\,\vec{p}^{\,2} + \frac{e}{m}\,\vec{A}\cdot\vec{p} - \frac{ie}{2m}\,(\vec{\nabla}\cdot\vec{A}) + \frac{e^2}{2m}\,\vec{A}^{\,2} - e\phi \\ &= \frac{1}{2m}\,\vec{p}^{\,2} + \frac{e}{m}\,\vec{A}\cdot\vec{p} + \frac{e^2}{2m}\,\vec{A}^{\,2} - e\phi \ . \end{aligned} \tag{68}$$

We will use this for doing first order perturbation theory, with

$$H_0 = \frac{1}{2m}\,\vec{p}^{\,2} - e\phi \tag{69}$$

and

$$V_{\rm S}(t) = \frac{e}{m}\,\vec{A}(\vec{x},t)\cdot\vec{p}\ . \tag{70}$$

Here I have dropped the term proportional to $\vec{A}^{\,2}$ because it contains two powers of e and we are doing first order perturbation theory in powers of e.

Let's use this for first order perturbation theory. First, we suppose that the radiation is turned on for a large but not infinite amount of time T, as described in the previous section by using a function $g(t)$, so that we use

$$V_{\rm S}(t) = \frac{e}{m}\,g(t)A_0\,\exp(i(\vec{k}\cdot\vec{x} - \omega t))\,\vec{\varepsilon}\cdot\vec{p} + \frac{e}{m}\,g(t)\,A_0\,\exp(-i(\vec{k}\cdot\vec{x} - \omega t))\,\vec{\varepsilon}\cdot\vec{p}\ . \tag{71}$$

Second, we simplify our calculation by using only one of these terms. We will want to calculate the amplitude for the atom to go from state $\big|1\big\rangle$ to state $\big|j\big\rangle$. This amplitude will have two terms, corresponding to the two terms in $V(t)$. The first term will be large for $\omega = \omega_j - \omega_1$ and nearly zero otherwise. The second will be large for $\omega = -\omega_j + \omega_1$ and nearly zero otherwise. Let

[3]My notation is that $\vec{k}$ is the photon momentum, a numerical vector, while $\vec{p}$ is the momentum operator for the electron.

us suppose that we are interested in $\omega_j > \omega_1$. Then only the first term will contribute. Thus we can simplify the calculation by taking

$$V_{\rm S}(t) = \frac{e}{m}\, g(t) A_0 \exp(i(\vec{k}\cdot\vec{x} - \omega t))\, \vec{\varepsilon}\cdot\vec{p}\ . \tag{72}$$

In first order perturbation theory using Eq. (72), the amplitude for the atom to go from state $\big|1\big\rangle$ to state $\big|j\big\rangle$ is

$$\big\langle j\big|U_{\rm I}^{(1)}(\infty,-\infty)\big|1\big\rangle = -i\int_{-\infty}^{+\infty} dt\ \big\langle j\big|V_{\rm S}(t)\big|1\big\rangle e^{i\omega_j t - i\omega_1 t}\ . \tag{73}$$

The probability for a transition per unit time that the potential was on is then

$$\begin{aligned}\frac{1}{T}\,|\big\langle j\big|U_I^{(1)}(\infty,-\infty)\big|1\big\rangle|^2 \sim\ & 2\pi\delta(-\omega+\omega_j-\omega_1) \\ & \times |\big\langle j\big|\frac{e}{m}\, A_0 \exp(i\vec{k}\cdot\vec{x})\, \vec{\varepsilon}\cdot\vec{p}\,\big|1\big\rangle|^2\ .\end{aligned} \tag{74}$$

(This follows from our expression in Eq. (58), with γ_j replaced by the matrix element that is the coefficient of $\exp(-i\omega t)$ in V.)

Cross section. Our atoms are being hit by a beam of photons. The flux of photons – the number per unit area per unit time – is

$$\mathcal{F} = \frac{\omega}{2\pi}\, A_0^2\ . \tag{75}$$

This is obtained by calculating the energy per unit area per unit time in the light beam and dividing by the energy per photon, ω.

One defines the cross section for absorbing photons as

$$\sigma(1\to j) = \frac{1}{\mathcal{F}T}\,|\big\langle j\big|U_I^{(1)}(\infty,-\infty)\big|1\big\rangle|^2\ . \tag{76}$$

Substituting Eqs. (74) and (75) into Eq. (76), we have for the cross section

$$\sigma(1\to j) \sim 2\pi\delta(-\omega+\omega_j-\omega_1)\,\frac{2\pi e^2}{\omega m^2}\,|\big\langle j\big|\exp(i\vec{k}\cdot\vec{x})\,\vec{\varepsilon}\cdot\vec{p}\,\big|1\big\rangle|^2\ . \tag{77}$$

Dipole approximation. The wavelength of light with $\omega = \omega_j - \omega_1$ for atomic states is generally much bigger than the size of an atom. For this reason, we can approximate

$$\exp(i\vec{k}\cdot\vec{x}) \to 1\ . \tag{78}$$

This gives

$$\sigma(1 \to j) \sim 2\pi\delta(-\omega + \omega_j - \omega_1)\, \frac{2\pi e^2}{\omega m^2}\, |\vec{\varepsilon} \cdot \langle j|\vec{p}|1\rangle|^2 \ . \tag{79}$$

We also have

$$p_i = im\,[H_0, x_i] \ . \tag{80}$$

Thus

$$\vec{\varepsilon} \cdot \langle j|\vec{p}|1\rangle = im(\omega_j - \omega_1)\ \vec{\varepsilon} \cdot \langle j|\vec{x}|1\rangle \ . \tag{81}$$

Because of the delta function, we can replace $(\omega_j - \omega_1) \to \omega$. This gives

$$\sigma(1 \to j) \sim 2\pi\delta(-\omega + \omega_j - \omega_1)\, 2\pi e^2 \omega\, |\vec{\varepsilon} \cdot \langle j|\vec{x}\,|1\rangle|^2 \ . \tag{82}$$

Note that the operator $-e\vec{x}$ measures contribution of the electron to the electric dipole moment of the atom. Also note that we have simplified the notation by considering just one electron, but we could have had lots of electrons. Then we would have a sum over electrons i of the electron positions $\vec{x}_i$.

10 Really absorbing and emitting photons

The presentation in the preceding section has been with an external classical electromagnetic field. One can also use the quantum electromagnetic field. Then we really have photons.

We can designate our photon states by $|\vec{k}, \lambda\rangle$. This represents a single photon with momentum $\vec{k}$ and polarization vector $\vec{\varepsilon}(\vec{k}, \lambda)$ specified by λ. We have $\vec{\varepsilon}(\vec{k}, \lambda)^2 = 1$ and $\vec{k} \cdot \vec{\varepsilon}(\vec{k}, \lambda) = 0$. Given $\vec{k}$, there are two linearly independent ways that $\vec{\varepsilon}$ can point and these are specified by $\lambda = 1$ or $\lambda = 2$. We choose to normalize states following the convention that we have been using for states of other particles:

$$\langle \vec{k}', \lambda'|\vec{k}, \lambda\rangle = \delta_{\lambda'\lambda}\, \delta(\vec{k}' - \vec{k}) \ . \tag{83}$$

The state with no photons is called $|0\rangle$, "the vacuum" or, in the case that we still have an atom, the photon vacuum.

We will need to know what is the flux of photons if there is just one of them. With our chosen normalization, the wave function for the photon is

$$\psi(\vec{x}, \lambda') = \langle \vec{x}, \lambda'|\vec{k}, \lambda\rangle = \frac{1}{(2\pi)^{3/2}}\, \delta_{\lambda'\lambda}\, e^{i\vec{k}\vec{x}} \tag{84}$$

The probability to find a photon in a box of size $d\vec{x}$ is then

$$\sum_{\lambda'} |\psi(\vec{x},\lambda')|^2 d\vec{x} = \frac{d\vec{x}}{(2\pi)^3} \quad . \tag{85}$$

That is, the density of photons is $1/(2\pi)^3$. The flux is then the density times the speed, which (with $c = 1$) is 1. Thus

$$\mathcal{F} = \frac{1}{(2\pi)^3} \quad . \tag{86}$$

Now that we know about photons, we need to know that the quantum electromagnetic potential destroys them or creates them. To see how that happens, we need quantum field theory. The development of quantum field theory is a rather long story, but the answer is easy to state. For photon destruction, we have

$$\big\langle 0\big|\vec{A}(\vec{x},t)\big|\vec{k},\lambda\big\rangle = \frac{1}{2\pi\sqrt{\omega}}\,\vec{\varepsilon}(\vec{k},\lambda)\,e^{i(\vec{k}\cdot\vec{x}-\omega t)} \quad , \tag{87}$$

where $\omega = |\vec{k}\,|$. In this formula, I have assigned the $e^{-i\omega t}$ time dependence to the operator $\vec{A}(\vec{x},t)$ instead of to the photon states. That is, I have $\vec{A}(\vec{x},t)$ and $\big|\vec{k},\lambda\big\rangle$ instead of $\vec{A}(\vec{x})$ and $\big|\vec{k},\lambda,t\big\rangle$. This puts us in the interaction picture with respect to the photons. This picture seems most natural if we want to connect with the formalism in which there is a time dependent classical field $\vec{A}(\vec{x},t)$. For photon creation, we have a similar formula,

$$\big\langle\vec{k},\lambda\big|\vec{A}(\vec{x},t)\big|0\big\rangle = \frac{1}{2\pi\sqrt{\omega}}\,\vec{\varepsilon}(\vec{k},\lambda)^*\,e^{-i(\vec{k}\cdot\vec{x}-\omega t)} \quad . \tag{88}$$

(We will normally take $\vec{\varepsilon}$ to be real, but it can also be complex, so we distinguish between $\vec{\varepsilon}$ and $\vec{\varepsilon}^{\,*}$.)

Now we use Eq. (70) for the perturbation $V_{\rm S}(t)$, including a "turn on" function $g(t)$,

$$V_{\rm S}(t) = \frac{e}{m}\,\vec{A}(\vec{x},t)\cdot\vec{p}\,g(t) \quad , \tag{89}$$

but now $\vec{A}(\vec{x},t)$ is the quantum operator. Next, we work out the amplitude for a transition from atom state $\big|1\big\rangle$ and photon state $\big|\vec{k},\lambda\big\rangle$ to atom state $\big\langle j\big|$ and photon state $\big\langle 0\big|$,

$$\begin{aligned}\big\langle j;0\big|U_{\rm I}^{(1)}(\infty,-\infty)\big|1;\vec{k},\lambda\big\rangle = & -i\frac{e}{m}\int_{-\infty}^{+\infty} dt\; g(t)\,e^{i\omega_j t - i\omega_1 t} \\ & \times\big\langle j;0\big|\vec{A}(\vec{x},t)\cdot\vec{p}\,\big|1;\vec{k},\lambda\big\rangle \quad .\end{aligned} \tag{90}$$

Using Eq. (87), this becomes

$$\begin{aligned}\left\langle j;0\middle|U_{\rm I}^{(1)}(\infty,-\infty)\middle|1;\vec{k},\lambda\right\rangle = & -i\frac{e}{m}\,\frac{1}{2\pi\sqrt{\omega}}\int_{-\infty}^{+\infty}dt\;g(t)\,e^{i(\omega_j-\omega_1-\omega)t} \\ & \times\left\langle j\middle|e^{i\vec{k}\cdot\vec{x}}\vec{\varepsilon}(\vec{k},\lambda)\cdot\vec{p}\middle|1\right\rangle\end{aligned} \tag{91}$$

Using the same derivation as previously, we see that the probability for a transition per unit time that the interaction was on is

$$\begin{aligned}\frac{1}{T}\left|\left\langle j;0\middle|U_I^{(1)}(\infty,-\infty)\middle|1;\vec{k},\lambda\right\rangle\right|^2 \sim & \,2\pi\delta(\omega_j-\omega_1-\omega)\,\frac{e^2}{m^2(2\pi)^2\omega} \\ & \times\left|\left\langle j\middle|\exp(i\vec{k}\cdot\vec{x})\,\vec{\varepsilon}\cdot\vec{p}\middle|1\right\rangle\right|^2\ .\end{aligned} \tag{92}$$

If we divide by the flux from Eq. (86), we have for the cross section

$$\sigma(1+\gamma\to j)\sim 2\pi\delta(-\omega+\omega_j-\omega_1)\,\frac{2\pi e^2}{\omega m^2}\left|\left\langle j\middle|\exp(i\vec{k}\cdot\vec{x})\,\vec{\varepsilon}\cdot\vec{p}\middle|1\right\rangle\right|^2\ . \tag{93}$$

This is the same as the result that we got in Eq. (77) by using the device of an external electromagnetic field. Thus we can do a problem involving exciting atoms with photons by either considering actual quantum photons or by using an external electromagnetic field. There is a difference if we start the atom in an excited state. It will decay by emitting a photon. Then there is no external electromagnetic field: the atom decays all by itself. To calculate the decay probability per unit time, we need quantum photons. If there is an external electromagnetic field, as in a laser cavity, that helps the atoms to decay. This is called stimulated emission. For stimulated emission, we can use either the classical picture or the photon picture, accounting for all of the photons in the laser cavity. I leave that for an optics course.

11 Photoelectric effect

Let us try out these ideas by calculating (approximately) the cross section for ionizing a hydrogen atom by hitting it with a beam of photons. We start with Eq. (77) with state $\left|j\right\rangle$ taken to be the state $\left|\vec{k}_{\rm f}\right\rangle$ consisting of an ionized electron with momentum $\vec{k}_{\rm f}$,

$$d\sigma(1\to\vec{k}_{\rm f})\sim 2\pi\delta(-\omega+\omega_{\rm f}-\omega_1)\,d\vec{k}_{\rm f}\,\frac{2\pi e^2}{\omega m^2}\left|\left\langle\vec{k}_{\rm f}\middle|\exp(i\vec{k}\cdot\vec{x})\,\vec{\varepsilon}\cdot\vec{p}\middle|1\right\rangle\right|^2\ . \tag{94}$$

Here there is a continuous range of states $\left|\vec{k}_{\rm f}\right\rangle$, so it doesn't make sense to talk about a cross section to go to exactly one of them. Rather, we consider integrating over a range of states. The factor $d\vec{k}_{\rm f}$ indicates that we are going to integrate over some set $\mathcal{S}$ of momenta $\vec{k}_{\rm f}$. The total cross section for finding our electron in the desired range is then written as

$$\int d\sigma\, \theta(\vec{k}_{\rm f} \in \mathcal{S}) = \int d\vec{k}_{\rm f}\ \theta(\vec{k}_{\rm f} \in \mathcal{S}) \cdots \tag{95}$$

For this to work consistently with the meaning of cross section as probability per unit time that the beam is on and per unit flux in the beam, we need to normalize the states $\left|\vec{k}_{\rm f}\right\rangle$ so that[4]

$$1 = \int d\vec{k}_{\rm f}\ \left|\vec{k}_{\rm f}\right\rangle\left\langle\vec{k}_{\rm f}\right| \quad . \tag{96}$$

One can take care of the restriction provided by the energy-conserving delta function by writing

$$\begin{aligned} d\vec{k}_{\rm f}\ \delta(-\omega+\omega_{\rm f}-\omega_1) &= k_{\rm f}^2 dk_{\rm f} d\Omega_{\rm f}\, \delta(-\omega+\omega_{\rm f}-\omega_1) \\ &= m k_{\rm f}\, d\left(\frac{k_{\rm f}^2}{2m}\right) d\Omega_{\rm f}\, \delta\left(-\omega+\frac{k_{\rm f}^2}{2m}-\omega_1\right) \\ &= m k_{\rm f}\, d\Omega_{\rm f} \quad , \end{aligned} \tag{97}$$

We have "integrated out" the energy-conserving delta function, thus fixing the absolute value of $\vec{k}_{\rm f}$ to

$$k_{\rm f} = \sqrt{2m(\omega_1+\omega)} \quad . \tag{98}$$

This gives

$$d\sigma(1 \to \vec{k}_{\rm f}) \sim d\Omega_{\rm f}\, \frac{4\pi^2 e^2 k_{\rm f}}{\omega m}\, |\left\langle\vec{k}_{\rm f}\right| \exp(i\vec{k}\cdot\vec{x})\, \vec{\varepsilon}\cdot\vec{p}\left|1\right\rangle|^2 \quad . \tag{99}$$

This is usually expressed using the notation

$$\frac{d\sigma}{d\Omega_{\rm f}} \sim \frac{4\pi^2 e^2 k_{\rm f}}{\omega m}\, |\left\langle\vec{k}_{\rm f}\right| \exp(i\vec{k}\cdot\vec{x})\, \vec{\varepsilon}\cdot\vec{p}\left|1\right\rangle|^2 \quad . \tag{100}$$

[4]Note that Sakurai takes the whole experiment to take place in a big box, so that the values of $\vec{k}_{\rm f}$ are discrete and one should sum over them instead of integrating. However, we are already used to states labeled by a continuous variable, so there is no real need to do this.

The quantity $d\sigma/d\Omega_{\rm f}$ is called a differential cross section because it is differential in the angles (θ, ϕ) of the final state electron.

You probably thought that $\left|\vec{k}_{\rm f}\right\rangle$ was a plane wave state. It is not, since it is supposed to be an eigenstate of H_0, which includes the Coulomb potential. Thus it should be an exact eigenstate of H_0 that looks like a plane wave with momentum $\vec{k}_{\rm f}$ far away from the origin. Thus, until we solve for the wave function $\left\langle\vec{x}\middle|\vec{k}_{\rm f}\right\rangle$ and perform all of the overlap integrals, we have not really solved our problem.

That's hard, so let's simplify some more. Let's just approximate $\left|\vec{k}_{\rm f}\right\rangle$ by a plane wave state. That should be a pretty good approximation if ω is much greater than the binding energy of the hydrogen atom. Taking $\left|\vec{k}_{\rm f}\right\rangle$ to be a plane wave state, our result is pretty simple. First of all, we can let the momentum operator $\vec{\varepsilon}\cdot\vec{p}$ operate on $\left\langle\vec{k}_{\rm f}\right|$, giving the eigenvalue $\vec{\varepsilon}\cdot\vec{k}_{\rm f}$. Now there is an intervening operator $\exp(i\vec{k}\cdot\vec{x})$, but the commutator of this operator with $\vec{\varepsilon}\cdot\vec{p}$ vanishes because $\vec{\varepsilon}\cdot\vec{k}=0$. Thus

$$\frac{d\sigma}{d\Omega_{\rm f}} \sim \frac{4\pi^2 e^2 k_{\rm f}}{\omega m}\,(\vec{\varepsilon}\cdot\vec{k}_{\rm f})^2 |\left\langle\vec{k}_{\rm f}\right|\exp(i\vec{k}\cdot\vec{x})\left|1\right\rangle|^2 \quad . \tag{101}$$

Next, let the operator $\exp(i\vec{k}\cdot\vec{x})$ operate on $\left\langle\vec{k}_{\rm f}\right|$. We have

$$\left\langle\vec{k}_{\rm f}\right|\exp(i\vec{k}\cdot\vec{x}) = \left\langle\vec{k}_{\rm f}-\vec{k}\right| \quad . \tag{102}$$

This is because, just as the operator $\vec{p}$ generates translations in position space, so also the operator $\vec{x}$ generates translations in momentum space. To see this in more detail, just take the inner product of the proposed equation above with any state $\left|\psi\right\rangle$ and write out the equation in terms of integrations of the wave function $\left\langle\vec{x}\middle|\psi\right\rangle$ over position.

The result is

$$\frac{d\sigma}{d\Omega_{\rm f}} \sim \frac{4\pi^2 e^2 k_{\rm f}}{\omega m}\,(\vec{\varepsilon}\cdot\vec{k}_{\rm f})^2 |\left\langle\vec{k}_{\rm f}-\vec{k}\middle|1\right\rangle|^2 \quad . \tag{103}$$

This is remarkably simple. The factor $\left\langle\vec{k}_{\rm f}-\vec{k}\middle|1\right\rangle$ is the hydrogen atom wave function in momentum space, which is simple. The interpretation is also simple: $|\left\langle\vec{k}_{\rm f}-\vec{k}\middle|1\right\rangle|^2$ is the probability that the electron in the hydrogen atom ground state has momentum

$$\vec{k}_{\rm i} = \vec{k}_{\rm f} - \vec{k} \quad . \tag{104}$$

The other factors give the probability that this electron gets momentum $\vec{k}_{\rm f}$ after absorbing the photon. The result is simple because we have neglected the effect of the Coulomb force on the outgoing electron.

We have approximated the state $\big|\vec{k}_{\rm f}\big\rangle$ by a plane wave state. This should be a pretty good approximation if ω is large, much larger than an atom binding energy. If ω is not so large, just large enough to ionize the atom, then the photon wavelength is much larger than the atom size and we would be better off to use the dipole approximation that we have discussed earlier. In this case, $\big|\vec{k}_{\rm f}\big\rangle$ is an exact Coulomb wave function. Other parts of the analysis remain the same.

12 Decaying states

Let us suppose that there is a discrete state $\big|1\big\rangle$ with unperturbed energy ω_1 and that there are other states $\big|n\big\rangle$, for $n \neq 1$, with unperturbed energies ω_n. A perturbing hamiltonian V links these states at first order: $\big\langle n\big|V\big|1\big\rangle \neq 0$. We consider the case of a V that is time independent in the Schrödinger picture except that we turn it off far in the past with a function $g(t)$ that is 0 far in the past and 1 for a wide range of times that includes $t = 0$. (Also, $g(t)$ is zero far in the future, but in this section, we are going to be looking at what is happening for times not far in the future, so this will not be relevant.) Thus

$$V_{\rm I}(t) = e^{iH_0 t} V e^{-iH_0 t} g(t) \ . \tag{105}$$

We will be particularly interested in the case that the states $\big|n\big\rangle$ for $n \neq 1$ include states with a continuous range of energies that includes $\omega_n = \omega_1$. In that case, we should really have integrals over some continuous quantum numbers rather than a sum over n, but we will retain the notation of a discrete index that we sum over just to keep the notation simple.

An important example of this is that $\big|1\big\rangle$ is an excited state of an atom and the states $\big|n\big\rangle$ are the atom in its ground state together with a photon. Then ω_n is the atom's ground state energy plus the energy of the photon, which can cover a continuous range.

We should note that we cannot really turn off the interaction that makes the excited state decay to the ground state plus a photon. We just pretend that we can in order to see what description we get. In a more sophisticated treatment, we would start with an atom in its ground state plus a photon

and let the interaction of these create the excited state of the atom, which can then decay. However, that is a more difficult derivation.

We will want to look at

$$C(t) = \big\langle 1 \big| U_{\rm I}(t, -\infty) \big| 1 \big\rangle \ . \tag{106}$$

This is the amplitude to find the system in state $\big|1\big\rangle$ after the interaction has been operating for a long time in the case that we started with state $\big|1\big\rangle$ far in the past. We will use second order perturbation theory to look at the rate at which $C(t)$ is changing,

$$\dot{C}(t) = -i \big\langle 1 \big| V_{\rm I}(t) U_{\rm I}(t, -\infty) \big| 1 \big\rangle \ . \tag{107}$$

Expanding U to first order, we get

$$\begin{aligned}
\dot{C}(t) = & -i \big\langle 1 \big| V_{\rm I}(t) \big| 1 \big\rangle \\
& - \int_{-\infty}^{t} d\tau \ \big\langle 1 \big| V_{\rm I}(t) V_{\rm I}(\tau) \big| 1 \big\rangle + \mathcal{O}(V^3) \\
= & -i \big\langle 1 \big| V_{\rm I}(t) \big| 1 \big\rangle - \int_{-\infty}^{t} d\tau \ \big\langle 1 \big| V_{\rm I}(t) \big| 1 \big\rangle \big\langle 1 \big| V_{\rm I}(\tau) \big| 1 \big\rangle \\
& - \int_{-\infty}^{t} d\tau \ \sum_{n \neq 1} \big\langle 1 \big| V_{\rm I}(t) \big| n \big\rangle \big\langle n \big| V_{\rm I}(\tau) \big| 1 \big\rangle + \mathcal{O}(V^3) \\
= & -ig(t) \big\langle 1 \big| V \big| 1 \big\rangle - g(t) \left| \big\langle 1 \big| V \big| 1 \big\rangle \right|^2 \int_{-\infty}^{t} d\tau \ g(\tau) \\
& - g(t) \int_{-\infty}^{t} d\tau \ g(\tau) \sum_{n \neq 1} \exp(i(\omega_1 - \omega_n)(t - \tau)) \ \left| \big\langle n \big| V \big| 1 \big\rangle \right|^2 \\
& + \mathcal{O}(V^3)
\end{aligned} \tag{108}$$

We are interested in times t that are not far in the past or in the future, so we set $g(t) = 1$. Thus

$$\begin{aligned}
\dot{C}(t) = & -i \big\langle 1 \big| V \big| 1 \big\rangle - \left| \big\langle 1 \big| V \big| 1 \big\rangle \right|^2 \int_{-\infty}^{t} d\tau \ g(\tau) \\
& - \int_{-\infty}^{t} d\tau \ g(\tau) \sum_{n \neq 1} \exp(i(\omega_1 - \omega_n)(t - \tau)) \ \left| \big\langle n \big| V \big| 1 \big\rangle \right|^2 \\
& + \mathcal{O}(V^3)
\end{aligned} \tag{109}$$

If we evaluate $C(t)$ (rather than $\dot{C}(t)$) in first order perturbation theory, we get

$$\begin{aligned} C(t) &= 1 - i\int_{-\infty}^{t} d\tau\ \big\langle 1\big|V_{\rm I}(\tau)\big|1\big\rangle + \mathcal{O}(V^2) \\ &= 1 - i\big\langle 1\big|V\big|1\big\rangle \int_{-\infty}^{t} d\tau\ g(\tau) + \mathcal{O}(V^2) \end{aligned} \tag{110}$$

Thus if we divide $\dot{C}(t)$ by $C(t)$ and expand to order V^2, there is a convenient cancellation and we are left with

$$\begin{aligned} \dot{C}(t)/C(t) = &-i\big\langle 1\big|V\big|1\big\rangle \\ &-\int_{-\infty}^{t} d\tau\ g(\tau) \sum_{n\neq 1} \exp(i(\omega_1-\omega_n)(t-\tau))\ \big|\big\langle n\big|V\big|1\big\rangle\big|^2 \\ &+\mathcal{O}(V^3) \end{aligned} \tag{111}$$

Now let's look at the integral

$$I = \int_{-\infty}^{t} d\tau\ g(\tau) \exp(i(\omega_1-\omega_n)(t-\tau)) \tag{112}$$

If we do this exactly, it depends on what our choice for $g(\tau)$ was. However, the only purpose of having $g(\tau)$ was to turn the interaction off gradually far in the past. That is, $g(\tau)$ is supposed to regulate the integral far in the past. We can accomplish the same thing by replacing the integral by

$$I = \int_{-\infty}^{t} d\tau\ \exp(i(\omega_1-\omega_n+i\epsilon)(t-\tau))\ , \tag{113}$$

where we understand that we want the limit $\epsilon \to 0$. We calculate

$$I = \frac{i}{\omega_1-\omega_n+i\epsilon}\ . \tag{114}$$

Thus

$$\begin{aligned} \dot{C}(t)/C(t) = &-i\big\langle 1\big|V\big|1\big\rangle \\ &-\sum_{n\neq 1} \frac{i}{\omega_1-\omega_n+i0}\ \big|\big\langle n\big|V\big|1\big\rangle\big|^2 \\ &+\mathcal{O}(V^3) \end{aligned} \tag{115}$$

where $\omega_1 - \omega_n + i0$ means $\omega_1 - \omega_n + i\epsilon$ with $\epsilon \to 0$.

What we have here is

$$\dot{C}(t) \sim (-i\Delta - \Gamma/2)C(t) \ , \tag{116}$$

where

$$-i\Delta - \Gamma/2 = -i\big\langle 1\big|V\big|1\big\rangle - \sum_{n\neq 1} \frac{i}{\omega_1 - \omega_n + i0} \ \big|\big\langle n\big|V\big|1\big\rangle\big|^2 \ . \tag{117}$$

My notation is that Δ and Γ are real. We can solve the differential equation to obtain

$$C(t) \sim e^{-i\Delta t}\, e^{-\Gamma t/2}\ C(0) \ . \tag{118}$$

The first factor represents a perturbative correction to the energy of state $\big|1\big\rangle$. The second factor gives exponential decay:

$$|C(t)|^2 \sim e^{-\Gamma t}\ |C(0)|^2 \ . \tag{119}$$

To see what Δ and Γ are, we use

$$\frac{1}{\omega_1 - \omega_n + i0} = \frac{1}{[\omega_1 - \omega_n]_{\rm P}} - i\pi\delta(\omega_1 - \omega_n) \ . \tag{120}$$

Here $1/[\omega_1 - \omega_n]_{\rm P}$ denotes the principle value of $1/[\omega_1 - \omega_n]$, which can be defined by

$$\begin{aligned} \int d\omega\ \frac{h(\omega)}{[\omega]_{\rm P}} &= \lim_{\epsilon\to 0} \int d\omega\ h(\omega)\, \frac{1}{2}\left\{ \frac{1}{\omega + i\epsilon} + \frac{1}{\omega - i\epsilon} \right\} \\ &= \lim_{\epsilon\to 0} \int d\omega\ h(\omega)\, \frac{\omega}{\omega^2 + \epsilon^2} \ . \end{aligned} \tag{121}$$

To prove Eq. (120), we take ϵ to be a positive number that will be arbitrarily small and write

$$\frac{1}{\omega + i\epsilon} = \frac{1}{2}\left\{ \frac{1}{\omega + i\epsilon} + \frac{1}{\omega - i\epsilon} \right\} + \frac{1}{2}\left\{ \frac{1}{\omega + i\epsilon} - \frac{1}{\omega - i\epsilon} \right\} \ . \tag{122}$$

The first term gives the principle value. For the second, we note that

$$I[h] \equiv \lim_{\epsilon\to 0} \int d\omega\ h(\omega)\, \frac{1}{2}\left\{ \frac{1}{\omega + i\epsilon} - \frac{1}{\omega - i\epsilon} \right\} \tag{123}$$

can get contributions only from ω near zero. Thus we can replace $h(\omega)$ by $h(0)$ giving

$$\begin{aligned} I[h] &= h(0) \lim_{\epsilon \to 0} \int d\omega \ \frac{1}{2} \left\{ \frac{1}{\omega + i\epsilon} - \frac{1}{\omega - i\epsilon} \right\} \\ &= \ - i\pi h(0) \\ &= \int d\omega \ h(\omega) \big[- i\pi \, \delta(\omega) \big] \ . \end{aligned} \tag{124}$$

This establishes Eq. (120).

Using Eq. (120), we have

$$\begin{aligned} \Delta &= \big\langle 1 \big| V \big| 1 \big\rangle + \sum_{n \neq 1} \frac{1}{[\omega_1 - \omega_n]_{\rm P}} \ \big| \big\langle n \big| V \big| 1 \big\rangle \big|^2 \ , \\ \Gamma &= \sum_{n \neq 1} 2\pi \delta(\omega_1 - \omega_n) \ \big| \big\langle n \big| V \big| 1 \big\rangle \big|^2 \ . \end{aligned} \tag{125}$$

Note that the expression for Δ is the result previously obtained for the energy shift using second order time independent perturbation theory, except that we now have the principle value prescription in the case that ω_n can equal ω_1. The expression for Γ is what we have derived previously for the rate at which states $\big| n \big\rangle$ get populated if we start with the system in the state $\big| 1 \big\rangle$ with probability 1. Our previous calculation was only in lowest order perturbation theory, so it did not account for the loss of probability from state $\big| 1 \big\rangle$.

13 Example of decaying states

Let's try an example. Let state $\big| 1 \big\rangle$ be the $2p$ state of a hydrogen atom with z-component of angular momentum m: $\big| 2, 1, m; 0 \big\rangle$. The $\{2, 1, m\}$ here is $\{n, l, m\}$, following our standard notation. The "0" denotes no photons. Let the states that we were denoting $\big\langle n \big|$ be the $1s$ state of the atom together with a photon: $\big\langle 1, 0, 0; \vec{k}, \lambda \big|$. Again, the atom state is denoted by $\{n, l, m\}$ with now $n = 1$ and $l = m = 0$. There is now one photon with momentum $\vec{k}$ and polarization labelled by λ, which takes two values. We are simply ignoring the spin of the electron in our hydrogen atom. For the photon, I use here the Schrödinger picture, in which the photon state evolves according to it's

energy ω and the field operator $\vec{A}(\vec{x})$ is time independent. This is a slightly different notation from that in section 10, in which we used the interaction picture for $\vec{A}(\vec{x},t)$.

The hamiltonian is H_0+V, where H_0 includes the hydrogen atom hamiltonian and the photon energy. Thus

$$H_0\big|2,1,m;0\big\rangle = -\frac{1}{4}\,E_0\big|2,1,m;0\big\rangle \quad , \tag{126}$$

where E_0 is an abbreviation for the binding energy of a hydrogen atom

$$E_0 = \frac{me^4}{2} \approx 13.6 \text{ eV} \quad , \tag{127}$$

and the $1/4$ is $1/n^2$ with $n=2$. Similarly,

$$H_0\big|1,0,0;\vec{k},\lambda\big\rangle = [-E_0+\omega(\vec{k})]\big|2,1,m;0\big\rangle \quad , \tag{128}$$

with $\omega(\vec{k}) = |\vec{k}\,|$. The perturbation is

$$V = \frac{e}{m}\,\vec{A}(\vec{x})\cdot\vec{p} \tag{129}$$

Here $\vec{x}$ is the position operator for the electron and $\vec{p}$ is the momentum operator for the electron. For the effect of the photon field in creating a photon, we use the secret result from quantum field theory,

$$\big\langle\vec{k},\lambda\big|\vec{A}(\vec{x})\big|0\big\rangle = \frac{1}{2\pi\sqrt{\omega}}\;\vec{\varepsilon}(\vec{k},\lambda)^*\,e^{-i\vec{k}\cdot\vec{x}} \quad . \tag{130}$$

Thus

$$\begin{aligned}\big\langle 1,0,0;\vec{k},\lambda\big|V\big|2,1,m;0\big\rangle = \\ \frac{e}{m}\,\frac{1}{2\pi\sqrt{\omega}}\;\vec{\varepsilon}(\vec{k},\lambda)^*\cdot\big\langle 1,0,0\big|e^{-i\vec{k}\cdot\vec{x}}\vec{p}\,\big|2,1,m\big\rangle \quad .\end{aligned} \tag{131}$$

In applying this result, one needs to square the amplitude, integrate over $\vec{k}$, and sum over the polarization choices λ. For the polarization sum, you need

$$\sum_{\lambda=1}^{2}(\vec{v}\cdot\varepsilon(\vec{k},\lambda))^2$$

for a certain vector $\vec{v}$. One way to handle that is to note that, letting $\hat{k}$ be a unit vector in the direction of $\vec{k}$,

$$\sum_{\lambda=1}^{2}(\vec{v}\cdot\varepsilon(\vec{k},\lambda))^2+(\vec{v}\cdot\hat{k})^2=\vec{v}^2 \quad , \tag{132}$$

since the vectors $\varepsilon(\vec{k},1)$, $\varepsilon(\vec{k},2)$ and $\hat{k}$ make a basis for the space of three dimensional vectors. Thus

$$\sum_{\lambda=1}^{2}(\vec{v}\cdot\varepsilon(\vec{k},\lambda))^2=\vec{v}^2-(\vec{v}\cdot\hat{k})^2=\vec{v}^2[1-\cos^2(\theta_{vk})] \quad . \tag{133}$$

Another approach is to choose the two polarization vectors so that $\varepsilon(\vec{k},1))$ is a unit vector in the plane of $\vec{v}$ and $\vec{k}$, orthogonal to $\vec{k}$, while $\varepsilon(\vec{k},2))$ is a unit vector orthogonal to both $\vec{v}$ and $\vec{k}$. Then only $\varepsilon(\vec{k},1)$ contributes. Then simple geometry gives

$$\sum_{\lambda=1}^{2}(\vec{v}\cdot\varepsilon(\vec{k},\lambda))^2=(\vec{v}\cdot\varepsilon(\vec{k},1))^2=\vec{v}^2\sin^2(\theta_{vk}) \quad . \tag{134}$$

Either way, one gets the same result.

Exercise 12.1 Calculate the total decay width Γ for the decay of the $2p$ state of hydrogen to the $1s$ state by photon emission. Use the dipole approximation from section 9. The answer should be independent of m, so you can just set $m=0$. Please find a numerical answer in units of eV.
