

Chapter 6

Methods of Approximation

So far we have solved the Schrödinger equation for rather simple systems like the harmonic oscillator and the Coulomb potential. There are, however, many situations where exact solutions are not available. In the present section we will discuss a number of approximation techniques: Time independent perturbation theory (Rayleigh–Schrödinger), the variational method (Riesz) and time dependent perturbation theory. As applications we work out the fine structure of hydrogen, discuss the Zeeman and the Stark effect, compute the ground state energy of helium and derive the selection rules for the dipol approximation for absorption and emission of light.

6.1 Rayleigh–Schrödinger perturbation theory

Time independent (Rayleigh–Schrödinger) perturbation theory applies to quantum mechanical systems for which a time-independent Hamiltonian H is the sum of an exactly solvable part H_0 and a small perturbation,

$$H = H_0 + \lambda V, \quad (6.1)$$

where λ serves as a formal infinitesimal parameter that organizes the orders of the perturbative expansion. Of course only the smallness, in some appropriate sense, of the product λV will be relevant for the quality of the resulting approximation.

Non-degenerate time-independent perturbation theory. We first assume that the eigenvalue problem for H_0 has been solved

$$H_0|a0\rangle = E_{a0}|a0\rangle \quad (6.2)$$

with discrete¹ non-degenerate eigenvalues E_{a0} for some index set $\{a\}$ and corresponding eigenstates $|a0\rangle$. The label 0 refers to the unperturbed Hamiltonian H_0 and to the 0th approximation.

¹ If appropriate, sums have to be augmented by integrals over continuous parts of the spectra as usual.

We now make the ansatz that the exact solution to the eigenvalue problem

$$H|a\rangle = E_a|a\rangle \quad (6.3)$$

can be expanded into a power series

$$E_a = E_{a0} + \lambda E_{a1} + \lambda^2 E_{a2} + \dots = \sum_{i=0}^{\infty} \lambda^i E_{ai}, \quad (6.4)$$

$$|a\rangle = |a0\rangle + \lambda|a1\rangle + \lambda^2|a2\rangle + \dots = \sum_{i=0}^{\infty} \lambda^i |ai\rangle. \quad (6.5)$$

Even if E_a and $|a\rangle$ are not analytic in λ so that the expansion does not converge to the exact solution we may get excellent approximations for sufficiently small λ . In that case (6.4)–(6.5) is called an asymptotic expansion. A well-known example is the anomalous magnetic moment of the electron (5.6), for which the perturbation series is known to diverge.

The next step is to insert the ansatz into the stationary Schrödinger equation

$$(H_0 + \lambda V) \left(\sum_{i=0}^{\infty} \lambda^i |ai\rangle \right) = \left(\sum_{j=0}^{\infty} \lambda^j E_{aj} \right) \left(\sum_{k=0}^{\infty} \lambda^k |ak\rangle \right), \quad (6.6)$$

or

$$\sum_{i=0}^{\infty} \lambda^i H_0 |ai\rangle + \sum_{i=0}^{\infty} \lambda^{i+1} V |ai\rangle = \sum_{i=0}^{\infty} \lambda^i \left(\sum_{l=0}^i E_{a,i-l} |al\rangle \right), \quad (6.7)$$

and to compare coefficients of λ^i . For λ^0 we obtain, of course, the unperturbed equation $H_0|a0\rangle = E_{a0}|a0\rangle$. For the higher order terms we bring the $|ai\rangle$ -terms to the left-hand-side and collect all other terms on the other side,

- for λ^1 :
$$(H_0 - E_{a0})|a1\rangle = (E_{a1} - V)|a0\rangle, \quad (6.8)$$

- for λ^2 :
$$(H_0 - E_{a0})|a2\rangle = (E_{a1} - V)|a1\rangle + E_{a2}|a0\rangle, \quad (6.9)$$

...

- for λ^i :
$$(H_0 - E_{a0})|ai\rangle = (E_{a1} - V)|a, i-1\rangle + \dots + E_{ai}|a0\rangle. \quad (6.10)$$

The best way to analyze the content of these equations is to compute their products with the complete set $\langle b0|$ of eigenstates of H_0 .

First order corrections. When evaluating $\langle b0|$ on (6.8) we can replace $\langle b0|H_0$ in the first term by $\langle b0|E_{b0}$ and thus obtain

$$\langle b0|(E_{b0} - E_{a0})|a1\rangle = \langle b0|(E_{a1} - V)|a0\rangle = \delta_{ab}E_{a1} - \langle b0|V|a0\rangle. \quad (6.11)$$

We thus have to distinguish two cases:

- for $a = b$ the equation can be solved for

$$E_{a1} = \langle a0|V|a0\rangle, \quad (6.12)$$

so that the first order energy correction is simply the expectation value of the perturbation.

- for $a \neq b$ the equation can be solved for

$$\langle b0|a1\rangle = -\frac{\langle b0|V|a0\rangle}{E_{b0} - E_{a0}}. \quad (6.13)$$

These scalar products are just the expansion coefficients of $|a1\rangle$ in the basis $|b0\rangle$ so that

$$|a1\rangle = \sum_{b \neq a} |b0\rangle \frac{\langle b0|V|a0\rangle}{E_{a0} - E_{b0}}, \quad (6.14)$$

where we omitted a potential contribution of $|a0\rangle$ for reasons that we now discuss in detail.

If we consider the products of $\langle b0|$ with the equations (6.8)–(6.10) we observe that the l.h.s. vanishes for $b = a$ and the r.h.s. contains a term $E_{ai}\delta_{ab}$. The resulting equations hence determine the energy corrections E_{ai} for each order λ^i . For $b \neq a$ the assumption of non-degenerate energy levels $E_{b0} \neq E_{a0}$ implies that the l.h.s. becomes $(E_{b0} - E_{a0})\langle b0|ai\rangle$ so that these equations determine the expansion coefficients of $|ai\rangle$ in the basis $|b0\rangle$ for $b \neq a$. But $\langle a0|ai\rangle$ remains completely undetermined! The reason for this is easily understood: The Schrödinger equation is linear, and so are all equations that we derived with our perturbative ansatz. Every solution $|a\rangle = |a0\rangle + \lambda|a1\rangle + \dots$ can hence be rescaled by an overall factor $f(\lambda) = 1 + f_1\lambda + \dots$, which would reorganize the perturbation series such that the coefficient of $|a0\rangle$ in the expansion of $|ai\rangle$ can be changed arbitrarily without impairing the orthonormalization $\langle a0|b0\rangle = \delta_{ab}$ of the unperturbed states. We are hence free to simply choose

$$\langle a0|ai\rangle = 0 \quad (6.15)$$

as is done, for example, in [Schwabl]. This is the most convenient way to fix the ambiguous coefficients. Then $|a1\rangle$ and all higher order corrections are orthogonal to $|a0\rangle$ so that (according to Pythagoras) the norm of $|a\rangle$ differs from the norm of $|a0\rangle$ by a positive correction term of order λ^2 . If we want to keep $|a\rangle$ normalized at each order in perturbation theory then we can divide $|a\rangle$ as obtained with the choice (6.15) by its norm, which will keep $\langle a0|a1\rangle = 0$ but change $\langle a0|ai\rangle$ for $i \geq 2$ such that $|a\rangle$ stays normalized. We are now ready to proceed with the

Second order energy correction. Multiplication of $\langle b0|$ with (6.9) yields

$$\langle b0|(E_{b0} - E_{a0})|a2\rangle = \langle b0|(E_{a1} - V)|a1\rangle + \delta_{ab}E_{a2}. \quad (6.16)$$

For $a = b$ we use $\langle a0|a1\rangle = 0$ and solve for $E_{a2} = \langle a0|V|a1\rangle$, thus obtaining the energy correction

$$E_{a2} = \sum_{b \neq a} \frac{|\langle a0|V|b0\rangle|^2}{E_{a0} - E_{b0}} \quad (6.17)$$

which shows that the contribution of other states to the energy correction is proportional to the squared matrix element of the perturbation, but suppressed by an energy denominator for states located at very different energy levels. It is straightforward to work out the second order corrections to the wave function² but since our interest is usually focused on energy spectra this would only be relevant for the computation of E_{a3} .

6.1.1 Degenerate time independent perturbation theory

In the above derivation we used that the energy levels are non-degenerate and the results (6.14) and (6.17) show that we get in trouble with vanishing energy denominators if $E_{a0} = E_{b0}$ for $a \neq b$. Indeed, equation (6.11) becomes inconsistent if there is an offdiagonal matrix element $\langle b0|V|a0\rangle$ between states with degenerate unperturbed energies. The reason is easily understood because the choice of basis is ambiguous within the eigenspace for a particular eigenvalue and when the degeneracy is lifted by λV then an arbitrarily small perturbation requires a non-infinitesimal change of basis. This is inconsistent with a perturbative ansatz. What we thus need to do is to diagonalize the matrix $\langle b0|V|a0\rangle$ within each degeneration space $E_{a0} = E_{b0}$ so that the perturbative ansatz becomes consistent. The vanishing r.h.s. of (6.11) thus implies

$$E_{\hat{a}1} = \langle \hat{a}0|V|\hat{a}0\rangle \quad \text{with} \quad \langle \hat{b}0|V|\hat{a}0\rangle = 0 \quad \text{for} \quad E_{\hat{a}0} = E_{\hat{b}0}, \quad \hat{a} \neq \hat{b} \quad (6.19)$$

so that the eigenvalues $E_{\hat{a}}$ of $\langle b0|V|a0\rangle$ yield the first order energy corrections.

6.2 The fine structure of the hydrogen atom

The fine-structure of the hydrogen atom, which partially lifts the degeneracies of the pure Coulomb interaction that we observed in chapter 4, is an important application of degenerate perturbation theory. The relevant Hamiltonian consists of the relativistic corrections

$$H = \underbrace{\frac{\vec{P}^2}{2m_e} + V(r)}_{H_0} - \underbrace{\frac{\vec{P}^4}{8m_e^3 c^2}}_{H_{RK}} + \underbrace{\frac{1}{2m_e^2 c^2 r} \frac{dV}{dr} \vec{L} \vec{S}}_{H_{SO}} + \underbrace{\frac{\hbar^2}{8m_e^2 c^2} \Delta V(r)}_{H_D} \quad (6.20)$$

² The solution $\langle b0|a2\rangle = (\langle b0|V|a1\rangle - E_{a1}\langle b0|a1\rangle)/(E_{a0} - E_{b0})$ of (6.16) leads to the second order wave function correction

$$|a2\rangle = \sum_{b \neq a} |b0\rangle \left(\sum_{c \neq a} \frac{\langle b0|V|c0\rangle \langle c0|V|a0\rangle}{E_{a0} - E_{b0} \quad E_{a0} - E_{c0}} - \frac{E_{a1} \langle b0|V|a0\rangle}{(E_{a0} - E_{b0})^2} \right) - \rho |a0\rangle, \quad (6.18)$$

where the normalization of $|a\rangle$ at order λ^2 requires $\rho = \frac{1}{2} \langle a1|a1\rangle$.

with $V(r) = -\frac{Ze^2}{r}$, where we dropped the unobservable zero-point energy mc^2 of the electron (as well as that of the nucleus).³ We now discuss the fine structure corrections

$$\begin{aligned} H_{RK} & \dots \text{relativistic kinetic energy correction} \\ H_{SO} & \dots \text{spin orbit coupling} \\ H_D & \dots \text{Darwin term} \end{aligned}$$

in turn. At the end we will observe a surprising simplification of the complete result.

Relativistic kinetic energy. Since the electron is much lighter than the nucleus, its velocity is much larger and the resulting relativistic corrections to the kinetic energy E_{RK} set the scale for the size of the fine structure corrections. According to the Bohr model the velocity of the ground state orbit is

$$\frac{v}{c} = \frac{e^2}{\hbar c} = \alpha \approx \frac{1}{137} \quad \Rightarrow \quad \frac{E_{RK}}{E_0} \sim \frac{v^2}{c^2} \sim \alpha^2 \sim 10^{-4} \quad (6.21)$$

so that first order perturbation theory is sufficient for the derivation of very precise values for the level splittings. For small velocities the relativistic expression

$$E = \sqrt{m_e^2 c^4 + P^2 c^2} = m_e c^2 \sqrt{1 + \left(\frac{P}{m_e c}\right)^2} \quad (6.22)$$

can be expanded into a power series using

$$\sqrt{1+x} = \sum_{n=0}^{\infty} \binom{1/2}{n} x^n = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \dots, \quad (6.23)$$

with the result

$$E = m_e c^2 + \frac{P^2}{2m_e} - \frac{1}{8} \frac{P^4}{m_e^3 c^2} + \dots \quad (6.24)$$

The second term is the non-relativistic kinetic energy, which is the kinetic part of the Schrödinger Hamiltonian H_0 . Thus the leading relativistic correction is

$$H_{RK} = -\frac{P^4}{8m_e^3 c^2}, \quad \Delta E_{nl}^{RK} = \langle nlm | H_{RK} | nlm \rangle \quad (6.25)$$

where ΔE denotes the first order energy correction (6.12), which depends on the principal quantum number n and the orbital quantum number l . Since the momentum commutes with the spin operator, we can ignore the spin quantum numbers. Since the states $|nlm\rangle$ satisfies the unperturbed Schrödinger equation $H_0 |nlm\rangle = E_n |nlm\rangle$ with energy

$$E_n = -\frac{Z^2 m_e e^4}{n^2 2\hbar^2} = -\frac{Z^2 e^2}{2n^2} \frac{1}{a_0} \quad (6.26)$$

we can use $P^2 = 2m_e(H_0 - V)$ with $V(r) = -\frac{Ze^2}{r}$ to trade expectation values of powers of P^2 for expectation values of powers of the Coulomb potential,

$$H_{RK} = -\frac{P^4}{8m_e^3 c^2} = -\frac{1}{2m_e c^2} \left(\frac{P^2}{2m_e}\right)^2 = -\frac{1}{2m_e c^2} \left(H_0 + \frac{Ze^2}{r}\right)^2. \quad (6.27)$$

³While the replacement of the electron mass m_e by the reduced mass μ with $\frac{1}{\mu} = \frac{1}{m_p} + \frac{1}{m_e}$ leads to a larger absolute shift in the total binding energies than the fine structure, it can be ignored for the relative level splittings.

Inserting the expectation values

$$\left\langle \frac{1}{r} \right\rangle_{nl} = \frac{Z}{a_0 n^2} = -\frac{2E_n}{Ze^2}, \quad a_0 = \frac{\hbar^2}{m_e e^2}, \quad (6.28)$$

$$\left\langle \frac{1}{r^2} \right\rangle_{nl} = \frac{Z^2}{a_0^2 n^3 (l + \frac{1}{2})} = \frac{4E_n^2}{Z^2 e^4} \frac{n}{l + \frac{1}{2}}, \quad (6.29)$$

where a_0 denotes the Bohr radius, the **relativistic kinetic energy** correction becomes

$$\Delta E_{nl}^{RK} = -\frac{m_e Z^4 e^8}{2n^4 c^2 \hbar^4} \left(\frac{n}{l + \frac{1}{2}} - \frac{3}{4} \right), \quad (6.30)$$

which is of order α^2 as expected.

Spin-orbit coupling. The Hamiltonian H_{SO} of the spin-orbit coupling

$$H_{SO} = \frac{1}{2m_e^2 c^2 r} \frac{dV}{dr} \vec{L} \vec{S} = \frac{1}{2m_e^2 c^2} \vec{L} \vec{S} \frac{Ze^2}{r^3} \quad (6.31)$$

has been discussed already in chapter 5 and will be derived from the Dirac equation in chapter 7. For the evaluation of the radial part of the expectation value we need

$$\left\langle \frac{1}{r^3} \right\rangle_{nl} = \frac{Z^3 m_e^3 e^6}{\hbar^6} \frac{1}{n^3 l(l+1)(l + \frac{1}{2})}. \quad (6.32)$$

For the evaluation of the angular part we recall that $J^2 = (L + S)^2 = L^2 + S^2 + 2\vec{L} \vec{S}$ so that the spin-orbit term can be written as

$$\vec{L} \vec{S} = \frac{1}{2} (J^2 - L^2 - S^2). \quad (6.33)$$

Since J^2 does not commute with L_z the spin-orbit Hamiltonian is not diagonal in the basis $|nlm_l\rangle \otimes |sm_s\rangle$ and we have to apply degenerate perturbation theory in the total angular momentum basis $|n j l s m_j\rangle$. The energy correction thus becomes

$$\langle n j l s m_j | H_{SO} | n j l s m_j \rangle = \frac{Ze^2}{4m_e^2 c^2} \hbar^2 (j(j+1) - l(l+1) - s(s+1)) \langle nl | \frac{1}{r^3} | nl \rangle \quad (6.34)$$

For an electron $s = \frac{1}{2}$ and therefor $j = l \pm \frac{1}{2}$. The spin-orbit coupling hence lifts, for example, the degeneracy of the 6 electron states $|21m\rangle$ in the p orbital with principal quantum number $n = 2$ and induces a fine-structure difference between the energy levels of the 4 states in the spin $\frac{3}{2}$ quartet $2p_{3/2}$ and the 2 states in the doublet $2p_{1/2}$. We hence need to distinguish the two cases $j = l \pm \frac{1}{2}$ and obtain

$$\Delta E_{njl}^{SO} = \frac{Ze^2}{4m_e^2 c^2} \left\langle \frac{1}{r^3} \right\rangle_{nl} \begin{cases} \hbar^2 ((l + \frac{1}{2})(l + \frac{3}{2}) - l(l+1) - \frac{1}{2} \frac{3}{2}) \\ \hbar^2 ((l - \frac{1}{2})(l + \frac{1}{2}) - l(l+1) - \frac{1}{2} \frac{3}{2}) \end{cases} \quad (6.35)$$

$$= \frac{Ze^2}{4m_e^2 c^2} \left\langle \frac{1}{r^3} \right\rangle_{nl} \begin{cases} \hbar^2 l \\ -\hbar^2 (l+1) \end{cases} \quad (6.36)$$

Inserting (6.32) the final result becomes

$$\Delta E_{njl}^{SO} = \frac{m_e Z^4 e^8}{4c^2 \hbar^4} \frac{1}{n^3 l(l+1)(l+\frac{1}{2})} \cdot \begin{cases} l & \text{for } j = l + \frac{1}{2}, \quad l > 0, \\ (-l-1) & \text{for } j = l - \frac{1}{2}, \quad l > 0. \end{cases} \quad (6.37)$$

For $l = 0$ the result is $\Delta E_{nj0}^{SO} = 0$ because the matrix element of $J^2 - L^2 - S^2$ vanishes.

The Darwin term. For the last contribution to the relativistic corrections we start with a heuristic argument that is based on the idea of a *Zitterbewegung*, i.e. an uncertainty in the position of an electron that is of the order of its Compton wave length

$$\lambda_e = \frac{\hbar}{m_e c} \approx 3.86 \times 10^{-13} m. \quad (6.38)$$

The effect of averaging over a slightly smeared out position of an electron in an electrostatic field would amount to an effective potential

$$V_{eff}(\vec{x}) = \left\langle V(\vec{x} + \delta\vec{x}) \right\rangle_{\delta\vec{x}} = V(\vec{x}) + \left\langle \delta x^i \right\rangle \partial_i V(\vec{x}) + \frac{1}{2} \left\langle \delta x^i \delta x^j \right\rangle \partial_i \partial_j V(\vec{x}) + \dots \quad (6.39)$$

Imposing rotational symmetry of the fluctuations we expect $\langle \delta x^i \rangle = 0$ and

$$\left\langle \delta x^i \delta x^j \right\rangle = \delta^{ij} \left\langle \delta x^1 \delta x^1 \right\rangle = \frac{1}{3} \delta^{ij} \left\langle \delta \vec{x} \delta \vec{x} \right\rangle = \frac{1}{3} \delta^{ij} (\delta |\vec{x}|)^2 = \frac{1}{3} \delta^{ij} (\delta r)^2. \quad (6.40)$$

If we set the expectation value of the fluctuation δr of the position equal to the Compton wavelength λ_e we obtain

$$V_{eff}(\vec{x}) = V(\vec{x}) + \frac{\hbar^2}{6m_e^2 c^2} \Delta V(\vec{x}) \quad (6.41)$$

where $\Delta = \delta^{ij} \partial_i \partial_j$ is the Laplace operator. This line of ideas leads to the correct functional form, but the correct prefactor ($\frac{1}{8}$ instead of $\frac{1}{6}$) of the Darwin term

$$H_D = \frac{\hbar^2}{8m_e^2 c^2} \Delta V(\vec{x}) = -\frac{\hbar^2 Z e^2}{8m_e^2 c^2} \Delta \frac{1}{r} \quad (6.42)$$

is obtained from the Dirac equation, as we will see in chapter 7). The Coulomb potential solves the Poisson equation with point-like source,

$$\Delta \frac{1}{r} = -4\pi \delta^3(\vec{x}). \quad (6.43)$$

Because of the δ -function only the s -waves contribute to the **Darwin term**

$$\Delta E_{nl}^D = \langle nlm | H_D | nlm \rangle = \frac{\pi \hbar^2 Z e^2}{2m_e^2 c^2} |\psi_{nlm}(0)|^2 = \frac{m_e Z^4 e^8}{2n^3 \hbar^4 c^2} \delta_{l,0}. \quad (6.44)$$

Note that the Darwin term exactly corresponds to the formal limit $l \rightarrow 0$ of the spin-orbit correction (6.37) for the case $j = l + \frac{1}{2}$

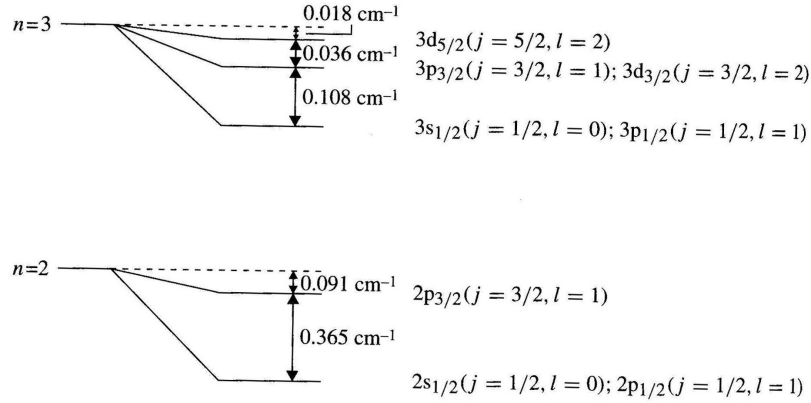


Figure 6.1: Fine structure splitting of the $n = 2$ and $n = 3$ levels of the hydrogen atom.

Fine structure. Putting everything together we obtain the complete fine structure energy correction

$$\Delta E_{nj}^{FS} = \frac{m_e Z^4 e^8}{2\hbar^4 n^4 c^2} \left(\frac{3}{4} - \frac{n}{j + \frac{1}{2}} \right). \quad (6.45)$$

As one can observe in figure 6.1 for the orbitals with principal quantum numbers $n = 2$ and $n = 3$, the energy shift is always negative because $j + \frac{1}{2} \leq n$. It is independent of the orbital quantum number l and only depends on n and the total angular momentum j , which leads to a degeneracy of the orbitals $2s_{1/2}$ and $2p_{1/2}$ that will be important in our discussion of the linear Stark effect.

6.3 External fields: Zeeman effect and Stark effect

We now analyze the level splittings that are due to external static electromagnetic fields. Such fields reduce the symmetry, at most, to rotations about some axis and hence can lead to further liftings of degeneracies that are otherwise protected by the 3-dimensional rotation symmetry of an isolated atom.

The Zeeman Effect. Taking into account the g -factor of the electron the Hamilton operator for the interaction with an external magnetic field is

$$H_Z = \vec{B}(\vec{m}_L + \vec{m}_S) = \frac{e}{2m_e c}(\vec{L} + 2\vec{S})\vec{B}. \quad (6.46)$$

For a constant magnetic field along the z -axis we have

$$\vec{B} = B_z \vec{e}_z \quad \Rightarrow \quad H_Z = \frac{e}{2m_e c}(L_z + 2S_z)B_z = \frac{e}{2m_e c}(J_z + S_z)B_z. \quad (6.47)$$

Taking the hydrogen atom without fine structure as the starting point, we note that the energy levels are degenerate for fixed n in the orbital quantum number $l < n$ and in the magnetic quantum numbers m_l, m_s . If we want to treat the spin-orbit coupling and the Zeeman Hamiltonian as perturbations, the problem is that $[H_{SO}, H_Z] \neq 0$, so that these operators cannot be diagonalized simultaneously (in the degeneration space of fixed n). We would therefore have to treat both interactions at once, thus diagonalizing much larger matrices. While this can be done (see [Schwabl] section 14.1.3), we rather consider the two limiting situations where one of the effects is dominant and the other is treated as a small perturbation on top of the larger one. Accordingly, there is a weak field and a strong field version of the Zeeman effect, where the latter is associated with the name Paschen-Back effect, as we will discuss below.

Weak field Zeeman effect. For weak external magnetic fields the spin-orbit coupling is dominant. We hence compute the matrix elements of H_Z between eigenstates $|njls m_j\rangle$ of H_{SO} and need to diagonalize within the degenerate subspace of fixed n, j, l . This will be a good approximation as long as the matrix elements of H_Z between states with different total angular momentum are small compared to the energy denominators (6.17) caused by the fine structure splitting due to H_{SO} so that second order perturbative corrections are small as compared to the leading order. This is the precise meaning of what we call a weak magnetic field.

Since we assume that H_{SO} is dominant, the first order energy correction is now computed for fixed J^2 , i.e. in the basis $|njls m_j\rangle$ and within the $2j + 1$ dimensional subspace $m_j = -j, \dots, j$. But in this subspace H_Z is already diagonal because $J_z + S_z$ commutes with J_z . Hence we only need to evaluate

$$\begin{aligned} \Delta E_{m_j}^Z &= \langle njls m_j | \frac{eB_z}{2m_e c} (J_z + S_z) | njls m_j \rangle \\ &= \frac{eB_z}{2m_e c} (\hbar m_j + \langle njls m_j | S_z | njls m_j \rangle). \end{aligned}$$

In order to evaluate S_z we expand $|njls m_j\rangle$ in the basis $|lsm_l m_s\rangle$,

$$|njls m_j\rangle = \sum_{m_s=\pm\frac{1}{2}} |lsm_l m_s\rangle \underbrace{\langle lsm_l m_s | jls m_j \rangle}_{\text{Clebsch-Gordan coeff. } C_{m_l m_s}^j}. \quad (6.48)$$

The Clebsch-Gordan coefficients $C_{m_l m_s}^j \equiv C_{m_l m_s}^{lsj}$ for $\vec{J} = \vec{L} + \vec{S}$ were computed in (5.69),

$$\begin{array}{c|cc} C_{m_l, m_s}^j & m_s = \frac{1}{2} & m_s = -\frac{1}{2} \\ \hline j = l + \frac{1}{2} & \sqrt{\frac{l+m_j+1/2}{2l+1}} & \sqrt{\frac{l-m_j+1/2}{2l+1}} \\ j = l - \frac{1}{2} & -\sqrt{\frac{l-m_j+1/2}{2l+1}} & \sqrt{\frac{l+m_j+1/2}{2l+1}} \end{array} \quad (6.49)$$

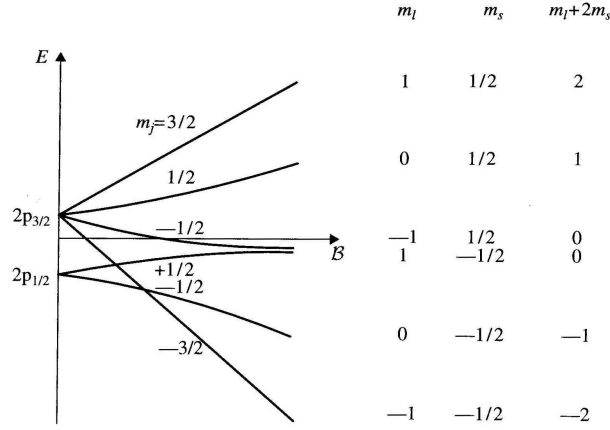


Figure 6.2: Schematic diagram for the splitting of the 2p levels of a hydrogen atom as a function of an external magnetic field B . For small B the degeneracy of the six 2p levels is completely removed, but as B becomes large the levels $j = \frac{3}{2}, m = -\frac{1}{2}$ and $j = \frac{1}{2}, m = \frac{1}{2}$ converge, so that the degeneracy is only partially removed in the limit of very large B (Paschen-Back effect).

with $m_j = m_l + m_s$. For the two cases $j = l \pm \frac{1}{2}$ the matrix elements of S_z thus become

$$\langle j l m_j | S_z | j l m_j \rangle = \left(C_{m_l, +\frac{1}{2}}^{l \pm \frac{1}{2}} \left\langle j l m_l, +\frac{1}{2} \right| + C_{m_l, -\frac{1}{2}}^{l \pm \frac{1}{2}} \left\langle j l m_l, -\frac{1}{2} \right| \right) \times \\ S_z \left(C_{m_l, +\frac{1}{2}}^{l \pm \frac{1}{2}} \left| j l m_l, +\frac{1}{2} \right\rangle + C_{m_l, -\frac{1}{2}}^{l \pm \frac{1}{2}} \left| j l m_l, -\frac{1}{2} \right\rangle \right) \quad (6.50)$$

$$= \frac{\hbar}{2} \left(\left| C_{m_l, +\frac{1}{2}}^{l \pm \frac{1}{2}} \right|^2 - \left| C_{m_l, -\frac{1}{2}}^{l \pm \frac{1}{2}} \right|^2 \right) = \pm \frac{\hbar}{2} \frac{2m_j}{2l+1}. \quad (6.51)$$

The energy shift induced by a weak external magnetic field is therefore

$$\Delta E_{j l m_j}^Z = \frac{e \vec{B}}{2m_e c} \left[\hbar m_j \pm \frac{\hbar m_j}{2l+1} \right] = \frac{e \hbar B}{2m_e c} m_j \cdot \begin{cases} \frac{2l+2}{2l+1} & j = l + 1/2 \\ \frac{2l}{2l+1} & j = l - 1/2 \end{cases}. \quad (6.52)$$

The spin-orbit coupling already removes the degeneracy in j . From equation (6.52) we see that a weak external magnetic field in addition lifts the degeneracy in m_j , thus explaining the name *magnetic quantum number*. A level with given quantum numbers n and j thus splits into $2j+1$ distinct lines. As an example consider the 2p orbitals of the hydrogen atom. The $2p_{3/2}$ level, with $j = l + 1/2$, splits into 4 levels according to $m_j = \frac{3}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{3}{2}$ with $\Delta E^Z = \frac{e \hbar B}{2m_e c} m_j \cdot \frac{4}{3}$. The $2p_{1/2}$ levels split into two with $m_j = \pm 1/2$ and $\Delta E^Z = \frac{e \hbar B}{2m_e c} m_j \cdot \frac{2}{3}$ (see figure 6.2).

Strong field and Paschen-Back effect. In the case of very strong magnetic fields the spin-orbit term becomes (almost) irrelevant and the Zeeman term H_Z forces the electrons into states that are (almost) eigenstates of $L_z + 2S_z = J_z + S_z$. The total angular momentum J^2 is hence no longer conserved, but L^2 and S^2 commute with $\hat{H}_Z$ so that we can use the original basis $|nlm_l\rangle \otimes |sm_s\rangle$ for the calculation. The fact that a strong magnetic field thus breaks up the coupling between spin and orbital angular momentum and makes them individually conserved quantities is called Paschen-Back effect.

The energy shift due to the external magnetic field B is now easily evaluated as

$$\Delta E_{nls m_l m_s}^Z = \frac{e\hbar B}{2m_e c} (m_l + 2m_s). \quad (6.53)$$

The magnetic field B does not remove the degeneracy of the hydrogenic energy levels in l and it removes the degeneracy in m_l and m_s only partially. Considering again the 2p level as shown in figure 6.2 we insert $m_l = -1, 0, 1$ and $m_s = \pm 1/2$ into equation (6.53). Due to the g -factor $g = 2$ the sum $m_l + 2m_s$ can now assume all 5 integral values between ± 2 , but the value $m_l + 2m_s = 0$ can be obtained in two different ways and hence corresponds to a degenerate energy level. In figure 6.2 one observes that the $m_j = -\frac{1}{2}$ line originating from $2p_{3/2}$ and the $m_j = \frac{1}{2}$ line originating from $2p_{1/2}$ converge for large B .

The Stark effect. A hydrogen atom in a uniform *electric field* $\vec{E} = E_z \vec{e}_z$ experiences a shift of the spectral lines that was first observed in 1913 by Stark. The interaction energy of the electron in the external field amounts to an external potential $V_S = -e\vec{E}\vec{x}$ and hence to an interaction Hamiltonian

$$H_S = -e\vec{E}\vec{X} = -eE_z z. \quad (6.54)$$

First of all we shall assume that E is large enough for the fine structure effects to be negligible. We hence work in the basis $|nlm\rangle$ and ignore the spin because the electric field does not couple to the magnetic moment of the electron. The matrix elements of H_S are strongly constrained by symmetry considerations. First we note that z is invariant under rotations about the z -axis so that J_z is conserved and $\langle l'm'|z|lm\rangle$ is proportional to $\delta_{m,m'}$. Moreover, under a parity transformation $\vec{X} \rightarrow -\vec{X}$ the interaction term H_S is odd, so that

$$\langle l'm'|z|lm\rangle \xrightarrow{\vec{x} \mapsto -\vec{x}} (-1)^{l-l'+1} \langle l'm'|z|lm\rangle. \quad (6.55)$$

Since the integral $\int d^3x |\psi(\vec{x})|^2 z$ is invariant under the change of variables $\vec{x} \mapsto -\vec{x}$ the matrix element can be non-zero only if $l - l'$ is odd. Moreover, one can show that $|l - l'| \leq 1$ for the *electric dipole* matrix element $\langle |\vec{X}| \rangle$, as we will learn in the context of tensor operators.⁴ Nonzero matrix elements therefore cannot be diagonal so that a linear Stark effect (i.e. a contribution in first order perturbation theory) can only occur if energy levels are degenerate for different orbital angular momenta. Such degeneracies only occur for excited states of the hydrogen atom (in the ground state one can only observe the *quadratic Stark effect*, i.e. a level splitting in second order perturbation theory).

The simplest situation for which we can hope for a linear Stark effect is for $l = 0, 1$ and $n = 2$, for which there may be a nonzero matrix element between the states $|200\rangle$ and $|210\rangle$,

⁴ A vector operator $\vec{X}$ corresponds to addition of spin 1. Adding spin one to a state of angular momentum l can only yield angular momentum l' with $|l' - l| \leq 1$ (see chapter 9). The same argument will apply to selection rules of in the dipol approximation for absorption and emission of electromagnetic radiation (see section 6.5).

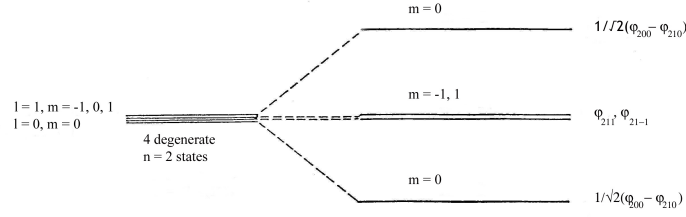


Figure 6.3: Splitting of the degenerate $n = 2$ levels of hydrogen due to the linear Stark effect.

which are degenerate and satisfy the selection rules $l - l' = 1$ and $m = m'$. Evaluation of the matrix element yields

$$\langle 210|z|200\rangle = \langle 200|z|210\rangle = -3eE_z a_0, \quad (6.56)$$

where a_0 is the Bohr radius. Since a matrix of the form $\begin{pmatrix} 0 & \lambda \\ \lambda & 0 \end{pmatrix}$ has eigenvalues $\pm\lambda$ the level shifts of the linear Stark effect in the hydrogen atom for $n = 2$, which affect the two states with magnetic quantum number $m = 0$, are

$$\Delta E_{n=2, m=0}^S = \pm 3eE_z a_0. \quad (6.57)$$

as shown in figure 6.3. Recall that the linear Stark effect can only occur if there are degenerate energy levels of different parity, which can only occur for hydrogen.

6.4 The variational method (Riesz)

The variational method is an approximation technique that is not restricted to small perturbations from solvable situations but rather requires some qualitative idea about how the ground state wave function looks like. It is based on the following fact:

Theorem: A wave function $|u\rangle$ is a solution to the stationary Schrödinger equation if and only if the energy functional

$$E(u) = \frac{\langle u|H|u\rangle}{\langle u|u\rangle} \quad (6.58)$$

is stationary, i.e.

$$H|u\rangle = E|u\rangle \quad \Leftrightarrow \quad \delta E = 0 \quad (6.59)$$

for arbitrary variations $u \rightarrow u + \delta u$, where we do not normalize u in order to have unconstrained variations.

For the *proof* of this theorem we compute the variation of the functional (6.58). Since variations are infinitesimal changes they obey the same rules as differentiation, including the formula $(f/g)' = f'/g - fg'/g^2$, i.e.

$$\delta E = \frac{\delta(\langle u|H|u\rangle)}{\langle u|u\rangle} - \frac{(\langle u|H|u\rangle)\delta(\langle u|u\rangle)}{(\langle u|u\rangle)^2} = \frac{\delta(\langle u|H|u\rangle)}{\langle u|u\rangle} - E \frac{\delta(\langle u|u\rangle)}{\langle u|u\rangle}. \quad (6.60)$$

Using the product rule $\delta\langle u|u\rangle = \langle\delta u|u\rangle + \langle u|\delta u\rangle$ stationarity of the energy implies

$$0 = ||u||^2 \cdot \delta E = \langle\delta u|H|u\rangle - E\langle\delta u|u\rangle + \langle u|H|\delta u\rangle - E\langle u|\delta u\rangle. \quad (6.61)$$

If the variations of u and u^* can be done independently then the first two terms (and the last two terms) on the r.h.s. have to cancel one another, i.e. $\langle\delta u|H|u\rangle - E\langle\delta u|u\rangle = 0$, for arbitrary variations $\langle\delta u|$ of $u^*(\vec{x})$, which is equivalent to the Schrödinger equation. To see that this is indeed the case we replace u by $v = iu$ in (6.61) so that $\delta v = i\delta u$ and $\delta v^* = -i\delta u^*$, implying

$$0 = -i(\langle\delta u|H|u\rangle - E\langle\delta u|u\rangle) + i(\langle u|H|\delta u\rangle - E\langle u|\delta u\rangle). \quad (6.62)$$

Adding i times (6.62) to (6.61) we find

$$\delta E = 0 \quad \Rightarrow \quad \langle\delta u|(H - E)|u\rangle = 0 \quad \forall \quad \langle\delta u|, \quad (6.63)$$

which implies the Schrödinger equation. This completes the proof since, in turn, $(H - E)|u\rangle = 0$ implies the vanishing of the variation (6.61).

If we expand $|u\rangle$ in a basis of states $|u\rangle = \sum_n c_n |e_n\rangle$ then our theorem tells us that the Schrödinger equation is equivalent to the equations $\frac{\partial E}{\partial c_n} = 0$. But for an infinite-dimensional Hilbert space we would have to solve infinitely many equations.

The variational method thus proceeds by introducing a family of trial wave functions

$$u(\alpha_1, \alpha_2, \dots, \alpha_n) \quad (6.64)$$

parametrized by a finite number of variables α_i and extremizes the energy functional within the subset of Hilbert space functions that are of the form (6.64) for some values of the parameters α_i , i.e. we solve

$$\frac{\partial E(u(\alpha_1, \dots))}{\partial \alpha_1} = \dots = \frac{\partial E(u(\alpha_n, \dots))}{\partial \alpha_n} = 0. \quad (6.65)$$

If the correct wave function $|u_0\rangle$ for the ground state happens to be contained in the family (6.64) of trial functions then the solution to the stationarity equations with the smallest value of $E(u)$ provides us with the exact solution to the Schrödinger equation. If we have, on the other hand, a badly chosen family that does not anywhere come close to $|u_0\rangle$ then our approximation to the ground state energy may be arbitrarily bad. Nevertheless, for an orthonormal energy eigenbasis $|e_n\rangle$

$$\langle u|H|u\rangle = \sum |c_n|^2 \langle e_n|H|e_n\rangle = \sum |c_n|^2 E_n \geq ||u||^2 E_{min} \quad \Rightarrow \quad E \geq E_{min} \quad (6.66)$$

so that we will always find an rigorous upper bound for the ground state energy.

6.4.1 Ground state energy of the helium atom

We now apply the variational method to improve a perturbative computation of the ground state energy of the helium atom, which is a system consisting of a nucleus with charge $Ze = 2e$ and two electrons. Treating the nucleus as infinitely heavy and neglecting relativistic effects like the spin-orbit interaction we consider the Hamiltonian

$$H = -\frac{\hbar^2}{2m_e}(\Delta_1 + \Delta_2) - \frac{2e^2}{r_1} - \frac{2e^2}{r_2} + \frac{e^2}{r_{12}}, \quad (6.67)$$

which consists of the kinetic energies $T_i = -\frac{\hbar^2}{2m_e}\Delta_i$, the Coulomb energies $V_i = -2e^2/|\vec{x}_i|$ due to the attraction by the nucleus and the mutual repulsion $V_{12} = e^2/|\vec{x}_1 - \vec{x}_2|$ of the electrons. If we omit the repulsive interaction among the electrons in a first step, the Hamiltonian becomes the sum of two commuting operators

$$H_0 = -\frac{\hbar^2}{2m_e}(\Delta_1 + \Delta_2) - \frac{2e^2}{r_1} - \frac{2e^2}{r_2} = (T_1 + V_1) + (T_2 + V_2) \quad (6.68)$$

for two independent particles and the Schrödinger equation $H_0|u\rangle = E_0|u\rangle$ is solved by product wave functions

$$u(\vec{x}_1, \vec{x}_2) = u_{n_1 l_1 m_1}(\vec{x}_1) u_{n_2 l_2 m_2}(\vec{x}_2). \quad (6.69)$$

The energy thus becomes the sum of the two respective energy eigenvalues,

$$E_0 = -Z^2 \mathcal{R} \left(\frac{1}{n_1^2} + \frac{1}{n_2^2} \right), \quad (6.70)$$

where $\mathcal{R} = \hbar^2/2m_e a_0^2 = m_e e^4/2\hbar^2 = 13.6 \text{ eV}$ is the Rydberg constant. For the approximate ground state wave function $u_0^{\text{pert}}(\vec{x}_1, \vec{x}_2) = u_{100}(\vec{x}_1) u_{100}(\vec{x}_2)$ this implies $E_0^{\text{pert}} \approx -108.8 \text{ eV}$ where the superscript refers to the Rayleigh–Schrödinger perturbation theory with

$$H = H_0 + V, \quad V = V_{12} = \frac{e^2}{r_{12}}. \quad (6.71)$$

For the wave function we have ignored so far the spin degree of freedom and the Pauli exclusion principle. In chapter 10 (many particle theory) we will learn that wave functions of identical spin 1/2 particles have to be anti-symmetrized under the simultaneous exchange of all of their quantum numbers (position and spin), which is the mathematical implementation of Pauli's exclusion principle. For the ground state of the helium atom the wave function is symmetric under the exchange $\vec{x}_1 \leftrightarrow \vec{x}_2$ and total antisymmetry implies antisymmetrization of the spin degrees of freedom so that

$$u_0(\vec{x}_1, \vec{x}_2, s_{1z}, s_{2z}) = u_{100}(\vec{x}_1) u_{100}(\vec{x}_2) |0, 0\rangle_{12}, \quad (6.72)$$

where $|0, 0\rangle_{12}$ is the singlet state in spin space. Since this will not influence any of our results we will, however, ignore the spin degrees of freedom for the rest of the calculation.

Taking into account now the repulsion term V between the electrons in (6.71) as a perturbation, the first order ground state energy correction becomes

$$E_1 = \langle u_0 | V_{12} | u_0 \rangle = e^2 \int d^3x_1 d^3x_2 \frac{|u_{100}(\vec{x}_1)|^2 |u_{100}(\vec{x}_2)|^2}{|\vec{x}_1 - \vec{x}_2|} \quad (6.73)$$

where

$$u_{100}(\vec{x}) = \frac{1}{\sqrt{\pi}} \left(\frac{Z}{a_0} \right)^{\frac{3}{2}} e^{-\frac{Z}{a_0} r} \quad (6.74)$$

is the wave function of a single electron in the Coulomb field of a nucleus with atomic number Z and a_0 is the Bohr radius.

The integrals for the energy correction (6.73) are best carried out in spherical coordinates,

$$E_1 = \left(\frac{Z^3}{a_0^3 \pi} \right)^2 \int_0^\infty dr_1 r_1^2 e^{-\frac{2Z}{a_0} r_1} \int_0^\infty dr_2 r_2^2 e^{-\frac{2Z}{a_0} r_2} \int d\Omega_1 d\Omega_2 \frac{e^2}{|\vec{x}_1 - \vec{x}_2|}. \quad (6.75)$$

If we first perform the angular integration $d\Omega_1$ it is useful to recall that a spherically symmetric charge distribution⁵ at radius r_1 creates a constant (force-free) potential $-q/r_1$ in the interior $r < r_1$ and a Coulomb potential $-q/r$ of a point charge located at the origin for $r > r_1$. Performing the Ω_1 -integration we thus obtain

$$E_1 = 4\pi \left(\frac{Z^3 e}{a_0^3 \pi} \right)^2 \int_0^\infty dr_2 r_2^2 \int_0^\infty dr_1 r_1^2 \int d\Omega_2 e^{-\frac{2Z}{a_0} r_1} e^{-\frac{2Z}{a_0} r_2} \cdot \begin{cases} \frac{1}{r_2} & r_2 > r_1 \\ \frac{1}{r_1} & r_2 < r_1 \end{cases} \quad (6.76)$$

$$= 2(4\pi)^2 \left(\frac{Z^3 e}{a_0^3 \pi} \right)^2 \int_0^\infty dr_1 r_1^2 \int_{r_1}^\infty dr_2 r_2 e^{-\frac{2Z}{a_0} r_1} e^{-\frac{2Z}{a_0} r_2}, \quad (6.77)$$

where the contribution of the domain $r_2 < r_1$ is accounted for by the prefactor 2 in the second line and the trivial Ω_2 -integration has also been done. With $\int r e^{-cr} = -\frac{1+cr}{c^2} e^{-cr}$ and $c = \frac{2Z}{a_0}$ we find

$$E_1 = 32 \left(\frac{Z}{a_0} \right)^6 e^2 \int_0^\infty dr_1 r_1^2 \left(\left(\frac{a_0}{2Z} \right)^2 + r_1 \frac{a_0}{2Z} \right) e^{-\frac{4Z}{a_0} r_1} \quad (6.78)$$

and with $\int_0^\infty dr r^{n-1} e^{-cr} = \Gamma(n)/c^n = (n-1)!/c^n$ the energy correction becomes

$$E_1 = 32 \frac{Z e^2}{a_0} \left(\frac{2!}{2^2 4^3} + \frac{3!}{2 \cdot 4^4} \right) = \frac{5}{8} \frac{Z e^2}{a_0} \approx 34.015 \text{ eV}. \quad (6.79)$$

With $E_0 = 8\mathcal{R}$ and $\mathcal{R} = 13.606 \text{ eV}$ our perturbative result for the ground state energy of the helium atom becomes

$$E_{He}^{(pert)} = E_0 + E_1 + \dots \approx -74.83 \text{ eV} \quad (6.80)$$

This is about 5% higher than the experimental value

$$E_{He}^{(exp)} \approx -79.015 \text{ eV}, \quad (6.81)$$

which is not too bad for our simple approach, in particular if we note that the first perturbative correction E_1 , with almost 1/3 of E_0 , is quite large.

⁵ Since all solutions to the homogeneous Laplace equation are superpositions of $r^l Y_{lm}$ and $r^{-l-1} Y_{lm}$ spherical symmetry implies $l = 0$ so that the potential is constant in the interior $r < r_{charge}$, as there is no singularity at the origin, and proportional to $1/r$ for $r > r_{charge}$, as the potential has to vanish for $r \rightarrow \infty$.

6.4.2 Applying the variational method and the virial theorem

In order to improve our perturbative result we note that the second electron partially screens the positive charge of the nucleus so that the electrons on average feel the attraction of an effective charge $q_{eff} < Ze$ and are less tightly bound. This suggests to use the ground state wave function of a hydrogen-like atom with the atomic number Z of the nucleus replaced by a continuous parameter b . Our starting point is thus the family $u(b)$ of normalized trial wave functions

$$u(\vec{x}_1, \vec{x}_2; b) = \frac{b^3}{\pi a_0^3} e^{-\frac{b}{a_0}(r_1+r_2)}, \quad (6.82)$$

where we recover the case (6.72) for $b = Z$ and expect to find $b < Z$ at the minimum of $E(b)$. The expectation value

$$\langle u(b) | V_{12} | u(b) \rangle = \frac{5}{8} \frac{be^2}{a_0} \quad (6.83)$$

directly follows from (6.79) by replacing Z by b . But in order to find the expectation value of $H_0 = T_1 + V_1 + T_2 + V_2$ we need to decompose the ground state energy $T + V$ into the kinetic contribution, for which we simply can replace Z by b , and the potential contribution, which is proportional to the charge Ze of the nucleus. The decomposition can be obtained as follows.

The virial theorem: If the potential $V(\vec{x})$ of a Hamiltonian of the form

$$H = T + V, \quad T = \frac{P^2}{2m} \quad (6.84)$$

is homogeneous of degree n , i.e.

$$V(\lambda \vec{x}) = \lambda^n V(\vec{x}) \quad \forall \lambda \in \mathbb{R}, \quad (6.85)$$

then the expectation values of T and V are related by⁶

$$2 \langle u | T | u \rangle = n \langle u | V | u \rangle \quad (6.89)$$

so that $\langle u | T | u \rangle = \frac{n}{n+2} E$ and $\langle u | V | u \rangle = \frac{2}{n+2} E$ for every bound state $|u\rangle$.

⁶ The proof of the quantum mechanical virial theorem is based on the Euler formula

$$\sum_i x^i \partial_i V(\vec{x}) = n V(\vec{x}) \quad (6.86)$$

for a homogeneous potential of degree n and on the fact that the expectation value of a commutator $[H, A]$ vanishes for bound states $|u_i\rangle$,

$$\langle u_i | [H, A] | u_i \rangle = \langle u_i | HA - AH | u_i \rangle = \langle u_i | E_i A - A E_i | u_i \rangle = 0. \quad (6.87)$$

The theorem then follows for $A = \vec{X} \vec{P}$ because

$$[\vec{X} \vec{P}, \frac{P^2}{2m}] = 2i\hbar \frac{P^2}{2m}, \quad [\vec{X} \vec{P}, V] = \frac{\hbar}{i} X^i \partial_i V \quad \Rightarrow \quad [\vec{X} \vec{P}, H] = i\hbar(2T - nV). \quad (6.88)$$

For further details see, for example, chapter 4 of [Gau].

Since the Coulomb potential is homogeneous of degree $n = -1$ we find

$$T_i = -\frac{1}{2}V_i = Z^2\mathcal{R} = \frac{Z^2e^2}{2a_0} \quad \mapsto \quad \langle u(b)|T_i|u(b)\rangle = \frac{b^2e^2}{2a_0}, \quad \langle u(b)|V_i|u(b)\rangle = -\frac{bZe^2}{a_0}. \quad (6.90)$$

The energy functional $E(b) = \langle u(b)|(T_1 + V_1 + T_2 + V_2 + V_{12})|u(b)\rangle$ thus becomes

$$E(b) = \frac{e^2}{a_0} \left(b^2 - 2bZ + \frac{5}{8}b \right) = \frac{e^2}{a_0} \left((b - Z + \frac{5}{16})^2 - (Z - \frac{5}{16})^2 \right). \quad (6.91)$$

The minimal value $E_{min} = -\frac{e^2}{a_0}(Z - \frac{5}{16})^2$ is obtained for $b = Z - \frac{5}{16}$. For the helium atom we thus obtain

$$E_{He}^{(var)} = -\frac{e^2}{a_0} \left(\frac{27}{16} \right)^2 \approx -77.5 \text{ eV}, \quad (6.92)$$

which is only 2% above the experimental value (6.81). The effective charge becomes $b \approx \frac{27}{16}$.

6.5 Time dependent perturbation theory

We now turn to non-stationary situations. In particular we will be interested in the response of a system to time dependent perturbations

$$H(t) = H_0 + W(t), \quad (6.93)$$

where the unperturbed Hamiltonian H_0 is not explicitly time dependent. For simplicity we assume that the unperturbed system has discrete and non-degenerate eigenstates

$$H_0|\varphi_n\rangle = E_n|\varphi_n\rangle. \quad (6.94)$$

If the perturbation is turned on at an initial time t_0 this implies

$$i\hbar \frac{\partial}{\partial t} |\varphi(t)\rangle = H_0 |\varphi(t)\rangle \quad \text{for} \quad t < t_0 \quad (6.95)$$

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = (H_0 + W(t)) |\psi(t)\rangle \quad \text{for} \quad t > t_0 \quad (6.96)$$

where the state $|\psi_i(t)\rangle$ is defined by the initial condition

$$|\psi_i(t = t_0)\rangle = |\varphi_i(t_0)\rangle \quad (6.97)$$

if the system is originally in the stationary state $|\varphi_i\rangle$. We first consider two limiting situations:

- In the **sudden approximation** we assume that a time independent perturbation is switched on very rapidly,

$$t_{switch} \ll t_{response} \quad \Rightarrow \quad W(t) \simeq \theta(t - t_0)W' \quad (6.98)$$

so that we can describe the time dependence by a step function $\theta(t - t_0)$. A physical example would be a radioactive decay, where the reorganization of the electron shell takes much longer than the nuclear reaction. Hence the Hamiltonian suddenly changes to a new time-independent form $H' = H_0 + W'$. For $t > t_0$ the system has a new set of stationary solutions $|\psi_f\rangle$, and since the wave function has no time to evolve under a time-dependent force the transition probability into a final state

$$\mathcal{P}_{i \rightarrow f} = |\langle \psi_f | \varphi_i \rangle|^2 \quad (6.99)$$

is determined by the overlap (scalar product) of the wave functions.

- The **adiabatic limit** is the other extremal situation,

$$t_{\text{switch}} \gg t_{\text{response}}, \quad (6.100)$$

for which the time variation of the external conditions is so slow that it cannot induce a transition and the system evolves by a continuous deformation of the energy eigenstate because we have, at each time, an almost stationary situation. More quantitatively, the transition probability will be negligible if the energy uncertainty that is due to the time variation of H is small in comparison to differences between energy levels.

In the rest of this section we will consider **small time-dependent perturbations** $W(t) = \lambda V(t)$, where a small parameter λ can be introduced to control the perturbative expansion, but it is equivalent to simply count powers of W . Since the perturbation is small we can, at each instant of time at which we perform a measurement, use eigenstates $|\varphi_f\rangle$ of H_0 to represent the possible outcomes of the reduction of the wave function. Our aim hence is to determine the probability

$$\mathcal{P}_{i \rightarrow f} = |\langle \varphi_f | \psi_i(t) \rangle|^2 \quad (6.101)$$

for finding the system in a final eigenstate $|\varphi_f\rangle$ after having evolved from $|\varphi_i\rangle$ under the influence of $H = H_0 + W$ according to (6.96) with boundary condition (6.97). For simplicity we set $t_0 = 0$. It is convenient to perform the perturbative computation of (6.101) in the interaction picture

$$|\psi(t)\rangle_I = e^{\frac{i}{\hbar} H_0 t} |\psi(t)\rangle = U_0^\dagger(t) |\psi(t)\rangle, \quad U_0(t) = e^{-\frac{i}{\hbar} H_0 t}, \quad (6.102)$$

so that

$$i\hbar \partial_t |\psi(t)\rangle_I = W_I(t) |\psi(t)\rangle_I \quad \text{with} \quad W_I(t) = e^{\frac{i}{\hbar} H_0 t} W(t) e^{-\frac{i}{\hbar} H_0 t} \quad (6.103)$$

as we found in (3.146–3.152).

Our next step is to transform the Schrödinger equation (6.103) of the interaction picture into an integral equation by integrating it over the interval from $t_0 = 0$ to t ,

$$|\psi_i(t)\rangle_I = |\varphi_i\rangle + \frac{1}{i\hbar} \int_0^t dt' W_I(t') |\psi_i(t')\rangle_I, \quad (6.104)$$

where we used the boundary condition (6.97). For small $W_I(t)$ we can solve this equation by iteration, i.e. we insert $|\psi_i\rangle_I = |\varphi_i\rangle + \mathcal{O}(W)$ on the r.h.s. and continue by inserting the resulting higher order corrections of $|\psi_i\rangle_I$. We thus obtain the **Neumann series**⁷

$$\begin{aligned} |\psi_i(t)\rangle = & \underbrace{|\varphi_i\rangle}_{\text{initial state}} + \underbrace{\frac{1}{i\hbar} \int_0^t dt' W_I(t') |\varphi_i\rangle}_{\text{first order correction}} + \\ & + \underbrace{\frac{1}{(i\hbar)^2} \int_0^t dt' \int_0^{t'} dt'' W_I(t') W_I(t'') |\varphi_i\rangle}_{\text{second order correction}} + \dots \end{aligned} \quad (6.107)$$

The transition amplitude now becomes

$$\mathcal{A}_{i \rightarrow f} = \langle \varphi_f | \psi_i(t) \rangle = \delta_{if} - \frac{i}{\hbar} \int_0^t dt' \langle \varphi_f | W_I(t') | \varphi_i \rangle + \mathcal{O}(W_I^2). \quad (6.108)$$

In the remainder of this section we focus on the leading contribution to the transition from an initial state $|\varphi_i\rangle$ to a final state $|\varphi_f\rangle$ with $f \neq i$,

$$\mathcal{A}_{i \rightarrow f}^{(1)} = -\frac{i}{\hbar} \int_0^t dt' \langle \varphi_f | e^{\frac{i}{\hbar} H_0 t'} W(t') e^{-\frac{i}{\hbar} H_0 t'} | \varphi_i \rangle \quad (6.109)$$

$$= -\frac{i}{\hbar} \int_0^t dt' e^{\frac{i}{\hbar} (E_f - E_i) t'} \langle \varphi_f | W(t') | \varphi_i \rangle \quad (6.110)$$

where we used $H_0 |\varphi_i\rangle = E_i |\varphi_i\rangle$ and $\langle \varphi_f | H_0 = \langle \varphi_f | E_f$ to evaluate the time evolution operators. For **first order transitions** we thus obtain the probability

$$\mathcal{P}_{i \rightarrow f}^{(1)} = \frac{1}{\hbar^2} \left| \int_0^t dt' e^{i\omega_{fi} t'} \langle \varphi_f | W(t') | \varphi_i \rangle \right|^2 \quad (6.111)$$

with the **Bohr angular frequency**

$$\omega_{fi} = \frac{E_f - E_i}{\hbar}. \quad (6.112)$$

⁷ Introducing the **time ordering** operator T by

$$TA(t_1)B(t_2) = \theta(t_1 - t_2)A(t_1)B(t_2) + \theta(t_2 - t_1)B(t_2)A(t_1) = \begin{cases} A(t_1)B(t_2) & \text{if } t_1 > t_2 \\ B(t_2)A(t_1) & \text{if } t_1 < t_2 \end{cases} \quad (6.105)$$

the Neumann series can be subsumed in terms of a formal expression for the time evolution operator

$$|\psi_i(t)\rangle_I = U_I(t) |\varphi_i\rangle, \quad U_I(t) = T e^{-\frac{i}{\hbar} \int_0^t dt' W_I(t')} \quad (6.106)$$

as is easily checked by expansion of the exponential.

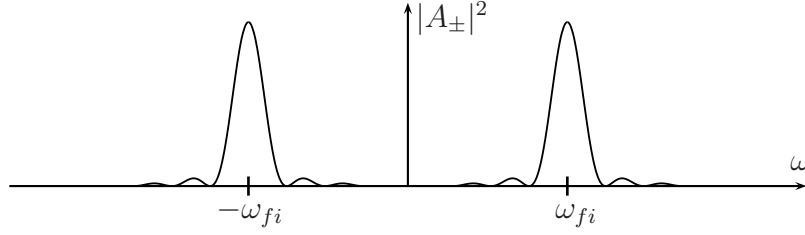


Figure 6.4: The functions $|A_{\pm}|^2 = \frac{\sin^2(\Delta\omega t/2)}{(\Delta\omega/2)^2} \rightarrow 2\pi t\delta(\Delta\omega)$ of height $t^2/2$ and width $4\pi/t$.

Note that the transition probability (6.111) is related to the Fourier transform at ω_{fi} of the matrix element $\langle\varphi_f|W(t')|\varphi_i\rangle$ restricted to $0 < t' < t$.

Periodic perturbations. In practice we will often be interested in the response to periodic external forces of the form

$$W(t) = \theta(t)(W_+e^{i\omega t} + W_-e^{-i\omega t}) \quad \text{with} \quad W_-^\dagger = W_+. \quad (6.113)$$

Then the time integration can be performed with the result

$$\mathcal{P}_{i \rightarrow f}^{(1)} = \frac{1}{\hbar^2} \left| A_+ \langle\varphi_f|W_+|\varphi_i\rangle + A_- \langle\varphi_f|W_-|\varphi_i\rangle \right|^2 \quad (6.114)$$

in terms of the integrals

$$A_{\pm} = \int_0^t dt' e^{i(\omega_{fi} \pm \omega)t'} = \frac{e^{i(\omega_{fi} \pm \omega)t} - 1}{i(\omega_{fi} \pm \omega)} = e^{\frac{i}{2}(\omega_{fi} \pm \omega)t} \frac{\sin((\omega_{fi} \pm \omega)t/2)}{(\omega_{fi} \pm \omega)/2}. \quad (6.115)$$

Figure 6.4 shows that the functions $A_{\pm}(\omega)$ are well-localized about $\omega = \mp\omega_{fi}$, respectively, and converge to δ -functions

$$|A_{\pm}|^2 = \left(\frac{\sin \Delta\omega t/2}{\Delta\omega/2} \right)^2 \rightarrow 2\pi t\delta(\Delta\omega) \quad \text{with} \quad \Delta\omega = \omega \pm \omega_{fi} \quad (6.116)$$

for late times $t \gg 1/\omega_{fi}$, where the prefactor follows from the integral

$$\int_{-\infty}^{\infty} d\xi \left(\frac{\sin(t\xi)}{\xi} \right)^2 = \pi t \quad \Rightarrow \quad \lim_{t \rightarrow \infty} \frac{1}{t} \left(\frac{\sin(t\xi)}{\xi} \right)^2 = \pi\delta(\xi). \quad (6.117)$$

For $t \rightarrow \infty$ the interference terms between A_+W_+ and A_-W_- in (6.114) can hence be neglected,

$$\mathcal{P}_{i \rightarrow f} \rightarrow \frac{2\pi t}{\hbar} \left(\delta(E_f - E_i - \hbar\omega) |\langle f|W_-|i\rangle|^2 + \delta(E_f - E_i + \hbar\omega) |\langle f|W_+|i\rangle|^2 \right). \quad (6.118)$$

For frequencies $\omega \approx \pm\omega_{fi}$ the transition probabilities become very large so that the contribution of A_-W_- is called *resonant term* (absorption of an energy quantum $\hbar\omega_{fi}$) while A_+W_+ is called *anti-resonant* (emission of an energy quantum $\hbar\omega_{fi}$). Since the probability becomes linear in t it is useful to introduce the **transition rate**

$$\Gamma_{i \rightarrow f} = \lim_{t \rightarrow \infty} \left(\frac{1}{t} \mathcal{P}_{i \rightarrow f} \right). \quad (6.119)$$

For discrete energy levels we thus obtain **Fermi's golden rule**

$$\Gamma_{i \rightarrow f} = \frac{2\pi}{\hbar} |\langle f | W_{\pm} | i \rangle|^2 \delta(E_f - E_i \pm \hbar\omega) \quad (6.120)$$

which was derived by Pauli in 1928, and called *golden rule* by Fermi in his 1950 book *Nuclear Physics*. The δ -function infinity for discrete energy levels is of course an unphysical artefact of our approximation. In reality spectral lines have finite width. For transitions to a continuum of energy levels we introduce the concept of a level density $\rho(E)$ by summing over transitions to a set F of final states,

$$\Gamma_{i \rightarrow F} = \sum_{f \in F} \Gamma_{i \rightarrow f} \rightarrow \int_{f \in F} df \Gamma_{i \rightarrow f} = \int_F dE \rho(E) \Gamma_{i \rightarrow f}. \quad (6.121)$$

Inserting this into the golden rule we can perform the energy integration and obtain the integrated rate

$$\Gamma_i(\omega) = \rho(E_f) \frac{2\pi}{\hbar} |\langle f | W_{\pm} | i \rangle|^2, \quad E_f = E_i \pm \hbar\omega. \quad (6.122)$$

If $f \in F$ is characterized by additional continuous quantum numbers β , like for example the solid angle covered by a detector, the level density can be generalized to $df(E, \beta) = \rho(E, \beta) dE d\beta$ and the integrated transition rate is obtained by integrating over the relevant range of β 's.

6.5.1 Absorption and emission of electromagnetic radiation

We now want to compute the rate for atomic transitions of electrons irradiated by an electromagnetic wave. The relevant Hamiltonian can be written as

$$H = \frac{1}{2m_e} \left(P - \frac{e}{c} \vec{A} \right)^2 + V(r) - \frac{e\hbar}{2m_e c} \vec{\sigma} \vec{B} \quad (6.123)$$

with

$\vec{A}$	...	vector potential of the electromagnetic radiation,
$V(\vec{r})$	...	central potential created by the nucleus,
$\frac{e\hbar}{2m_e c} \vec{\sigma} \vec{B}$	...	magnetic interaction with the radiation field.

We split the Hamiltonian as

$$H = H_0 + W(t) \quad \text{with} \quad H_0 = \frac{P^2}{2m_e} + V(r) \quad (6.124)$$

and interaction term

$$W(t) = - \underbrace{\frac{e}{2m_e c} (\vec{P} \vec{A} + \vec{A} \vec{P})}_{W_A} - \underbrace{\frac{e\hbar}{2m_e c} \vec{\sigma} \vec{B}}_{W_B} + \frac{e^2}{2m_e c^2} \vec{A}^2. \quad (6.125)$$

The last term can be ignored because it is quadratic in the perturbation $\vec{A}$. For the vector potential of the electromagnetic field we take a plane wave

$$\vec{A}(\vec{x}, t) = \vec{\varepsilon} A_0 \left(e^{-i(\omega t - \vec{k}\vec{x})} + e^{i(\omega t - \vec{k}\vec{x})} \right) \quad (6.126)$$

with frequency ω and polarization vector $\vec{\varepsilon}$. Since $\vec{B} = \text{curl } \vec{A} = i\vec{k} \times \vec{A}$ the magnetic interaction term W_B can be neglected for optical light, for which $\hbar k \ll |\langle f | \vec{P} | i \rangle| \sim \hbar/a_0$. In the radiation gauge

$$\text{div } \vec{A} = \frac{i}{\hbar} \vec{P} \vec{A} = 0 \quad \Leftrightarrow \quad \vec{\varepsilon} \vec{k} = 0, \quad (6.127)$$

hence $\vec{P} \vec{A} + \vec{A} \vec{P} = 2\vec{A} \vec{P}$ and the relevant matrix element becomes (up to a phase)

$$\langle f | W_A | i \rangle = -A_0 \frac{e}{m_e c} \vec{\varepsilon} \langle f | \vec{P} e^{\pm i\vec{k}\vec{x}} | i \rangle. \quad (6.128)$$

For optical transitions expectation values of $\vec{x}$ are of the order of $a_0 \ll 1/k$ so that we can drop contributions from the exponential

$$\langle f | \vec{P} e^{\pm i\vec{k}\vec{x}} | i \rangle \approx \langle f | \vec{P} | i \rangle. \quad (6.129)$$

This is called **electric dipole approximation** because the matrix element of $\vec{P}$ can be related to the matrix element of the dipole

$$\vec{D}_{fi} = \langle f | \vec{X} | i \rangle \quad (6.130)$$

using

$$[H_0, \vec{X}] = \frac{\hbar}{im_e} \vec{P} \quad \Rightarrow \quad \langle f | \vec{P} | i \rangle = i \frac{m_e}{\hbar} \langle f | [H_0, \vec{X}] | i \rangle = i \frac{m_e}{\hbar} (E_f - E_i) \langle f | \vec{X} | i \rangle. \quad (6.131)$$

Inserting everything into Fermi's golden rule we obtain the transition rate

$$\Gamma_{i \rightarrow f} = \frac{2\pi}{\hbar} \left| \frac{A_0 e}{m_e c} \frac{m_e}{\hbar} (E_f - E_i) \vec{\varepsilon} \langle f | \vec{X} | i \rangle \right|^2 \delta(E_f - E_i \pm \hbar\omega) \quad (6.132)$$

$$= \frac{2\pi e^2 \omega^2}{\hbar c^2} \delta(E_f - E_i \pm \hbar\omega) A_0^2 \left| \vec{\varepsilon} \langle f | \vec{X} | i \rangle \right|^2. \quad (6.133)$$

The **selection rules** for the dipole approximation are obtained by considering the matrix elements

$$\langle n' l' m' | \vec{X} | n l m \rangle. \quad (6.134)$$

Under a parity transformation $\vec{X} \rightarrow -\vec{X}$ the spherical harmonics transform as

$$Y_{lm}(\pi - \theta, \varphi + \pi) = (-1)^l Y_{lm}(\theta, \varphi). \quad (6.135)$$

The matrix element (6.134) hence transforms with a factor $(-1)^{l-l'+1}$ so that spherical symmetry implies that $l - l'$ must be odd. Since

$$[L_z, X_3] = 0, \quad [L_z, X_1 \pm iX_2] = \pm \hbar (X_1 \pm iX_2) \quad (6.136)$$

the components X_3 and $X_{\pm} = X_1 \pm iX_2$ of $\vec{X}$ change the magnetic quantum number by $m' - m \in \{0, \pm 1\}$. More generally, we will learn in chapter 9 that the vector operator $\vec{X}$ corresponds to addition of angular momentum 1, so that $|l' - l| \leq 1$. Combining all constraints we find

$$l' - l = 1, -1, \quad m'_l - m_l = 1, 0, -1. \quad (6.137)$$

Moreover, since we neglected magnetic interactions, spin is conserved $m'_s = m_s$. These selection rules translate to

$$j' - j = 1, 0, -1, \quad j = 0 \Rightarrow j' = 1 \quad (6.138)$$

in the total angular momentum basis $|jlm_j\rangle$.

In the present chapter we could only compute induced absorption and emission. Spontaneous emission will be discussed in chapter 10.