

Envelope functions and effective mass approximation

4

In this chapter we are interested in the quantum mechanical motion of electrons in the crystal if the periodic lattice potential is perturbed. This can occur as a result of the presence of lattice defects, impurities, or doping atoms. It can also arise due to the incorporation of interfaces between different layers of materials. Other reasons could be the presence of external electric or magnetic fields, or internal fields arising from time-dependent lattice distortions or vibrations such as those caused by phonons or surface acoustic waves. In this chapter, we will restrict ourselves to static perturbations small enough to be treated in lowest order perturbation theory, and of a spatial range much larger than the lattice constant of the underlying material. We will see that this restriction leads to considerable simplifications leading us to an effective mass Schrödinger equation for electrons in conduction bands with parabolic dispersion.

4.1 Quantum mechanical motion in a parabolic band	53
4.2 Semiclassical equations of motion, electrons and holes	59
Further reading	60
Exercises	61

4.1 Quantum mechanical motion in a parabolic band

Weak and long-range perturbations of perfect crystal symmetry can be caused, for example, by an external electric field, or by the presence of a charged doping atom. Figure 4.1 shows schematically the perturbed

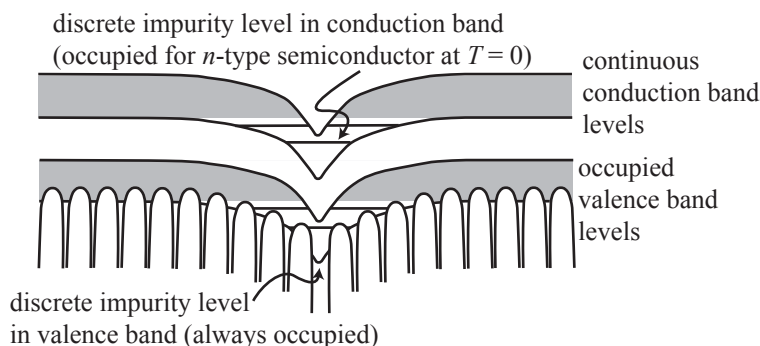


Fig. 4.1 Continuum and discrete energy levels in the vicinity of a doping atom in a semiconductor. E_1 is the energy of a discrete level below the conduction band edge; E_2 is the energy of a state in the continuum. (Reprinted with permission from Slater, 1949. Copyright 1949 by the American Physical Society.)

lattice potential in the presence of a positively charged doping atom.

There are a number of different ways of solving this quantum mechanical problem for the electronic motion. The methods differ essentially in the set of basis functions used as a starting point for a perturbation treatment. People have used Bloch-states (Enderlein and Schenk, 1992), band edge states from $\mathbf{k} \cdot \mathbf{p}$ -theory (Luttinger and Kohn, 1955), and the so-called Wannier states (Wannier, 1937; Zinman, 1972; Kittel, 1970). In order to give some insight into the derivation of the equation of motion, we will work in the Bloch-state basis and restrict the discussion to a perturbation of a parabolic conduction band with minimum at Γ as it is found, for example, in GaAs.

The problem on the basis of Bloch-states. Assume that we have solved Schrödinger's equation for the unperturbed crystal. The corresponding dispersion relations $E_n(\mathbf{k})$ and the Bloch-functions $\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\mathbf{r}}u_{n\mathbf{k}}(\mathbf{r})$ are known. Now we seek the solution of the perturbed Schrödinger equation

$$[H_0 + U(\mathbf{r})] \Psi(\mathbf{r}) = E\Psi(\mathbf{r}), \quad (4.1)$$

where H_0 is the hamiltonian of the unperturbed lattice and $U(\mathbf{r})$ is the perturbing potential. We expand the wave function $\Psi(\mathbf{r})$ on the basis of Bloch-states:

$$\Psi(\mathbf{r}) = \sum_{n,\mathbf{k}} F_n(\mathbf{k})\psi_{n\mathbf{k}}(\mathbf{r}).$$

Inserting this expansion into Schrödinger's equation gives

$$\sum_{n\mathbf{k}} \psi_{n\mathbf{k}}(\mathbf{r}) [E_n(\mathbf{k}) - E + U(\mathbf{r})] F_n(\mathbf{k}) = 0.$$

Multiplying by $\psi_{n'\mathbf{k}'}^*(\mathbf{r})$ and integrating over \mathbf{r} leads to

$$\sum_{n,\mathbf{k}} [(E_n(\mathbf{k}) - E) \delta_{n\mathbf{k},n'\mathbf{k}'} + U_{n'\mathbf{k}',n\mathbf{k}}] F_n(\mathbf{k}) = 0, \quad (4.2)$$

where we have used the orthogonality of Bloch-states and introduced the matrix elements of the perturbing potential

$$U_{n'\mathbf{k}',n\mathbf{k}} = \int d^3r \psi_{n'\mathbf{k}'}^*(\mathbf{r})U(\mathbf{r})\psi_{n\mathbf{k}}(\mathbf{r}).$$

The matrix elements of the perturbation. We will now further simplify the matrix elements of the perturbation. To this end we introduce the Fourier transform of $U(\mathbf{r})$ (see Appendix A.2) and obtain

$$U_{n'\mathbf{k}',n\mathbf{k}} = \int d^3q U(\mathbf{q}) \int d^3r e^{i(\mathbf{k}-\mathbf{k}'+\mathbf{q})\mathbf{r}} u_{n'\mathbf{k}'}^*(\mathbf{r})u_{n\mathbf{k}}(\mathbf{r}).$$

In this expression we can expand the lattice periodic function $u_{n'\mathbf{k}'}^*(\mathbf{r})u_{n\mathbf{k}}(\mathbf{r})$ into a Fourier series and obtain for the matrix element

$$U_{n'\mathbf{k}',n\mathbf{k}} = \int d^3q U(\mathbf{q}) \sum_{\mathbf{K}} C_{n\mathbf{k}}^{n'\mathbf{k}'}(\mathbf{K}) \int d^3r e^{i(\mathbf{k}-\mathbf{k}'+\mathbf{q}+\mathbf{K})\mathbf{r}}$$

with the so-called Bloch integral

$$C_{n\mathbf{k}}^{n'\mathbf{k}'}(\mathbf{K}) = \frac{1}{V_0} \int_{EZ} d^3r e^{-i\mathbf{K}\mathbf{r}} u_{n'\mathbf{k}'}^*(\mathbf{r}) u_{n\mathbf{k}}(\mathbf{r}).$$

The spatial integral in the expression for the matrix element $U_{n'\mathbf{k}',n\mathbf{k}}$ contributes only if the exponent vanishes, i.e., if $\mathbf{q} = \mathbf{k}' - \mathbf{k} - \mathbf{K}$. As a matter of fact, the integral is a representation of Dirac's delta function. Therefore the matrix element simplifies to

$$U_{n'\mathbf{k}',n\mathbf{k}} = (2\pi)^3 \sum_{\mathbf{K}} U(\mathbf{k}' - \mathbf{k} - \mathbf{K}) C_{n\mathbf{k}}^{n'\mathbf{k}'}(\mathbf{K}). \quad (4.3)$$

So far we have used the periodicity of the crystal lattice without using any approximation.

Simplifying approximations. For further simplifications to the problem we make the following assumptions about the perturbation:

- (1) We assume that the perturbing potential changes slowly on the scale of the lattice constant, i.e., $U(\mathbf{q})$ is significant only for $q \ll \pi/a$.
- (2) We assume that the perturbation is small compared to typical energy separations of bands in the crystal.
- (3) We assume that the coefficients $F_n(\mathbf{k})$ have significant values only for small values of \mathbf{k} .

According to the third assumption, we consider only states near the nondegenerate Γ -minimum. As a consequence of this and the first assumption, in the sum over \mathbf{K} only $\mathbf{K} = 0$ is retained and the matrix element simplifies to

$$U_{n'\mathbf{k}',n\mathbf{k}} \approx (2\pi)^3 U(\mathbf{k}' - \mathbf{k}) C_{n\mathbf{k}}^{n'\mathbf{k}'}(0).$$

Now we would like to simplify the Bloch integral $C_{n\mathbf{k}}^{n'\mathbf{k}'}(0)$. Based on the third assumption, we employ the expansion of the Bloch-functions near the conduction band minimum, eq. (3.19). We obtain

$$C_{n\mathbf{k}}^{n'\mathbf{k}'}(0) = \frac{1}{V_0} \int_{EZ} d^3r u_{n'\mathbf{k}'}^*(\mathbf{r}) u_{n\mathbf{k}}(\mathbf{r}) \approx \frac{1}{(2\pi)^3} \delta_{nn'} + \mathcal{O}(k^2),$$

and therefore

$$U_{n'\mathbf{k}',n\mathbf{k}} \approx U(\mathbf{k}' - \mathbf{k}) \delta_{nn'}.$$

This means that, given our assumptions, the perturbation does not mix states of neighboring bands, but only states of different \mathbf{k} near the Γ -minimum. With the above result for the matrix element, the equation of motion (4.2) simplifies to

$$\sum_{\mathbf{k}} [(E_n(\mathbf{k}) - E) \delta_{\mathbf{k},\mathbf{k}'} + U(\mathbf{k}' - \mathbf{k})] F_n(\mathbf{k}) = 0.$$

Simplification of the wave function. The wave function in real space now reads

$$\Psi(\mathbf{r}) = \sum_{\mathbf{k}} F_n(\mathbf{k}) e^{i\mathbf{k}\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}).$$

Only small wave vectors \mathbf{k} are important here, due to the long-range nature of $U(\mathbf{r})$. We therefore approximate $u_{n\mathbf{k}}(\mathbf{r}) \approx u_{n0}(\mathbf{r})$ and obtain for the wave function

$$\Psi(\mathbf{r}) = u_{n0}(\mathbf{r}) \sum_{\mathbf{k}} F_n(\mathbf{k}) e^{i\mathbf{k}\mathbf{r}} = u_{n0}(\mathbf{r}) F_n(\mathbf{r}).$$

In the last step we have interpreted the sum over \mathbf{k} as the Fourier series of a real space function $F_n(\mathbf{r})$. This function is of long range compared to the lattice period and is called the *envelope function* of the wave function.

Approximating the dispersion. We now approximate the dispersion relation $E_n(\mathbf{k})$ accordingly by using an approximation for small \mathbf{k} . Near the Γ -minimum we have [cf. eq. (3.22)]

$$E_c(\mathbf{k}) = E_c + \frac{\hbar^2 k^2}{2m^*},$$

where m^* is the effective mass of electrons in the conduction band. With these simplifications the equation of motion for electrons reads

$$\frac{\hbar^2}{2m^*} k^2 F_c(\mathbf{k}) + \sum_{\mathbf{k}'} U(\mathbf{k} - \mathbf{k}') F_c(\mathbf{k}') = (E - E_c) F_c(\mathbf{k}).$$

Equation of motion in real space. This equation determines the Fourier components of the envelope function $F_c(\mathbf{r})$. Transformation from Fourier space into real space is straightforward. The first term on the left-hand side corresponds to the second derivative of the envelope function in real space. The second term is a convolution integral which transforms into the product of the two corresponding functions in real space. We therefore obtain the following differential equation determining the envelope function $F_c(\mathbf{r})$:

$$\left[-\frac{\hbar^2}{2m^*} \Delta + \underbrace{E_c + U(\mathbf{r})}_{:=E_c(\mathbf{r})} \right] F_c(\mathbf{r}) = E F_c(\mathbf{r}). \quad (4.4)$$

This is exactly Schrödinger's equation (4.1) where the periodic lattice potential hidden in H_0 has disappeared, but the free electron mass in H_0 has been replaced by the effective mass of the conduction band electrons. Introducing the local band edge energy $E_c(\mathbf{r})$, this function acts as the effective potential in which the conduction band electrons move.

The envelope function $F_c(\mathbf{r})$ brings about very convenient simplifications. For example, matrix elements of a quantum mechanical quantity, which have to be calculated using the complete electronic wave function,

can usually be expressed as integrals over the envelope function alone. As an example, we consider the electron density. Assume that the envelope functions $F_i(\mathbf{r})$ are solutions of eq. (4.4) with energies E_i . The electron density of the system is then given by

$$n(\mathbf{r}) = \sum_i |\psi_i(\mathbf{r})|^2 f(E_i) = |u_{c0}(\mathbf{r})|^2 \sum_i |F_i(\mathbf{r})|^2 f(E_i),$$

where $f(E)$ is the Fermi distribution function. The envelope function and the lattice periodic function $u_{c0}(\mathbf{r})$ vary on different length scales. Within a primitive cell at position \mathbf{R} of the lattice $F_i(\mathbf{r}) \approx F_i(\mathbf{R})$ is essentially constant. If we are interested only in the mean density in the primitive cell at \mathbf{R} , it is given by

$$n(\mathbf{R}) = \underbrace{\frac{1}{V_{EZ}} \int_{EZ} dV |u_{c0}(\mathbf{r})|^2}_{=1} \sum_i |F_i(\mathbf{R})|^2 f(E_i) = \sum_i |F_i(\mathbf{R})|^2 f(E_i).$$

On a length scale that is large compared to the lattice constant, the electron density is given by the envelope function alone and we can neglect the lattice periodic function $u_{n0}(\mathbf{r})$.

Hydrogen-like impurities. A simple application of the concept of the envelope function is the determination of the energy levels of a hydrogen-like impurity in a semiconductor. It has indeed been shown that modern fabrication techniques have the potential to allow a precise incorporation of single doping atoms at predefined locations. Figure 4.2 shows scanning tunneling microscope images of a hydrogen passivated Si(001) surface. Using the tip of the scanning tunneling microscope, hydrogen atoms can be locally desorbed. Such a spot of about 1 nm size is shown in Fig. 4.2(a). If the surface is then exposed to PH_3 , the molecules are preferentially adsorbed at those positions, where the hydrogen passivation has been removed. A thermal annealing step lets the P atom diffuse into the top layer of the Si substrate where it forms a substitutional doping site as shown in Fig. 4.2(b).

As an example for the use of the effective mass equation, we consider a silicon atom sitting on the Ga site in a GaAs lattice. The silicon atom can satisfy all bonds with neighboring arsenic atoms using only three of its four valence electrons. As a consequence, one excess electron and an excess positive elementary charge in the silicon nucleus remain. Such a silicon atom is called a donor, because it can give away the excess electron. However, the positively charged donor ion will bind the excess electron, and the Coulomb interaction between them will appear in the equation for the envelope function:

$$\left[-\frac{\hbar^2}{2m^*} \Delta - \frac{e^2}{4\pi\epsilon\epsilon_0 r} \right] F_c(\mathbf{r}) = (E - E_c) F_c(\mathbf{r}).$$

The important point is that the relative dielectric constant of the host crystal, in our case GaAs, enters in the Coulomb potential. It accounts

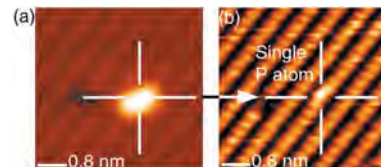


Fig. 4.2 STM images of atomically controlled single phosphor atom incorporation into Si(001). (a) Hydrogen terminated Si(001) surface with a hydrogen desorption point. (b) The same area after PH_3 dosing and annealing showing a single P atom incorporated at the location defined by the H-desorption point. (Reprinted with permission from Schofield, 2003. Copyright 2003 by the American Physical Society.)

for the polarization of the lattice by the charged donor, which effectively reduces the interaction strength. The solution of this quantum problem is that of the hydrogen problem, in which the Rydberg energy $E_{\text{Ry}} = 13.6 \text{ eV}$ is replaced by an effective Rydberg energy E_{Ry}^* and Bohr's radius $a_{\text{B}} = 0.53 \text{ \AA}$ by an effective radius a_{B}^* :

$$E_{\text{Ry}}^* = \frac{e^4 m^*}{2(4\pi\epsilon\epsilon_0)^2 \hbar^2} = E_{\text{Ry}} \frac{m^*}{m_e} \frac{1}{\epsilon^2}$$

$$a_{\text{B}}^* = \frac{4\pi\epsilon\epsilon_0 \hbar^2}{m^* e^2} = a_{\text{B}} \frac{m_e}{m^*} \epsilon.$$

For GaAs, with $\epsilon = 12.53$ and $m^* = 0.067m_e$, we find $E_{\text{Ry}}^* = 5.7 \text{ meV}$ and $a_{\text{B}}^* = 100 \text{ \AA}$. The energy levels of the hydrogen-like impurity are then

$$E_n = E_c - \frac{E_{\text{Ry}}^*}{n^2}.$$

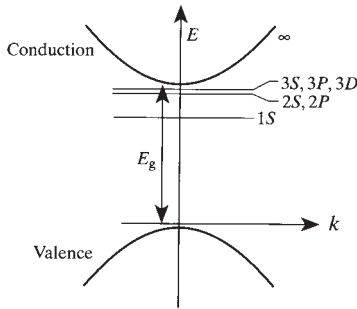


Fig. 4.3 Energy levels of a hydrogen-like impurity in GaAs (Yu and Cardona, 2001).

These states are discrete and lie below the conduction band edge of the unperturbed crystal as schematically shown in Fig. 4.3. As in the hydrogen atom, the excitation energy E_{Ry}^* from the ground state to the lower edge of the conduction band (continuum) is called the binding energy. Measured binding energies of donors in GaAs are 5.789 meV for Se_{As} , 5.839 meV for Si_{Ga} , 5.870 meV for S_{As} , 5.882 meV for Ge_{Ga} , and 5.913 meV for C_{Ga} . These values agree quite well with the theoretical prediction for E_{Ry}^* .

Figure 4.4 shows the total wave function of the ground state including the Bloch part emphasizing that the envelope function determines the shape of the probability density distribution on length scales large compared to the lattice constant.

Equation of motion at the Γ -minimum of the conduction band in the presence of a magnetic field. The equation of motion of an electron at the conduction band minimum under the influence of a

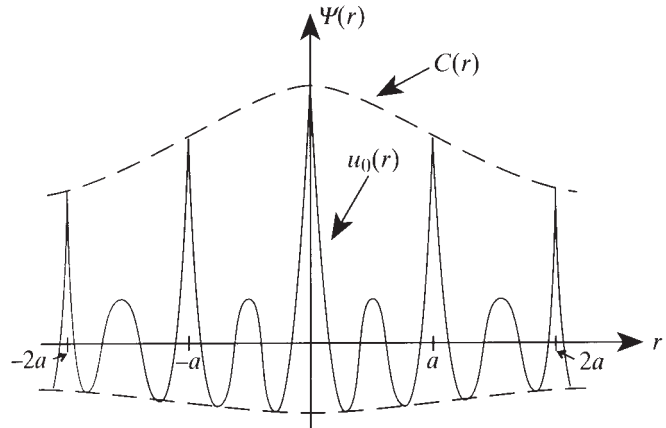


Fig. 4.4 Total wave function of the hydrogen-like impurity in GaAs including the Bloch contribution (Yu and Cardona, 2001).

magnetic field has been derived by Luttinger (1951), and by Luttinger and Kohn (1955) using similar methods. It was also found that, in this case, the equation for the envelope function is identical to the effective mass Schrödinger equation for a free particle in a magnetic field. Under the simultaneous influence of a vector potential $\mathbf{A}(\mathbf{r})$ and an electrostatic potential $U(\mathbf{r})$ the equation of motion for electrons at the Γ -minimum of the conduction band (see, e.g., Winkler 2003) reads

$$\left[\frac{1}{2m^*} \left(\frac{\hbar}{i} \nabla + |e| \mathbf{A}(\mathbf{r}) \right)^2 + U(\mathbf{r}) + \frac{1}{2} g^* \mu_B \boldsymbol{\sigma} \mathbf{B} \right] F_c(\mathbf{r}) = (E - E_c) F_c(\mathbf{r}). \quad (4.5)$$

Here, the elementary charge $|e| = 1.6 \times 10^{-19}$ C is taken to be a positive number. In the following chapters of the book we will frequently call the envelope function $F_c(\mathbf{r})$ simply the *wave function* of the electron, because its equation of motion is identical with that of an electron with mass m^* in vacuum. We will further use the convention that all energies are measured from the conduction band edge of the unperturbed crystal, such that $E_c = 0$ in the above equation. The effective mass m^* and the effective g^* -factor entering in the above equation can be calculated from the knowledge of the band edge parameters given in Table 3.6 using eqs. (3.30) and (3.31).

Equation (4.5) is of great importance for semiconductor nanostructures. Methods of structuring and patterning materials allow the fabrication of tailored potential landscapes $U(\mathbf{r})$. Magnetic fields can be created in the laboratory that influence the electronic motion as they do in the free electron case. Solving the equations of motion is greatly facilitated by the existence of many analytical solutions and approximative schemes from quantum mechanics textbooks.

The considerations leading to eq. (4.5) for conduction band electrons near Γ can be extended to semiconductors with conduction band minima at other points in the first Brillouin zone (e.g., silicon or germanium). In this case, the wave function is expanded at the corresponding conduction band minima rather than at Γ . More complicated equations of motion result due to the valley degeneracy and the anisotropic effective masses. The theory for valence band holes is also much more demanding, because there are degenerate states at Γ .

4.2 Semiclassical equations of motion, electrons and holes

Conduction band electrons. With the validity of the effective mass Schrödinger equation (4.5) for the crystal electron, the semiclassical limit of quantum mechanics (i.e., the motion of wave packets) must have its range of application in semiconductor physics. Wave packets can be constructed from the envelope functions $F_c(\mathbf{r})$ and the dynamics of its center of mass can be investigated. The result is Newton's equation of

motion

$$m^* \ddot{\mathbf{r}} = -|e|(\mathbf{E} - \dot{\mathbf{r}} \times \mathbf{B}), \quad (4.6)$$

where \mathbf{E} is the electric field and \mathbf{B} is the magnetic field at the location of the electron. As a consequence, there is a variety of possibilities in the physics of semiconductor nanostructures to investigate the borderlines between classical and quantum physics. Examples are investigations of the relation between classical and quantum chaos, or the transition from quantum to classical mechanics in the presence of decoherence.

Valence band holes. We will now briefly discuss the dynamics of holes, i.e., missing electrons near a maximum of the valence band, in the classical limit. The convex curvature of the valence band could be interpreted using a negative effective mass. Newton's equation of motion reads in this case

$$-m^* \ddot{\mathbf{r}} = -|e|(\mathbf{E} - \dot{\mathbf{r}} \times \mathbf{B}).$$

However, a negative effective mass is physically not very intuitive. We can reinterpret this equation of motion by multiplying it by -1 :

$$m^* \ddot{\mathbf{r}} = +|e|(\mathbf{E} - \dot{\mathbf{r}} \times \mathbf{B})$$

This can be interpreted as the equation of motion for particles with positive mass m^* , but with *positive charge* $+|e|$. The occurrence of a positive charge at the top of the valence band is also intuitive from another point of view. In the electrically neutral, uncharged semiconductor crystal the valence band is completely filled. Removing an electron from the top of the valence band, an initially localized positive charge remains. Such a missing electron is called a hole. According to the above equation of motion, the effective mass m^* and the charge $+e$ are properties of this hole which appears to move through the crystal like a classical particle.

Further reading

- Papers: Slater 1949; Luttinger 1951; Luttinger and Kohn 1955.
- Effective mass from $\mathbf{k} \cdot \mathbf{p}$ -theory: Davies 1998; Kittel 1970; Yu and Cardona 2001.
- Effective mass from quasi-classical considerations with group velocity and Newton's equation of motion: Kittel 2005; Kittel 1970; Singleton 2001; Ashcroft and Mermin 1987.
- Effective mass from the hydrogen problem in semiconductors, doping: Davies 1998.
- Band structure of semiconductors: Winkler 2003.

Exercises

- (4.1) Consider the differential equation for the envelope function, eq. (4.5), with a magnetic field $\mathbf{B} = (0, 0, B)$ and the Coulomb potential $U(\mathbf{r}) = e^2/4\pi\epsilon\epsilon_0 r$.

- (a) Give reasons why the solution of the problem can be separated in that of the orbital motion and that of the spin dynamics.
- (b) Discuss qualitatively the effects of the magnetic field on the spin dynamics.
- (c) Discuss qualitatively how the magnetic field affects the orbital energy levels and wave functions.

- (4.2) In silicon, the hamiltonian for the conduction band envelope function in the effective mass equation is given by

$$H = \frac{\hbar^2}{2m_L} \frac{\partial^2}{\partial x^2} + \frac{\hbar^2}{2m_T} \left(\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + V_c(r),$$

where $V_c = e^2/4\pi\epsilon\epsilon_0 r$ is the Coulomb potential, and m_L and m_T are the longitudinal and transverse effective masses, respectively. Consider the case $m_L = m_T + \Delta m$, where $\Delta m/m_T \ll 1$. Calculate the effect of the presence of Δm on the energies of the 1s-, 2s-, and 2p-states of a hydrogen-like impurity using perturbation theory.