Inserting this relation in (10.53) and integrating over χ, we have finally

$$\sum_k P_k = \frac{2\pi t}{\hbar} g(E_{\rm out})\,|\langle {\rm out}|V|{\rm in}\rangle|^2. \tag{10.57}$$

This equation establishes **Fermi's golden rule**[6] of perturbation theory: *a perturbation $V{\rm e}^{{\rm i}\omega t}$ causes a system to transition to a new state lower in energy by $\hbar\omega$ at a rate equal to $2\pi/\hbar$ times the mod-square of the matrix element of V between the initial and final states times the density of relevant states at the final energy.* It is easy to see that if the time-dependence of the perturbation were ${\rm e}^{-{\rm i}\omega t}$, it would cause transitions at the same rate to states higher in energy by $\hbar\omega$.

10.3.2 Radiative transition rates

We now use equation (10.51) to calculate the rate at which electromagnetic waves induce an atom to make radiative transitions between discrete stationary states. Our treatment is valid when the quantum uncertainty in the electromagnetic field may be neglected, and the field treated as a classical object. This condition is satisfied, for example, in a laser, or at the focus of the antenna of a radio telescope.

Whereas in our derivation of Fermi's golden rule, we took the frequency ω of the perturbation to be fixed and assumed a continuum of final states, now that we are considering the case of a discrete final state, we argue that the electromagnetic field is a superposition of plane waves of various frequencies, and that we should sum the transition probability (10.51) that each wave independently contributes. Thus we write

$$\sum_{\rm waves} P_k(t) = 4 \sum_{\rm waves} |\langle E_k|\delta V_0|E_N\rangle|^2 \frac{\sin^2\left((\hbar\omega + E_k - E_N)t/2\hbar\right)}{(\hbar\omega + E_k - E_N)^2}, \tag{10.58}$$

where δV_0 and ω relate to an individual wave.

In vacuo the electric field of an electromagnetic wave is divergence free, being entirely generated by Faraday's law, $\nabla \times \mathbf{E} = -\partial\mathbf{B}/\partial t$. It follows that the whole electromagnetic field of the wave can be described by the vector potential $\mathbf{A}$ through the equations

$$\mathbf{B} = \nabla \times \mathbf{A} \quad \text{and} \quad \mathbf{E} = -\partial\mathbf{A}/\partial t. \tag{10.59}$$

We are considering a superposition of plane waves, which individually have a potential

$$\delta\mathbf{A}(\mathbf{x}, t) = \delta\mathbf{A}_0 \cos(\mathbf{k}\cdot\mathbf{x} - \omega t), \tag{10.60}$$

where $\delta\mathbf{A}_0$ is a constant vector and $\mathbf{k}$ is the wavevector. From equations (10.59) and (10.60) we have that the wave's contribution to the electric field is

$$\delta\mathbf{E}(\mathbf{x}, t) = -\omega\delta\mathbf{A}_0 \sin(\mathbf{k}\cdot\mathbf{x} - \omega t), \tag{10.61}$$

[6] The golden rule was actually first given by P.A.M. Dirac, *Proc. Roy. Soc. A*, **114**, 243 (1927).

so $\delta\mathbf{E}$ is parallel to $\delta\mathbf{A}_0$. From $\boldsymbol{\nabla}\cdot\delta\mathbf{E}=0$ it follows that

$$\mathbf{k}\cdot\delta\mathbf{A}_0=0, \tag{10.62}$$

so $\mathbf{k}\cdot\delta\mathbf{E}=0$ and the wave is transverse. As a consequence of the orthogonality of $\mathbf{A}_0$ and $\mathbf{k}$, $\delta\mathbf{A}_0\cdot\mathbf{p}$ commutes with $\mathbf{k}\cdot\mathbf{x}$ because a component of momentum always commutes with a perpendicular component of position.

In §10.1.3 we saw that an external electromagnetic field adds to an atom's Hamiltonian the perturbing term (10.27) for each electron. In the present case the perturbation is

$$\begin{aligned}\delta V(\mathbf{x},t)&=\frac{e}{m_\mathrm{e}}\cos(\mathbf{k}\cdot\mathbf{x}-\omega t)\delta\mathbf{A}_0\cdot\mathbf{p}\\&=\frac{e}{2m_\mathrm{e}}\,\delta\mathbf{A}_0\cdot\mathbf{p}\left(\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{x}-\omega t)}+\mathrm{e}^{-\mathrm{i}(\mathbf{k}\cdot\mathbf{x}-\omega t)}\right).\end{aligned} \tag{10.63}$$

We now make the approximation that the electromagnetic wavelength is much bigger than the characteristic size of the atom or molecule. This is a good approximation providing the atom or molecule moves between states that are separated in energy by much less than $\alpha m_\mathrm{e}c^2$ (Problem 10.21), as will be the case for waves with frequencies that are less than those of soft X-rays. In this case we will have $\mathbf{k}\cdot\mathbf{x}\ll 1$ for all locations $\mathbf{x}$ in the atom or molecule at which there is significant probability of finding an electron. When this condition is satisfied, it makes sense to expand the factors $\mathrm{e}^{\pm\mathrm{i}\mathbf{k}\cdot\mathbf{x}}$ in equation (10.63) as a power series and discard all but the constant term. We then have

$$\delta V(\mathbf{x},t)=\frac{e}{2m_\mathrm{e}}\,\delta\mathbf{A}_0\cdot\mathbf{p}\left(\mathrm{e}^{-\mathrm{i}\omega t}+\mathrm{e}^{\mathrm{i}\omega t}\right), \tag{10.64}$$

where we have retained the exponentials in time because large values of t cannot be excluded in the way that we can exclude large values of $\mathbf{x}$. Finally, we note that in the gross-structure Hamiltonian H_0, $\mathbf{p}$ occurs only in the term $p^2/2m_\mathrm{e}$, so $[\mathbf{x},H_0]=\mathrm{i}(\hbar/m_\mathrm{e})\mathbf{p}$. When we use this relation to eliminate $\mathbf{p}$ from equation (10.64), we have

$$\delta V(\mathbf{x},t)=-\mathrm{i}\frac{e\delta A_0}{2\hbar}[\mathbf{n}\cdot\mathbf{x},H_0](\mathrm{e}^{-\mathrm{i}\omega t}+\mathrm{e}^{\mathrm{i}\omega t}), \tag{10.65}$$

where $\mathbf{n}$ is the unit vector in the direction of $\delta\mathbf{A}_0$. Thus a plane electromagnetic wave gives rise to perturbations with both positive and negative frequencies. Above we derived the frequency condition $\omega=(E_N-E_k)/\hbar$ for transitions from $|E_N\rangle$ to $|E_k\rangle$, so the negative frequency perturbation is associated with excitation of the system ($E_k>E_N$), while the positive frequency perturbation is associated with radiative decays.

We identify the time-independent part of δV as the operator δV_0 that occurs in equation (10.58) and then have that the net transition probability is

$$\begin{aligned}\sum_{\text{waves}}P_k(t)&=\frac{e^2}{\hbar^2}\sum_{\text{waves}}(\delta A_0)^2|\langle E_k|[\mathbf{n}\cdot\mathbf{x},H_0]|E_N\rangle|^2\frac{\sin^2((\hbar\omega+E_k-E_N)t/2\hbar)}{(\hbar\omega+E_k-E_N)^2}\\&=\frac{e^2}{4\hbar^4}(E_k-E_N)^2\sum_{\text{waves}}|\langle E_k|\mathbf{n}\cdot\mathbf{x}|E_N\rangle|^2(\delta A_0)^2\frac{\sin^2(\chi t)}{\chi^2},\end{aligned} \tag{10.66a}$$

where

$$\chi \equiv (\hbar\omega + E_k - E_N)/2\hbar. \tag{10.66b}$$

Even though the expression

$$\rho = \frac{1}{2\mu_0}\{(E/c)^2 + B^2\} \tag{10.67}$$

for the energy density of an electromagnetic field is quadratic in the field amplitudes E and B, the volume-averaged energy density of a superposition of plane waves is just the sum of the energy densities of each individual wave. Moreover, the electric and magnetic energy densities of a plane wave are equal, so the energy density contributed by our plane wave is just twice its electric energy density, and from equations (10.61) and (10.67) we infer that one wave contributes the time-averaged energy density

$$\delta\rho = \frac{\omega^2(\delta A_0)^2}{2\mu_0 c^2} = \tfrac{1}{2}\omega^2\epsilon_0(\delta A_0)^2, \tag{10.68}$$

where the second equality uses $\mu_0 c^2 = 1/\epsilon_0$. Using this expression to eliminate δA_0 from equation (10.66a), we obtain

$$\sum_{\text{waves}} P_k(t) = \frac{e^2}{2\epsilon_0\hbar^4}(E_k - E_N)^2 \sum_{\text{waves}} |\langle E_k|\mathbf{n}\cdot\mathbf{x}|E_N\rangle|^2 \delta\rho \frac{\sin^2(\chi t)}{\omega^2\chi^2}. \tag{10.69}$$

Let $\rho(\omega)$ be the power contained in all waves that have frequencies less than ω. Then if the radiation is isotropic (no preferred polarisation) we have

$$\sum_{\text{waves}} \delta\rho = \int \mathrm{d}\omega \, \frac{\mathrm{d}\rho}{\mathrm{d}\omega} \int \frac{\mathrm{d}^2\Omega}{4\pi}, \tag{10.70}$$

where $\mathrm{d}^2\Omega$ is an element of the solid angle that embraces a sheaf of the waves' polarisation directions $\mathbf{n}$. When we use this expression to replace the sum on the right side of equation (10.69) by an integral, and we use equation (10.66b) to replace $\mathrm{d}\omega$ with $2\mathrm{d}\chi$, we obtain

$$\sum_{\text{waves}} P_k(t) = \frac{e^2}{\epsilon_0\hbar^4}(E_k - E_N)^2 I \int \mathrm{d}\chi \frac{\mathrm{d}\rho}{\mathrm{d}\omega} \frac{\sin^2(\chi t)}{\omega^2\chi^2}, \tag{10.71a}$$

where

$$I \equiv \frac{1}{4\pi} \int \mathrm{d}^2\Omega \, |\langle E_k|\mathbf{n}\cdot\mathbf{x}|E_N\rangle|^2. \tag{10.71b}$$

Now

$$\begin{aligned} |\langle E_k|\mathbf{n}\cdot\mathbf{x}|E_N\rangle|^2 &= \langle E_k|\mathbf{n}\cdot\mathbf{x}|E_N\rangle\langle E_N|\mathbf{n}\cdot\mathbf{x}|E_k\rangle \\ &= \sum_{ij} n_i n_j \langle E_k|x_i|E_N\rangle\langle E_N|x_j|E_k\rangle. \end{aligned} \tag{10.72}$$

When we integrate this expression over the sphere of unit vectors $\mathbf{n}$, the only terms in the sum over ij to contribute are those with $i = j$ because when $i \neq j$ the product $n_i n_j$ is equally often negative as positive.

Dropping the terms with $i \neq j$, we write $\mathbf{n}$ in terms of the usual polar coordinates θ, ϕ: $n_x = \sin\theta\cos\phi$, $n_y = \sin\theta\sin\phi$, $n_z = \cos\theta$. Then

$$\begin{aligned} I &= \frac{1}{4\pi}\int_0^{\pi} \mathrm{d}\theta\, \sin\theta \int_0^{2\pi} \mathrm{d}\phi \left\{ \sin^2\theta \left(\cos^2\phi |\langle E_k|x|E_N\rangle|^2 + \sin^2\phi |\langle E_k|y|E_N\rangle|^2 \right) \right. \\ &\qquad \left. + \cos^2\theta |\langle E_k|z|E_N\rangle|^2 \right\} \\ &= \tfrac{1}{3}\left(|\langle E_k|x|E_N\rangle|^2 + \langle E_k|y|E_N\rangle|^2 + |\langle E_k|z|E_N\rangle|^2 \right). \end{aligned} \tag{10.73}$$

After inserting this expression for I into equation (10.71a), we let t become large and exploit equation (10.56) to evaluate the integral. The result is

$$\sum_{\text{waves}} P_k(t) = t\frac{\pi e^2}{3\epsilon_0 \hbar^2} \sum_i |\langle E_k|x_i|E_N\rangle|^2 \frac{\mathrm{d}\rho}{\mathrm{d}\omega}\bigg|_{\omega=|E_N-E_k|/\hbar}. \tag{10.74}$$

The coefficient of t on the right of this equation gives the rate R at which transitions occur. When we express the cumulative energy density of the wave-field in terms of frequency ν rather than angular frequency ω and use equation (8.15) to eliminate ϵ_0 in favour of the Bohr radius, the rate becomes

$$R = \frac{2\pi}{3a_0 m_\mathrm{e}} \sum_i |\langle E_k|x_i|E_N\rangle|^2 \frac{\mathrm{d}\rho}{\mathrm{d}\nu}\bigg|_{\nu=|E_N-E_k|/h}. \tag{10.75}$$

When $E_k > E_N$, the negative-frequency term in equation (10.64) gives rise to excitations at an identical rate. Thus we have recovered from a dynamical argument Einstein's famous result that stimulated emission of photons occurs, and that the coefficient B that controls the rate of stimulated emission is equal to the absorption coefficient (Box 10.2). Einstein's prediction of stimulated emission led 38 years later to the demonstration of a maser (§5.2.1) and 44 years later to the construction of the first laser by Theodore Maiman.[7] In view of this history, it's a remarkable fact that a laser operates in the regime in which the electromagnetic field can be treated as a classical object, as we have done here. Emission of light by a humble candle, by contrast, is an inherently quantum-mechanical phenomenon because it occurs through spontaneous emission. Our treatment does not include spontaneous emission because we have neglected the quantum uncertainty in the electromagnetic field. This uncertainty endows the field with zero-point energy (§3.1), and spontaneous emission can be thought of as emission stimulated by the zero-point energy of the electromagnetic field.

Using the argument given in Box 10.2, Einstein was able to relate the coefficient A of spontaneous emission to B. Einstein's argument does not yield a numerical value for either A or B. Our quantum mechanical

[7] The word 'laser' is an acronym for "light amplification by stimulated emission". Curiously Maiman's paper (*Nature*, **187**, 493 (1960)) about his laser was rejected by the *Physical Review*.

Box 10.2: Einstein A and B coefficients

In 1916, when only the merest fragments of quantum physics were known, Einstein showed (*Verh. Deutsch. Phys. Ges.* **18**, 318) that systems must be capable of both spontaneous and stimulated emission of photons, and that the coefficient of stimulated emission must equal that for absorption of a photon. He obtained these results by requiring that in thermal equilibrium there are equal rates of absorption and emission of photons of a given frequency ν by an ensemble of systems. He considered a frequency ν for which $h\nu = \Delta E$, the energy difference between two states $|1\rangle$ and $|2\rangle$ of the systems. The rate of absorptions he assumed to be $N_{\rm abs} = B_{\rm a}N_1(\mathrm{d}\rho/\mathrm{d}\nu)$, where $B_{\rm a}$ is the absorption coefficient, N_1 is the number of systems in the state $|1\rangle$, and $\mathrm{d}\rho/\mathrm{d}\nu$ is the energy density in radiation of frequency ν. The rate of emissions he assumed to be $N_{\rm em} = B_{\rm e}N_2(\mathrm{d}\rho/\mathrm{d}\nu) + AN_2$, where $B_{\rm e}$ is the coefficient for stimulated emission and A is that for spontaneous emission. Equating $N_{\rm abs}$ to $N_{\rm em}$ yields

$$0 = (B_{\rm e}N_2 - B_{\rm a}N_1)\frac{\mathrm{d}\rho}{\mathrm{d}\nu} + AN_2.$$

In thermal equilibrium $N_1 = N_2\mathrm{e}^{h\nu/kT}$ and $\mathrm{d}\rho/\mathrm{d}\nu$ is given by the Planck function. Using these relations to eliminate N_1 and $\mathrm{d}\rho/\mathrm{d}\nu$ and then cancelling N_2, we find

$$0 = (B_{\rm e} - B_{\rm a}\mathrm{e}^{h\nu/kT})\frac{8\pi h\nu^3}{c^3(\mathrm{e}^{h\nu/kT} - 1)} - A.$$

In the limit of very large T, $\mathrm{e}^{h\nu/kT} \to 1$, so the factor multiplying the bracket with the Bs becomes large, and the contents of this bracket tends to $B_{\rm e} - B_{\rm a}$. It follows that these coefficients must be equal. We therefore drop the subscripts on them, take B out of the bracket, cancel the factors with exponentials, and finally deduce that

$$A = 8\pi h(\nu/c)^3 B. \qquad (1)$$

treatment has yielded a value for B, and with Einstein's relation (eq. 1 in Box 10.2) between B and A we can infer the value of A:

$$A = \frac{16\pi^2 h\nu^3}{3c^3 a_0 m_{\rm e}} \sum_i |\langle E_k|x_i|E_N\rangle|^2. \qquad (10.76)$$

From this we can estimate the typical lifetime for radiative decay from an excited state of an atom.

When the radiation density ρ is very small, the number N_2 of atoms in an excited state obeys $\dot{N}_2 = -AN_2$ (Box 10.2), so N_2 decays exponentially with a characteristic time A^{-1}. Unless some symmetry condition causes the matrix element in equation (10.76) to vanish, we expect the sum

of squared matrix elements to be $\sim a_0^2$. So the characteristic radiative lifetime of a state is

$$\tau = A^{-1} = \frac{m_e c^2}{h\nu} \frac{\lambda}{a_0} \frac{3}{16\pi^2 \nu}. \tag{10.77}$$

For an optical transition, $h\nu \sim 2\,\mathrm{eV}$, $\lambda \sim 650\,\mathrm{nm} \sim 1.2 \times 10^4 a_0$, and $\nu \sim 4.6 \times 10^{14}\,\mathrm{Hz}$, so $\tau \sim 10^{-7}\,\mathrm{s}$. It follows that $\sim 4 \times 10^7$ oscillations of the atom occur before the radiation of energy causes the atom to slump into the lower state.

10.3.3 Selection rules

Equation (10.75) states that the rate of radiative transitions is proportional to the mod-square of the electric dipole operator ez. For this reason the approximation we made, that $\mathbf{k} \cdot \mathbf{x} \ll 1$, is called the **electric dipole approximation**.

There are important circumstances in which symmetry causes the matrix element of the dipole operator to vanish between the initial and final states. Transitions between such states are said to be **forbidden** in contrast to **allowed transitions**, for which the matrix element does not vanish. Some approximations were involved in our derivation of equation (10.75), so the transition rate does not necessarily vanish completely when the matrix element is zero. In fact, forbidden transitions often *do* occur, but at rates that are much smaller than the characteristic rate of allowed transitions (eq. 10.77) because the rate of a forbidden transition is proportional to terms that we could neglect in the derivation of equation (10.75). We now investigate relations between the initial and final states that must be satisfied if the states are to be connected by an allowed transition. Such relations are called **selection rules**. The slower rate of forbidden transitions must be determined by either including the next term of the Taylor expansion of $\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}$, or taking into account the perturbation $\mu_B \mathbf{S} \cdot \mathbf{B}$ that arises from the interaction of the intrinsic magnetic moment of an electron with the wave's magnetic field.

We are interested in matrix elements between states that are eigenstates of operators that commute with the Hamiltonian H that the atom would have if it were decoupled from electromagnetic waves. The Hamiltonian should include spin–orbit coupling as well as interaction with whatever steady external electric or magnetic fields are being applied. The operator in the matrix element is the component of the position operator parallel to the electric field of the radiation that is being either absorbed or emitted.

Even in the presence of an external field, the angular momentum parallel to the field, which we may call J_z, commutes with H, so the kets of interest are labelled with m. Since $[J_z, z] = 0$, the ket $z|E, m\rangle$ is an eigenket of J_z with eigenvalue m. It follows that $\langle E, m|z|E', m'\rangle = 0$ unless $m = m'$. This gives us the first selection rule listed in Table 10.1, namely that when the electric vector of the radiation is parallel to the imposed field, the quantum number m is unchanged by radiation.