

Chapter 2

Semiconductor Surfaces and Interfaces

The experiments that we will focus on here are using nanoscale devices formed in 2D electron gases in GaAs/AlGaAs heterostructures, consisting of several layers that are brought into contact at interfaces. Surfaces and interfaces therefore play a very important role and are the topic of this chapter.

At the surface of a semiconductor crystal to vacuum, the bands of the solid get related to the vacuum energy level. It takes a finite amount of energy to remove electrons from the crystal to the vacuum, since formation of the crystal from far separated atoms lowered the energy. The *work function* Φ_A is the energy to transfer one electron at the chemical potential from the crystal into the vacuum. In pure semiconductors and in insulators, there are no states at the chemical potential. Hence, the *electron affinity* χ_e is introduced as the energy difference from the bottom of the conduction band to the vacuum energy level. Both the affinity and the work function are dependent on the bulk crystal/material properties.

2.1 Electronic Surface States

The periodic pattern of chemical bonds in the crystal is interrupted at the surface, resulting in unsaturated (dangling) bonds, which can rearrange themselves (surface reconstruction) and/or which might be saturated by a (mono)layer of adatoms (sometimes oxygen). This results in a change of both the surface crystal structure and the allowed energies that depend sensitively on the materials and bulk crystal structures involved. Often, the electronic surface structure has little to do with the bulk structure. The surface states can be probed, for example with scanning tunneling techniques or with photo emission spectroscopy. For simplicity, we consider here the case without surface reconstruction and without passivation, assuming essentially the perfect periodic crystal is simply cut off in a plane. This will give a qualitatively correct picture.

While not giving a derivation here, a simple motivation for formation of electronic surface states goes as follows: Bloch's theorem in principle allows the wave vector in the wave function $\psi_k(r) = e^{ik \cdot r} u_k(r)$ to be a complex number. An imaginary wave vector corresponds to an exponentially damped wave function, an "evanescent wave". In an infinite,

perfect crystal, that is not physical, since it would specify a special location in the crystal, violating periodicity or translational symmetry. But the surface breaks translational symmetry, therefore allowing wave vectors with nonzero imaginary component. It can be shown that the resulting states are localized at the surface, with a wave function amplitude that decays exponentially over a few lattice constants when going from the surface into the bulk. It turns out that the energies of these states are usually located inside the band gap, forming a separate band of surface states (see figure 2.1).

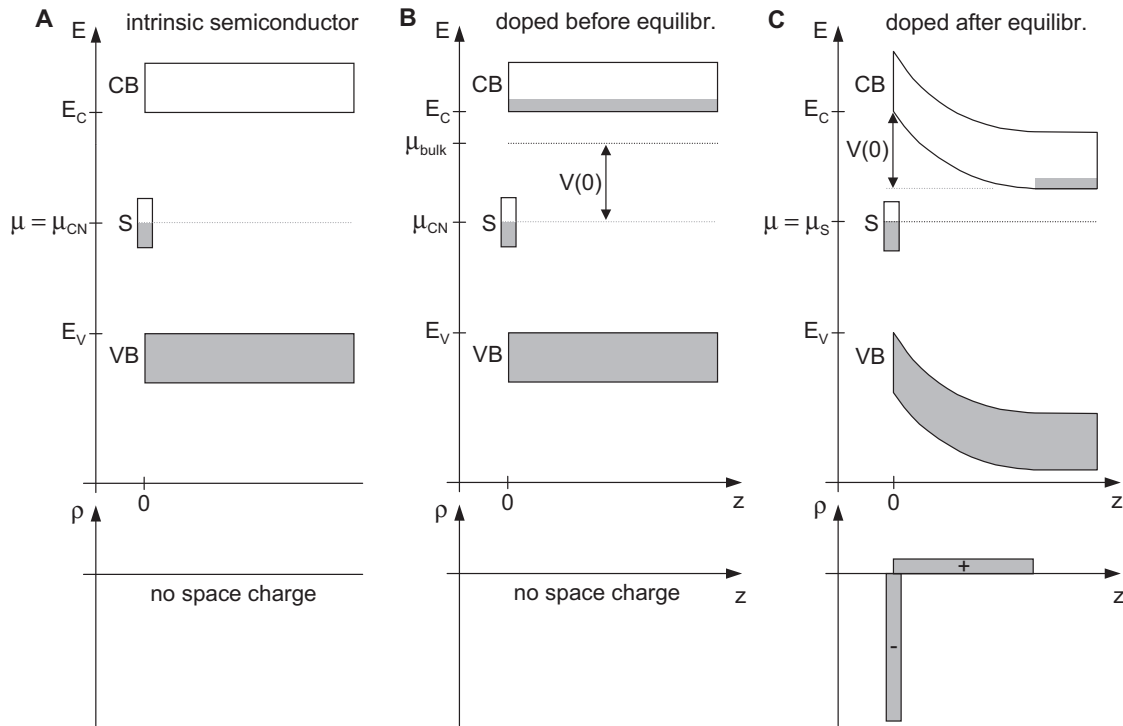


Figure 2.1: Band structure at the surface of an intrinsic semiconductor (A, left), n-doped semiconductor before (B, middle) and after (C, right) equilibration. Gray indicates occupied states. To match chemical potentials at the surface and in the bulk, charge builds up on the surface depleting the donor charges a depth into the bulk and bending the bands accordingly.

The number of surface states per area is essentially given by the number of atoms per area at the surface. These states can be comprised of states that in the infinite, boundary less crystal would have contributed to either the conduction or the valence bands and can be an admixture of both types of bands. Due to charge neutrality, in the case of the intrinsic semiconductor the number of filled surface states is equal to the number of electrons that were removed from the bulk valence band due to surface formation, resulting in a neutral, uncharged surface. The remaining surface states are empty. Filled surface states have electrons that in principle could be given into empty available states and can therefore be considered donor-like states. Vice-versa: empty surface band states are called acceptor-like. The energy up to which the surface states (within the surface band) are filled in the intrinsic, un-doped semiconductor is dictated by charge neutrality and is sometimes referred to as the charge neutrality level or *charge neutrality chemical potential* μ_{CN} .

2.2 The Semiconductor-Vacuum interface

For a surface of a doped semiconductor, electrons occupying the conduction band (originating from the dopants) can lower their energy by filling empty surface band states (if available, which is often the case). This gives a net surface charge n_S (charge per area, in this case negative) which is balanced to maintain charge neutrality by a region inside the crystal of equal total charge of opposite sign (positive)—referred to as *depletion region*. The charge in the depletion region is due to holes left behind from conduction band electrons that dropped into the surface band. The surface charge n_S is located within a few lattice constants around the surface, forming a 2D electron gas. The dopant density is to some extent arbitrary and can be externally controlled. It is often chosen by various other considerations, see the heterostructure section, usually resulting in a dopant density that is much lower than the available surface states, typically of the order of one dopant atom for every 10'000 crystal atoms. Therefore, this results a depletion region that can reach many lattice constants into the bulk.

2.2.1 Band bending

For a bit more quantitative consideration, let z_{dep} denote the extent of the depletion region into the bulk starting from the surface at $z = 0$ (see figure 2.1) and N_D the dopant density (per m^3). All donors are ionized in the depletion region giving a space charge density of $\rho = eN_D$. The Poisson equation for the z -dependence of the potential $V(z)$ within the depletion region $0 \leq z \leq z_{dep}$ is

$$\frac{d^2V}{dz^2} = -\frac{e^2N_D}{\epsilon\epsilon_0} \quad \Rightarrow \quad V(z) = -\frac{e^2N_D}{2\epsilon\epsilon_0}(z - z_{dep})^2, \quad (2.1)$$

where the normalization of V was chosen as $V = 0$ in the bulk (for $z > z_{dep}$) and the constant of integration was chosen to match $V(z_{dep}) = 0$ accordingly. According to Eq. 2.1, the bands are therefore *bending* quadratically, with a total shift of $V(z \rightarrow 0^+) = -e^2N_Dz_{dep}^2/(2\epsilon\epsilon_0)$. More generally, the local curvature of a band is proportional to the local space charge density, according to Eq. 2.1. At the surface $z = 0$, $V(z)$ will jump (over the narrow extent of the surface charge accumulation layer), due to the charge n_S accumulated strongly localized at the surface. This is not indicated in figure 2.1 since it is very narrow compared to the depth z_{dep} of the depletion region. The surface accumulation charge is of equal size but opposite sign as the total depletion region charge: $n_S = -N_Dz_{dep}$ (charge neutrality) and creates a further change in V that we neglect. In this approximation, the bands have bent by a total amount of $\Delta V = V(0)$.

At the surface, the chemical potential is in the surface band (since that band is partly filled for sufficiently low doping density), and the surface band energy and width depends on material properties. In the bulk, the chemical potential is in the gap, usually closer to the conduction band (again, depending of the doping density). However, in equilibrium, the chemical potential has to be the same everywhere, and in particular needs to be the same at the surface and in the bulk. This condition therefore dictates the value of ΔV , i.e. the amount by which the bands need to bend, originating from the material dependent surface properties, and therefore determines the depletion depth z_{dep} .

2.2.2 Fermi Level pinning

To make an example, let's assume that the surface band is centered in the middle of the gap, with a width of a fraction of the gap size, and is half filled in the intrinsic material. Doping will fill the surface band slightly more. Let's take the case of *GaAs* with gap $E_g = 1.4\text{ eV}$, a typical doping of $N_D = 10^{24}\text{ m}^{-3}$ (corresponding to about 1 dopant for every 10^4 crystal atoms, or about 1 dopant for every 20 atoms along a crystal axis). Then, $\Delta V \sim E_g/2$, and with $\epsilon \sim 13$ and using

$$z_{dep} = \sqrt{\frac{2\epsilon\epsilon_0\Delta V}{e^2N_D}} \quad (2.2)$$

it follows that $z_{dep} \sim 30\text{ nm}$, or about 60 lattice constants. This results in a surface charge of $n_S = z_{dep}N_D \sim 3 \times 10^{16}\text{ m}^{-2}$, much smaller than the total surface density of states $\sim 2/a_0^2 \sim 6 \times 10^{18}\text{ m}^{-3}$, with lattice constant $a_0 \sim 5.5\text{ \AA}$. Therefore, the chemical potential μ_S at the surface is essentially independent of the doping level (at least for typical doping densities as above). One says that the Fermi level is *pinned* at the surface. Particularly considering that the surface atoms make up a very small fraction of the total number of atoms in the crystal, the role of the surface is quite important. Also, note that the depletion depth z_{dep} can be changed with the dopant density as formulated in Eq. 2.2.

2.3 Metal-Semiconductor Interface: Schottky Barrier

One important type of interface is the one between a metal and a semiconductor. Among various possible scenarios depending on the respective alignment of the bands and chemical potentials, two important cases are the *Schottky barrier* and *Ohmic contacts*. If a Schottky barrier forms, charge can move from the metal into the semiconductor only by tunneling through a barrier; we first discuss this situation. The most relevant case is the situation when the chemical potential in the conduction band of the metal lies inside the gap of the semiconductor. At the interface, it can be shown that presence of the metal creates *induced gap states* (IGS) within the gap of the semiconductor, as shown in figure 2.2A, displaying the interface before before charge transfer occurred.

Let's start with this situation and let's further assume an intrinsic (undoped) semiconductor. The common energy scale is clearly the vacuum energy level. In general, the chemical potential in the metal μ_M is not aligned with either the surface chemical potential μ_S nor the bulk semiconductor chemical potential μ_{SC} , but often one finds $\mu_M \geq \mu_S$. Now we allow charge transfer. Electrons from the metal can lower their energy by filling empty surface states, thereby leaving behind holes in the metal. A charge dipole located at the interface results and aligns μ_M and μ_S , via the dipole potential obeying the Poisson equation, similar to Eq. 2.1, and corresponding band bending in the semiconductor by an amount $\Delta\mu = \mu_M - \mu_S$. This dipole and band bending is strongly localized at the surface since the surface states decay exponentially over a few lattice constants into the bulk semiconductor. The chemical potential μ_M of the metal is essentially unchanged, since the number of available electrons in the bulk metal is very large compared to the necessary number of electrons to fill the surface states up to μ_M . The resulting situation is depicted in figure 2.2B.

In an intrinsic semiconductor, the entire system is then in a thermodynamic equilibrium, since in the intrinsic case $\mu_{SC} = \mu_S$. In a doped semiconductor, however, there is

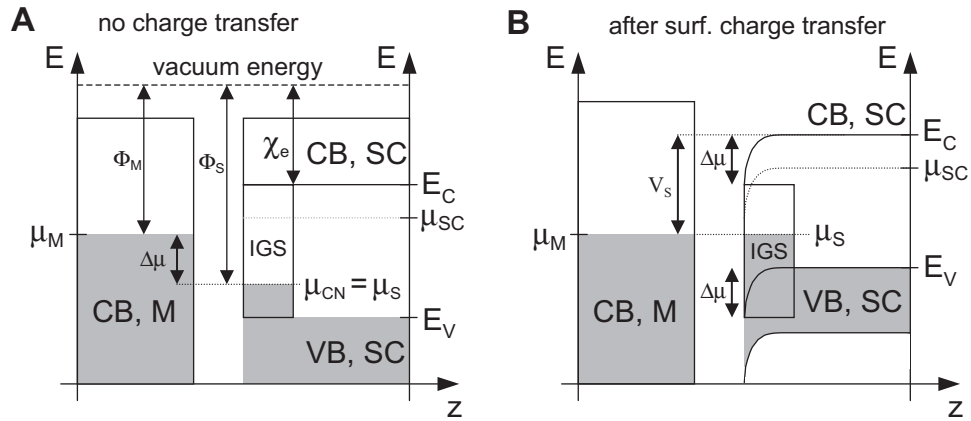


Figure 2.2: Band structure of a metal-semiconductor interface before (left, A) and after (partial) charge transfer (right, B). Shown here is case where the chemical potential of the metal is situated inside the semiconductor gap, which results in a Schottky barrier V_S increased by $\Delta\mu = \mu_M - \mu_S$ due to surface charge accumulation when the bands bend to match $\mu_M = \mu_S$. IGS denotes induced gap states. For an intrinsic semiconductor, $\mu_{SC} = \mu_S$, and B represents thermodynamic equilibrium. For the doped semiconductor $\mu_{SC} \neq \mu_S$ and further charge is transferred, as detailed in Figure 2.3.

still a mismatch $\mu_S \neq \mu_{SC}$. By a further charge transfer of donor electrons onto the surface and corresponding band bending, $\mu_S = \mu_{SC}$ will be achieved, directly analogous to the situation described before for the semiconductor-vacuum interface, resulting in the band diagram in figure 2.3.

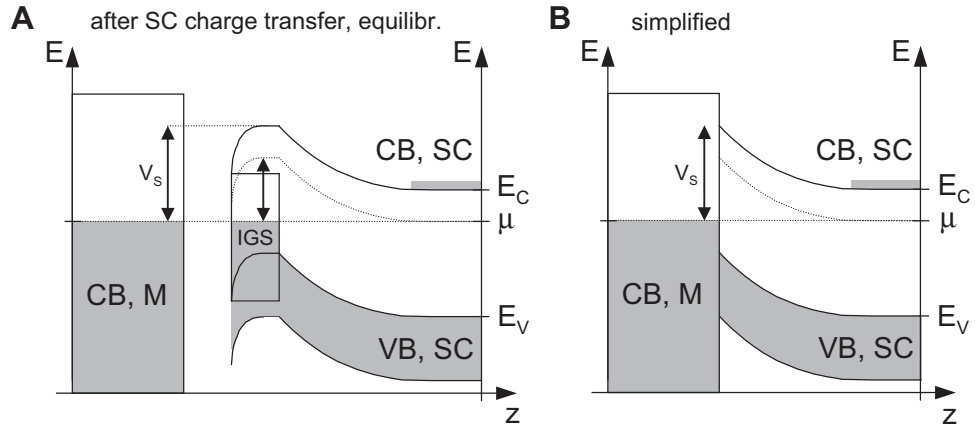


Figure 2.3: Band structure of a metal-semiconductor interface where the semiconductor is doped and the system is in thermodynamic equilibrium (left, A). The right panel shows a simplified view omitting the semiconductor surface. V_S denotes the Schottky barrier.

2.3.1 Schottky Barrier and Model

To transfer electrons from the metal into the *bulk* semiconductor, a barrier of energy $V_S = \Phi_M - \chi_e + \Delta\mu$ —the Schottky barrier—has to be overcome (see figure 2.2 and 2.3), which was increased by $\Delta\mu$ due to the surface charge accumulation. The Schottky barrier height V_S depends on the (often complicated and difficult to calculate) surface properties via the surface chemical potential μ_S in $\Delta\mu = \mu_M - \mu_S$, while the metal work function Φ_M

and the electron affinity in the semiconductor χ_e are bulk parameters. The Schottky model neglects the surface effects and takes $V_S = \Phi_M - \chi_e$. This is equivalent to arguing that removing electrons with affinity χ_e from the semiconductor and adding them to the metal with work function Φ_M will lead to a depletion region in the semiconductor establishing equilibrium and forming a barrier of height $V_S = \Phi_M - \chi_e$. A corresponding band diagram is shown in Figure 2.4 together with some measured barrier heights for metals on GaAs and Si.

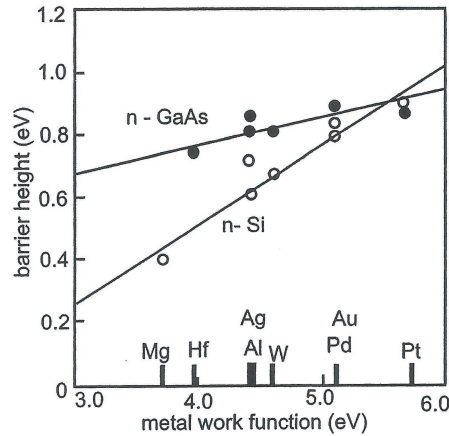


Figure 2.4: Schottky barrier height of Si and GaAs in contact with different metals as a function of the metal work function. From Sze, 1985.

2.3.2 Schottky diode

The metal-semiconductor interface with a Schottky barrier acts as a diode, the Schottky diode, which has been discussed at length in the literature. Suffice it here to state that conduction through this diode depends on the tunneling current through the barrier, which varies exponentially with barrier width and height. Applying a voltage V to the metal with respect to the grounded semiconductor changes the chemical potential of the metal and therefore can be used to control the barrier, resulting in an the exponential I-V curve of a diode. A positive Voltage will reduce the barrier height (and therefore also the barrier width) via $V_S = \Phi_M - eV - \chi_e$. When a Voltage of $V \sim V_S$ is applied, large currents will flow. Typical Schottky barrier heights for metals on GaAs are about 0.8 eV. On the other hand, the currents can be extremely small for negative voltages, corresponding to the diodes reverse bias direction, and the resulting current is called *leakage current*.

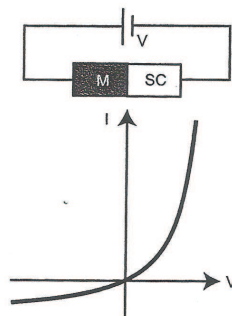


Figure 2.5: I-V characteristic of the Schottky diode. (from Heinzel)

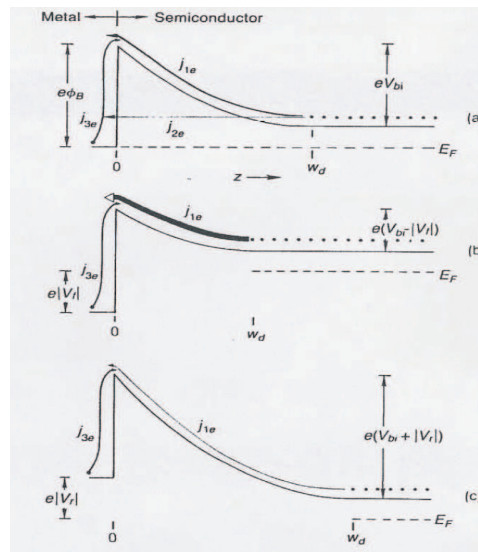


Figure 2.6: left: Schottky barriers under a) no bias, b) forward bias ($V > 0$) and c) reverse bias ($V < 0$). Also indicated are various electron currents components in a Schottky barrier: j_{1e} and j_{3e} are injection currents, which are thermally driven/excited, j_{2e} is the tunneling current. At low temperatures, the termionic currents can be neglected. [Look 1988]