

Impurities in semiconductors

15.1 Bohr's hydrogen atom model

Shallow impurities are of great technological importance in semiconductors since they determine the conductivity and the carrier type of the semiconductor. Shallow impurities are defined as impurities which are ionized at room temperature. This condition limits the ionization energy of such impurities to values $\ll 100$ meV. Shallow impurities can be either acceptors or donors, *i. e.* 'accept' electrons from the valence band or 'donate' electrons to the conduction band.

The hydrogen atom model can serve as the basis for the calculation of many properties of shallow impurities such as ionization energy and state wave functions. In this chapter, the hydrogen atom is analyzed in terms of Bohr's semi-classical model and in terms of a quantum mechanical approach. The hydrogen atom model is then applied to shallow impurities. Properties such as ionization energies, wave functions, central cell correction terms, and screening of impurity potentials by free carriers are summarized.

Impurities in semiconductors can be incorporated on substitutional sites, interstitial sites, or as impurity complexes. Here, we restrict ourselves to *substitutional, shallow* impurities. Examples for such impurities are Be, Zn, Si, and Sn. These impurities are shallow, *i. e.* their ionization energy is comparable to the thermal energy kT at room temperature. As a consequence, shallow impurities are fully ionized at room temperature. The hydrogen atom model has proven to predict accurately many properties of shallow impurities.

Coulomb potential

The electrostatic potential of a point charge is called the Coulomb potential or the $1/r$ potential. The Coulomb potential of a positive point charge ($+e$) in vacuum located at $r=0$ is obtained from Poisson's equation and is given in spherical coordinates by

$$V(r) = \frac{e}{4\pi\epsilon_0 r} \quad (15.1a)$$

where e is the elementary charge and ϵ_0 is the permittivity of vacuum. Analogously, the Coulomb potential of a positively charged impurity located at $r=0$ in a semiconductor with the dielectric constant $\epsilon_r = \epsilon/\epsilon_0$ is given by

$$V(r) = \frac{e}{4\pi\epsilon r} \quad (15.1b)$$

where ϵ is the permittivity of the semiconductor.

Binding energy and Bohr radius

The hydrogen atom model developed by Bohr is based on (i) classical mechanics of an electron in the Coulomb potential of a positive point charge and on (ii) the quantization of the electron angular momentum. The Bohr model predicts many of the physical properties of the hydrogen atom most notably the emission spectra of the atom. The model is a fascinating example of the

simplicity and the power of quantum mechanics. For the classical motion of an electron in the Coulomb potential, the potential energy is given by

$$E_{\text{pot}} = -\frac{e^2}{4\pi\epsilon_0 r} \quad (15.2)$$

For the hydrogen atom, the permittivity is that of vacuum since a vacuum is assumed between the proton and the electron. The schematic Coulomb potential and an electron orbiting the proton at a distance r are shown in **Fig. 15.1**. The *attractive* electrostatic Coulomb force F_C on the proton and the electron towards each other is given by

$$F_C = e\mathcal{E}_C = \frac{1}{4\pi\epsilon_0} \frac{e^2}{r^2} \quad (15.3)$$

where $\mathcal{E}_C = dE_{\text{pot}}/dr$ is the Coulomb field. The *repulsive* centrifugal force F_z on the orbiting electron with mass m_0 and velocity v is given by

$$F_z = \frac{m_0 v^2}{r} = m_0 r \omega^2 \quad (15.4)$$

where ω is the angular frequency of the electron. Equation (15.4) is valid, only if the electron mass, m_0 , is much smaller than the proton mass, m_p . This condition is fulfilled, since $m_0/m_p \approx 1/1840$. The first classical condition in Bohr's hydrogen atom model is $F_C = F_z$, *i. e.* the balance of attractive and repulsive forces.

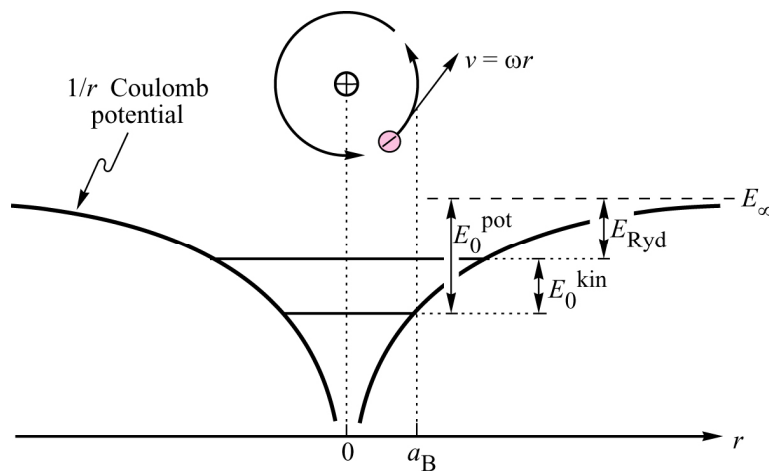


Fig. 15.1. Illustration of Bohr's hydrogen atom model. The electron orbits the positive proton with a $1/r$ Coulomb potential at the radius a_B . The ionization energy of the atom is the Rydberg energy $E_{\text{Ryd}} = E_\infty - |E_0^{\text{pot}}| + |E_0^{\text{kin}}|$.

The second condition for the Bohr atom is the quantization of the angular momentum of the electron which is given by

$$m_0 v r = m_0 \omega r^2 = n \hbar \quad (n = 1, 2, 3 \dots) \quad (15.5)$$

The validity of the angular momentum quantization can be visualized by recalling the wave character of the electron. An electron wave around the positive proton is shown in **Fig. 15.2**. The electron wave is stable, only if the wave is interfering constructively with itself, *i. e.* when the length of the electron orbit equals integer multiples of the electron wavelength, λ ,

$$2\pi r = n\lambda \quad (n = 1, 2, 3 \dots) \quad (15.6)$$

The reader can easily verify that Eqs. (15.5) and (15.6) are identical by recalling that the kinetic energy of a particle is given by $E = \hbar^2 k^2 / 2m_0$ where $k = 2\pi/\lambda$ and m_0 are the electron wave vector and mass, respectively.

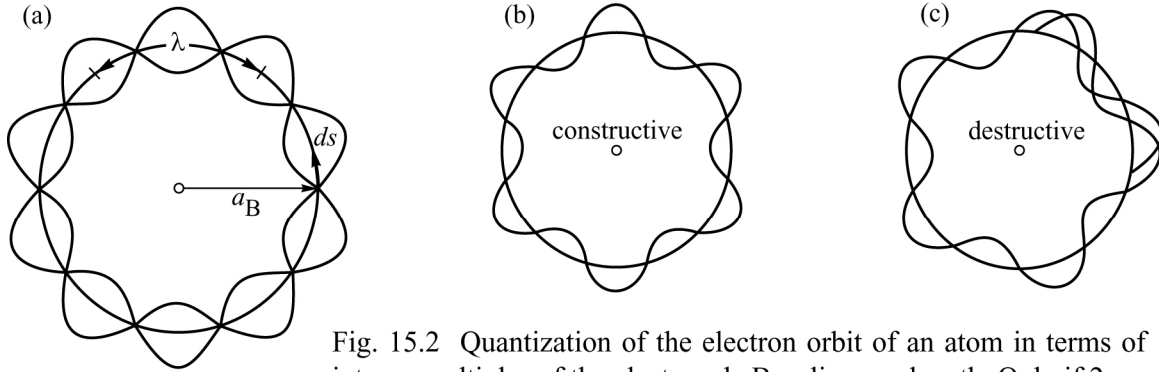


Fig. 15.2 Quantization of the electron orbit of an atom in terms of integer multiples of the electron de Broglie wavelength. Only if $2\pi r_B = n\lambda$, does constructive interference of the electron wave occur.

Elimination of force and velocity from Eqs. (15.3) – (15.5) yields the radii of the allowed electron orbits in the hydrogen atom

$$a_{B,n} = \frac{4\pi \epsilon_0 n^2 \hbar^2}{m_0 e^2} \quad (n = 1, 2, 3 \dots) \quad (15.7)$$

The radius of the ground state orbit ($n = 1$) is given by

$$a_B = 0.053 \text{ nm} \quad (15.8)$$

which is called the **Bohr radius** of the hydrogen atom.

Insertion of the Bohr radii into Eq. (15.2) yields the potential energy of the electron

$$E_{\text{pot},n} = \frac{-1}{(4\pi \epsilon_0)^2} \frac{e^4 m_0}{n^2 \hbar^2} \quad (n = 1, 2, 3 \dots) \quad (15.9)$$

Furthermore, the kinetic energy is obtained via the electron velocity of Eq. (15.5) according to

$$E_{\text{kin},n} = \frac{1}{2} \frac{1}{(4\pi \epsilon_0)^2} \frac{e^4 m_0}{n^2 \hbar^2} \quad (n = 1, 2, 3 \dots) \quad (15.10)$$

Comparison of Eq. (15.9) with Eq. (15.10) reveals that the kinetic energy is just half of the potential energy, *i. e.*

$$E_{\text{kin},n} = \frac{1}{2} |E_{\text{pot},n}| \quad (15.11)$$

The energy required to move the electron from the n th state energy, E_n , to the vacuum level at infinite distance from the proton, *i. e.* $E_\infty = E(r \rightarrow \infty)$, is given by

$$E_{\text{Ryd},n} = \frac{1}{2} \frac{1}{(4\pi\epsilon_0)^2} \frac{e^4 m_0}{n^2 \hbar^2} \quad (n = 1, 2, 3, \dots) \quad (15.12a)$$

which, for $n = 1$, is called the **Rydberg energy**. This energy is required to ionize a hydrogen atom. For $n = 1$ the Rydberg energy is given by

$$E_{\text{Ryd}} = 13.6 \text{ eV}. \quad (15.12b)$$

In the classical orbital motion of the electron around the proton, the ratio of the electron velocity and the velocity of light c can be calculated from the Bohr model. The ratio is obtained as

$$\alpha = \frac{v}{c} = \frac{e^2}{4\pi\epsilon_0 \hbar c} \approx \frac{1}{137} \quad (15.13)$$

which is called **the Sommerfeld fine structure constant**. Evaluation of Eq. (15.13) yields that the electron orbits the proton with a velocity of approximately 2200 km/s.

The magnetic field generated by the circular current of the orbiting electron can be calculated from Bohr's model using the Maxwell equations. It is given by

$$\mu_B = \frac{\mu_0 \hbar e}{2 m_0} \quad (15.14)$$

and is called the **Bohr magneton**.

The above calculation demonstrates that the relatively simple Bohr model, *i. e.* classical mechanics and angular momentum quantization, provides many physical quantities of the hydrogen atom. The calculated state energies of the hydrogen atom were found to agree with hydrogen emission spectra. The Bohr model and its prediction of the electron energies was one of the first successes of the quantum theory. Further refinement of the model is obtained by considering not only circular orbits but also elliptical orbits. On such an elliptical orbit the velocity of the electron is a function of the position, *i. e.* the velocity is not constant as in the circular orbit. (The position dependence of the velocity is analogous to the planetary motion around the sun).

Using the momentum $p = \hbar k$ and $k = 2\pi/\lambda$, the angular momentum quantization condition, which for circular motion is given by Eq. (15.6), can be written for any orbit as

$$\frac{1}{2\pi} \int_0^{2\pi} p_\phi d\phi = n_\phi \hbar \quad (n_\phi = 1, 2, 3 \dots) \quad (15.15)$$

where $p_\phi = m \omega_\phi r^2$ is the position-dependent (that is angle-dependent) angular momentum. For a circular orbit, the angular momentum is a constant and Eq. (15.15) reduces to Eq. (15.5).

The condition of classical mechanics for motion on an elliptical orbit and the angular momentum quantization condition lead to the total energy of the electron. The total energy is given by

$$E_{n_\phi} = \frac{1}{2} \frac{1}{(4\pi \varepsilon_0)^2} \frac{e^4 m_0}{n_\phi^2 \hbar^2} \quad (n_\phi = 1, 2, 3 \dots) \quad (15.16)$$

which is identical to Eq. (15.12). Thus, elliptical orbits for the electron exist and have the same energy as electrons on circular orbits. The total energy of a particle on an elliptical orbit in a $1/r$ potential can be calculated by classical mechanics and depends only on the main axis a of the ellipsis. The main axis is then given by

$$a = \frac{4\pi \hbar^2 \varepsilon_0}{m_0 e^2} n_\phi^2. \quad (15.17)$$

The angular momentum of the particle on the elliptical orbit is given by

$$p_\phi^2 = \frac{a(1-\varepsilon^2)e^2 m_0}{4\pi \varepsilon_0} \quad (15.18a)$$

where ε is the eccentricity of the ellipsis. Since the angular momentum is quantized according to $p_\phi = n_\phi \hbar$, one obtains with Eq. (15.18a)

$$n_\phi^2 \hbar^2 = \frac{a(1-\varepsilon^2)e^2 m_0}{4\pi \varepsilon_0}. \quad (15.18b)$$

Inserting the main axis a , given by Eq. (15.17), into Eq. (15.18b) yields

$$\frac{b}{a} = \frac{n_\phi}{n} \quad (15.19)$$

where a and b are the axes of the ellipsis. The angular momentum quantum number n_ϕ can assume values of 1, 2, 3 ..., n , which represents a family of ellipses. If the angular quantum number coincides with the principal quantum number, *i. e.* $n_\phi = n$, the previously calculated circular orbit is obtained. The ellipses for the $n = 1, 2$, and 3 states are shown in **Fig. 15.3(a)**.

If the angular momentum quantum number is formally introduced as $l = n_\phi - 1$, then l can assume values of

$$l = 0, 1, 2 \dots n - 1. \quad (15.20)$$

The value of $l = 0$ corresponds to the ellipsis with the largest eccentricity. The value of $l = n - 1$ represents the circular orbit.

Each orbit of an electron around the proton of a hydrogen atom is fully determined by the principle quantum number, n , and the angular quantum number, l . The orbits of the electron for different quantum numbers and the corresponding energies are shown in **Fig. 15.3**. Frequently, the $l = 0, 1, 2$ and 3 orbitals are historically denoted as the s, p, d and f orbitals (White, 1934). If several quantum states have the same energy the states are called **degenerate states**. For example, the two states determined by $n = 2, l = 0$ and $n = 2, l = 1$ are degenerate. In addition to the principal quantum number and the angular momentum number, the quantum number m

describes the quantization of the azimuthal angular momentum in units of \hbar . For a discussion of the azimuthal quantum number, we refer to the literature (see, for example, Bohm, 1951).

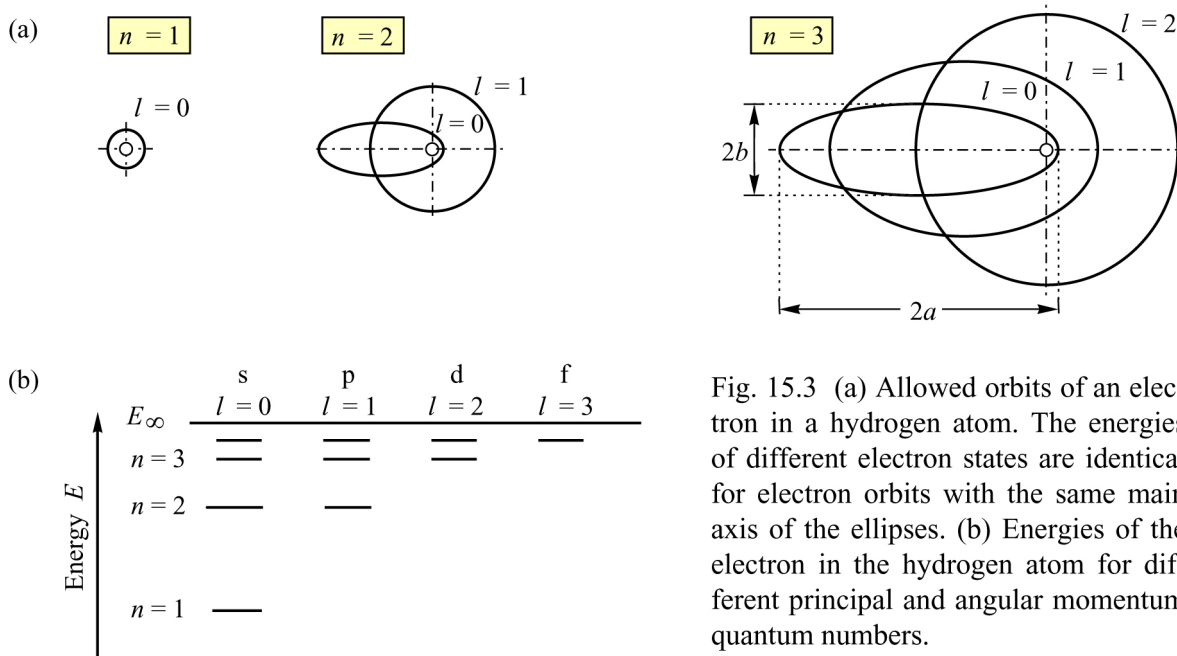


Fig. 15.3 (a) Allowed orbits of an electron in a hydrogen atom. The energies of different electron states are identical for electron orbits with the same main axis of the ellipses. (b) Energies of the electron in the hydrogen atom for different principal and angular momentum quantum numbers.

Upon any perturbation of the hydrogen atom, different electron orbits respond in different ways to the perturbation. Therefore, degenerate electron states will split and become non-degenerate upon a suitable perturbation. The perturbation of the hydrogen atom can be achieved, for example, by an electric field (Stark effect) or a magnetic field (Zeeman effect).

Even though the Bohr model explains many characteristics of the hydrogen atom it is limited in its applicability. For example, if the principles of the Bohr model are applied to the helium atom, incorrect results are obtained for the energy levels in that atom. In addition, the deterministic Bohr hydrogen model violates the quantum mechanical uncertainty principle. That is, momentum and position of the electron are exactly determined at all times in Bohr's model, which contradicts $\Delta x \Delta p \approx \hbar$. Nevertheless, due to its simplicity and clarity, the Bohr model has not lost its attractiveness.

The Bohr model was refined in 1925 by inclusion of the electron spin. The spin of an electron is also called its intrinsic angular momentum and can be visualized as the rotation of an electron around its own symmetry axis. Goudsmit and Uhlenbeck postulated the spin when conducting Zeeman effect experiments on hydrogen. Due to angular momentum quantization the difference between intrinsic angular momenta is \hbar . The intrinsic angular momentum is then given by

$$p_s = s \hbar \quad \text{with} \quad s = \pm 1/2 \quad (15.21)$$

where s is the spin quantum number which can assume values of $s = \pm 1/2$. If an electron with spin is subjected to an external magnetic field, the spin-axis will orientate (align) itself in a parallel or antiparallel manner, as shown in **Fig.** 15.4. All other orientations are of transient nature since the repulsive and attractive magnetic forces of the spin and external field tend to reorient the spin to the parallel or antiparallel orientation.

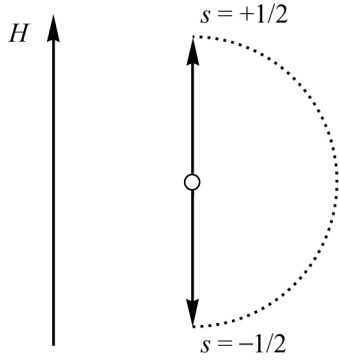


Fig. 15.4. Alignment of the spin of an electron in an external magnetic field H . The spin is either parallel (spin 'up') or antiparallel (spin 'down') to the magnetic field, as represented by the spin quantum number $s = \pm 1/2$. The difference in angular momentum between the two orientations is $\hbar/2$.

Wave functions of the hydrogen atom

The Bohr model predicts the energy levels of the hydrogen atom with amazing accuracy. However, the wave functions ψ_{nlm} and the probability distributions $\psi_{nlm} \psi_{nlm}^*$ cannot be obtained from the Bohr model. The Schrödinger equation must be solved in order to obtain the exact solutions. Although the exact solution of the hydrogen atom model goes beyond the scope of this book, the results of the exact hydrogen atom are summarized in the following. For further study, the reader is referred to textbooks on quantum mechanics (Bohm, 1951; Sherwin, 1959; Davydov, 1965; Borowitz, 1967; Saxon, 1968; Merzbacher, 1970).

The wave functions of the hydrogen atom can be obtained by solving the Schrödinger equation. For a particle with mass m and charge e in a potential V , the Schrödinger equation is given by

$$\frac{\hbar^2}{2m} \Delta \Psi + eV \Psi = -\frac{\hbar}{i} \frac{\partial \Psi}{\partial t} \quad (15.22)$$

where $\Psi = \Psi(x, y, z, t) = \psi(x, y, z) e^{i\omega t}$ is the time-dependent wave function and $\Delta = (\partial^2 / \partial x^2 + \partial^2 / \partial y^2 + \partial^2 / \partial z^2)$ is the delta operator. Not being interested in the time dependence of the solution, we use the kinetic energy operator $E = -(\hbar/i)(\partial / \partial t)$ to obtain the time-independent Schrödinger equation

$$\boxed{-\frac{\hbar^2}{2m} \Delta \psi + (eV - E) \psi = 0} \quad (15.23)$$

where $\psi = \psi(x, y, z)$ is the time-independent wave function. The Coulomb potential is a spherically symmetric potential and can be expressed solely as a function of the radius r , *i. e.* $V = V(r)$ (see Eq. (15.1)). It is useful to convert the Δ -operator into spherical coordinates (r, θ, ϕ) , *i. e.*

$$\begin{aligned} \Delta &= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \\ &= \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \end{aligned} \quad (15.24)$$

where r , θ and ϕ are the radius, the polar angle, and azimuthal angle, respectively. The Schrödinger equation is a separable linear differential equation and can be solved by employing the product method. The wave functions can then be written as

$$\psi(r,\theta,\phi) = R(r) \Theta(\theta) \Phi(\phi) . \quad (15.25)$$

Since $\psi \psi^*$ is the quantum mechanical probability density of the particle, the wave function ψ must satisfy the condition

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi \psi^* dx dy dz = \int_0^{\infty} \int_0^{2\pi} \int_0^{\pi} \psi \psi^* r^2 \sin\theta d\theta d\phi dr = 1 . \quad (15.26)$$

The solution of the Schrödinger equation in spherical coordinates for the Coulomb potential is a set of orthogonal functions which are usually classified by the four quantum numbers n , l , m_l , and s . The quantum numbers are

- n = the principal quantum number
- l = the orbital angular momentum number
- m_l = the azimuthal quantum number
- s = the spin quantum number

The quantum number can assume values of

$$n = 1, 2, 3 \dots \quad (15.27a)$$

$$l = 0, 1, 2 \dots n-2, n-1 \quad (15.27b)$$

$$m_l = -l, -l+1 \dots l-1, +l \quad (15.27c)$$

$$s = -\frac{1}{2}, +\frac{1}{2} . \quad (15.27d)$$

Two electrons with different spin can occupy an orbit defined by the three quantum numbers n , l , and m_l (Pauli principle).

The wave functions corresponding to the three quantum numbers n , l , and m_l are designated as ψ_{nlm_l} . Correspondingly, the radial parts of the wave functions are denoted as R_{nlm_l} (see Eq. 15.25). The wave functions for some of the lowest states of the hydrogen atom (Sherwin, 1959) are given by

$$\psi_{100} = \frac{1}{\sqrt{\pi}} \left(\frac{1}{a_B} \right)^{3/2} e^{-r/a_B} \quad (15.28a)$$

$$\psi_{200} = \frac{1}{4\sqrt{2\pi}} \left(\frac{1}{a_B} \right)^{3/2} \left(2 - \frac{r}{a_B} \right) e^{-r/2a_B} \quad (15.28b)$$

$$\psi_{210} = \frac{1}{4\sqrt{2\pi}} \left(\frac{1}{a_B} \right)^{3/2} (\cos\theta) \frac{r}{a_B} e^{-r/2a_B} \quad (15.28c)$$

$$\psi_{211} = \frac{1}{4\sqrt{2\pi}} \left(\frac{1}{a_B} \right)^{3/2} \frac{e^{i\phi}}{\sqrt{2}} (\sin\theta) \frac{r}{a_B} e^{-r/2a_B} \quad (15.28d)$$

$$\psi_{21-1} = \frac{1}{4\sqrt{2\pi}} \left(\frac{1}{a_B} \right)^{3/2} \frac{e^{-i\phi}}{\sqrt{2}} (\sin\theta) \frac{r}{a_B} e^{-r/2a_B} \quad (15.28e)$$

Further solutions of the hydrogen atom state can be found in the literature (Bohm, 1951). The radial parts R_{nl0} ($m_l = 0$) for some of the lowest hydrogen atom states are shown in **Fig. 15.5**. States of ‘s-type’ (*i. e.* $l = 0$) symmetry have a maximum of the wave function at $r = 0$. States of ‘p-type’ (*i. e.* $l = 1$) symmetry have a node at $r = 0$.

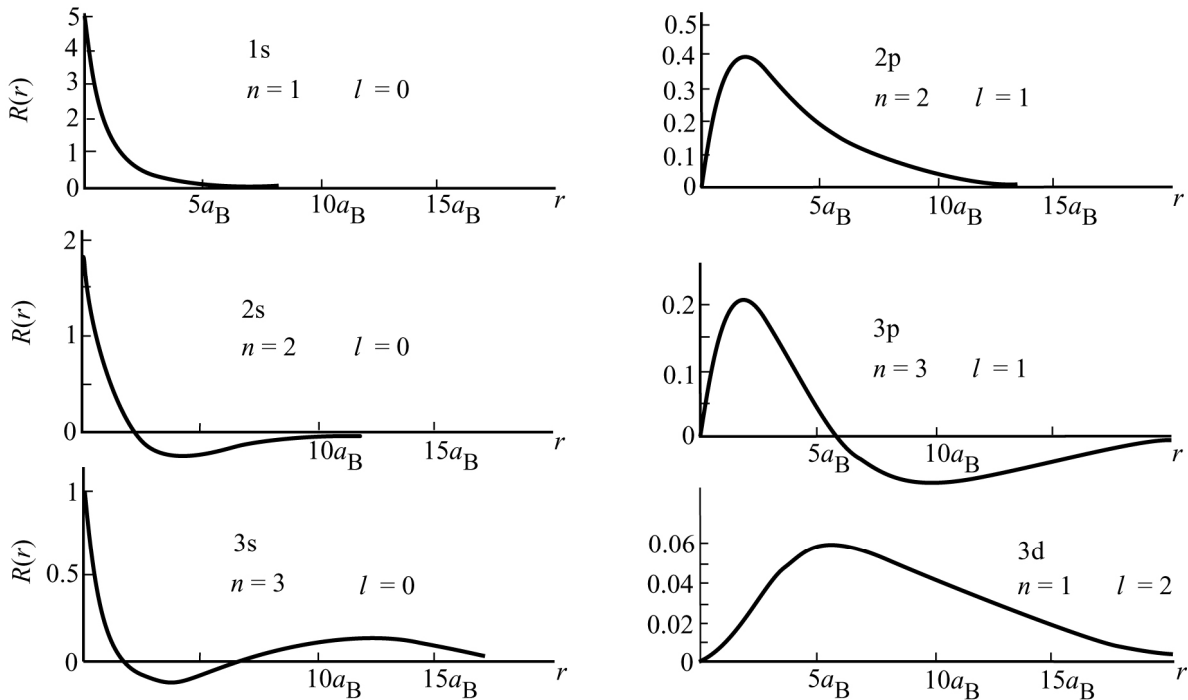


Fig. 15.5. Calculated radial parts of the wave functions for some elementary states of the hydrogen atom.

The probability of finding an electron at radius r can be obtained by integration over all angles θ and ϕ

$$p(r) dr = \int_0^{2\pi} \int_0^\pi \psi \psi^* r^2 \sin\theta d\theta d\phi dr \quad (15.29)$$

For wave functions of s-type symmetry, ψ does not depend on θ and ϕ . Eq. (15.29) then yields

$$p(r) = \psi \psi^* 4\pi r^2 . \quad (15.30)$$

Using $\psi = \psi_{100}$ according to Eq. (15.28) one obtains

$$p(r) = 4 \frac{r^2}{a_B^3} e^{-2r/a_B} . \quad (15.31)$$

The probability $p(r)$ has a maximum at $r = a_B$. Thus, the classical Bohr radius is the radius of maximum probability in the quantum mechanical picture.

15.2 Hydrogenic donors

The hydrogen atom model can be applied to shallow donors in III–V semiconductors. Properties predicted by the hydrogen atom model agree amazingly well with experimentally determined properties of shallow donors. Such donors are called *effective-mass-like donors*, *hydrogen-like donors*, or briefly *hydrogenic donors*. Ionization energy, wave functions, and effective Bohr radius are well predicted for such donors. The similarity of the hydrogen atom and donor impurities originates in the $1/r$ coulombic potential of both entities.

Two modifications are required in order to apply Bohr's hydrogen atom model to shallow donors. These corrections are related to the effective mass of carriers and of the dielectric constant of the semiconductor (Bethe, 1942). First, the effective mass of electrons in semiconductors, m_e^* , differs from the electron mass quite significantly. The dispersion relation of a semiconductor with a spherically symmetric band structure is given by

$$E = \frac{\hbar^2 k^2}{2 m_e^*} . \quad (15.32)$$

Thus, the dispersion relation $E(k)$ allows one to determine the effective mass according to

$$m_e^* = \hbar^2 \left(\frac{d^2 E}{dk^2} \right)^{-1} . \quad (15.33)$$

The dispersion relation of Eq. (15.32) differs from the dispersion relation of a free electron just by the magnitude of the electron mass. In order to apply the hydrogen atom model, the electron mass must be replaced by the effective electron mass. The second correction arises from the dielectric properties of semiconductors. The Coulomb potential of a positive point-charge in a semiconductor located at $r = 0$ is given by

$$V(r) = \frac{e}{4\pi \epsilon r} \quad (15.34)$$

which differs from Eq. (15.1a) by the static dielectric constant $\epsilon_r = \epsilon / \epsilon_0$. How does the potential change in the presence of an electron orbiting the impurity charge? The polarization of the valence electrons is then more complicated and cannot be taken into account by the substitution of ϵ for ϵ_0 . To answer the question we first make the simplifying assumption that the electronic charge can be described by a diffuse electron cloud with a spatial extent much larger than the lattice constant. In the limit of an infinitely large electron cloud, the potential of the positive

impurity is correctly described by Eq. (15.34). The situation changes, however, if the electron were to orbit the impurity atom with a radius comparable to the lattice constant. In this case the polarization of the lattice depends on the donor as well as on the electron charge. The true potential is then not given by the dielectrically screened potential of Eq. (15.34). For such a small electron orbit, the polarization of lattice atoms is overestimated by Eq. (15.34). A smaller dielectric constant $\epsilon_r^* < \epsilon_r$ can be used to account for the reduction in polarization. It should therefore be noted that Eq. (15.34) assumes a Coulomb potential screened by the dielectric properties of the semiconductor (*i. e.* by polarization of tightly bound valence electrons and nuclei of the lattice) and that the equation can only be used if the electron can be described by a diffuse electron cloud with a large spatial extent. The true potential $V(r)$ which arises from the positive donor ion, the electron bound to the donor ion, and the polarization of the surrounding semiconductor is rather complicated and cannot be expressed in terms of a simple $1/r$ potential (Kohn, 1957a, 1957b). Employing the approximate $1/r$ potential, the Schrödinger equation is given by

$$-\frac{\hbar^2}{2m_e^*} \Delta\psi + \frac{e}{4\pi\epsilon r} \psi = E \psi \quad (15.35)$$

which is called the effective-mass equation for a hydrogenic impurity.

Using the effective-mass and the dielectric constant corrections, the following properties of hydrogenic impurities can be derived. The **effective Bohr radius** is obtained from Eq. (15.7) and is given by

$$a_{B,n}^* = \frac{4\pi\epsilon n^2 \hbar^2}{m_e^* e^2} \quad (n = 1, 2, 3 \dots) \quad (15.36)$$

The radius of the donor ground state ($n = 1$) is then given by

$$a_B^* = \frac{4\pi\epsilon \hbar^2}{m_e^* e^2} = \frac{\epsilon_r}{m_e^*/m_0} a_B = \frac{\epsilon_r}{m_e^*/m_0} 0.53 \text{ \AA} \quad (15.37)$$

The effective Bohr radius is also called the *donor Bohr radius*. As an example, we consider a hydrogenic donor in GaAs with $\epsilon_r = 13.1$ and $m_e^* = 0.067 m_0$. Insertion of these values into Eq. (15.37) yields $a_B^* = 103 \text{ \AA}$ which is the effective Bohr radius of donors in GaAs.

The **effective Rydberg energy** is obtained by applying the effective-mass and the dielectric constant corrections to Eq. (15.12).

$$E_{\text{Ryd},n}^* = \frac{1}{2} \frac{1}{(4\pi\epsilon)^2} \frac{e^4 m_e^*}{n^2 \hbar^2} \quad (n = 1, 2, 3 \dots) \quad (15.38)$$

The **donor ionization energy** is required for a transition from $n = 1$ to $n \rightarrow \infty$ and is given by

$$E_d = \frac{e^4 m_e^*}{2(4\pi\epsilon \hbar)^2} = \frac{m_e^*/m_0}{\epsilon_r^2} E_{\text{Ryd}} = \frac{m_e^*/m_0}{\epsilon_r^2} 13.6 \text{ eV} \quad (15.39)$$

The donor ionization energy is occasionally also referred to as donor Rydberg energy. As an example, we consider a hydrogenic donor in GaAs and obtain $E_d = 5.3$ meV which is in agreement with experimental results.

Finally, the wave functions of hydrogenic donors can be obtained from Eq. (15.28) by substituting the effective Bohr radius for the Bohr radius. The ground-state envelope wave function is then obtained as

$$\psi_{100}(r) = \frac{1}{\sqrt{\pi}} \left(\frac{1}{a_B^*} \right)^{3/2} e^{-r/a_B^*} . \quad (15.40)$$

It should be noted that Eq. (15.40) describes the donor *envelope* function rather than the donor wave function. The actual donor ground-state wave function is given by (Kohn, 1957a)

$$\Psi_{d,100}(\mathbf{r}) = \psi_{100}(r) u_k(\mathbf{r}) \quad (15.41)$$

where $u_k(\mathbf{r})$ is the lattice-periodic factor of the well-known Bloch function of conduction band electrons. The function $u_k(\mathbf{r})$ has translational symmetry with respect to the semiconductor lattice constant. The ground-state wave function according to Eq. (15.41) is schematically shown in **Fig. 15.6**. The dashed curve represents the impurity envelope function of Eq. (15.40).

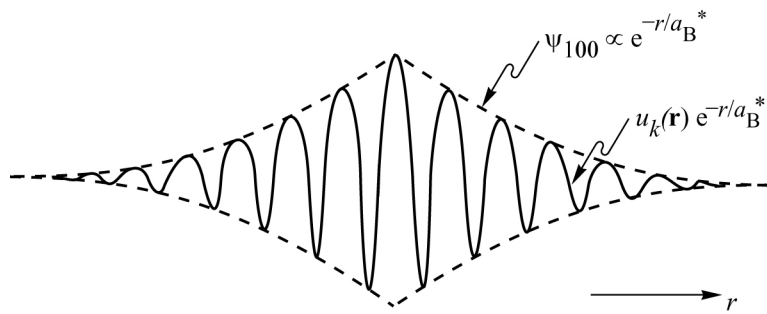


Fig. 15.6. Schematic illustration of a hydrogenic donor wave function with the quantum numbers $n = 1$ and $l = m_l = 0$. The wave function is the product of the lattice-periodic Bloch function $u_k(\mathbf{r})$ and the envelope function ψ_{100} .

It should be noted that the use of Bohr's hydrogen atom model for shallow impurities is not self-sufficient. The *ab initio* assumption of a 'large' electron cloud and the substitution of effective electron mass and dielectric constant cannot be justified solely on the basis of the hydrogenic model. Even though the hydrogenic model yields a relatively large electron orbit, this result does not justify the initial assumptions. However, more rigorous calculations (Kohn 1957a, 1957b; Madelung, 1978; Altarelli and Bassani, 1982) indeed demonstrate that the electron distribution has a spatial extent much larger than the lattice constant. The substitution of the electron mass, m_0 , by the effective mass, m_e^* , is therefore justified since the electron orbit around the donor extends over many lattice constants. The effective mass of electrons in semiconductors is a direct consequence of the periodic potential of the lattice. Thus, electrons bound to donors are subject to the periodic potential, since the effective Bohr radius is much larger than the lattice constant, $a_B^* \gg a_B$. If, in contrast, electrons were tightly bound ($a_B^* \approx a_B$) the effective-mass correction could not be applied. Similar arguments apply to the substitution of the permittivity of the semiconductor, ϵ , for the permittivity of vacuum, ϵ_0 . The lattice atoms are polarized by the Coulomb field which results in its reduction as compared to the field without polarization. The effect of the polarization is taken into account via the dielectric constant. Since the effective Bohr radius extends over many lattice constants, the use of the dielectrically

screened Coulomb potential is justified. Despite the simplicity of the hydrogen atom model for shallow donors, the model yields quite accurate results.

The *degeneracy* of the donor ground state is a quantity required for the occupancy probability of the donor state (see Chap. 3). The ground state has the quantum numbers $n = 0$, $l = 0$, $m_l = 0$ and $s = \pm 1/2$. Thus, since the donor ground state can be occupied by an electron with spin $+ 1/2$ or $- 1/2$, the ground state degeneracy is $g = 2$.

Exercise 1: Dopant ionization energies. In GaAs, the dielectric constant is $\epsilon_r = 13.1$ and the effective electron mass is $m_e^* = 0.067 m_0$. In Si, the dielectric constant is $\epsilon_r = 11.9$ and the effective electron mass is $m_e^* = 0.98 m_0$. The experimental values for donor ionization energies for Si donors in GaAs and As donors in Si are 6 meV and 54 meV, respectively. Calculate the donor ionization energies and effective Bohr radii for donors in the two materials material by using the hydrogen model. Are hydrogenic theory and experimental values for the ionization energy in reasonable agreement?

Compare the donor ionization energies in GaAs and Si with the thermal energy kT at room temperature and at 77 K. Would you expect donors in the two materials to be ionized at 77 K? Would you expect donors in the two materials to be ionized at 300 K?

The hydrogen atom model can be also applied to *acceptors*, by using the *effective hole mass* rather than the effective electron mass. In GaAs, the dielectric constant is $\epsilon_r = 13.1$ and the effective heavy-hole mass is $m_{hh}^* = 0.45 m_0$. In GaN, the dielectric constant is $\epsilon_r = 9.0$ and the effective heavy-hole mass is $m_{hh}^* = 0.8 m_0$. The experimental values for acceptor ionization energies for Be acceptors in GaAs and Mg acceptors in GaN are 26 meV and 200 meV, respectively. Calculate the ionization energies and the effective Bohr radii for acceptors in the two materials by using the hydrogen model. Are hydrogenic theory and experimental values for the ionization energy in reasonable agreement?

Compare the ionization energies for acceptors in GaAs and acceptors in GaN with the thermal energy kT at room temperature. Are hydrogenic acceptors in GaAs mostly *ionized* or *neutral* at room temperature? Are hydrogenic acceptors in GaN mostly *ionized* or *neutral* at room temperature?

What are “shallow” dopants and why is it important that dopants are shallow?

Solution:

Given: GaAs: $\epsilon_r = 13.1$; $m_e^* = 0.067 m_0$; $E_d = 6 \text{ meV}$ (Si donor)
 Si: $\epsilon_r = 11.9$; $m_e^* = 0.98 m_0$; $E_d = 54 \text{ meV}$ (As donor)

We use Eq. (15.39) $E_d = [(m_e^* / m_0) / \epsilon_r^2] 13.6 \text{ eV}$ and Eq. (15.37) $a_B^* = [\epsilon_r / (m_e^*/m_0)] 0.53 \text{ \AA}$ and obtain:

For Si in GaAs: Donor ionization energy $E_d = 5.3 \text{ meV}$ Effective Bohr radius $a_B^* = 103.6 \text{ \AA}$
 For As in Si: Donor ionization energy $E_d = 94 \text{ meV}$ Effective Bohr radius $a_B^* = 6.4 \text{ \AA}$

This result shows that for Si in GaAs, the hydrogen model is a good approximation. However the model does not work well for As donors in Si.

The larger mass in Si causes effective Bohr radius to be smaller, which means that the electron is closer to the donor atom core, so that a_B^* is not much larger than lattice constant. Thus hydrogen model does not work well.

Let us compare the donor ionization energy in GaAs and Si with the thermal energy kT at 77 K and 300 K.

At $T = 77$ K, the thermal energy is given by $kT = 6.63$ meV
 At $T = 300$ K, the thermal energy is given by $kT = 25.9$ meV

Comparing the thermal energy with the ionization energies in GaAs and Si, it can be concluded that Si in GaAs will be ionized at 77 K, but that As in Si will not be ionized. At 300 K, Si donors in GaAs will be fully ionized, whereas As donors in Si may not be fully ionized. (This conclusion is solely based on $E_d > kT$. However, due to the high density of states in the Si conduction band, it turns out that most of As donors are in fact ionized.)

Given:	GaAs	$\epsilon_r = 13.1$;	$m_{hh}^* = 0.45 m_0$;	$E_a = 26$ meV	(Be acceptor)
	GaN	$\epsilon_r = 9.0$;	$m_{hh}^* = 0.8 m_0$;	$E_a = 200$ meV	(Mg acceptor)

For acceptor in GaAs

$$E_a = [(m_{hh}^*/m_0)/\epsilon_r^2] 13.6 \text{ eV} = 35.6 \text{ meV}; \quad a_B^* = [\epsilon_r/(m_{hh}^*/m_0)] 0.53 \text{ \AA} = 15.4 \text{ \AA}$$

For acceptor in GaN

$$E_a = [(m_{hh}^*/m_0)/\epsilon_r^2] 13.6 \text{ eV} = 134 \text{ meV}; \quad a_B^* = [\epsilon_r/(m_{hh}^*/m_0)] 0.53 \text{ \AA} = 5.96 \text{ \AA}$$

We see that the theoretical results are close to the experimental values.

At $T = 300$ K, $kT = 25.9$ meV. Thus at 300 K, acceptors in GaAs will be definitely ionized, but acceptors in GaN will not.

“Shallow” dopants are dopants that have a low ionization energy (on the order of kT at 300 K or lower), so that at room temperature all dopants are ionized. This is important, because full dopant activation at room temperature is desired.

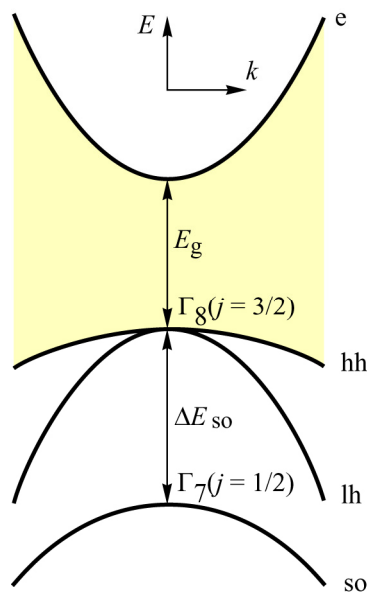


Fig. 15.7. Schematic electron, heavy hole (hh), light hole (lh), and split-off (so) dispersion relations near the center of the first Brillouin zone ($k = 0$).

15.3 Hydrogenic acceptors

The application of the hydrogenic model to acceptors in III–V semiconductors is complicated by their degenerate valence band structure. For hydrogenic donors, the effective electron mass was substituted for the electron mass. The substitution was possible, since the conduction band was assumed to be parabolic, isotropic, and non-degenerate (as for most III–V semiconductors). Such

a simple substitution is not possible for acceptors, since the valence band structure of III–V semiconductors is much more complicated than the conduction band structure. The electronic band structure of several III–V semiconductors with zinblende structure was calculated by Chelikowsky and Cohen (1976). The band structure near the center of the Brillouin zone is schematically shown in **Fig. 15.7**. The highest point of the valence band is located at $k=0$. Without spin-orbit coupling this point would be sixfold degenerate with three dispersion relations and twofold spin degeneracy (Kohn, 1957a). The simplest way to understand this degeneracy is to consider the tight binding limit, in which the wave functions corresponding to the highest point go over into atomic 3p functions. The spin-orbit coupling lifts the degeneracy partially and leads to the situation shown in **Fig. 15.7**. The top of the valence band remains fourfold degenerate at $k=0$. The corresponding dispersion relations are called the heavy hole (hh) and light hole (lh) dispersion relations. The top of the twofold (spin) degenerate split-off (so) band is at $k=0$ at an energy E_{so} below the valence band maximum where E_{so} is the spin-orbit coupling energy. The top of the valence band corresponds to atomic $j=3/2$ states (Kohn, 1957a), where j is the total angular momentum (orbit + spin), *i. e.* $j=l+s=1+1/2=3/2$. The *inner quantum number* j is formally not necessary since it can be expressed by l and s . The introduction of j by Sommerfeld (1920) has historic reasons (Finkelburg, 1958). The top of the valence band has Γ_8 symmetry. The split-off band corresponds to atomic $j=l-s=1-1/2=1/2$ states which have Γ_7 symmetry.

For strong spin-orbit coupling, the split-off band is far removed from the top of the valence band, *i. e.* $\Delta E_{so} \gg E_a$ where E_a is the acceptor binding energy. In this case, both the heavy hole and light hole band must be taken into account and the Hamiltonian is, therefore, a 4×4 matrix (Kohn, 1957a). In the limit of weak spin-orbit coupling, *i. e.* $\Delta E_{so} \approx E_a$, all three valence bands must be taken into account and the Hamiltonian is a 6×6 matrix.

Kohn (1957a) used a 6×6 Hamiltonian matrix to calculate acceptor energies in cubic semiconductors. The author used variational envelope wave functions for acceptors of the form

$$\psi_i(r) = A_i e^{-r/r_i} \quad (15.42)$$

where r_i is a variational parameter. Subsequently, Baldareschi and Lipari (1973) developed a now widely accepted model for shallow acceptor states in cubic semiconductors with degenerate valence bands. In their approach, the Hamiltonian is written as the sum of a spherical term and a cubic correction, thus pointing out the relevance of the spherical symmetry in the acceptor problem and the strong similarity to the case of atoms with spin-orbit interaction. Neglecting the cubic term, Hamiltonians with radial symmetry were obtained. Variational wave functions were used to calculate acceptor ionization energies.

In the limit of strong spin-orbit interaction, that is for spin-orbit splittings (ΔE_{so}) much larger than the acceptor energy (E_a), Baldareschi and Lipari (1973) calculated the effective Bohr radius. For an effective hole mass m_h^* , they obtained the effective Bohr radius

$$a_B^* = \frac{4 \pi \epsilon \hbar^2 \gamma_1}{e^2 m_h^*}, \quad (15.43)$$

the effective Rydberg energy

$$E_{Ryd}^* = \frac{e^4 m_h^*}{2 (4 \pi \epsilon \hbar)^2 \gamma_1}, \quad (15.44)$$

and the acceptor ionization energy

$$E_a = E_{\text{Ryd}}^* f(\mu), \quad (15.45)$$

with

$$\mu = \frac{6\gamma_3 + 4\gamma_2}{5\gamma_1}. \quad (15.46)$$

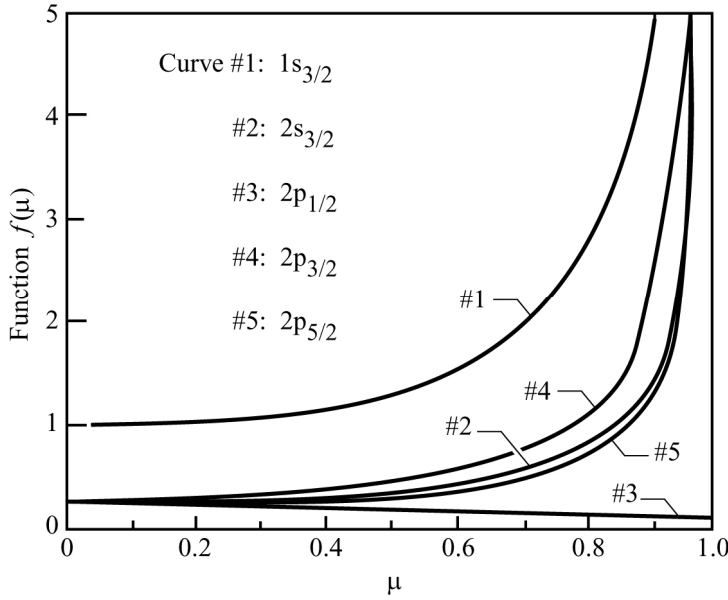


Fig. 15.8. Calculated function $f(\mu)$ versus μ in the limit of strong spin-orbit coupling for the spherical acceptor model (after Baldareschi and Lipari, 1973).

The parameters γ_1 , γ_2 , and γ_3 are the so-called Luttinger parameters which describe the hole dispersion relation near the center of the Brillouin zone (Luttinger, 1956). The function $f(\mu)$ relates the acceptor energy with the effective Rydberg energy (see Eq. (15.45)). The function $f(\mu)$ is shown in **Fig. 15.8**. For the Luttinger parameters of III–V semiconductors, μ assumes values of $\mu \approx 0.6 - 0.9$ and the function $f(\mu)$ assumes values of $f(\mu) \approx 1.5 - 4$ for the ground state energy of acceptors. Baldareschi and Lipari (1973) obtained the numerical values of $f(\mu)$ using a variational approach and using spherical trial functions (Kohn, 1957a) for the acceptor wave functions (see Eq. 15.42). The Luttinger parameters, the variable μ , the effective Rydberg energy, and the ground state ($1s_{3/2}$) and excited state energies of hydrogenic acceptors in several III–V semiconductors were given by Luttinger (1956) and by Baldareschi and Lipari (1973). The calculated and experimental acceptor energies agree reasonably well.

In subsequent work, Baldareschi and Lipari (1974) investigated the contribution of cubic symmetry terms of the Hamiltonian to the spherical model for acceptor states. The effects of the cubic symmetry were studied using perturbation theory which allowed the authors to reproduce all the details of acceptor spectra. The quantitative changes in acceptor energy caused by the cubic term are small, *i. e.* less than 1 meV for the III–V semiconductors.

The **degeneracy** of the acceptor ground states in III–V semiconductors is $g = 4$. Typical acceptor energies are much smaller than the spin-orbit splitting energy, *i. e.* $E_a \ll \Delta E_{\text{so}}$. The top of the valence band is fourfold degenerate due to heavy and light hole dispersion and due to twofold spin degeneracy. Since the acceptor wave functions are composed of valence band wave functions near the top of the band (Kohn, 1957a), the acceptor degeneracy is $g = 4$ as well.

15.4 Central cell corrections

The ionization energy of hydrogenic donors and acceptors as calculated from effective-mass theory does not depend on the chemical nature of the impurity atom. On the other hand, experimental values of the ionization energy do depend on the chemical nature especially for acceptors in III–V semiconductors. The difference in ionization energy between chemically different impurities is attributed to *central cell potentials*. The central cell potential is assumed to be due to the chemical characteristics (*e. g.* electronegativity) of the impurity atom and thus the potential leads to a correction of the hydrogenic ionization energy. This correction is frequently referred to as *chemical shift* (Pantelides, 1975).

The total impurity potential is the sum of the Coulomb potential and the central cell potential V_{cc} and can be written as

$$V(r) = \frac{e}{4\pi\epsilon r} + V_{cc}(r) . \quad (15.47)$$

The central cell potential is a short range potential and has a spatial extent of no more than the unit cell (central cell) of the host semiconductor. *Donor* wave functions are usually quite delocalized in III–V semiconductors. Therefore, the central cell corrections play a minor role for donors and their ionization energy is well described by the hydrogen model. In contrast, the *acceptor* Bohr radius is usually much smaller resulting in a significant central cell correction. This difference between donor and acceptor states is indeed observed experimentally, for example in GaAs. Several model potentials have been used for the central potential including a constant potential extending over the unit cell (Abarenkov and Heine, 1965), and δ -function-like potentials. Various models for central cell potentials were reviewed by Pantelides (1978), Stoneham (1975, 1986), and Altarelli and Bassani (1982). Finally, it is worthwhile to note that *isoelectronic* impurities lack the Coulomb term in Eq. (15.47). For isoelectronic impurities, the central cell potential is the only potential that can bind electrons (Thomas and Hopfield, 1966).

A simple model explaining the chemical shift of impurity ionization energies was proposed by Phillips (1970a, 1970b). The model is based on the local strain around the impurity atom. The strain is caused by the mismatch of the valence bonds of the impurity with valence bonds of the host lattice. Using this model, the chemical trend in donor ionization energies of Te, S, and Se in GaP were qualitatively explained (Phillips, 1970b). Phillips (1973) developed a second model in which the chemical shift in ionization energy is based on the difference in electronegativity, ΔX , between the impurity atom and the host lattice. The author showed that a large difference in electronegativity between impurity and the atom replaced by the impurity results in a large chemical shift. The chemical shift was assumed to be proportional to the heat of formation, *i. e.*

$$\Delta E \propto (\Delta X)^2 \quad (15.48)$$

where ΔE is the difference between the calculated effective mass impurity energy and the actual impurity energy (*i. e.* ΔE is the chemical shift). Using Eq. (15.48), differences in chemical shifts of impurities occupying cation and anion sites in GaP were explained.

Note that chemical shifts are expected to be larger for impurity states with s-type symmetry as compared to states with p-type symmetry. The central cell potential is spatially restricted to the atomic vicinity of the impurity atom. In this region the amplitude of s-type wave functions is large while p-type wave functions have a node. Thus, perturbation theory predicts greater corrections for s-type symmetry states as compared to p-symmetry states.

15.5 Impurities associated with subsidiary minima

The conduction band structure of III–V semiconductors consists of three local minima, which occur at the L , Γ , and X -point of the Brillouin zone. The Γ minimum is located at the center of the Brillouin zone at $k=0$, while the L and X minima occur at finite wave vectors. Donors, which by their very nature are associated with the conduction band, can form donor levels with all local minima of the conduction band. Impurity states associated with subsidiary minima of the conduction band were first analyzed by Bassani *et al.* (1969). The theoretical study revealed that the impurity states are generally formed from Bloch functions of many Brillouin zones and their respective contributions depend on the particular band structure and on the strength and nature of the impurity potential. The existence of several conduction band minima with large effective masses can increase the number of bound states as compared to the single valley hydrogenic model. It was further shown that resonant states in the continuum of one local minimum can be produced by impurity states associated with another minimum. This situation is shown in **Fig. 15.9** which depicts two donor levels associated with two conduction band minima and a resonant state in the continuum of the energetically lower minimum (Altarelli and Bassani, 1982).

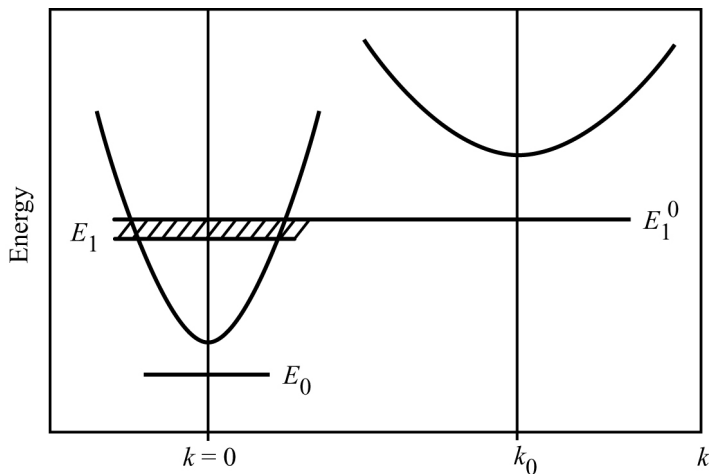


Fig. 15.9. Schematic representation of a bound state E_0 and a resonant state E_1^0 associated with the absolute minimum and a subsidiary minimum of the band structure, respectively. The resonant state E_1 (originating from the unperturbed state E_1) has a finite width due to its degeneracy with band states (after Altarelli and Bassani, 1982).

Donor levels associated with subsidiary conduction band minima were experimentally observed in III–V semiconductors. Adler (1969) used hydrostatic pressure to study the effect of donors on the electron transfer in n-type GaAs. Onton *et al.* (1972) directly observed a subsidiary conduction band minimum and its associated donor levels in InP by optical absorption measurements.