# Chapter 7

# Schottky Diode and Ohmic Contact

version 2.0 (updated: July 31, 2012), ©Prentice Hall

**Outline**

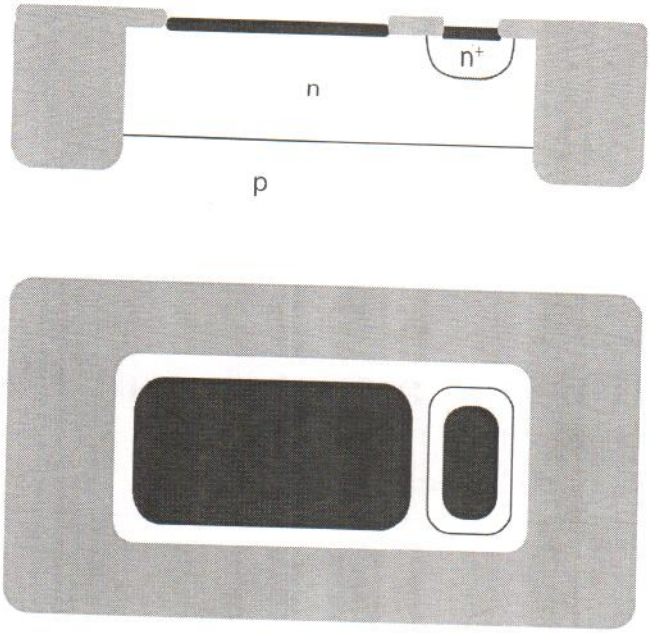



Figure 7.1: Cross section and top view of integrated Schottky diode.

This chapter is about the metal-semiconductor junction and two important applications: Schottky barrier diodes (or simply Schottky diodes) and ohmic contacts. Metal-semiconductor junctions are present in virtually all microelectronic devices. Their most pervasive form is ohmic contacts which are used to provide electrical access to devices from the outside world. However, under the right conditions, metal-semiconductor junctions can display rectifying behavior, just like PN junctions. This is exploited to make Schottky diodes. The uniqueness of the Schottky diode, in comparison with the PN diode, is its fast dynamic response. This arises from the fact that the Schottky diode is largely a majority-carrier-type device. Because of this, Schottky diodes have found use in many analog, digital, power and communications applications.

A top view and a cross-sectional view of an integrated Schottky diode is shown in Fig. 7.1. It consists of an n-type well on a p-type substrate. A metal layer is in direct contact with the n-well. If appropriately designed, this junction exhibits rectifying characteristics. The contact to the body is made through an $n^+$region with a metal contact applied to it. The design here is optimized to make an ohmic contact that exhibits minimum contact resistance. One of the goals of this chapter is to understand when a metal-semiconductor junction shows rectifying characteristics and when it exhibits ohmic behavior. For the diode shown in this figure, the metal is referred to as the *anode* and the n-type semiconductor is the *cathode*. It is also possible to make metal/p-type Schottky diodes. In this case, the metal is the cathode and the p-type semiconductor is the anode.

This chapter studies in detail the physics of the metal-semiconductor junction and its use in Schottky diodes and ohmic contacts. It is organized as follows. We start by defining the notion of an "ideal Schottky diode." This is a hypothetical device of simplified geometry and physics that helps us focus on the most important issues. The metal-semiconductor junction in thermal

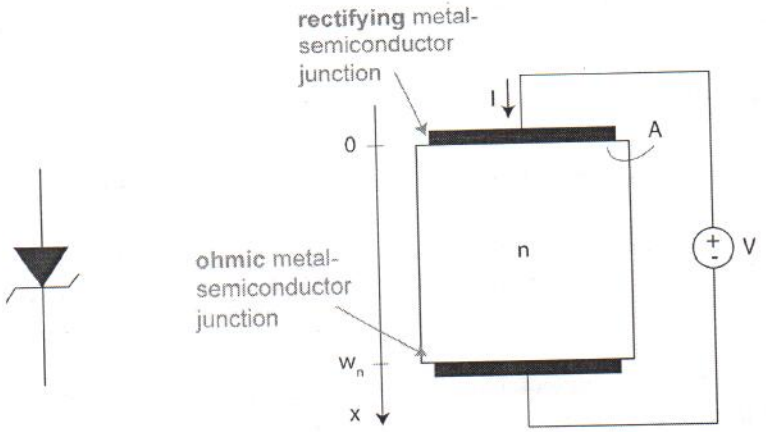



Figure 7.2: Left: circuit symbol of Schottky diode. Right: sketch of ideal Schottky diode.

equilibrium is discussed next. The treatment relies heavily on the energy band diagram view of semiconductors and metals. The consequences of applying a voltage to a Schottky diode are discussed next, first qualitatively and then quantitatively. We study both the current-voltage and the charge-voltage characteristics of Schottky diodes. We then discuss equivalent circuit models for Schottky diodes and some of the most significant non-ideal and second order effects. We devote a few pages to a few important technology and design issues in practical Schottky diodes. The chapter finishes with a detailed discussion of ohmic contacts, their physics and key design issues. A number of Advanced Topics at the end of the chapter allow the student to deepen understanding in several areas.

## 7.1 The ideal Schottky diode

We start this Chapter by defining the concept of an ideal Schottky diode. This is a device with simplified geometry and physics and no parasitics. The ideal Schottky diode captures the essence of the rectifying metal-semiconductor junction and hides some of its complexities. Later on in this chapter, we will relax some of the assumptions that we make here and we will also study the most significant non-idealities.

The circuit symbol and a sketch of an ideal Schottky diode are shown in Fig. 7.2. In this book, the ideal Schottky diode consists of an n-type semiconductor region with a Schottky metal on top (we use this term to refer to the metal of a metal-semiconductor junction that exhibits rectifying characteristics). The semiconductor is contacted by means of an ideal ohmic contact at the bottom. As a short-hand, when we simply talk about "the metal" in the context of a Schottky diode, we refer to the Schottky metal that yields the rectifying characteristics. The metal that makes the ohmic contact is usually referred to as "ohmic metal" or simply "contact."

In the analysis of the ideal Schottky diode, we are going to make the following assumptions:

- All carrier flow is one dimensional. There are no 2D or 3D effects.

- The metal-semiconductor interface is smooth with ideal bonding for the semiconductor atoms preserved.
- The doping level in the semiconductor is uniform throughout.
- We assume that non-degenerate carrier statistics apply in all situations.
- We assume that the Schottky barrier height is given by the difference between the work function of the metal and the electron affinity of the semiconductor (for n-type semiconductor, this is explained below). This assumption is discussed in Sec. AT7.1.
- We assume that the alignment between metal and semiconductor work functions is such that in equilibrium, there is a depletion region on the semiconductor side.
- We treat the metal-semiconductor junction under the depletion approximation. We consider the rest of the semiconductor as quasi-neutral.
- We disregard the minority carriers and we therefore neglect generation and recombination.
- We ignore any resistance effects associated with the Schottky metal, the semiconductor or the ohmic contact (we study the impact of parasitic resistance in Sec. 7.6.1).
- We assume ideal ohmic contact as defined in Sec. 5.2.2.
- We ignore any edge effects or other effects associated with the sidewalls of the device.

Fig. 7.2 defines the axis that we will use in our analysis of the Schottky diode. We place its origin at the metallurgical interface between the metal and the semiconductor. The extent of the semiconductor is $w_n$. The area of the device is $A$.

For a metal/n-semiconductor junction, such as the one depicted in Fig. 7.2, the voltage polarity is usually defined as in the figure. With this polarity, when $V > 0$, the Schottky diode is in forward bias. When $V < 0$, the Schottky diode is in reverse bias. The metal is the anode and the semiconductor is the cathode. Everything is reversed for a metal/p-semiconductor junction.

## 7.2 Ideal Schottky diode in thermal equilibrium

A metal-semiconductor junction is an artificial structure made out of two rather dissimilar materials. Whenever two materials with different properties are brought together in intimate contact, charge redistribution takes place. This has often fascinating consequences. The PN junction that was studied in the previous chapter was an example of this. The metal-semiconductor junction is unique in a number of ways. Before studying in detail the metal-semiconductor junction, a number of important concepts can be better grasped in a simpler metal-metal junction.

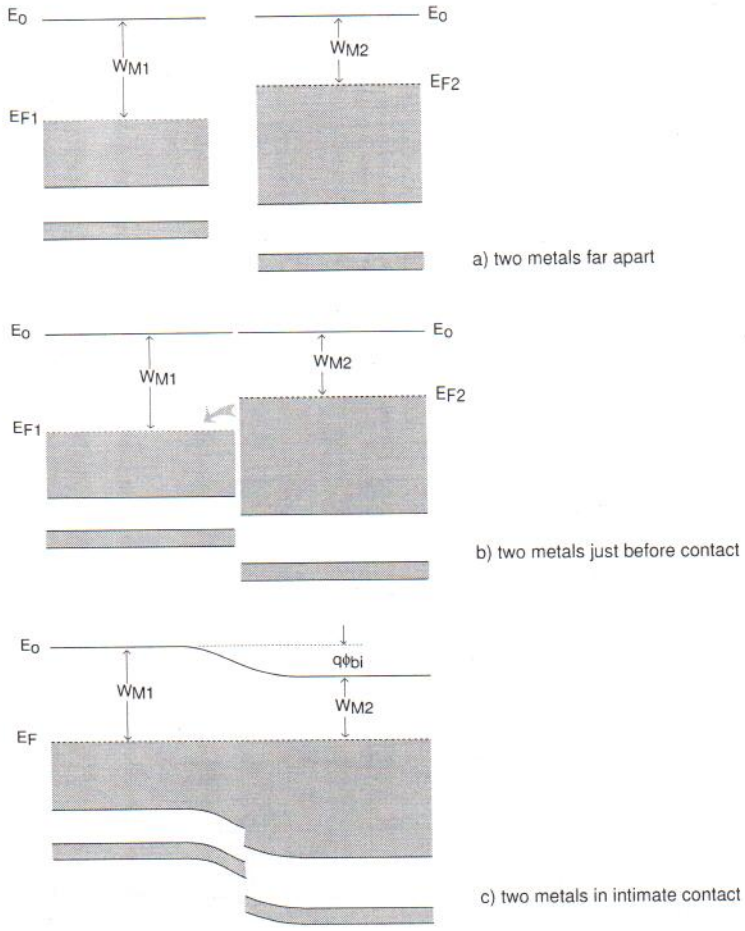



Figure 7.3: Energy band diagrams illustrating two different metals with different work functions: a) far apart from each other, b) just before contact is established, c) in intimate contact.

### 7.2.1 A simpler system: a metal-metal junction

Fig. 7.3 illustrates the evolution of the electronic structure of two different metals as they are brought in contact to form a metal-metal junction. Fig. 7.3a) shows the energy band diagram of each metal when placed far apart from each other. Each metal has its own peculiar energy distribution of bands and bandgaps. The work functions (the minimum energy required to free up an electron) is also different in the two materials.

As the two metals are brought together, some electrons in the low work function metal, metal 2 on the right of Fig. 7.3, are presented with empty states at lower energy in the high work function metal 1 (Fig. 7.3b). As soon as contact is established, electrons from metal 2 rush to metal 1 to lower their energy. This makes metal 1 negatively charged and metal 2 positively charged. As time goes on and electron flow continues, an electric field builds up across the interface that in due time stops further electron migration. Equilibrium is eventually reached.

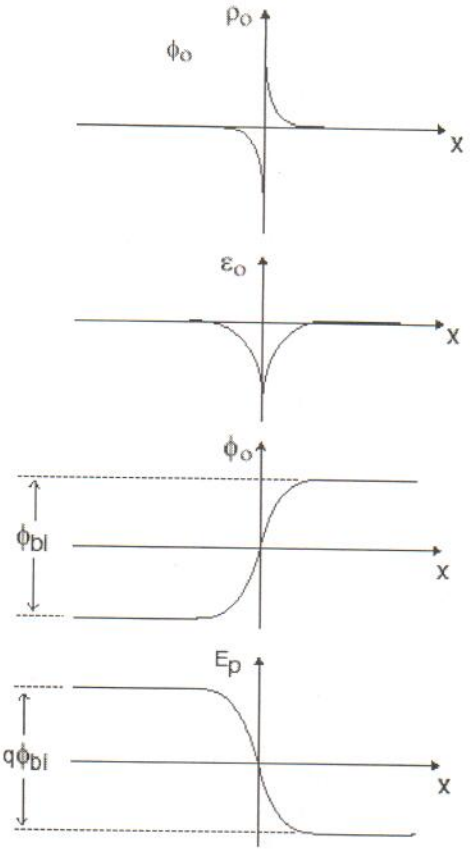



Figure 7.4: Volume charge density, electric field, electrostatic potential, and potential energy corresponding to the metal-metal junction of Fig. 7.3 in equilibrium.

This is a dynamic equilibrium situation. The initial tendency of electrons to flow from metal 2 to metal 1 in order to lower their energy is precisely counterbalanced by the electric field that gets set up at the interface. This electric field originates on the *dipole of charge* that is formed as a consequence of electron migration. This is graphed in Fig. 7.4. The electric field has a sign such that it opposes further electron transfer from metal 2 to metal 1. In Fig. 7.4, $\mathcal{E}$ is negative as a consequence of the choice of axis. The electric field in turn results in a potential difference between the two materials that is called the *built-in potential*, $\phi_{bi}$. The potential energy associated with the electric field must be added to the energy band diagram. As a consequence, all the bands bend with a shape that is identical to the electrostatic potential with a minus sign (Fig. 7.3c).

When the two metals are sufficiently far apart, they can be considered as two different electronic systems with their own separate Fermi levels. The moment the two materials are brought in intimate contact and electrons are allowed to freely flow from one to the other, they constitute a new unified electronic system. In consequence, when thermal equilibrium is established, the Fermi level must be flat everywhere. This is the situation depicted in Fig. 7.3c).

At this point it is legitimate to ask: what happens to the work function? How much energy does it take to extract an electron from the new combined electronic system? The best way to answer these questions is to carry out a fictitious photoelectric effect experiment, as sketched in Fig. 7.5. Suppose we have a tunable light source that can emit a spatially narrow beam of photons of any arbitrary energy. Suppose also that we have a way to detect when electrons are extracted from any one of the metals. If the beam of light has a narrow spatial spread, we can measure the *local work function* by shining light at a spot of the surface of the metal and observing the light energy at which electrons start escaping from that location.

When the two metals are apart, the energy required to free up an electron is well defined, *i.e.*, $W_{M1}$ for metal 1 and $W_{M2}$ for metal 2. When the metals are in contact the situation is not so clear. Sufficiently far away from the interface, each metal should not be aware of the presence of the other. Since the nature of each one has not been modified by bringing them into contact, it is then reasonable that the threshold light energy that frees up an electron remains unchanged. Far away from the interface, then, the work function of metal 1 is still $W_{M1}$ and for metal 2 it is $W_{M2}$. As we approach the interface, electrons have migrated from metal 2 to metal 1. In consequence, the threshold photon-energy of metal 2 slowly *increases* as we approach metal 1, that is, we have to "dig deeper" into the electronic structure of metal 2 to find electrons that can be freed up. When we cross the interface into metal 1, previously empty electronic states are now full. As a result, the energy required to free up some of these electrons will be smaller than $W_{M1}$. As we go further and further away from the interface, the local work function approaches $W_{M1}$. This is sketched in Fig. 7.5.

We can then define, without any ambiguity, a *local work function*. We find that this parameter changes smoothly from one metal to the other and sufficiently far away from the interface it recovers the value that the metal exhibits when in isolation. It is clear that this local work function cannot change abruptly in space. If it did, electrons at the location where the discontinuity exists could move a small distance and easily lower their energy. This would not be an equilibrium situation.

Through the local work function, we can also define a *local vacuum energy*, $E_o$ as in in Fig. 7.3. At any point in space, $E_o$ is located at an energy above the Fermi level equal to the local work function. The shape of the local vacuum energy is identical to the potential energy of Fig. 7.4 and the total difference in energy of the local vacuum level across the structure is $W_{M1} - W_{M2}$. This allows us to conclude that the built-in potential of this structure is $\phi_{bi} = (W_{M1} - W_{M2})/q$.

A rigorous solution of the electrostatics of a metal-metal junction in thermal equilibrium, although not being very difficult to carry out, is of no great use to us. The one aspect of it that is important is the length scale of the space charge regions at the interface between the two metals. Because of the high electron density of a metal (about $\sim 10^{22}\ cm^{-3}$), a metal does not tolerate net charge in its bulk. If there is no overall charge neutrality in a metal, its net charge is confined to a thin sheet at the surface. In our case, this is the metal/metal interface. The space charge region at the interface is only a fraction of a *nm* thick, much smaller than all scale lengths of interest in microelectronic devices. For all purposes, it is then safe to consider this as a delta function and that is what we will do in this book.

We can now turn to the metal-semiconductor junction, the main topic of this chapter.

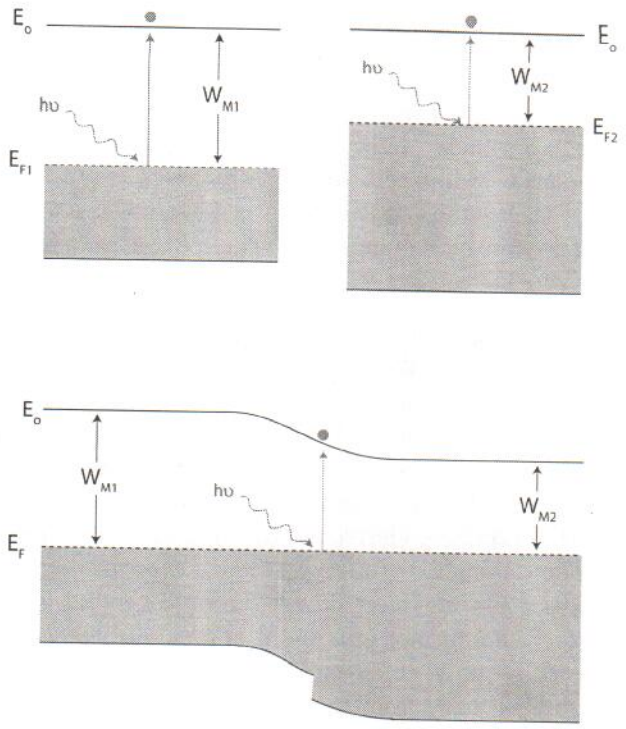



Figure 7.5: The photoelectric effect can be used to define a local work function in a metal-metal junction. Top: with the metals isolated, the threshold energy for extracting an electron is equal to the regular work functions $W_{M1}$ and $W_{M2}$. Bottom: in a metal-metal junction, the local work function changes smoothly across the interface from $W_{M1}$ to $W_{M2}$.

### 7.2.2 Energy band lineup of metal-semiconductor junction

A metal-semiconductor junction shares a lot of similarities with the metal-metal junction. While a metal is characterized by a partially full band, the most prominent feature of the energy band structure of a semiconductor is a complete band separated by a bandgap from the next empty band. Depending on the doping concentration and type of the semiconductor, the Fermi level can be located just about anywhere in the bandgap; it can even penetrate to some extent into the conduction or valence bands at high doping levels. In consequence, the detailed charge redistribution that takes place between a metal and a semiconductor when they are brought in intimate contact depends on the doping type and doping level of the semiconductor. So will the electrical properties of the resulting metal-semiconductor junctions.

Fig. 7.6 shows the energy band diagram of a typical metal and a non-degenerate n-type semiconductor. As in a metal, a semiconductor is characterized by a work function $W_S$. In a semiconductor, the work function is defined as the energy difference between the vacuum level and the Fermi level in equilibrium. Because of this, $W_S$ is a function of the doping level. In order to completely specify the relationship between the work function and the doping level in a semiconductor, an additional parameter is required. This is the *electron affinity*, $\chi$, which measures the energy difference between the vacuum level and the edge of the conduction band. $\chi$

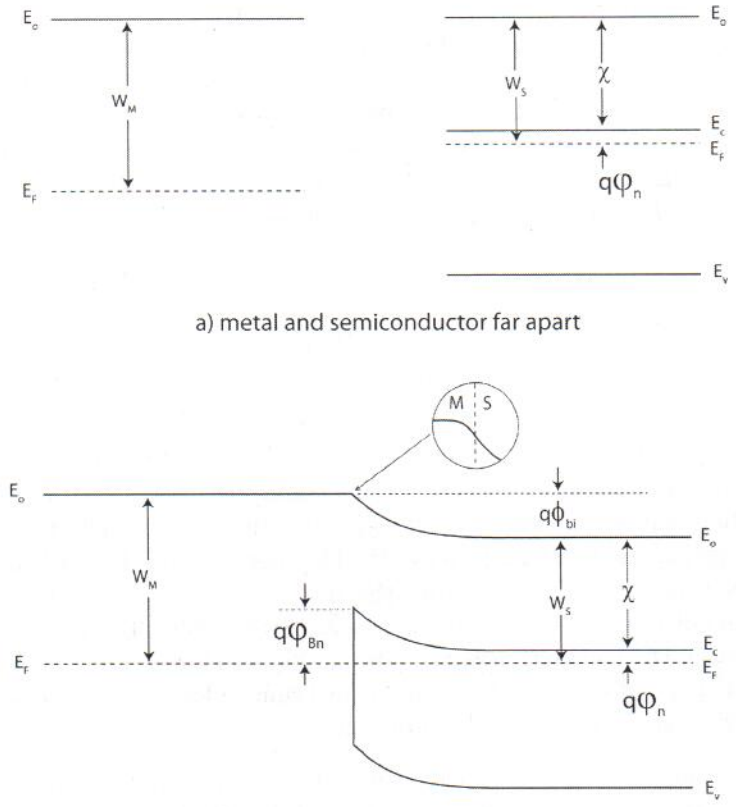



Figure 7.6: Energy band diagrams illustrating a metal and an n-type semiconductor: a) far apart from each other, b) in intimate contact.

is a property of the host semiconductor. For Si, for example, $\chi = 4.04\ eV$ at room temperature. The relationship between the electron affinity and the semiconductor work function is given by:

$$W_S = \chi + q\varphi_n \tag{7.1}$$

where $q\varphi_n$ is the energy difference between $E_c$ and $E_F$ and is calculated from the doping level as discussed in Ch. 2.

Figure 7.6 illustrates a case in which the work function of the metal is higher than the work function of the semiconductor. This is the most typical situation for Si. In this instance, when the metal and the semiconductor are brought together, the difference in work functions results in electron redistribution in which electrons preferentially flow from the semiconductor to the metal. As in the metal-metal junction, this creates a charge dipole at the interface of the two materials. The semiconductor side becomes positively charged as a consequence of the exposed ionized donors, while the metal gets flooded by additional electrons and becomes negatively charged. The resulting energy band diagram in equilibrium is shown at the bottom of Fig. 7.6.

This energy band diagram is qualitatively identical to the one drawn in Fig. 7.3 for the metal-metal junction. Fundamental parameters of the semiconductor, such as $\chi$ and $E_g$, are not affected by the presence of the metal. Additionally, sufficiently far away from the interface, the semiconductor and the metal are not upset by the presence of the junction. In consequence, their properties must remain the same as when they were isolated. For the semiconductor, that means that far away on the right, $E_F$ is located at a distance $W_S$ below the vacuum level and the energy difference between $E_F$ and $E_c$ is unchanged from the isolated case. In consequence, the built-in potential of this junction is given by:

$$\phi_{bi} = \frac{1}{q}(W_M - W_S) \tag{7.2}$$

What is substantially different in the case of the metal-semiconductor junction with respect to the metal-metal one is the relative band bending in the metal and the semiconductor. Since a metal has a carrier concentration several orders of magnitude higher than a semiconductor, the potential distribution across a metal-semiconductor junction is extremely asymmetric, with the semiconductor absorbing the great majority of it. The metal energy bands bend a negligible fraction of $q\phi_{bi}$ over a very small distance from the interface. In contrast, the bands in the semiconductor bend up substantially over a much longer length scale. The closer we get to the interface, the further away the conduction band is from the Fermi level and the fewer electrons there are. The detailed shape of the band bending in the semiconductor will be calculated in the following section. We will see that it is nearly parabolic.

At the metal-semiconductor interface, an electron with an energy equal to the Fermi level has an abrupt barrier to overcome in order to get from the metal to the semiconductor. This energy barrier plays a crucial role in the operation of metal-semiconductor junctions out of equilibrium and is called the *Schottky barrier height*. From Fig. 7.6, it is easy to see that:

$$q\varphi_{Bn} = W_M - \chi \tag{7.3}$$

This equation says that the Schottky barrier height is a property of the metal-semiconductor pair and it is not a function of the doping level of the semiconductor. This is often referred to as the *Schottky-Mott relation.*

Combining Eqs. 7.3, 7.1 and 7.2, we obtain an expression for the built-in potential of a Schottky diode in terms of the Schottky barrier height:

$$q\phi_{bi} = q\varphi_{Bn} - q\varphi_n \tag{7.4}$$

Fig. 7.7 shows the energy band diagram for a typical metal/p-semiconductor junction. With the two materials far apart, the Fermi level in the metal is higher than the Fermi level in the semiconductor. In this instance, as we bring the materials together, electrons flow from the metal to the semiconductor. On the metal side, this leads to a surface with a deficiency of electrons, that is, positively charged. Since the semiconductor is p-type, metal electrons rush to the semiconductor and preferentially occupy the holes at the top of the valence band. In

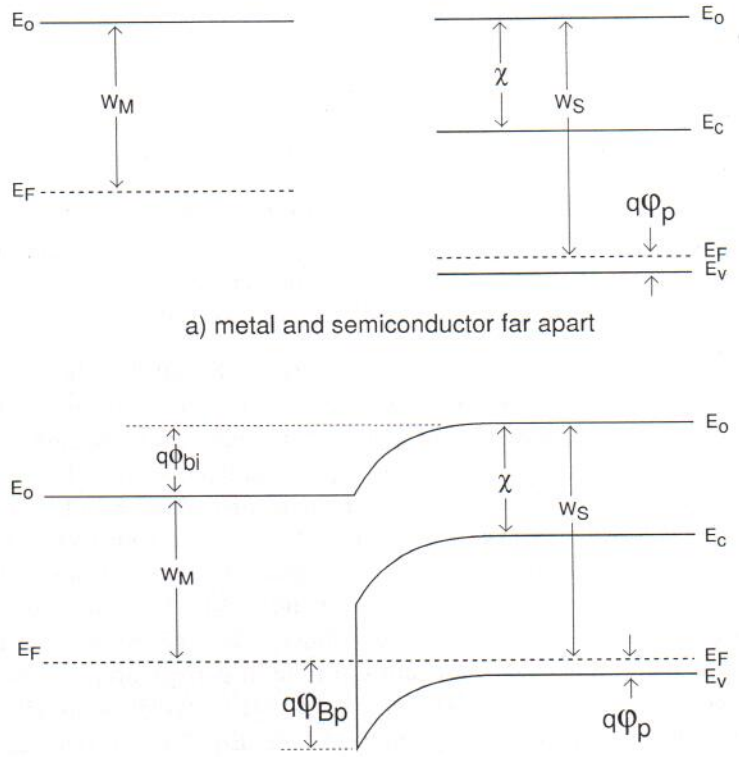



Figure 7.7: Energy band diagrams illustrating a metal and an p-type semiconductor: a) far apart from each other, b) in intimate contact.

consequence, there are fewer holes there when equilibrium is established and the semiconductor becomes negatively charged.

In thermal equilibrium, the bands in the semiconductor bend down reflecting the reduction of holes on the semiconductor side as the interface is approached. There is also an energy barrier that appears at the metal-semiconductor interface. It is customary to refer to it as the Schottky barrier height for holes $q\varphi_{Bp}$ (always a positive quantity). From the figure, we can see that $q\varphi_{Bp}$ is given by:

$$q\varphi_{Bp} = \chi + E_g - W_M \tag{7.5}$$

This is also independent of doping. In terms of $q\varphi_{Bp}$, $\phi_{bi}$ is:

$$q\phi_{bi} = q\varphi_{Bp} - q\varphi_p \tag{7.6}$$

where $q\varphi_p$ is the energy distance between $E_F$ and $E_v$.

An interesting property of a metal-semiconductor pair is obtained by adding Eqs. 7.3 and 7.5:

$$q\varphi_{Bn} + q\varphi_{Bp} = E_g \tag{7.7}$$

The Schottky barrier heights of a metal on a semiconductor when p-type doped and n-type doped add up to the bandgap of the semiconductor. This is a handy relationship that can be exploited in practice if we are missing the Schottky barrier height for a certain metal-semiconductor pair but we have its value for the same system with the contrary polarity of the semiconductor.

Before we move on, it is important to note that Eqs. 7.3 and 7.5 should not be used to estimate the Schottky barrier height of a metal/semiconductor pair. These equation suggest that the Schottky barrier height is a fundamental property of the bulk materials involved. In practice, it is found that for a given metal-semiconductor pair, $q\varphi_{Bn}$ and $q\varphi_{Bp}$ depend on the details of the fabrication process and the crystalline orientation of the semiconductor. This strongly suggests that the Schottky barrier height is affected to some extent by the interfacial chemistry and is not just a property of the bulk materials. For many metals, the predictions of Eq. 7.3 can be significantly off. More details of this are given in Appendix AT7.1. The practical approach is not to rely on Eq. 7.3 or 7.5 but rather design experiments to measure the actual value of the Schottky barrier height of interest. Eqs. 7.4 and 7.6 remain correct and can be safely used to relate the built-in potential, the Schottky barrier height and the relative location of the Fermi level in the semiconductor. As is shown in Appendix AT7.1, Eq. 7.7 also remains correct.

In the next section, we study the electrostatics of the metal-semiconductor junction in thermal equilibrium in a more quantitative way.

### 7.2.3 Electrostatics of metal-semiconductor junction in equilibrium

The goal of this section is to develop a first-order model for the electrostatics of a metal/semiconductor junction in thermal equilibrium, that is, to calculate the volume charge density, electric field, electrostatic potential and carrier concentration distributions in space across the structure. This understanding is essential to deal with situations out of equilibrium later on in this chapter. This section treats the metal/n-semiconductor junction. The procedure is similar to the metal/p-type semiconductor junction and leads to very similar equations.

The problem to be solved is a typical one in semiconductor devices and it resembles the PN diode. In most metal-semiconductor junctions, the net volume charge density in the semiconductor in the vicinity of the metal is high enough that it cannot be considered quasi-neutral any longer. Because of this, the region close to the metal-semiconductor interface is called the *space-charge region*, or SCR. Unlike the relatively gradual non-uniformly doped distributions that we studied in Ch. 4, a space charge region many times appears in regions in which abrupt transitions occur. It is the case, for example, in a PN junction and also that of a metal-oxide-semiconductor structure, as we will describe later on in this book. Sufficiently far away from the metal-semiconductor interface, in the bulk of the semiconductor, charge neutrality should prevail. This region is called the *quasi-neutral region* or QNR. The boundary between the SCR and the

QNR is fairly sharp.

The first step in thinking about the electrostatics of this problem is to consider the volume charge density, $\rho_o$. In general, for a metal/n-semiconductor junction, and in the absence of a substantial concentration of compensating acceptors, $\rho_o$ is given by:

$$\rho_o = q(p_o - n_o + N_D) \tag{7.8}$$

Deep in the quasi-neutral bulk of the semiconductor $\rho_o \simeq 0$. As we advance towards the metal-semiconductor interface, $n_o$ drops and $p_o$ consequently rises (the $n_o p_o$ product must remain constant in equilibrium). This makes $\rho_o$ positive. Since $p_o$ starts from an extremely small value in the QNR, the drop in $n_o$ is most significant. If the built-in potential of the junction is not too small, as we get closer to the interface, we will reach a point in which $n_o$ has become negligible in front of $N_D$. Similarly, if $\phi_{bi}$ is not too high either, $p_o$ might never become significant next to $N_D$. When all this happens, $\rho_o \simeq qN_D$. Because of the exponential dependence of $n_o$ on $\phi_o$, the transition between $\rho_o \simeq 0$ and $\rho_o \simeq qN_D$ is actually quite sharp. This suggests a simple approximation to the charge distribution in the semiconductor:

$$\rho_o(x) \simeq qN_D \qquad \text{in SCR: } 0 \leq x < x_d \tag{7.9}$$

$$\rho_o(x) \simeq 0 \qquad \text{in QNR: } x_d < x \tag{7.10}$$

This assumption of a box-like shape for the charge distribution is the *depletion approximation*, something that we have already become familiar with in the analysis of the PN diode. Establishing the location of the boundary between the SCR and the QNR, $x_d$ in Eqs. 7.9 and 7.10, is one of the goals of the calculations that follow. This charge distribution is shown in Fig. 7.8.

Integration of this volume space-charge density distribution yields the electric field profile:

$$\mathcal{E}_o(x) \simeq \frac{qN_D}{\epsilon}(x - x_d) \qquad \text{in SCR: } 0 \leq x \leq x_d \tag{7.11}$$

$$\mathcal{E}_o(x) \simeq 0 \qquad \text{in QNR: } x_d \leq x \tag{7.12}$$

This field is negative because it points from the semiconductor into the metal, contrary to our choice of axis, as shown in Fig. 7.8.

One more integration yields the electrostatic potential:

$$\phi_o(x) = -\frac{qN_D}{2\epsilon}(x^2 - 2x_d x + x_d^2) \qquad \text{in SCR: } 0 \leq x \leq x_d \tag{7.13}$$

$$\phi_o(x) = 0 \qquad \text{in QNR: } x_d \leq x \tag{7.14}$$

where we have selected $\phi_o(bulk) = 0$ as the reference for potentials. The potential is also graphed in Fig. 7.8.

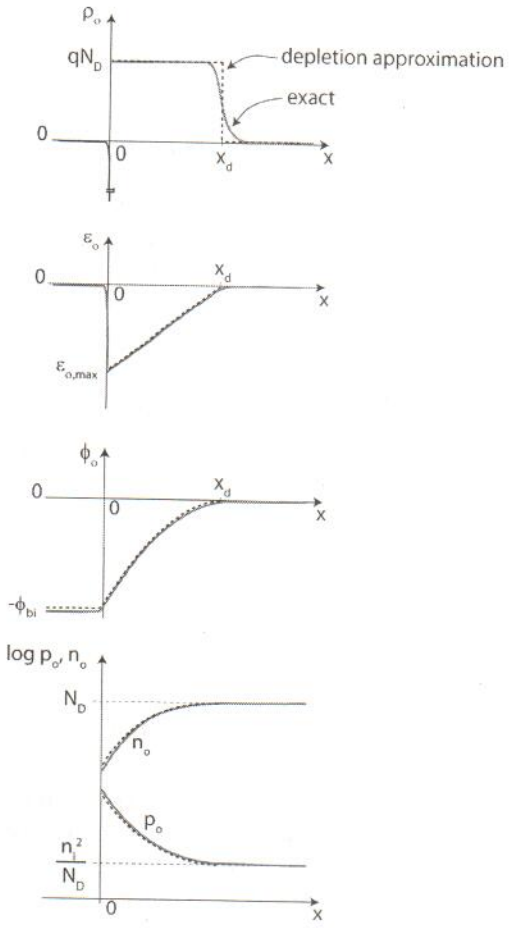



Figure 7.8: Volume charge density, electric field, electrostatic potential, and equilibrium carrier concentrations across metal/n-semiconductor junction.

The extension of the depletion region, $x_d$, is obtained by demanding that the total potential difference across the semiconductor be $\phi_{bi}$ which we know from energy arguments (Eq. 7.4). With our selected potential reference, that implies that $\phi_o(0) = -\phi_{bi}$. Solving for $x_d$ in Eq. 7.13, we get:

$$x_d = \sqrt{\frac{2\epsilon\phi_{bi}}{qN_D}} \tag{7.15}$$

The maximum electric field occurs at the interface and it is given by:

$$|\mathcal{E}_{o,max}| = \frac{qN_D x_d}{\epsilon} = \sqrt{\frac{2qN_D\phi_{bi}}{\epsilon}} \tag{7.16}$$

These two equation state that the higher the doping level, the thinner the depletion region is and the higher the electric field is at the metal/semiconductor interface. This is a consequence of the electrostatics of a charge dipole. To attain a certain potential build-up, a more spatially compact charge dipole requires a higher charge and results in a bigger field inside the dipole. This is what happens in the depletion region of a metal-semiconductor junction when the doping level in the semiconductor increases.

It is interesting to realize that Eqs. 7.15 and 7.16 are identical to that of a highly asymmetric $P^+N$ junction (Eqs. 6.21 and 6.22, respectively). This makes sense as in that case, we learned that the electrostatic potential drops entirely in the lowly doped n-type semiconductor. That is also the case in the metal-semiconductor junction studied here.

**Exercise 7.1**: *Calculate the built-in potential, the depletion-region thickness and the maximum electric field at 300 K of an Al-Si junction with $q\varphi_{Bn} = 0.68\ eV$ in thermal equilibrium. The doping level of the Si is $N_D = 10^{17}\ cm^{-3}$.*

For $N_D = 10^{17}\ cm^{-3}$ Si at 300 $K$, the distance between the the conduction band and the Fermi level is $q\varphi_n = 0.15\ eV$. Using Eq. 7.4, $\phi_{bi}$ is:

$$\phi_{bi} = \varphi_{Bn} - \varphi_n = 0.53\ V$$

The depletion region thickness can be obtained from Eq. 7.15:

$$x_d = \sqrt{\frac{2\epsilon\phi_{bi}}{qN_D}} = \sqrt{\frac{2 \times 1.04 \times 10^{-12}\ F/cm \times 0.53\ V}{1.6 \times 10^{-19}\ C \times 10^{17}\ cm^{-3}}} = 8.3 \times 10^{-6}\ cm = 83\ nm$$

The maximum electric field occurs at the metal-semiconductor interface. From Eq. 7.16, it is given by:

$$|\mathcal{E}_{o,max}| = \frac{qN_D x_d}{\epsilon} = \frac{1.6 \times 10^{-19}\ C \times 10^{17}\ cm^{-3} \times 8.3 \times 10^{-6}\ cm}{1.04 \times 10^{-12}\ F/cm} = 1.4 \times 10^5\ V/cm$$

With the electrostatic potential now completely determined, we are in a position to calculate the equilibrium carrier concentrations and verify our initial assumptions. For this, we use the Boltzmann relations, which for our choice of potential reference are written as:

$$n_o(x) = N_D \exp\frac{q\phi_o(x)}{kT} \tag{7.17}$$

$$p_o(x) = \frac{n_i^2}{N_D} \exp\frac{-q\phi_o(x)}{kT} \tag{7.18}$$

with $\phi_o(x)$ given by Eqs. 7.13-7.14. $n_o(x)$ and $p_o(x)$ are also sketched in Fig. 7.8.

We can now verify the assumptions that were made in formulating the depletion approximation. First, in going from 7.8 to 7.9, we neglected the equilibrium hole concentration. The location where $p_o$ is highest, as can be seen in Fig. 7.8, is at the metal-semiconductor interface. At that point,

$$p_o(0) = \frac{n_i^2}{N_D} \exp \frac{q\phi_{bi}}{kT} = N_v \exp \frac{-(E_g - q\varphi_{Bn})}{kT} \tag{7.19}$$

where we have also used Eqs. 7.4 and 2.39. This equation implies that for $p_o(0)$ to be negligible next to $N_D$, $q\phi_{Bn}$ must not get too close to $E_g$. As the exercise below shows, for typical doping levels and common metals, this condition is readily satisfied.

We also neglected $n_o$ next to $N_D$ everywhere in the depletion region. We must make sure that at the interface, $n_o$ is sufficiently small for this assumption to hold everywhere else. At $x = 0$, we have:

$$n_o(0) = N_D \exp \frac{-q\phi_{bi}}{kT} = N_c \exp \frac{-q\varphi_{Bn}}{kT} \tag{7.20}$$

where we have also used Eqs. 7.4 and 2.39. This condition is satisfied if $\phi_{Bn}$ is at least several $kT/q$'s.

**Exercise 7.2**: *Specify the range of values of Schottky barrier height that at 300 K allow the use of the depletion approximation in equilibrium for metal-semiconductor junctions built on* $N_D = 10^{17}\ cm^{-3}$ *Si.*

The upper limit of $q\varphi_{Bn}$ is set by the maximum tolerable $p_o$ at the metal-semiconductor interface. This is given by Eq. 7.19:

$$p_o(0) = N_v \exp \frac{-(E_g - q\varphi_{Bn})}{kT} \ll N_D$$

Solving for $q\varphi_{Bn}$, we get:

$$q\varphi_{Bn} \ll E_g - kT \ln \frac{N_v}{N_D}$$

For $N_D = 10^{17}\ cm^{-3}$ at 300 $K$, this demands that $q\varphi_{Bn}$ be a few $kT$'s smaller than 0.98 $eV$.

The lower limit of $q\varphi_{Bn}$ is set by the maximum tolerable $n_o$ at $x = 0$ given by Eq. 7.20:

$$n_o(0) = N_c \exp \frac{-q\varphi_{Bn}}{kT} \ll N_D$$

For $N_D = 10^{17}\ cm^{-3}$ at 300 $K$, this demands that $q\varphi_{Bn}$ be at least several $kT$'s over 0.15 $eV$.

Since most metals on Si fall within these two limits, the depletion approximation is of very wide applicability.

## 7.3 Current-voltage characteristics of ideal Schottky diode

When a voltage is applied to the two terminals of a metal-semiconductor junction, current flows. As in the case of the PN diode, the current through a Schottky diode exhibits rectifying behavior. Before we can understand the origin of this, we must study the modifications that the electrostatics of the junction undergo upon the application of a voltage. This is the topic of the next subsection.

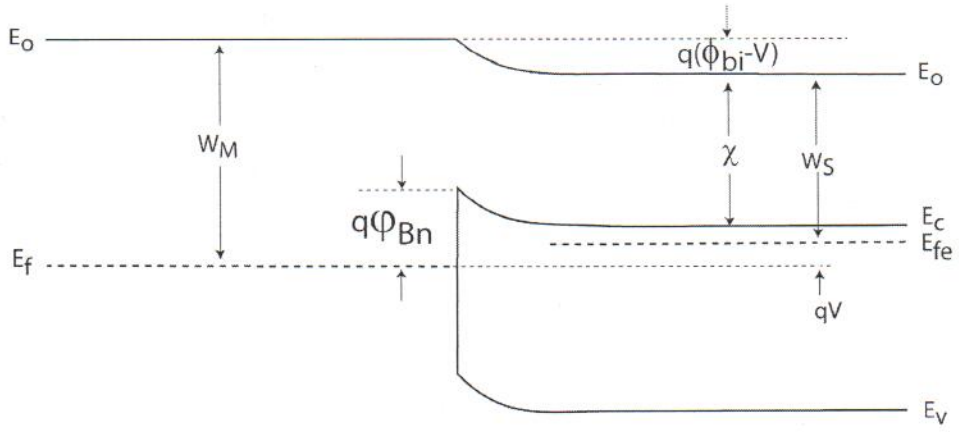



Figure 7.9: Energy band diagram of a metal/n-semiconductor junction in forward bias ($V > 0$).

### 7.3.1 Electrostatics under bias

Let us connect a battery with voltage $V$ across a Schottky diode. Through the contact to the Schottky metal, the positive side of the battery "grabs" on the Fermi level in the metal. Through the ohmic contact, as we will see in Section 7.8, the negative side of the battery grabs on the majority carrier quasi-Fermi level of the semiconductor. The application of a voltage $V$ between the Schottky metal and the ohmic contact therefore results in a split between the Fermi level of the metal and the quasi-Fermi level for electrons in the semiconductor of a magnitude $qV$, as shown in Fig. 7.9 for forward bias and Fig. 7.10 for reverse bias. What is the resulting energy band diagram? Do the electrostatics change across the structure? If so, how do they change?

A complete answer to these questions must be deferred until we have a way to calculate the current that flows when a voltage is applied. This is a chicken and egg problem. We cannot calculate the current without knowing the potential distribution through the structure, which we cannot compute without knowing the currents. To break this circle, let us first assume that the current has a negligible impact on the electrostatics. We will then use this to compute the currents. At that point, we can estimate the order of magnitude of the required corrections and perform them if they are not too large, or ignore them if they are very small.

In the case of a resistor, when a voltage was applied across its terminals, we made the implicit assumption that the voltage dropped in a uniform way along its length. That is why the energy band diagram is tilted with a constant slope everywhere. This is reasonable if the resistor has uniform properties (resistivity and cross section) along its length. In the case of the Schottky diode, the situation is rather different. This structure has basically five rather different regions where the voltage can drop: the Schottky metal bulk, the interface of the Schottky metal with the semiconductor, the space-charge region in the semiconductor, the quasi-neutral region or bulk of the semiconductor and the ohmic contact to the semiconductor. Since the Schottky metal is very conducting in comparison with the semiconductor, the voltage drop there is negligible. At

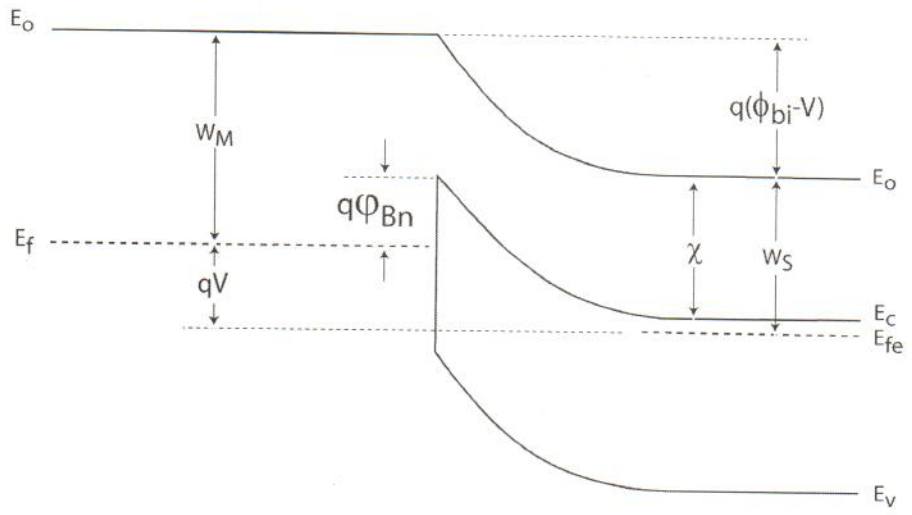



Figure 7.10: Energy band diagram of a metal/n-semiconductor junction in reverse bias ($V < 0$).

its interface with the semiconductor, the Schottky metal can sustain some electrical charge. This is a very thin region and cannot absorb any significant voltage. Moving into the semiconductor, if we compare the QNR with the SCR, we can see that the QNR has many more carriers than the SCR; the SCR is a lot more "resistive" than the QNR and it is reasonable to assume that most of the voltage drops there. In an ideal Schottky diode, we also neglect the ohmic drop in the ohmic contact. This all leads to assuming in a first pass that the voltage that is applied to a Schottky diode drops entirely across the SCR. We will review this after we discuss the I-V characteristics of the junction.

Fig. 7.9 shows the energy band diagram in forward bias. In this case, the quasi-Fermi level for electrons in the QNR of the semiconductor is raised over the Fermi level of the metal by an amount $qV$. Since we assume that there is no ohmic drop across the QNR in the semiconductor, $E_{fe}$ remains flat throughout. The difference in the local vacuum level across the structure is reduced from a value of $q\phi_{bi}$ in equilibrium, to a value $q(\phi_{bi} - V)$ in forward bias. The total built-in potential has accordingly shrunk from $\phi_{bi}$ to $\phi_{bi} - V$. How can this be accommodated? By reducing the magnitude of the dipole of charge that is set on both sides of the Schottky interface. Since the volume charge density, $qN_D$, cannot be changed, the only way to accomplish this is by shrinking the extent of the space-charge region.

In reverse bias, the situation is similar. In this case, the battery lowers the Fermi level in the semiconductor with respect to the Fermi level in the metal. This *increases* the total band bending across the structure to $q(\phi_{bi} - V)$ ($V$ is negative). In order to accommodate this, the space charge region must widen. The energy band diagram is shown in Fig. 7.10. In both cases, forward and reverse, continuity of the local vacuum potential prevents any changes on the Schottky barrier height as a result of the bias application.

Fig. 7.11 shows the distribution of volume charge density, electric field, and electrostatic po-

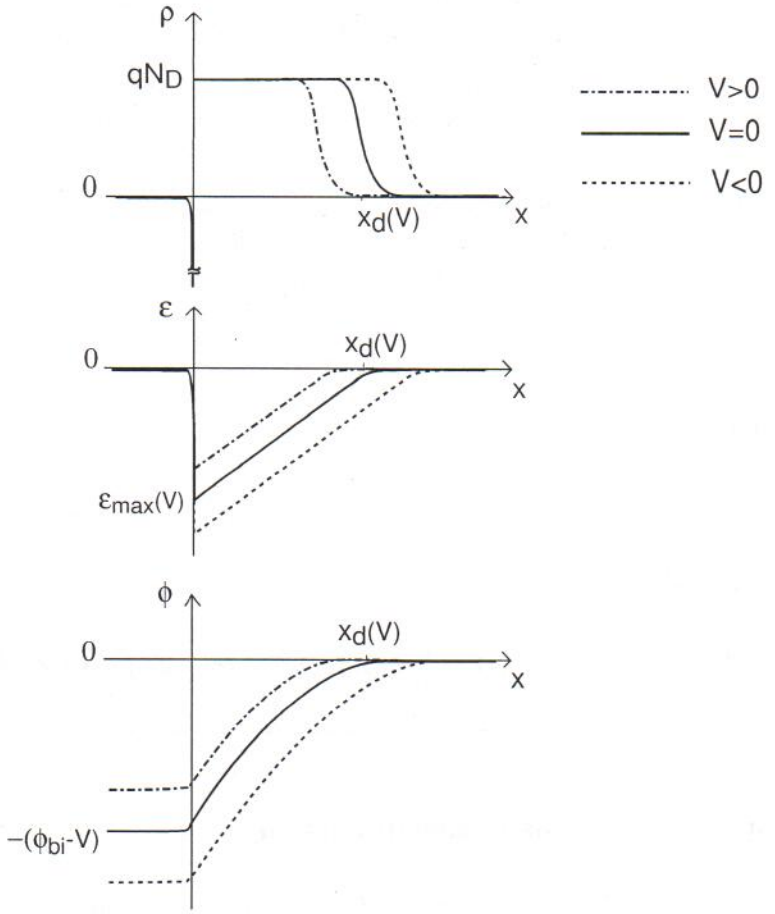



Figure 7.11: Sketches of volume charge density, electric field, and electrostatic potential in equilibrium and under bias in a metal/n-semiconductor junction.

tential under forward and reverse bias. The electrostatics are essentially identical to equilibrium, except that the potential difference across the structure is not $\phi_{bi}$ but $\phi_{bi} - V$. This allows us to reuse the solutions to the electrostatics of the equilibrium situation solved above, except where we wrote $\phi_{bi}$, we must now write $\phi_{bi} - V$. The depletion region thickness is, for example:

$$x_d(V) = \sqrt{\frac{2\epsilon(\phi_{bi} - V)}{qN_D}} = x_d(V = 0)\sqrt{1 - \frac{V}{\phi_{bi}}} \tag{7.21}$$

The maximum electric field under bias is:

$$|\mathcal{E}_{max}(V)| = \frac{qN_D x_d(V)}{\epsilon} = \sqrt{\frac{2qN_D(\phi_{bi} - V)}{\epsilon}} = |\mathcal{E}_{max}(V = 0)|\sqrt{1 - \frac{V}{\phi_{bi}}} \tag{7.22}$$

These expressions are valid provided that the depletion approximation continues to apply. This is the case if the forward voltage does not get too close to $\phi_{bi}$.

Sketching the carrier concentration distribution and the location of the Fermi level across the structure requires solving the current-voltage characteristics. This is postponed until Section 7.3.3 below.

**Exercise 7.3**: *Calculate the extent of the depletion region and the peak electric field in an Al/n-Si Schottky diode with* $N_D = 10^{17}\ cm^{-3}$ *at a forward voltage of* $0.5\ V$ *at 300K.*

We use the results of Exercise 7.1. The built-in potential of this structure is $\phi_{bi} = 0.53\ V$. Then,

$$\sqrt{1 - \frac{V}{\phi_{bi}}} = \sqrt{1 - \frac{0.5\ V}{0.53\ V}} = 0.24$$

The extent of the depletion region is:

$$x_d(0.5\ V) = x_d(V = 0)\sqrt{1 - \frac{V}{\phi_{bi}}} = 8.3 \times 10^{-6}\ cm \times 0.24 = 20\ nm$$

The peak electric field is:

$$|\mathcal{E}_{max}(0.5\ V)| = |\mathcal{E}_{max}(V = 0)|\sqrt{1 - \frac{V}{\phi_{bi}}} = 1.4 \times 10^5\ V/cm \times 0.24 = 3.4 \times 10^4\ V/cm$$

### 7.3.2 I-V characteristics: qualitative discussion

A Schottky diode is essentially a *majority-carrier device.* Minority carriers only play a secondary role in its behavior. This is not hard to understand since in equilibrium there is not a significant amount of minority carriers anywhere in the semiconductor that can flow and recombine.

It is not uncommon in semiconductor devices under bias to find that a specific region constitutes the bottleneck to current flow while the rest of the device adapts as needed. In these kinds of situations, the strategy to derive a model for the current-voltage characteristics is to focus on this bottleneck and construct a model for transport there. In a Schottky diode, the lowest concentration of carriers is found in the space-charge region. This is where we will center our attention. Fig. 7.12 qualitatively helps us to consider what happens there.

Fig. 7.12 shows three sketches of energy band diagrams for our Schottky diode in three different situations. In equilibrium, Fig. 7.12a, electrons face an energy barrier of height $q\varphi_{Bn}$ as they attempt to flow from the metal to the semiconductor or vice versa. The electron flow from the metal to the semiconductor is balanced out by the flow from the semiconductor to the metal. The net flow of electrons across the barrier is zero and the net current is zero.

In forward bias, Fig. 7.12b, the Fermi level on the semiconductor is raised by an amount $qV$ with respect to the Fermi level in the metal. Looking from the point of view of the electrons, the energy barrier facing metal electrons attempting to enter into the semiconductor is $q\varphi_{Bn}$,

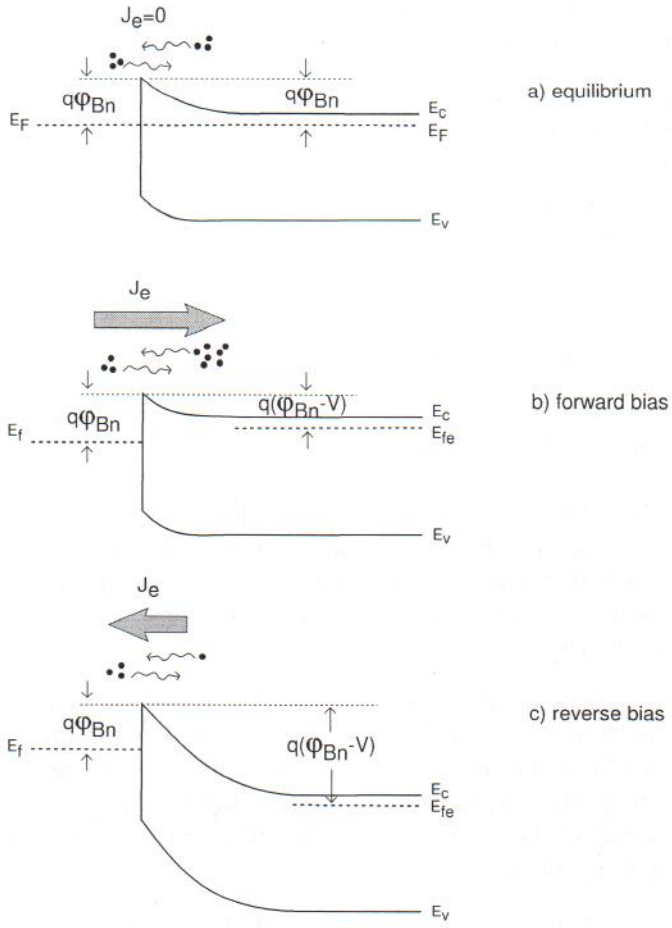



Figure 7.12: Sketch of electron flow across metal-semiconductor junction in a) equilibrium, b) forward bias, and c) reverse bias. In equilibrium there is no net flow of electrons. In forward bias, there is a net flow of electrons from the semiconductor to the metal. In reverse bias, there is net electron flow from the metal to the semiconductor.

just as in equilibrium. In contrast, the energy barrier preventing the flow of electrons from the semiconductor to the metal has *decreased* to $q(\varphi_{Bn} - V)$. Clearly, this results in a net flow of electrons from the semiconductor to the metal producing a current in the contrary sense.

In reverse bias, Fig. 7.12c, the Fermi level on the semiconductor is lowered by an amount $qV$ with respect to the Fermi level in the metal. The energy barrier facing the metal electrons is still unchanged from equilibrium, while the energy barrier in front of semiconductor electrons has now *increased* to $q(\varphi_{Bn} - V)$ ($V$ is negative in reverse bias). In consequence, there is a net flow of electrons from the metal to the semiconductor and a current in the contrary sense.

The sketches of Fig. 7.12 also allow us to recognize the functional dependence of the current on the voltage applied across the junction. In forward bias, as $V$ is increased, the number of electrons in the semiconductor that have enough energy to overcome the energy barrier at the Schottky interface increases exponentially with $V$. This is a consequence of the exponentially

decaying distribution of electrons in the conduction band of the semiconductor. The electron flux from the semiconductor to the metal should therefore increase as $e^{qV/kT}$.

In reverse bias, on the other hand, the number of electrons in the semiconductor with enough energy to overcome the interfacial barrier is also exponentially suppressed as the reverse bias $|V|$ increases. When $|V|$ exceeds a few $kT/q$'s, the injection of electrons from the semiconductor to the metal is completely suppressed and the current saturates to the value given by the injection of electrons from the metal to the semiconductor which is unchanged from equilibrium.

Our qualitative arguments have brought us quite far. We have identified the fact that overcoming the energy barrier at the Schottky interface is the bottleneck to electron flow. Thinking in terms of the electron population at that bottleneck qualitatively explains the rectifying behavior of the Schottky diode. We are just a step away from developing a simple model for the current-voltage characteristics. What we are missing is a way to figure out the *rate* at which electrons flow through the SCR.

Let us start by thinking about the situation in thermal equilibrium. With $V = 0$, the net current anywhere is zero. In the SCR, where an electric field is present, that means that there is a drift current that must be perfectly balanced everywhere by a diffusion current. This originates in the gradient of electron concentration that exists from the bulk, where it is highest, to the metal-semiconductor interface where it is lowest.

Under forward bias, the field distribution in the SCR shrinks in magnitude and spatial extent. This breaks the balance of drift and diffusion with diffusion prevailing over drift. This yields a net electron flow from the semiconductor towards the metal. As these electrons enter the metal, they quickly lose their energy thermalizing with the many electrons there. Somehow, what limits the forward bias current is the ability of electrons to reach the metal-semiconductor interface. Once there, they are immediately "sucked" by the metal.

We can contemplate two extreme situations here. In one, the rate limiting step to the current is the process of electron diffusion against the strong electric field in the SCR. In this limit, the electron concentration drops as we advance from the edge of the SCR towards the Schottky interface. At the metal-semiconductor interface, the electron concentration approaches the equilibrium value as these electrons are in a quasi-equilibrium state with those in the metal. A model that describes the current in this limit is known as the *drift-diffusion model.* This situation is more likely to arise in semiconductors with small mobilities and for low forward bias when the opposing electric field is relatively high.

At the other limit, if the mobility is high and the field is small as a result of strong forward bias, electrons readily reach the metal-semiconductor interface. In this case, it is the process of "emission" into the metal that limits the current. What is exactly this process? We focus at the tip of the energy barrier on the semiconductor side of the the metal-semiconductor interface because that is where the electron concentration is the lowest anywhere in the semiconductor. At that location, the electron concentration is finite and the velocity at which they flow into the metal is also finite. This limits the current to some amount. So, in the limit in which electrons can easily arrive to the Schottky interface from the bulk, it is this maximum "emission current" at the tip of the barrier that sets the diode current. This is known as the *thermionic emission model.*

As it turns out, for Si, GaAs and many common semiconductors under most conditions, the thermionic emission current is found to be the rate limiting step. This is because the mobilities are high enough and carriers can readily reach the Schottky interface. The next section formulates the thermionic emission model. Appendix AT7.2 describes the drift-diffusion model.

### 7.3.3 I-V characteristics: thermionic emission model

As we have discussed in the previous section, in the thermionic emission model, the current is limited by the electron emission process over the tip of the energy barrier at the Schottky interface. We then focus at the semiconductor side of the Schottky interface located at $x = 0^+$ and we write the diode current as the net electron current at that location. This is given by the balance between current due to electron flow from the semiconductor to the metal minus the reverse electron flow from the metal to the semiconductor:

$$J_t \simeq J_e(0^+) = J_{e,SM}(0^+) - J_{e,MS}(0^+) \tag{7.23}$$

In forward bias, $J_{e,MS}(0^+)$ is very small next to $J_{e,SM}(0^+)$. We then neglect this component for the time being and focus on this first term that accounts for electrons emitted from the semiconductor into the metal. This current can be approximately expressed as:

$$J_{e,SM}(0^+) \simeq -qn(0^+)v_e(0^+) \tag{7.24}$$

Here, the minus sign comes from the fact that with our definition of axis, the velocity of these electrons is against $x$, or negative ($J_{e,SM}(0^+)$ ends up positive). Note also that we assume that all electrons at $x = 0^+$ contribute to this current component. The reverse current is very small and, for the time being, we can ignore the few electrons that support it.

To proceed, we need to exploit the assumption that electrons have no difficulty reaching the Schottky interface from the bulk. This is in essence, the *quasi-equilibrium approximation* in which electrons across the SCR are all in equilibrium with each other and also with electrons in the bulk of the semiconductor. Without further thinking, the Boltzmann relation would allow us to express $n(0^+)$ in terms of the electron concentration in the bulk, $N_D$, and the difference of potentials between the surface and the bulk, $\phi_{bi} - V$. We would be then tempted to write:

$$n(0^+) = N_D \exp \frac{-q(\phi_{bi} - V)}{kT} \tag{7.25}$$

But there is a problem here. At $x = 0^+$ the electron distribution cannot be close to equilibrium. In fact, it has to be very different from an equilibrium Maxwellian distribution. A carrier distribution can only be considered in thermal equilibrium when electrons undergo many collisions with the lattice that randomize their velocity. Right at the metal-semiconductor interface, any electron that has velocity components pointing towards the metal will be "sucked" by the metal but no electron in the metal can make it into the semiconductor because there is a large

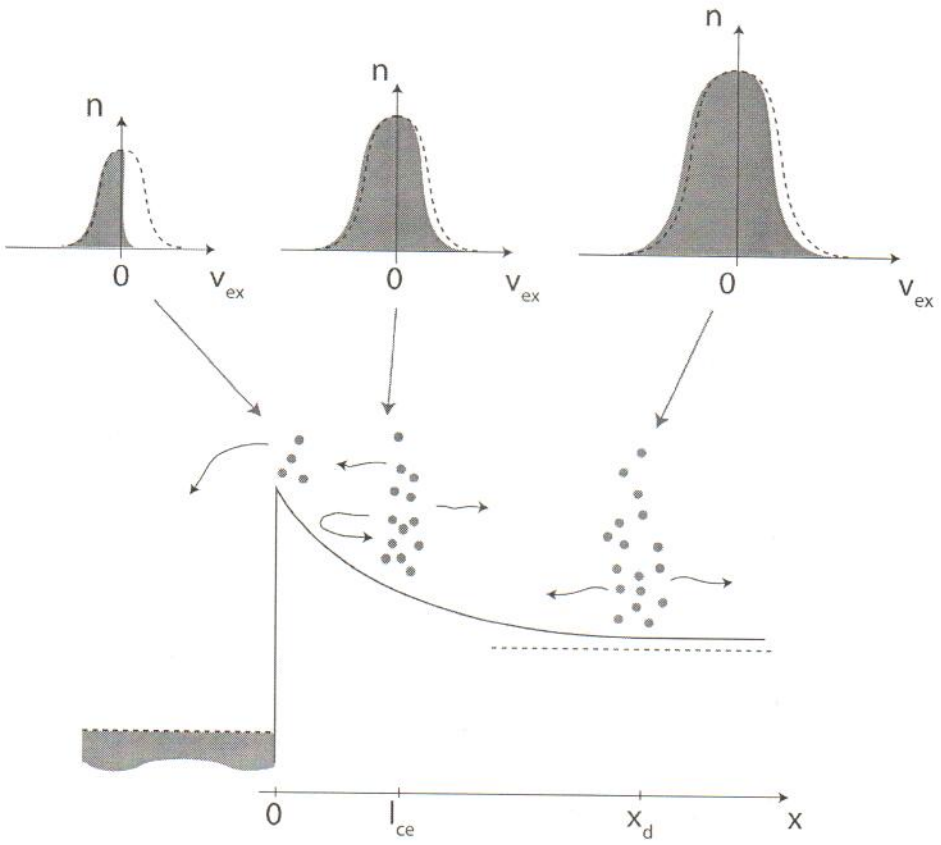



Figure 7.13: Sketch of conduction band across SCR in Schottky diode. Also shown are the distributions of x-component of electron velocity at three locations. Across most of the SCR, electrons are in quasi-equilibrium with the bulk and the electron distribution is Maxwellian. On the semiconductor side of the Schottky interface, the velocity distribution is hemi-maxwellian as only electrons with velocity components pointing towards the metal are present.

energy barrier at the interface that prevents it. So, right at the tip of the barrier at $x = 0^+$, there are very few electrons with velocity components pointing towards the semiconductor (this, to the extent that we can neglect $J_{e,MS}(0^+)$ next to $J_{e,SM}(0^+)$). In effect, at $x = 0^+$ we have an *hemi-maxwellian distribution.* This is sketched in Fig. 7.13.

This figure sketches the conduction band diagram of a Schottky diode in forward bias. It also sketches the distribution of the x-component of the electron velocity at three different locations. In the bulk of the semiconductor, there is a large electron concentration that can be considered to be in near thermal equilibrium with the lattice. The distribution of electron concentration in energy and the x component of velocity are very close to perfect equilibrium (dashed lines). Only a small difference exists between the distribution of electron velocities pointing towards the metal and away from the metal that gives rise to a net electron flow towards the Schottky interface. Under the quasi-equilibrium approximation, this situation persists across nearly the entire SCR.

On average, a mean free path away from the metal-semiconductor interface, at $x = l_{ce}$, the electrons that ultimately are emitted into the metal suffer their last collision. From there on, those that have enough energy to overcome the energy barrier at the interface and that after their last collision have a component of velocity pointing towards the metal succeed in getting emitted into the metal and contribute to diode current. Those that have the right velocity component but do not have enough energy eventually are reflected by the energy barrier and do not contribute to current. All those that scatter back into the semiconductor also do not contribute to current. The end result is that the distribution of electron velocities right at the Schottky interface at $x = 0^+$ is very far from thermal equilibrium as only the negative velocity half of the Maxwellian distribution is present.

This understanding allows now to correctly estimate $n(0^+)$. We can easily express $n(0^+)$ in terms of $n(l_{ce})$, the electron concentration a mean free path away from the interface. On average, only half of the electrons at $x = l_{ce}$ have a velocity that points towards the interface. Of those, only a fraction have enough kinetic energy to make it over the potential barrier that exists between $x = 0$ and $x = l_{ce}$. The height of this potential barrier is given by the electrostatic potential difference between $x = 0^+$ and $x = l_{ce}$, that is $q[\phi(l_{ce}) - \phi(0^+)]$. Then, the electron concentration at $x = 0$ can be written as:

$$n(0^+) = \frac{n(l_{ce})}{2} \exp \frac{q[\phi(0^+) - \phi(l_{ce})]}{kT} \tag{7.26}$$

The electron concentration at $x = l_{ce}$ can be computed by solving the drift-diffusion problem across the rest of the SCR. As we have mentioned before, in semiconductors with relatively high mobilities, such as Si and GaAs, a quasi-equilibrium situation prevails over the SCR from the bulk up to $x = l_{ce}$. We can then use the Boltzmann relation to relate $n(l_{ce})$ and $n(x_d) \simeq N_D$:

$$n(l_{ce}) \simeq N_D \exp \frac{q\phi(l_{ce})}{kT} \tag{7.27}$$

We can now substitute 7.27 into 7.26 and use the fact that $\phi(0^+) = -(\phi_{bi} - V)$ obtained in Section 7.3 to get:

$$n(0^+) = \frac{N_D}{2} \exp \frac{-q(\phi_{bi} - V)}{kT} = \frac{N_c}{2} \exp \frac{-q(\varphi_{Bn} - V)}{kT} \tag{7.28}$$

Equation 7.28 has a simple and intuitive physical interpretation. The electron concentration at $x = 0^+$ is exactly half of what one would have if it was in thermal equilibrium with the bulk (Eq. 7.25). The factor of 1/2 originates from the lack of scattering to the left of $x = 0$ to bring any electrons back. All electrons at $x = 0^+$ are injected into the metal. Notice that, as expected, $n(0^+) \sim e^{qV/kT}$.

The second step in the construction of a model for current is the computation of the electron velocity at the interface pointing towards the metal. In Ch. 4 we introduced the notion of thermal velocity of electrons, $v_{the}$. This is the average instantaneous velocity of electrons over distances shorter than a mean free path. At the metal-semiconductor interface, however, most electrons

point into the metal with velocities at an angle, therefore contributing a smaller net forward velocity. When the statistics of the electron velocity distribution are properly taken into account, the velocity at which electrons are emitted from the semiconductor into the metal (and vice versa in reverse bias) is actually $v_{the}/2$. Using Eq. 4.5, we can then write:

$$v_e(0^+) = -\frac{v_{the}}{2} = -\sqrt{\frac{2kT}{\pi m^*_{ce}}} \tag{7.29}$$

where the minus sign simply indicates that electrons are flowing in the contrary sense to our choice of axis. For Si at room temperature, $v_e(0^+)$ is just about $10^7\ cm/s$.

Plugging Eqs. 7.28 and 7.29 into 7.24, we get:

$$J_{e,SM}(0^+) = qN_D\sqrt{\frac{kT}{2\pi m^*_{ce}}}\exp\frac{-q(\phi_{bi}-V)}{kT} = qN_c\sqrt{\frac{kT}{2\pi m^*_{ce}}}\exp\frac{-q\varphi_{Bn}}{kT}\exp\frac{qV}{kT} \tag{7.30}$$

where we have used Eq. 7.4 and the relationship given in Ch. 4 between $\varphi_n$, $N_D$ and $N_c$.

Using now the expression for $N_c$ given in Eq. 2.26, we can write:

$$J_{e,SM}(0^+) = A^*T^2\exp\frac{-q\varphi_{Bn}}{kT}\exp\frac{qV}{kT} \tag{7.31}$$

where $A^*$ is:

$$A^* = \frac{4\pi qk^2m_o}{h^3}\sqrt{\frac{(m^*_{de}/m_o)^3}{(m^*_{ce}/m_o)}} \tag{7.32}$$

$A^*$ is called Richardson's constant and is a characteristic of the material and the type of carrier. The quantity in front of the square root has a value of 120 $A\cdot cm^{-2}\cdot K^{-2}$.

We have solved half of the problem. We still need to get $J_{e,MS}(0^+)$ to complete the calculation. However, we know that the current due to electron flow from the metal to the semiconductor is unaffected by the application of a voltage. This is again due to the fact that the energy barrier that is presented to electrons in the metal does not change with bias. Therefore, we can get an expression for $J_{e,MS}(0^+)$ by equating it to $J_{e,SM}(0^+)$ in Eq. 7.31 at zero volts when the net current should be zero. This yields:

$$J_{e,MS}(0^+) = A^*T^2\exp\frac{-q\varphi_{Bn}}{kT} \tag{7.33}$$

The final net diode current is then:

$$J_t = A^*T^2\exp\frac{-q\varphi_{Bn}}{kT}(\exp\frac{qV}{kT}-1) \tag{7.34}$$

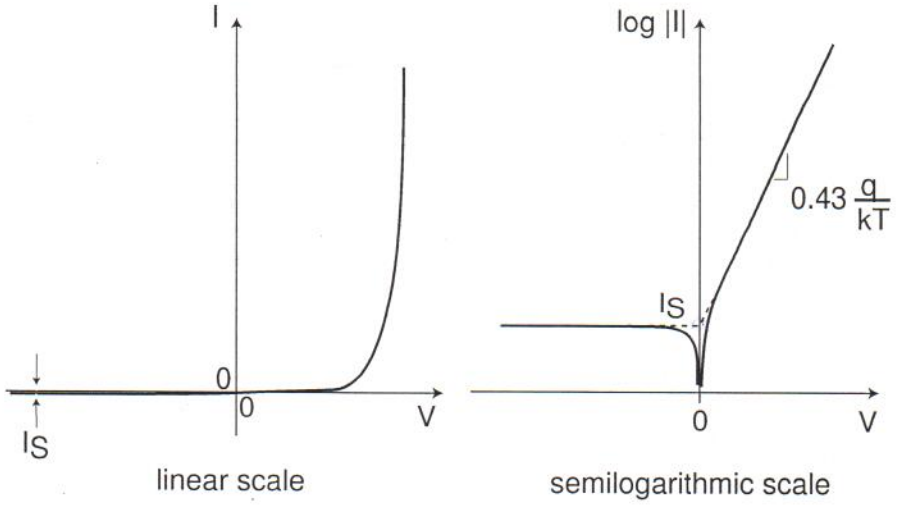



Figure 7.14: Sketch of $I-V$ characteristics of metal-semiconductor junction, in a linear scale on the left, and in a semilogarithmic scale on the right.

This equation is valid in forward as well as in reverse bias since in reverse bias, the current saturates to Eq. 7.33 (with a minus sign) as the forward emission of electrons from the semiconductor to the metal is completely suppressed.

In a diode that has a junction area $A_j$, the current flowing through the diode is then:

$$I = I_S(\exp\frac{qV}{kT} - 1) \tag{7.35}$$

where $I_S$ is called the *saturation current.* This is identical to a PN diode and is plotted again in Fig. 7.14 in linear and semilog scales. In a semilog scale, the forward bias current appears as a straight line with a slope of $(q/kT)\log\ e = 0.43q/kT$ or 60 $mV/dec$ at room temperature.

Fig. 7.15 shows experimental $I-V$ characteristics of a PtSi/n-Si junction. The experimental data follows the shape predicted by Eq. 7.35 very accurately.

The saturation current $I_S$ has a very peculiar set of dependencies in itself:

$$I_S = A_j A^* T^2 \exp\frac{-q\varphi_{Bn}}{kT} \tag{7.36}$$

$I_S$ in this equation depends exponentially with a minus sign on the Schottky barrier height. This is reasonable since the higher the barrier, the lower (exponentially) the flow of electrons over it. The pre-exponential factor depends on the square of the absolute temperature. Since $A^*$ does not depend on temperature, $I_S/T^2$ is thermally activated with an activation energy equal to $q\varphi_{Bn}$. This suggests an experimental procedure to determine $q\varphi_{Bn}$. First, $I-V$ characteristics are measured as a function of temperature and $I_S$ is extracted at all temperatures. Then $I_S/T^2$ is graphed in an Arrhenius plot. The slope of the resulting straight line gives $q\varphi_{Bn}$. This is

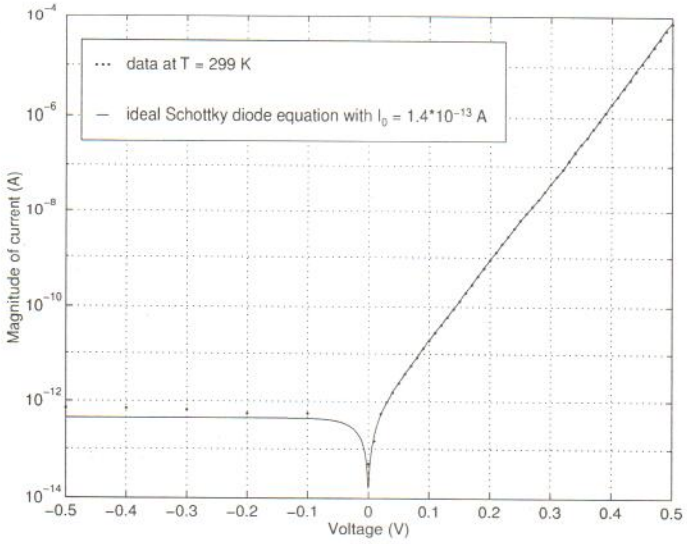



Figure 7.15: Experimental $I-V$ characteristics of PtSi/n-Si Schottky diodes at room temperature [*diode courtesy of B. Scharf (Analog Devices), data and model courtesy of S. Krupenin (MIT)*].

experimentally illustrated in Fig. 7.16 for the same diode of Fig. 7.15.

A word about Richardson constant $A^*$. As seen in Eq. 7.32, the thermionic emission theory predicts it to depend only on the semiconductor and to be independent of the metal. For n-Si, if one plugs in the appropriate values of effective masses, one obtains $A^* \simeq 258\ A \cdot cm^{-2} \cdot K^{-2}$. A more refined theory predicts that $A^*$ depends somehow on the crystalline orientation of the material. It is not the same to make a Schotky diode on (100) Si than on (111) Si. In practice, one can experimentally extract $A^*$ from actual measurements on Schottky diodes. What is found is that $A^*$ depends a lot on the nature of the metal itself, its thickness, and the details of the process: surface preparation prior to metal deposition, the deposition parameters and any post-deposition treatments. The reasons for these dependencies are various: surface roughness, local strain, the formation of metal-semiconductor compounds, interfacial oxides, etc. The bottom line is that for a given metal-semiconductor system, we cannot simply take the theoretically predicted value of $A^*$ given by the thermionic emission theory and expect to accurately predict $I_S$ in Schottky diodes. The proper action is to build some test diodes and measure $A^*$ and then use this in the model.

Let us now discuss the key assumption that we have made in the development of our theory. In deriving the previous equations, we assumed that thermionic emission of electrons over the Schottky barrier is the rate-limiting mechanism for electron transport, and that drift and diffusion through the SCR are in near perfect balance from $x = l_{ce}$ to the end of the depletion region towards the body. This assumption allowed us to focus on the emission process over the tip of the energy barrier at the metal-semiconductor interface in order to compute the current. Elsewhere in the SCR, we used the Boltzmann relation to calculate the electron concentration.

For this quasi-equilibrium assumption to apply, the net current that flows through the junction

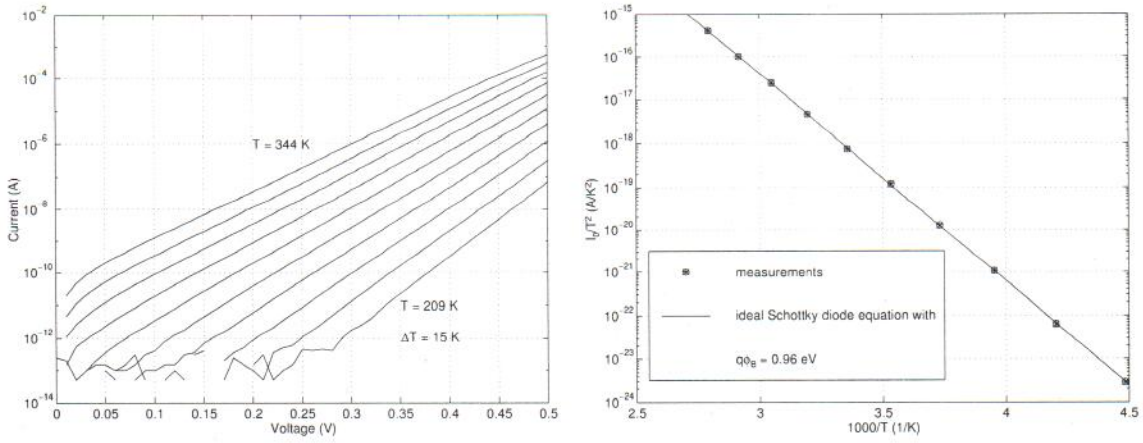



Figure 7.16: Left: Forward I-V characteristics of PtSi/n-Si Schottky diode for different temperatures. Right: Arrhenius plot of $I_S/T^2$ for the same diode. The slope of the straight line gives $q\varphi_{Bn}$ [*diode courtesy of B. Scharf (Analog Devices), data and fit courtesy of S. Krupenin (MIT)*].

must be much smaller than either the drift or the diffusion currents. In this way, only a small imbalance between these two currents needs to exist. It is particularly intuitive to focus on the drift current. Quasi-equilibrium is guaranteed up to $x = l_{ce}$ if the thermionic emission current is much smaller than the drift current at that location, that is,

$$|J_{e,SM}(0^+)| \ll |J_{e,drift}(l_{ce})| \tag{7.37}$$

Using Eqs. 7.24, 7.26 and 7.29 on the left hand side, and the usual expression for the drift current on the right hand side ($J_{e,drift} = q\mu_e n\mathcal{E}$), we can rewrite this as:

$$\frac{1}{4}v_{the}\exp(-\frac{q\Delta\phi_{lce}}{kT}) \ll \mu_e|\mathcal{E}(l_{ce})| \tag{7.38}$$

where we have defined $\Delta\phi_{lce}$ as the potential build-up over the first mean-free path of the SCR starting from the metal-semiconductor interface.

Expressions 4.5, 4.8 and 4.10 allow us to express the mobility in terms of the velocity and the mean free path as: $\mu_e = ql_{ce}/(m^*_{ce}v_{the})$. Plugging this on the right hand side of Eq. 7.38, and noting that $l_{ce}|\mathcal{E}(l_{ce})| < \Delta\phi_{lce}$, we can rewrite this expression as:

$$\frac{q\Delta\phi_{lce}}{kT}\ \exp(\frac{q\Delta\phi_{lce}}{kT}) \gg \frac{2}{\pi} \tag{7.39}$$

This is readily satisfied if:

$$\Delta\phi_{lce} > 1.5\frac{kT}{q} \tag{7.40}$$

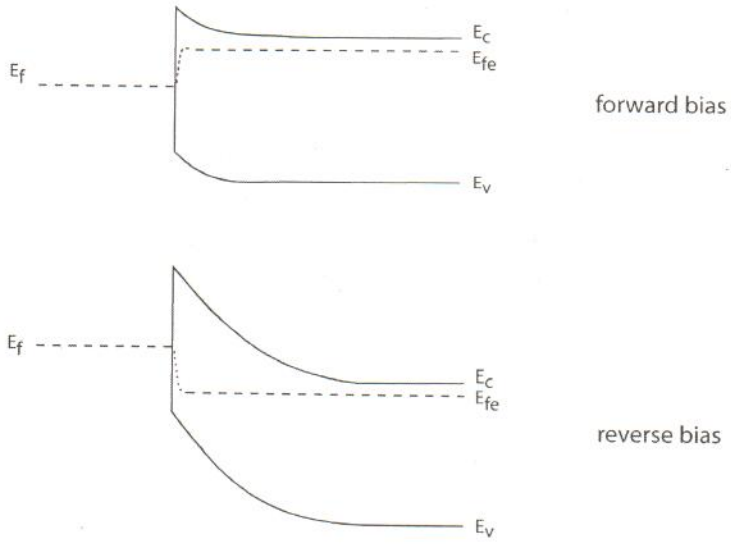



Figure 7.17: Energy band diagrams in metal-semiconductor junction in forward and reverse bias showing the location of the quasi-Fermi level for electrons.

or, in other words, if the potential drop in the first mean free path from the interface exceeds $1.5kT/q$. This is called the *Bethe condition*. In Si this condition is easy to satisfy at all practical voltages because the mean-free paths are rather long. An example is given in Exercise 7.4.

**Exercise 7.4**: *Evaluate the validity of the Bethe condition for an Al/n-Si Schottky diode with* $N_D = 10^{17}\ cm^{-3}$ *at a forward voltage of* $0.5\ V$ *at 300K.*

For this calculation, we need the mean free path in the SCR. If we assume that the electron mobility in an SCR is identical to that of a QNR with an identical doping level, we can use the result of Exercise 4.1 where we obtained for this same doping level $l_{ce} \simeq 22\ nm$.

From Exercise 7.3 we know that for this Schottky diode at this forward voltage, the peak field at the Schottky interface is $|\mathcal{E}_{max}| \simeq 3.4 \times 10^4\ V/cm$. Then, the potential drop in the first mean free path away from the Schottky interface is:

$$\Delta\phi_{lce} \simeq l_{ce}|\mathcal{E}_{max}| \simeq 22 \times 10^{-7}\ cm \times 3.4 \times 10^4\ V/cm = 0.075\ V$$

This is comfortably bigger than $1.5kT/q \simeq 39\ mV$ and the Bethe condition is fulfilled. For this Schottky diode, even for such a high forward voltage, the thermionic emission model is the appropriate one.

We are finally in a position to draw the electron quasi-Fermi level, $E_{fe}$, across the device. Discussion of the hole quasi-Fermi level still needs to be postponed until we discuss the behavior of minority carriers. Drawing $E_{fe}$ is now easy. The assumption of quasi-equilibrium across the entire device with the exception of the very first mean-free path implies that the quasi-Fermi level for electrons is flat in this entire region. It location is set by the doping level in the QNR. This is sketched in Fig. 7.17.

Within the very first mean-free path away from the metal-semiconductor interface, the concept

of quasi-Fermi level is not appropriate as the electron distribution is very far from Maxwellian. We indicate this in Fig. 7.17 by means of a dotted line that suggests that the electron distribution comes to equilibrium with the electrons in the metal at the metal-semiconductor interface.

## 7.4 Charge-voltage characteristics of ideal Schottky diode

As we learned in the PN diode, in order to describe the dynamic behavior of a Schottky diode in a circuit, we need a model for the stored charge. If the voltage across the diode changes, the stored charge must change too and the outside circuit must deliver it. This is a real current that needs to be accounted for.

The ideal Schottky diode is a majority carrier device. The current that flows upon the application of a voltage arises from the flow of majority carriers over the energy barrier at the metal-semiconductor interface. In an ideal Schottky diode, unlike a PN diode, the minority carrier concentration is unperturbed from its equilibrium value. There is then no minority carrier storage to contend with. In an ideal Schottky diode, the only stored charge to keep track of is that of the space-charge region which thickness is modulated with the applied voltage as we discussed in Sec. 7.3.1.

The situation is similar to that of a PN diode. Fig. 7.18 sketches the volume charge density in the space-charge region of a Schottky diode with a certain voltage $V$ applied. Whether the voltage is forward or reverse, there is a dipole of charge of a magnitude $Q$ across the metal-semiconductor interface. For the example depicted in Fig. 7.18, there is a charge $+Q$ in the SCR of the semiconductor that is imaged in an equal amount of charge $-Q$ due to a pile up of electrons at the metal surface.

If the voltage that is applied across the diode is increased (made more positive) by a small amount, $\Delta V$, the SCR shrinks a small amount and the pile up of electrons at the metal surface is also reduced somehow. This means that a small quantity of positive charge, $+\Delta Q$, is delivered to the metal (the positive plate) and the same amount of negative charge, $-\Delta Q$, is delivered to the semiconductor (the negative plate). The overall charge dipole shrinks, as required.

The stored charge in the SCR can easily be obtained within the depletion approximation. From Eq. 7.21, we have:

$$Q(V) = -AqN_D x_d(V) = -A\sqrt{2qN_D\epsilon(\phi_{bi} - V)} \tag{7.41}$$

As for the PN diode, this equation can be rewritten as:

$$Q(V) = Q(V=0)\sqrt{1 - \frac{V}{\phi_{bi}}} \tag{7.42}$$

It should not be surprising that we obtain an identical $Q - V$ relationship for the Schottky diode and the PN diode. Essentially, the electrostatics of the SCR for both devices is identical.

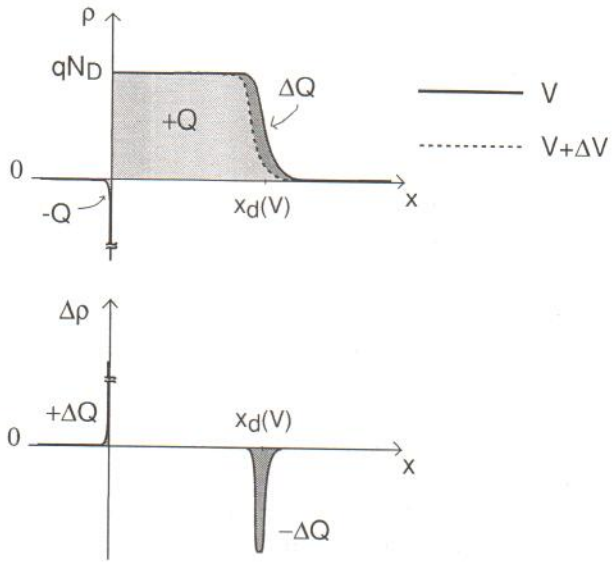



Figure 7.18: Change in the volume charge density in a metal/n-semiconductor junction as a result of increasing the voltage across.

The graph of $Q$ vs. $V$ for the PN diode of Fig. 6.21 applies here.

## 7.5 Equivalent circuit models for the ideal Schottky diode

A device equivalent circuit model is a circuit-like description of its electrical behavior. With the physics of the Schottky diode and the PN diode related in so many ways, their equivalent circuit models are also bound to be very close. In fact, topologically, the equivalent circuit models for these two devices are identical.

We studied the equivalent circuit models for the ideal PN diode in Sec. 6.5. We saw that the large-signal ideal PN diode model contains a "charge-less" ideal diode element that embodies its rectifying I-V characteristics plus a charge storage element in parallel that represents the SCR and the minority carrier charge. The equivalent circuit model for the Schottky diode should look exactly the same except for the absence of the minority carrier charge in the charge storage element. For convenience, we graph it again on the left of Fig. 7.19. In this figure $I_S$ represents the saturation current of the Schottky diode (Eq. 7.36) and $Q$ represents the SCR charge given in Eq. 7.42.

When we discussed the PN diode equivalent circuit models, we also noted the interest in developing a small-signal model that represents in a circuit form the behavior of the device to small excursions around a bias point. For the PN diode, the ideal-diode element is linearized into a resistor and the charge storage element turns into a capacitor. The small-signal equivalent circuit model for an ideal Schottky diode is identical. As graphed on the right of Fig. 7.19, it contains a resistor in parallel with a capacitor. The small-signal resistor is given by an identical

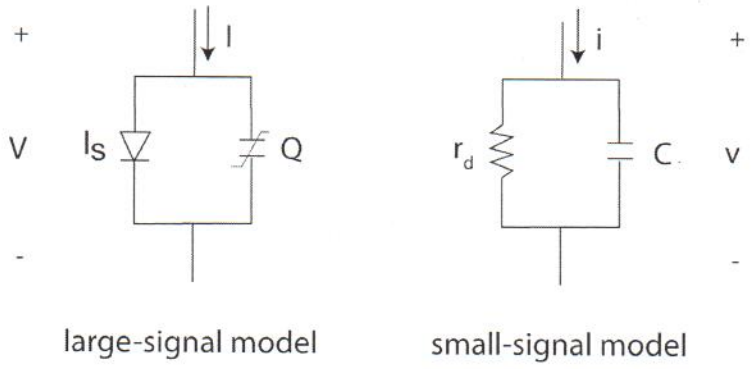



Figure 7.19: Large-signal and small-signal equivalent circuit models for the Schottky diode.

expression to that of the PN diode, reproduced here for convenience:

$$r_d = \frac{kT}{q(I+I_S)} \simeq \frac{kT}{qI} \tag{7.43}$$

where the approximation applies for moderate and strong forward bias. This is the dynamic resistance of the ideal Schottky diode.

The capacitor is associated with SCR charge storage. As in the case of the PN diode, we can obtain it by differentiating Eq. 7.41 or Eq. 7.42 with respect to voltage:

$$C(V) = A\sqrt{\frac{\epsilon q N_D}{2(\phi_{bi}-V)}} = \frac{C(V=0)}{\sqrt{1-\frac{V}{\phi_{bi}}}} \tag{7.44}$$

This is an identical functional expression to that of a PN diode (sketched in Fig. 6.25). As in that case, this expression for the SCR capacitance can also be obtained in analogy with a parallel plate capacitor by using $C(V) = A\epsilon/x_d(V)$ and then using Eq. 7.21.

Also in a close parallelism with a strongly asymmetric PN diode, the inverse square of $C(V)$ in Eq. 7.44 is linearly proportional to $N_D$, the doping level in the semiconductor. This is a standard way to measure this plus the Schottky barrier height in Schottky diodes, as shown in the example of Fig. 7.20.

## 7.6 Non-ideal and second order effects

In this section we chip away at the long list of assumptions that was made in the definition of the ideal Schottky diode and discuss some important non idealities and second-order effects.

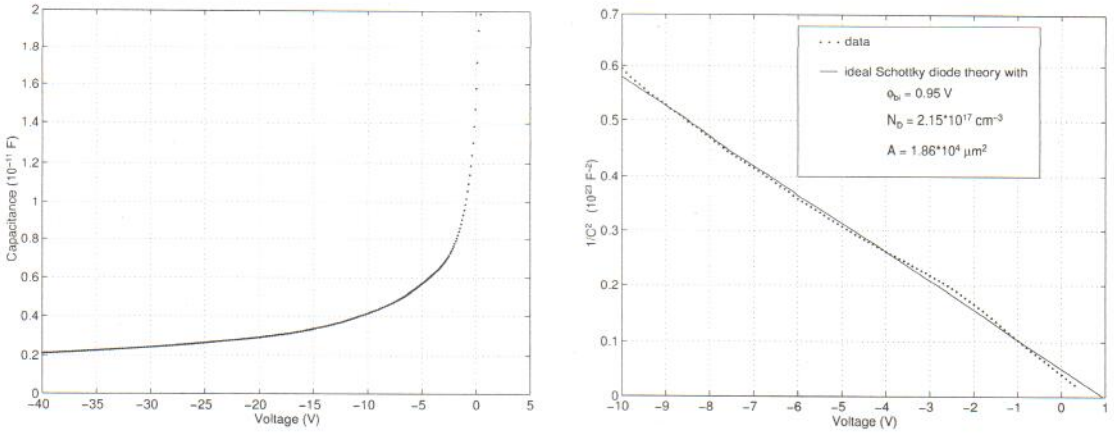



Figure 7.20: $C$ vs. $V$ and $1/C^2$ vs. $V$ for a GaN Schottky diode [*diode courtesy of R. J. Molnar (MIT Lincoln Laboratory), data and fit courtesy of S. Krupenin (MIT)*].

### 7.6.1 Series resistance

In our analysis of the ideal Schottky diode of Fig. 7.2 we neglected the presence of any series resistance. When we derived the expressions for the I-V characteristics, we assumed that the entire voltage applied to the terminals of the device appears across the SCR. But we know that there are ohmic drops at the bottom contact of the diode and in the quasi-neutral body of the device. Just like in the PN diode, these ohmic drops can affect the I-V characteristics of the device at high currents in a significant way.

Since the I-V characteristics of the Schottky diode are functionally identical to those of the PN diode, the role of series resistance is also the same in both devices. We refer to Eq. 6.86 for a modified diode equation that incorporates series resistance, and to Fig. 6.32 for a manifestation of the series resistance in its I-V characteristics. Series resistance is also captured in an equivalent circuit model of the Schottky diode in the same way as for the PN diode (see Fig. 6.33).

There are two components to the series resistance of the Schottky diode: the contact resistance and the body resistance. Between the bottom metal and the semiconductor, there is a contact resistance. We study the physics of ohmic contacts in detail later on in this chapter. As we discussed in the context of the PN diode, a metal-semiconductor contact resistance is characterized through its contact resistivity, $\rho_c$. The contact resistance is given by the product of $\rho_c$ and the area of the contact. Therefore, for the Schottky diode of Fig. 7.2, we have:

$$R_c = \frac{\rho_c}{A} \tag{7.45}$$

Then there is the resistance of the body. This is associated with the QNR as majority carrier electrons drift through this region to support the diode current. The resistance is simply the geometrical resistance of this region given by:

$$R_n = \frac{1}{A}\frac{w_n - x_d}{q\mu_e N_D} \tag{7.46}$$

The voltage dependence of $R_n$ through that of $x_d$ is generally negligible. The total resistance of the diode is given by the sum of $R_c$ and $R_n$.

Schottky diodes are frequently utilized in applications in which it is important to minimize any parasitic ohmic drop as this adds to power dissipation. In high frequency applications, $R_s$ degrades the dynamic response of the diodes, as we will discuss in Section AT7.4. For these reasons, a key consideration in Schottky diode design is to minimize series resistance.

### 7.6.2 Breakdown voltage

In a Shottky diode in reverse bias, a small trickle current flows due to electrons in the metal that have enough energy to overcome the Schottky barrier height at the metal-semiconductor interface and get injected into the semiconductor. As the reverse voltage increases, the magnitude of the electric field in the depletion region increases too. For sufficiently high reverse bias, the flowing electrons might gain enough energy from the electric field and suffer impact ionizing collisions as they travel through the depletion region. The additional carriers that are generated contribute to the total current which in this way increases over the ideal value given by Eq. 7.36 above. For high enough reverse bias, avalanche breakdown might take place and the current increases abruptly to very high values. The manifestation of breakdown in the I-V characteristics of a Schottky diode is similar to that of a PN diode (Fig. 6.34). The voltage at which avalanche breakdown occurs is the *breakdown voltage* of the device. This is a very important parameter since it limits the highest reverse bias that the device can support.

As in so many other aspects, the physics of avalanche breakdown in Schottky diodes and PN diodes are closely related. So, in principle, we could use the theory presented in Sec. 6.6.4 to model the breakdown voltage of a Schottky diode. From our study of PN diode breakdown, we would then conclude that the breakdown voltage in a Schottky diode is mostly set by the doping level of the semiconductor and should be rather independent of the metal that is used. We can also expect that the higher the doping in the semiconductor, the lower the breakdown voltage.

In practice what is found is that for the same doping level in the semiconductor, the breakdown voltage of a Schottky diode is significantly smaller than what is typically obtained in a PN diode. The reason for this has to do with the presence of a high electric field at the sharp edges of the metal that leads to premature breakdown. The situation is illustrated in Fig. 7.21. This problem does not exist in PN diodes because the high field region is inside the semiconductor and away from the metal contacts.

Several solutions to this problem are sketched in Fig. 7.21. One of them relies on the *metal overlap* over an $SiO_2$ window that defines the active area of the diode. The edges of the metal in this way are not in direct contact with the semiconductor. A *moat* structure sometimes is feasible in certain processes. In this approach, as shown, the semiconductor is slightly etched in the active area of the diode. The curvature of the periphery of the moat smoothes out the electric field from the edge of the metal.

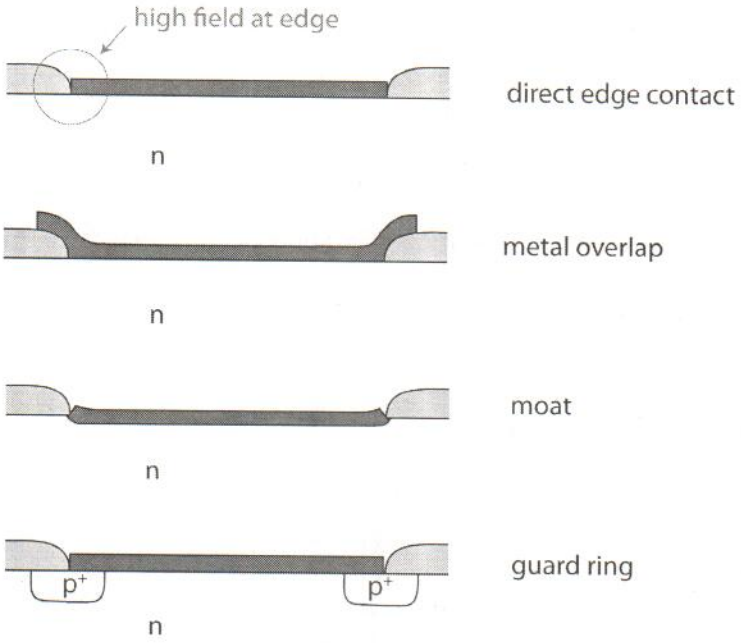



Figure 7.21: A direct overlap of the metal edge and the semiconductor results in high electric fields at the corner that lead to premature reverse breakdown (top). This can be mitigated through the use of a metal/oxide overlap, a moat, or a p-n junction guard ring.

A p-n junction *guard ring* is also effective. In a design on an n-type well, a p-region is inserted right underneath the periphery of the diode, as sketched in Fig. 7.21. At the edges of the metal, this produces a PN junction in parallel with the Schottky junction. This combination can support a higher reverse voltage. This approach has the drawback of the extra space needed in the lateral dimension by the p-region. This makes the diode bigger and in consequence less attractive economically. There is also the danger in this design that the p-n junction might turn on in forward bias. If that happens, the minority carrier charge associated with the p-n diode considerably slows down the combined device. This can occur, in particular, if a metal with a high Schottky barrier height is being used. Careful design of the p-doping level is required to avoid this problem. This is typically done using a device CAD tool since the problem is intrinsically two-dimensional.

## 7.7 Integrated Schottky diode

Schottky diodes have been in use in a variety of applications for many years. For a long time, they were used as "clamps" in the bipolar transistor TTL digital logic family. When placed in parallel with the base-collector junction of the BJT, the forward branch of the I-V characteristics of the Schottky diode limits the maximum forward voltage in the base-collector junction and prevents the bipolar transistor from entering saturation. In modern analog circuits, integrated Schottky diodes are very frequently used in such diverse applications as track and hold circuits

in analog-to-digital converters, and in the pin driver circuits of IC test equipment. The Schottky diode is the preferred choice for such applications because of its small forward voltage and its fast on-off switching speed. Schottky diodes are particularly useful as detectors and mixers in communications and radar applications at the very high frequency end of the spectrum where active devices (that is, devices with gain, such as transistors) have poor performance or are not available altogether. Their low intrinsic capacitance allows them to handle signals at frequencies beyond most other devices.

In spite of their great usefulness, in modern microelectronics, cost considerations dictate that Schottky diodes be engineered by taking advantage of process modules that have been developed for other circuit components. Rarely are new process modules introduced for the sole purpose of fabricating Schottky diodes. For example, the metal used for the Schottky barrier is often the ohmic contact metal. Device engineers must show resourcefulness and imagination to design Schottky diodes with low parasitics that meet the required specifications based on existing process modules.

Fig. 7.22 shows the cross section of two typical integrated Schottky diodes. One is fabricated in a CMOS process (left) and the other one is made on a bipolar process (right). In the CMOS process example, the n-well of the p-MOSFET is used as cathode and a metal placed in direct contact to the well serves as anode (the Schottky junction itself). The n-well is contacted using the body contacts of the p-MOSFET. In the bipolar design, the collector of a Si npn bipolar transistor serves as cathode and a metal placed in direct contact to it is the anode. The $n^+$-subcollector or buried layer and the $n^+$-collector plug provide a low resistance path for the current from the cathode to the surface of the semiconductor.

An integrated Schottky diode does not have the isolation problems of an integrated PN diode that were discussed in Sec. 6.7.1. Because minority carrier injection does not take place, there is no "hidden" parasitic bipolar transistor to worry about. Still, the device needs to be properly isolated. In both examples in Fig. 7.22 the Schottky diode incorporates a parasitic substrate PN diode. Recognizing this is important for two reasons. To insure proper circuit operation, the n-well should never become forward biased with respect to the p-type substrate. Under DC, this can be insured by connecting the substrate to the most negative power supply that is available. Under transient conditions, making sure of this is harder and requires appropriate circuit simulations. The second reason to be aware of regarding the parasitic substrate PN diode is its associated junction capacitance. Depending on the circuit configuration, this can seriously affect the dynamics of the diode and needs to be part of the design process.

An additional consideration in an integrated Schottky diode is its series resistance. $R_s$ affects the performance of a Schottky diode in just about any circuit application. All components of the series resistance must be well understood and modeled. In Sec. 6.7.2 we showed how to compute the series resistance of a PN diode in a typical case. The procedure is very much geometry dependent but there is no fundamental difference between the Schottky diode and the PN diode. Refer to Secs. 6.7.2 and 7.6.1 for more details.

Given that in a typical IC process there is some freedom to choose the body and the metal of an integrated Schottky diode, it is important to critically think through this and be mindful of the trade-offs involved. To discuss this, consider the I-V characteristics of a Schottky diode in

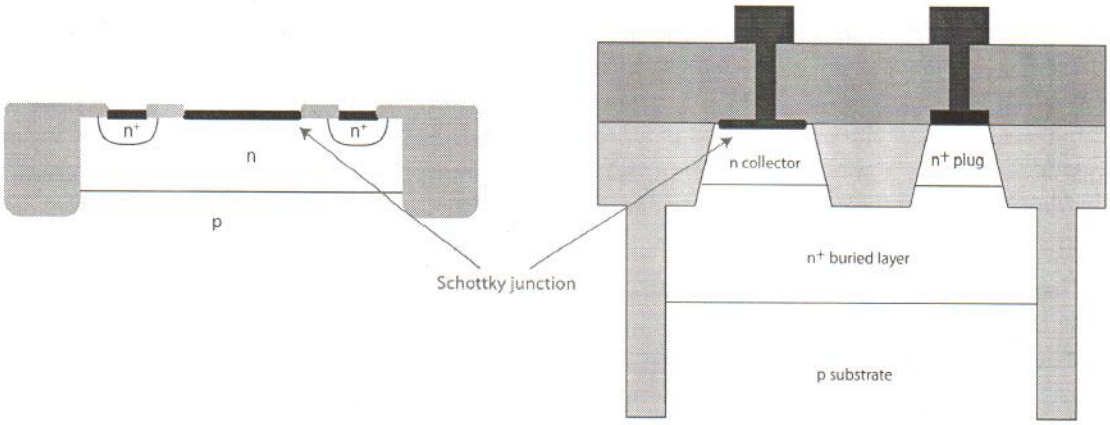



Figure 7.22: Schematic cross sections of two integrated Schottky diodes fabricated on a CMOS process (left) and on a bipolar process (right).

Fig. 7.23. For the purpose of this discussion, let us assume that what is desired is a diode that switches from an ON state where it must support a certain amount of forward bias current $I_f$ while dropping a very low forward voltage $V_f$, and an OFF state where the diode must be capable of blocking a sufficient reverse voltage $V_r$ while conducting a very small amount of reverse bias current.

The choice of Schottky metal is of prime importance. The higher the Schottky barrier height of the metal, the more forward voltage in the ON state is needed to supply a given current, but the smaller the reverse leakage current in the OFF state. Also, a high Schottky barrier height results in increased temperature sensitivity of the I-V characteristics. The doping level in the bulk is also a key design parameter. A high doping level minimizes series resistance and switching speed but results in more capacitance and a smaller breakdown voltage.

The vertical extension of the bulk is also important. In designs with a buried layer, such as the one shown on the right of Fig. 7.22, the vertical thickness of the bulk of the diode should be enough to entirely accommodate the depletion region at breakdown. Otherwise, premature breakdown will occur. Excessive body thickness, on the other hand, adds series resistance.

Left in the design process is selecting the area of the diode. This is the prerogative of the circuit designer who does this taking into consideration the specific requirements of each diode in its circuit context. The smaller the diode area, the lower the capacitance and the lower the reverse leakage current. A smaller diode footprint is also more economical. However, in small diodes, the series resistance and the forward voltage that is required to deliver a certain current are larger.

## 7.8 Ohmic contacts

Electrical current flows in and out of microelectronic devices through its ohmic contacts. A perfect ohmic contact does not present any resistance to current. Real ohmic contacts, on the

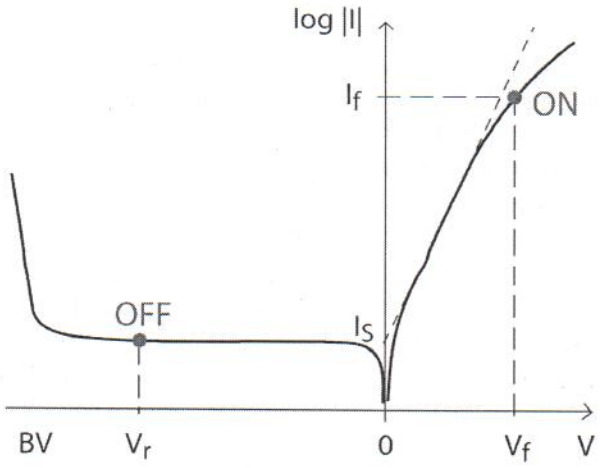



Figure 7.23: I-V characteristics of Schottky diode showing the two typical operating points in a switching application. For small-signal applications, the device is usually biased in the ON state.

other hand, drop a small voltage when current flows and therefore present a parasitic resistance. If excessive, this can degrade device performance.

Ohmic contacts are in essence metal/semiconductor junctions with a very large saturation current density. They can support substantial current in forward or reverse bias with little voltage across. When $V \ll kT/q$ in a metal-semiconductor junction, the exponential in the $J-V$ characteristics can be linearized ($e^{qV/kT} \simeq 1 + qV/kT$), to obtain:

$$J \simeq A^*T^2 \exp\frac{-q\varphi_{Bn}}{kT}\frac{qV}{kT} = \frac{V}{\rho_c} \tag{7.47}$$

In the limit of small $V$, the current through the ohmic contact is *linear* in $V$ for forward and reverse bias. This allows us to define a figure of merit for an ohmic contact called the *ohmic contact resistivity*, $\rho_c$, as in Eq. 7.47. The units of $\rho_c$ are $\Omega \cdot cm^2$. A good ohmic contact is characterized by a very small value of $\rho_c$. For Si, good ohmic contacts display $\rho_c \leq 10^{-7}\ \Omega{\cdot}cm^2$ (or $10\ \Omega.\mu m^2$, an easy number to remember). This means that when a current density of $10^5\ A/cm^2$ flows through them, the ohmic drop across the contact is only $10\ mV$.

How does a metal-semiconductor junction become a good ohmic contact? Eq. 7.47 seems to indicate that the only degree of freedom provided to the engineer is the selection of a metal with a very small Schottky barrier height. While that certainly helps, something else can also be done that is far more effective, *i.e.*, increasing the doping level right below the metal. To understands the reasons for this, we need to look beyond thermionic emission limited transport.

As the doping level increases in a metal-semiconductor junction, the extent of the depletion region decreases. This can be seen in equilibrium in Eq. 7.15. If you put numbers to this expression, you will realize that for a typical $\phi_{bi}$ of $0.5\ V$, at a doping level of about $N_D \simeq 10^{19}\ cm^{-3}$, the depletion region thickness in thermal equilibrium is of a magnitude comparable to the de Broglie wavelength at room temperature (which is about $7.6\ nm$, see Section 1.1.2). At

these dimensions, electron tunneling suddenly becomes rather likely.

Tunneling is a quantum mechanical phenomenon that enables carriers with insufficient energy to "bore" through energy barriers. While classically forbidden, the quantum mechanical nature of the electron allows tunneling to happen with a finite probability. The energy band diagrams of Fig. 7.24 schematically illustrate tunneling in n-type and p-type ohmic contacts. In this figure, the Fermi level has been drawn inside the conduction band for n-type semiconductor and the valence band for the p-type semiconductor to indicate the fact that tunneling becomes prevalent when the doping levels are degenerate. Examining an n-type ohmic contact first (left of Fig. 7.24), the application of a voltage across the contact misaligns the Fermi level on the metal with respect to the semiconductor. For both polarities of the voltage, this presents electrons on the semiconductor or the metal with empty states on the other side of the metal-semiconductor barrier. By tunneling through the barrier, electrons can then lower their energy. In this manner current flows. The thinner the barrier, the higher the tunneling probability and the higher the current that will flow through the contact for a given voltage.

For a p-type ohmic contact (right of Fig. 7.24) we need to examine this picture in somehow more detail since conduction in the metal takes place through electrons, but conduction through the semiconductor involves holes. When the quasi-Fermi level for holes on the semiconductor is slightly above the Fermi level on the metal, electrons in the valence band of the semiconductor can tunnel to empty states into the metal. This leaves holes behind on the semiconductor that can now support the current away from the contact into the body of the semiconductor. For the reverse polarity, electrons in the metal tunnel into the valence band of the semiconductor consuming holes and terminating a hole current that flows from the body of the semiconductor into the contact.

A property of tunneling is that the tunneling current depends exponentially on the inverse thickness of the barrier. Once tunneling becomes significant, the tunneling probability increases very quickly with further enhancements of the doping level. This is then by far the most effective and practical way of engineering low resistivity ohmic contacts. Using metals with small Schottky barrier heights also helps, but it is not required. Fig. 7.25 shows ohmic contact resistance as a function of doping level for PtSi and W ohmic contacts on n-Si. As seen, $\rho_c$ decreases steeply at high doping levels and reaches values of the order of $10^{-8}\ \Omega.cm^2$ at about a doping level in the mid $10^{20}\ cm^{-3}$ regime.

In a well designed device, wherever an ohmic contact is needed, a heavily doped region is introduced. The contact resistance of an ohmic contact of area $A_c$ is to the first order:

$$R_c = \frac{\rho_c}{A_c} \tag{7.48}$$

The bigger the contact area, the smaller the contact resistance.

Eq. 7.48 applies only if the current flow is normal to the metal-semiconductor interface across the entire contact area. This equation, however, is not suitable for lateral contacts, that is, contacts in which the current in the semiconductor ends up flowing parallel to the semiconductor surface. These are very common contacts in semiconductor devices and require special treatment.

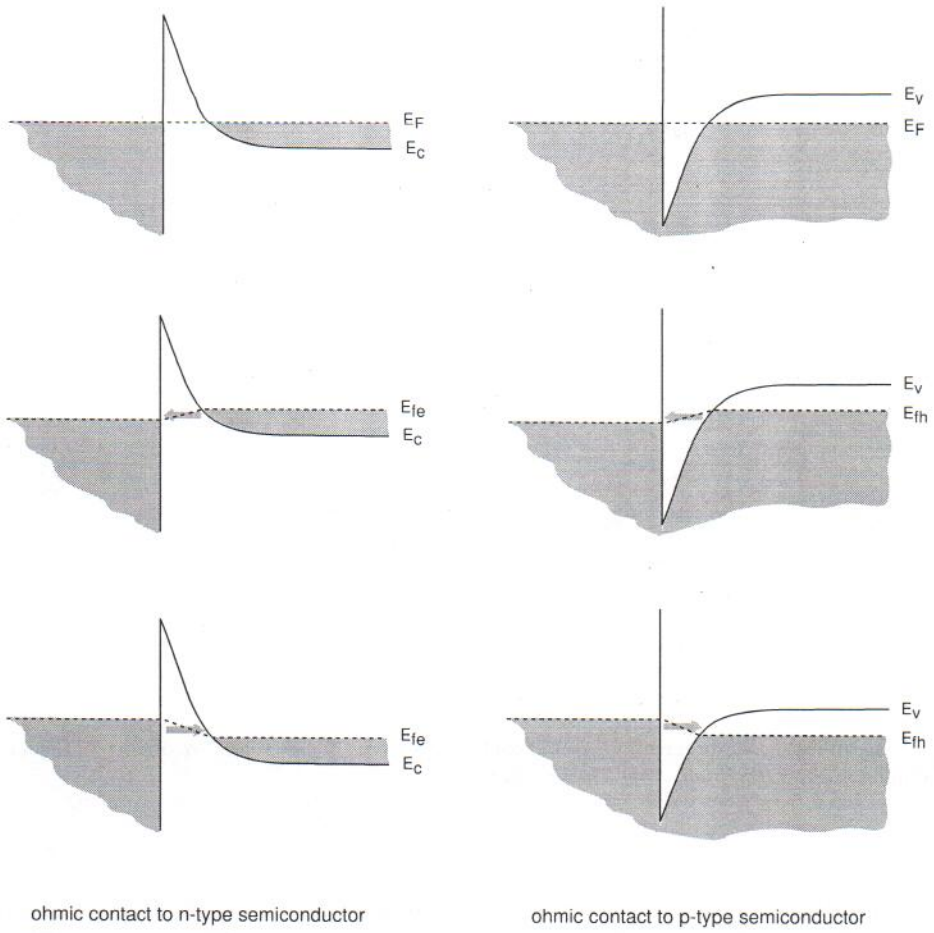



Figure 7.24: Energy band diagrams illustrating electron tunneling in n-type (left) and p-type (right) ohmic contacts. Top row represents thermal equilibrium. Second row and third row represent situations for $V > 0$ and $V < 0$, respectively. The arrow represents direction of electron tunneling.

This is presented next.

### 7.8.1 Lateral ohmic contact: transmission-line model

Consider the situation of Fig. 7.26. It depicts two ohmic contacts applied to a thin $n^+$ layer on a p-type substrate. When a voltage is applied to one ohmic contact with respect to the other, current flows down one contact, then laterally through the $n^+$ layer and then up the other contact. The current is confined to the $n^+$ layer due to the presence of the PN junction. This kind of lateral contact structure is widely used in devices. Of this type are, for example, the source and drain contacts of a MOSFET and the contacts to the base of a bipolar transistor.

An equivalent circuit model of this structure is shown in the bottom diagram of Fig. 7.26. We can think of this as three resistances in series. Two are associated with the contacts and one with

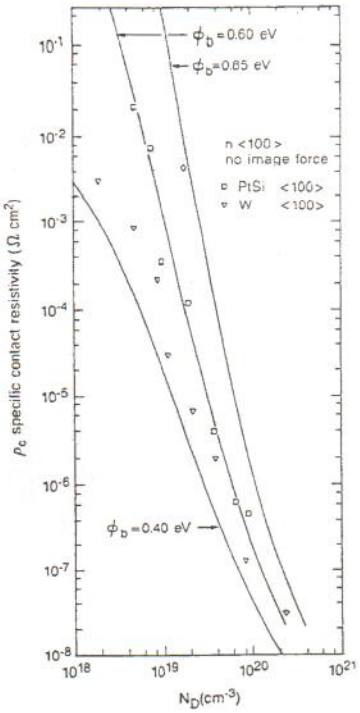



Figure 7.25: Experimental measurements of ohmic contact resistivity for PtSi and W ohmic contacts on n-Si as a function of doping level [*S. Swirhun, PhD Thesis, 1986*].

the lateral $n^+$ region in between. The lateral contact resistance, $R_c$, lumps the resistance all the way from the ohmic metal to the leading edge of the contact (dashed line). We wish to derive an expression for $R_c$ in terms of the relevant parameters. The ohmic contacts are characterized by a contact resistivity $\rho_c$. The length of the contacts is $L_c$ and their width is $W_c$ (into the paper). The $n^+$ layer has a sheet resistance $R_{sh}$.

It is clear that $R_c$ is not given by an expression like Eq. 7.48. That expression implicitly assumes that the current density is uniform across the metal/semiconductor interface. Looking at Fig. 7.26, the situation here is quite different. Because of the finite resistance of the semiconductor layer, the current will tend to bunch on the leading edge of the contact. The current density will be higher there and drop as we proceed further away from that edge. If the contact is long enough, the current density will become negligible sufficiently far away from the leading edge of the contact. Clearly, the resistance of this contact (as defined above) is higher than given by Eq. 7.26 on two counts. First, current only flows through a fraction of the contact area. Second, there is a contribution to the contact resistance that comes from semiconductor.

This is an interesting problem that is best solved through what is called a *transmission-line model*. This is an effective way to analyze distributed problems such as this one. A transmission-line equivalent circuit of the contact structure on the right (from the dashed line on towards the right) is shown in Fig. 7.27.

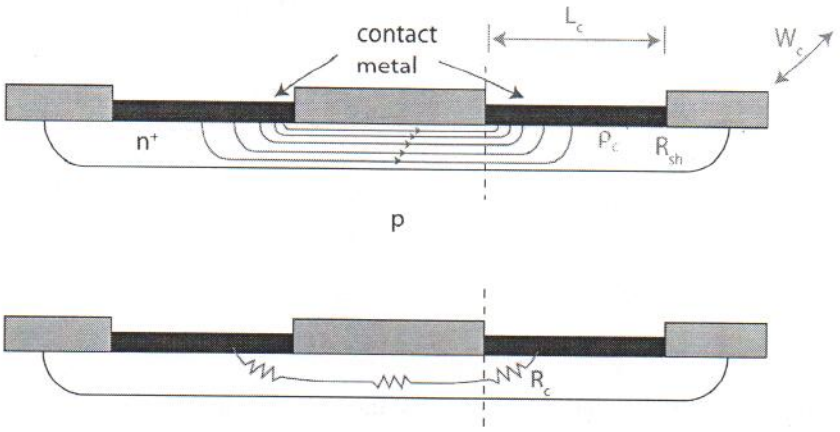



Figure 7.26: Top: schematic of current flow in lateral ohmic contact structure. Bottom: equivalent resistance model defining the contact resistance.

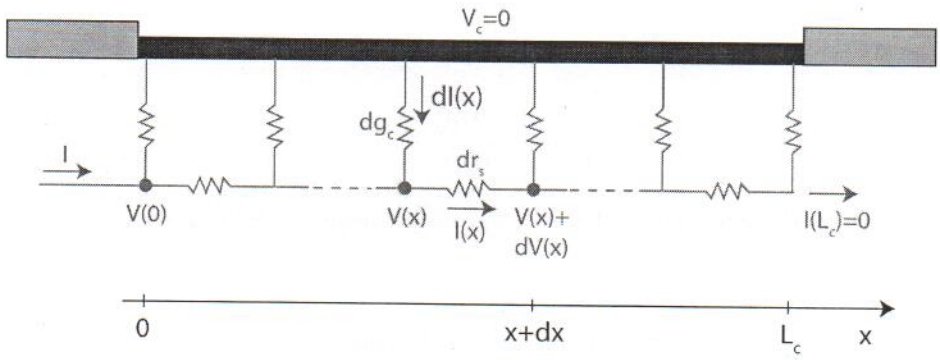



Figure 7.27: Transmission-line equivalent circuit model for ohmic contact on right (from dashed line) of Fig. 7.26.

This electrical problem can be described by a distributed resistance network as shown in Fig. 7.27. A pair of resistors is associated with each elemental length of the contact $dx$. In each segment, we can define a differential contact *conductance* as $dg_c = (W_c/\rho_c)dx$ (the conductance of the elemental contact vanishes as $dx$ goes to zero) and a differential lateral semiconductor resistance as $dr_s = (R_{sh}/W_c)dx$. We assume that a current $I$ flows into this contact structure from the left. $x = 0$ is the leading edge of the contact and we place its voltage at $V(0)$. The contact has a length $L_c$. At the other end of the contact, no current can emerge and $I(L_c) = 0$. With these definitions, the contact resistance is $R_c = V(0)/I$. In order to get $R_c$, we need to solve for $V(x)$ and $I(x)$.

At location $x$, Kirchoff voltage and current laws yield, respectively:

$$dV(x) = -I(x)dr_s = -I(x)\frac{R_{sh}}{W_c}dx \tag{7.49}$$

$$dI(x) = -V(x)dg_c = -V(x)\frac{W_c}{\rho_c}dx \tag{7.50}$$

From this, we can write:

$$\frac{dV(x)}{dx} = -I(x)\frac{R_{sh}}{W_c} \tag{7.51}$$

$$\frac{dI(x)}{dx} = -V(x)\frac{W_c}{\rho_c} \tag{7.52}$$

Taking a derivative on both sides of Eq. 7.52 and substituting Eq. 7.51 yields:

$$\frac{d^2I(x)}{dx^2} - \frac{I(x)}{L_t^2} = 0 \tag{7.53}$$

Here, we have defined the *transfer length* as:

$$L_t = \sqrt{\frac{\rho_c}{R_{sh}}} \tag{7.54}$$

It is best to write the solutions to this second order linear differential equation as a sum of hyperbolic functions:

$$I(x) = I_1 \sinh(\frac{x}{L_t}) + I_2 \cosh(\frac{x}{L_t}) \tag{7.55}$$

We obtain coefficients $I_1$ and $I_2$ by matching boundary conditions $I(0) = I$ and $I(L_c) = 0$ to finally obtain:

$$I(x) = I[\cosh(\frac{x}{L_t}) - \coth(\frac{L_c}{L_t})\sinh(\frac{x}{L_t})] \tag{7.56}$$

Inserting this result into Eq. 7.52 yields the voltage distribution along the semiconductor:

$$V(x) = \frac{\sqrt{\rho_c R_{sh}}}{W_c} I[\coth(\frac{L_c}{L_t})\cosh(\frac{x}{L_t}) - \sinh(\frac{x}{L_t})] \tag{7.57}$$

At $x = 0$, this equation yields:

$$V(0) = \frac{\sqrt{\rho_c R_{sh}}}{W_c} I \coth(\frac{L_c}{L_t}) \tag{7.58}$$

The contact resistance of this contact, defined as $V(0)/I$ is:

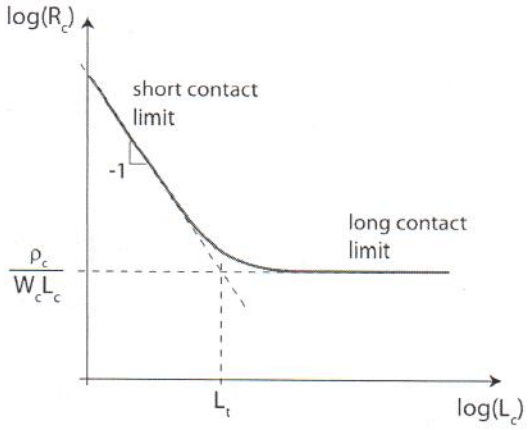



Figure 7.28: Evolution of contact resistance with contact length in a lateral ohmic contact.

$$R_c = \frac{\sqrt{\rho_c R_{sh}}}{W_c} \coth(\frac{L_c}{L_t}) \tag{7.59}$$

We find that, in general, the contact resistance of a lateral contact depends on the contact resistivity itself, the sheet resistance of the semiconductor layer underneath and the length and width of the contact. Note how $R_c \sim 1/W_c$, as it should be.

Fig. 7.28 sketches the dependence of $R_c$ on $L_c$, the length of the contact. There are two interesting limits to this expression. If the contact length is much shorter than the transfer length, $L_c \ll L_t$, we can consider this a *short contact* and $R_c$ becomes (see Taylor series expansion of $\coth x$ around $x = 0$ in Appendix D):

$$R_c \simeq \frac{\rho_c}{W_c L_c} \tag{7.60}$$

In this limit, the resistivity of the contact dominates the contact resistance and the contribution to this from the resistance of the semiconductor is negligible because the contact is very short. Notice how in this case, the current density is uniform across the contact.

The other limit is that of the *long contact*. In this case, $L_c \gg L_t$. Since for $x \gg 1$, $\coth x \simeq 1$, we have:

$$R_c \simeq \frac{\rho_c}{W_c L_t} \tag{7.61}$$

In this case, the contact resistance scales with the transfer length $L_t$ and is independent of the actual contact length. In this limit, the transfer length becomes the effective contact length. For long contacts, the current and the voltage drop exponentially away from the leading edge of the contact. After a few transfer lengths, there is no more current left and the actual length of the contact is irrelevant.

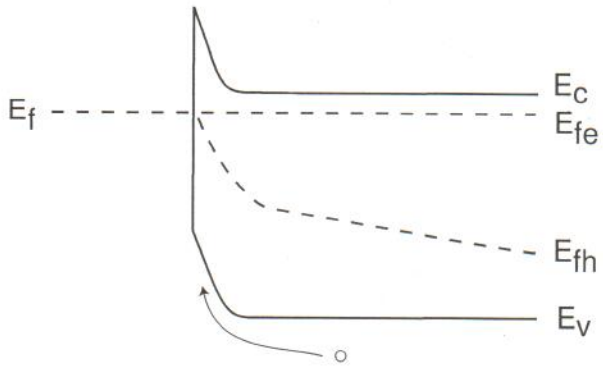



Figure 7.29: Energy band diagram of an ohmic contact to an n-type semiconductor with excess hole concentration in its vicinity. The electric field associated with the metal-semiconductor interface effectively "pulls" holes towards it where they recombine.

Understanding the scalability of lateral ohmic contacts is important in device design. It is obvious that one does not want to design a contact that is longer than two or three transfer lengths as this does not reduce the contact resistance but it increases the area that the contact consumes.

### 7.8.2 Boundary conditions imposed by ohmic contacts

We are now ready to justify two assumptions made elsewhere in this book about ohmic contacts. The first one refers to our statement that through an ohmic contact, we can "grab" on the majority carrier Fermi level at the semiconductor surface. Fig. 7.24 justifies this assertion. It is clear in this figure that if the contact resistance is very small, only a small split of the Fermi level between the metal and the semiconductor is required to support a sizable current. In consequence, if the Fermi level of the metal is pulled up or down by a battery, the Fermi level of the semiconductor is brought along with it.

The second assumption refers to the excess minority carrier concentration at an ohmic contact. We assumed in Ch. 5 that it is always zero, or in other words, that the minority carrier concentration is in equilibrium at an ohmic contact. This is now easy to understand. Fig. 7.29 shows the energy band diagram of an ohmic contact on an n-type semiconductor in a situation where there are excess holes in the neighborhood. The electric field associated with the metal-semiconductor junction pulls holes from the semiconductor towards its interface with the metal where they recombine due to the high density of defects and dangling bonds. In consequence, in the vicinity of the ohmic contact, the hole concentration is never too far away from its equilibrium value. The same applies for electrons in a p-type ohmic contact.

## 7.9 Summary

- At the interface of a typical metal-semiconductor junction, charge redistribution takes place. A consequence of this is that a depletion region is created on the semiconductor side. A second consequence is that an energy barrier to majority carrier flow appears between the metal and the semiconductor.
- Under bias, the barrier that blocks carrier flow from the semiconductor to the metal is modulated by the applied voltage, while the barrier in the reverse direction is unchanged. This results in the rectifying behavior of the Schottky diode.
- In most circumstances, current in a Schottky diode in forward bias is dominated by the process of emission of carriers into the metal.
- Under bias, the extension of the depletion region in the semiconductor of a Schottky diode is changed by the applied voltage. This produces a capacitive effect.
- Under most circumstances, a Schottky diode under bias does not store minority carriers. This is the key reason for their fast response to dynamic signals. The dominant time constant of a Schottky diode is the RC time constant of the depletion capacitance and the total series resistance.
- Practical design of Schottky diodes attempts to achieve a balance among a small time constant, sufficient forward conduction, small reverse current, high maximum reverse voltage and small area. The challenge to the microelectronics device designer is to accomplish this constrained by the use of established processes. This is because dedicated processes are rarely available to the fabrication of Schottky diodes.
- Good ohmic contacts present a negligible barrier to current flow in and out of the semiconductor. The key to accomplish this is to heavily dope the semiconductor directly underneath the metal. The higher the doping level, the lower the contact resistivity, the most important figure of merit of an ohmic contact.
- In lateral ohmic contacts, the sheet resistance of the semiconductor layer underneath the contact plays a role. There is a transfer length beyond which longer contacts do not result in smaller contact resistance.

## 7.10 Further reading

**Metal-Semiconductor Contacts and Devices** by S. S. Cohen and G. Sh. Gildenblat, VLSI Electronics Microstructure Science, Vol. 13, Academic Press 1986 (ISBN 0-12-234113-9, TK7874.V56 vol. 13). This high-level book is entirely dedicated to Schottky structures. Several chapters are of interest to us. Ch. 2 describes a generalized theory of transport in the metal-semiconductor junction including drift and diffusion, tunneling, and minority-carrier injection. Ch. 3 presents experimental techniques to determine the Schottky barrier height. Ch. 4, 5, and 6 deal with ohmic contact characterization and technology. Although the technology portions are quite dated, the book is still a useful reference for fundamental concepts.

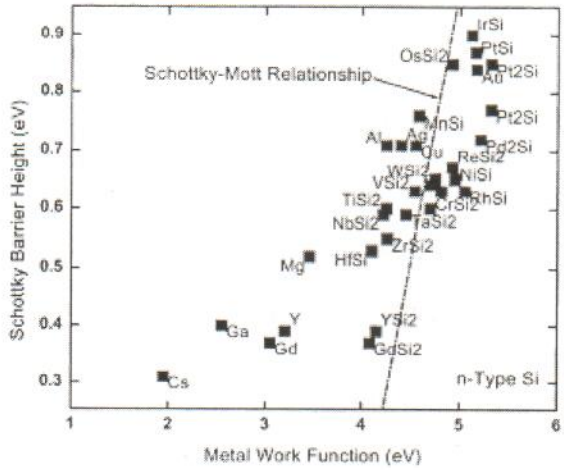



Figure 7.30: Schottky barrier height of various metals on n-type Si as a function of the metal work function. The Schottky-Mott relationship (Eq. 7.3) is indicated [from *R. T. Tung, Mat. Sci. and Eng., R35, 1 (2001)*].

## AT7.1 Non-ideal Schottky barrier height of metal-semiconductor junctions

When discussing the band line up of a metal/n-type semiconductor junction in thermal equilibrium in Sec. 7.2.2, we concluded that the Schottky barrier height is given by the difference in the work function of the metal and the electron affinity of the semiconductor (Eq. 7.3). This is very relevant from an engineering point of view as it suggests that by appropriate selection of the metal, a wide range of Schottky barrier heights should be available to the device designers.

The reality is somehow different. This can be seen in Fig. 7.30 that graphs the experimental Schottky barrier height of metal-semiconductor junctions prepared with different metals on n-type Silicon. This figure shows that $q\varphi_{Bn}$ tends to increase as $W_M$ increases but the dependence is significantly weaker than expected from simple theory (line). A wide distribution of Schottky barrier heights is available by using different metals but the range is much narrower than expected from the Schottky-Mott relation.

Understanding at a fundamental level the situation depicted in Fig. 7.30 is an active topic of research in modern solid-state physics. In this book, we cannot do justice to the various theories that are currently under consideration. Nevertheless, it is important for device designers to develop an appreciation of some of the critical issues involved and how to think about energy band line ups that are not ideal (as defined in the main body of this chapter). This is the purpose of this Advanced Topic.

When one stops to think about it, it is not surprising that the simple approach to predicting Schottky barrier heights described in Sec. 7.2.2 fails. Through Eq. 7.3, the Schottky barrier height, which is a purely interfacial property of a metal and a semiconductor, is calculated using information from the bulk characteristics of its constituents. This approach ignores the very

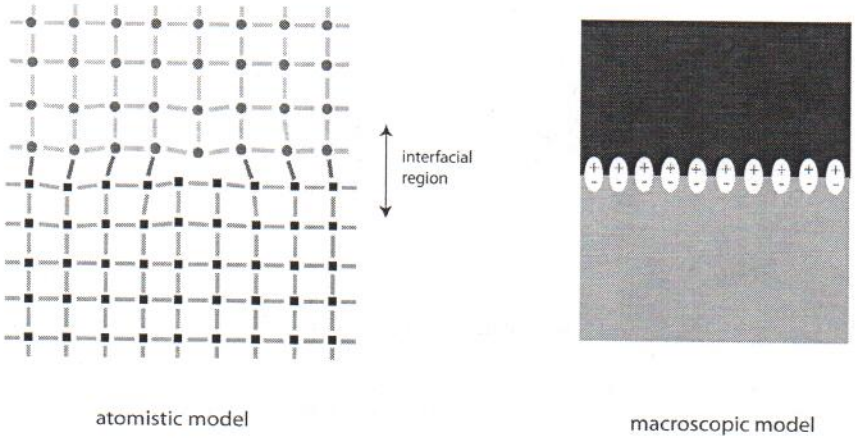



Figure 7.31: Models for metal-semiconductor interface. Left: atomistic model showing interfacial bonding, dangling bonds on either side and lattice deformation that can extend a few monolayers into each material. Right: macroscopic model showing a fixed dipole of charge at the interface.

presence of an interface between these two materials. It is right at that interface where everything happens.

Just about every metal under almost any circumstance reacts in some form when deposited over a semiconductor. In careful experiments it has been observed that under the right conditions, the Schottky barrier height of a given metal-semiconductor junction depends on the cleaning process, the deposition and post-deposition annealing techniques that are used as well as the surface orientation of the semiconductor and the thickness of the metal. This strongly suggests that the details of the atomic bonding and interfacial structure of the metal-semiconductor interface should affect the Schottky barrier height. Atomistic calculations of metal-semiconductor interfaces actually confirm this. How can we understand this?

The interface between a metal and a semiconductor is actually a messy affair. A conceptual sketch is shown on the left of Fig. 7.31. This figure illustrates how, in general, the semiconductor and the metal lattices are different (in fact, metals are often polycrystalline) and as a result, atoms do not precisely line up across the interface. Right at the boundary, some metal-semiconductor atomic bonds are formed leaving some dangling bonds on either side. Also, the crystal lattices on both sides get deformed to a few monolayers in depth into each material. It is then reasonable to expect that the interfacial region of a metal semiconductor junction ends up with different electronic properties than the bulk of either material.

From an electrical point of view, this relatively disorganized interface region when averaged over enough area is very likely to produce a charge dipole across the interface (right of Fig. 7.31). This does not mean that there is an overall loss of charge neutrality but that there is a slight separation of positive and negative charge at the interface (electrons and core ions) with a few more electrons ending up on one side than on the other. The sign and strength of this dipole depends on the atomic details of the interface.

A dipole of charge at the interface results in a sharp electric field confined to a small region

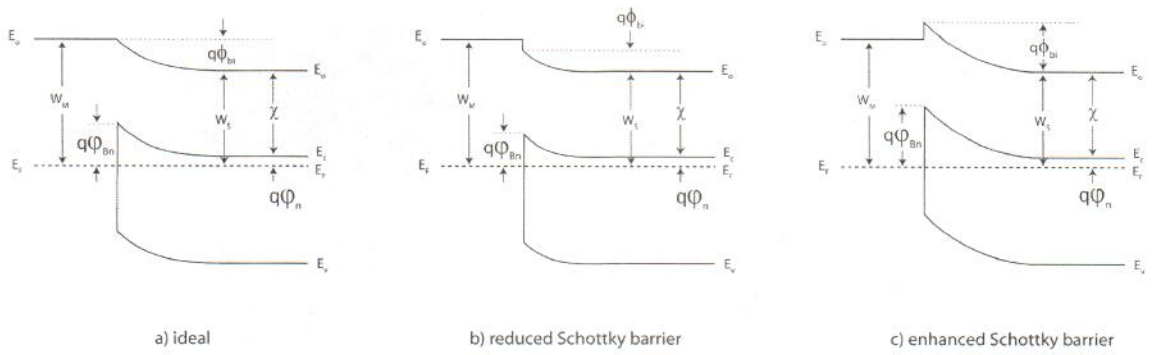



Figure 7.32: Energy band diagrams of metal-semiconductor interfaces in the presence of interfacial charge dipole. Left: ideal case; middle: charge dipole reduces Schottky barrier height; right: charge dipole enhances Schottky barrier height.

and a small but finite electrostatic potential buildup that affects the overall band line up. From simple electrostatics, a charge dipole with a density $N_\delta$ and an average charge separation $d$ gives rise to a potential buildup given by:

$$\Delta\phi_\delta = \frac{qN_\delta d}{\epsilon} \tag{7.62}$$

We can put some numbers to this expression. If we have a charge dipole density of $10^{13}\ cm^{-2}$ with an average charge separation of 1 $nm$ and we use the permittivity of Si, the potential build-up at the interface ends up being 0.16 V. This is well in line with the discrepancies that we see between the simple theory and the experiments in Fig. 7.30. Note that a dipole density of $10^{13}\ cm^{-2}$ is not that large when compared with the atomic density of a typical semiconductor surface. For example, in (100) Si, the atomic density is $6.7 \times 10^{14}\ cm^{-2}$.

How is the band diagram modified if there is a charge dipole at the metal-semiconductor interface? This is shown in Fig. 7.32. Depending on the sign of the interfacial dipole, the Schottky barrier height can be reduced or enhanced with respect to the ideal value. As we have studied before, in the absence of a charge dipole, the difference between vacuum levels in the metal and the semiconductor of a metal-semiconductor junction in thermal equilibrium is given by the difference in work functions. The Schottky barrier height is then the difference between the metal work function and the electron affinity (left figure). Depending on the sign of the interfacial charge dipole, the potential build up across the semiconductor will have to adapt so that the overall vacuum level difference ends up unchanged. This can then reduce (middle figure) or enhance (right picture) the Schottky barrier height with respect to the ideal value.

The key lesson from this section is that the Schottky barrier height of a metal-semiconductor system is strongly affected by the details of the interfacial region and should be measured in a practical situation. This understanding also opens the possibility for interface engineering as a way to tune the Schottky barrier height of a metal-semiconductor pair.

A final observation. Remember how for a given metal-semiconductor pair, we concluded in Sec. 7.2.2 that the Schottky barrier height on the n-type semiconductor and on the p-type

semiconductor add up to the bandgap (Eq. 7.7). The presence of a dipole of charge at the metal-semiconductor interface is evidently a property of the metal and the semiconductor and the details of fabrication and has nothing to do with the doping type of the semiconductor. As a result, the shifts that the interfacial charge dipole introduces are identical for both semiconductor types and they cancel out when the Schottky barrier heights are added. Eq. 7.7 remains valid.

## AT7.2 Drift-diffusion model for I-V characteristics

This Advanced Topic presents the drift-diffusion model for the I-V characteristics of the Schottky diode. Here we think in terms of a metal/n-type semiconductor Schottky diode that is in forward bias. The drift-diffusion model assumes that the rate limiting step to the current in the device is the process of electron transport across the SCR and not the process of electron emission into the metal.

Under forward bias, as discussed in the main body of this chapter, diffusion prevails over drift in the SCR and there is a net flow of electrons from the body of the device, across the SCR towards the metal/semiconductor interface. If we continue to neglect the role of minority carrier holes, the diode current is all supported by electrons and, in general, can be written as the sum of its drift and a diffusion components:

$$J_t \simeq J_e = q\mu_e n\mathcal{E} + qD_e\frac{dn}{dx} = qD_e(-\frac{qn}{kT}\frac{d\phi}{dx} + \frac{dn}{dx}) \tag{7.63}$$

We can multiply both sides of this expression by $\exp(-\frac{q\phi}{kT})$, to get:

$$\begin{aligned} J_t \exp(-\frac{q\phi}{kT}) &= qD_e[-\frac{qn}{kT}\frac{d\phi}{dx}\exp(-\frac{q\phi}{kT}) + \frac{dn}{dx}\exp(-\frac{q\phi}{kT})] \\ &= qD_e\frac{d}{dx}[n\exp(-\frac{q\phi}{kT})] \end{aligned} \tag{7.64}$$

We now integrate across the SCR:

$$J_t\int_{0^+}^{x_d}\exp(-\frac{q\phi}{kT})dx = qD_e n\exp(-\frac{q\phi}{kT})|_0^{x_d} \tag{7.65}$$

where we have brought $J_t$ out of the integral because in steady state, the total current is constant in space.

To perform the integral on the left-hand side, we need an expression for $\phi(x)$. For a general case under bias, starting from Eq. 7.13, we can easily write:

$$\phi(x) = -(\phi_{bi} - V)(\frac{x^2}{x_d^2} - \frac{2x}{x_d} + 1) \qquad \text{for } 0 \le x \le x_d \tag{7.66}$$

Integrating this does not yield an analytical solution. However, since the biggest contribution to the integral is around $x = 0$, we can expand $\phi(x)$ in that region:

$$\phi(x) \simeq -(\phi_{bi} - V)(1 - \frac{2x}{x_d}) \qquad \text{around } x = 0 \tag{7.67}$$

The integral can now be easily performed:

$$\int_{0^+}^{x_d} \exp(-\frac{q\phi}{kT})dx \simeq \frac{kTx_d}{2q(\phi_{bi} - V)} \exp\frac{q(\phi_{bi} - V)}{kT} \tag{7.68}$$

For the right-hand side, we note that $x = x_d$ is the edge of the depletion region where $\phi(x_d) = 0$ and $n(x_d) = N_D$. $x = 0$ is the metal-semiconductor interface where $\phi(0^+) = -(\phi_{bi} - V)$. Under the drift-diffusion approximation, the electron concentration at $x = 0^+$ is unperturbed from equilibrium. This is consistent with the notion that the rate limiting step is electrons reaching the interface. Electron emission into the metal can happen at much faster rate. We can then use the result from thermal equilibrium that was given in Eq. 7.20. all together, we have:

$$qD_e n \exp(-\frac{q\phi}{kT})|_0^{x_d} = qD_e N_D(1 - \exp\frac{-qV}{kT}) \tag{7.69}$$

Assembling now Eqs. 7.65, 7.68 and 7.69 and solving for $J_t$, we get:

$$J_t = q\mu_e\sqrt{\frac{2q(\phi_{bi} - V)N_D}{\epsilon}}N_c \exp\frac{-q\varphi_{Bn}}{kT}(\exp\frac{qV}{kT} - 1) \tag{7.70}$$

where we have again used Eq. 7.19 as well as 7.21.

Multiplying by the area $A_j$ allows us to formulate again the I-V characteristics in the classical rectifying form of Eq. 7.35.

The key dependencies of our result in Eq. 7.70 are very much similar to those of the thermionic emission model given by 7.34. It has the same rectifying voltage dependence and the same exponential dependence on the Schottky barrier height. Actually, if we look at the reverse bias current predicted by this model the result is very intuitive. For large enough reverse bias, Eq. 7.70 simplifies to:

$$J_t \simeq -q\mu_e\sqrt{\frac{2q(\phi_{bi} - V)N_D}{\epsilon}}N_c \exp\frac{-q\varphi_{Bn}}{kT} = -q\mu_e|\mathcal{E}_{max}|n(0) \tag{7.71}$$

The reverse bias current is entirely supported by drift at the metal-semiconductor interface and it then takes the classic drift current expression with the electron concentration and the electric field computed at $x = 0$.

It is interesting to look at the ratio of the current predicted by the drift-diffusion model and the thermionic emission model. Using Eqs. 7.70 and 7.34, after some simple algebra we can find:

$$\frac{J_t(DD)}{J_t(TE)} = 4\frac{\mu_e|\mathcal{E}_{max}|}{v_{the}} \tag{7.72}$$

This is really interesting and also makes good sense. The ratio of the currents can be phrased as a ratio of velocities. In the case of the drift-diffusion model, the velocity that matters is the drift velocity at the metal semiconductor interface. For the thermionic emission model, it is a quarter of the thermal velocity of electrons that matters. Whichever one of these velocities is the smallest limits current transport. Since the peak electric field in a Schottky diode can be quite high even under strong forward bias, in materials in which $\mu_e$ is relatively high, it is often the case that thermionic emission is the most restrictive process. The thermionic emission process is therefore the appropriate model to use to describe the I-V characteristics of Schottky diodes.

**Exercise 7.5**: *Evaluate the ratio of the current obtained in an Al/n-Si Schottky diode with* $N_D = 10^{17}\ cm^{-3}$ *at a forward voltage of* 0.5 *V at 300K by the drift-diffusion and thermionic emission models.*

We assume that the electron mobility in the SCR is identical to that of a QNR of the same doping level. Therefore, for this doping level, $\mu_e \simeq 700\ cm^2/V.s$.

From Exercise 7.3 we know that for this Schottky diode at this forward voltage, the peak electric field at the Schottky interface is $|\mathcal{E}_{max}| \simeq 3.4 \times 10^4\ V/cm$.

We also know that the thermal velocity for electrons at 300 $K$ is about $v_{the} \simeq 2.0 \times 10^7\ cm/s$.

All together, then:

$$\frac{J_t(DD)}{J_t(TE)} = 4\frac{\mu_e\mathcal{E}_{max}}{v_{the}} = 4 \times \frac{700\ cm^2/V.s \times 3.4 \times 10^4\ V/cm}{2.0 \times 10^7\ cm/s} = 4.8$$

Under these conditions, the drift-diffusion model predicts a current that is about 4.8 times higher than the thermionic emission model. Therefore, the process of emission from electrons into the metal is the rate limiting step and the thermionic emission model is the correct one to use here. This finding is consistent with the verification of the validity of the Bethe condition that was carried out in Exercise 7.4.

## AT7.3 Equivalent circuit model of Schottky diode for circuit design

In the modeling of Schottky diodes for circuit CAD, the equivalent circuit model of a PN diode is often used. This works well because the behavior of these two devices is very similar. Referring to Fig. 6.33, in its simplest form, this circuit consists of an ideal diode, a capacitor in parallel with it, and a series resistance.

Several of the model parameters in the diode CAD model have a similar meaning when describing the PN diode and the Schottky diode. Referring to Sec. AT6.2 that describes the equivalent circuit model for a PN diode for circuit CAD, this is the case for model parameters **N**, **RS**, **IS**, **CJO**, **VJ**, **M** and **BV**.

Other model parameters take on a different meaning due to the subtle difference in physics

| name | parameter description | units | ideal value |
|---|---|---|---|
| **IS** | saturation current | $A$ | - |
| **N** | ideality factor | - | 1 |
| **EG** | Schottky barrier height | V | - |
| **RS** | series resistance | $\Omega$ | - |
| **CJO** | zero bias depletion capacitance | $F$ | - |
| **VJ** | built-in potential | $V$ | - |
| **M** | grading coefficient | - | 0.5 |
| **XTI** | saturation current temperature exponent | - | 2 |
| **TT** | transit time | $s$ | 0 |
| **BV** | breakdown voltage | $V$ | - |

Table 7.1: Summary of simplest set of Schottky diode model parameters for circuit CAD. Ideal value refers to the simple Schottky diode described in these pages.

between both diodes. For example, looking at the expression of **IS** of Eq. 6.141, and comparing it with $I_S$ for the Schottky diode derived in Eq. 7.36, it is clear that model parameter **EG** plays the role here of the Schottky barrier height. Its value should therefore be smaller than in a PN diode where it plays the role of the bandgap. Also, model parameter **XTI** ideally should have a value of 2, which tends to be smaller than the values needed in PN diodes.

The biggest difference between the Schottky diode model and the PN diode model is in parameter **TT** (transit time). This captures minority carrier storage in the PN diode which does not exist in an ideal Schottky diode. Therefore **TT** should be set to zero. Actual Schottky diodes in strong forward bias do show some minority carrier effects. Describing this accurately will require the introduction of a finite value to **TT**.

Table 7.1 summarizes the basic set of model parameters used to describe Schottky diodes in a circuit CAD environment. When appropriate, their ideal values are also given.

As in the case of the PN diode, additional model parameters are found in more advanced models to describe non-ideal leakage currents in reverse bias, temperature dependence, parasitic diode isolation, and the noise characteristics. Also, scaling parameters are introduced to scale the diode geometry from the "nominal" device that was experimentally characterized to extract the model, to any other device size.

## AT7.4 Switching characteristics of Schottky diode

A real uniqueness of Schottky diodes is their fast switching characteristics. In many large-signal applications, Schottky diodes must quickly switch from a strong forward bias with substantial current flow to a high reverse bias characterized by negligible current, and vice versa. How quickly a Schottky diode is able to switch between these two states is a key figure of merit for many large signal applications.

We studied these kinds of transients in the PN diode in Sec. AT6.3. We learned that they

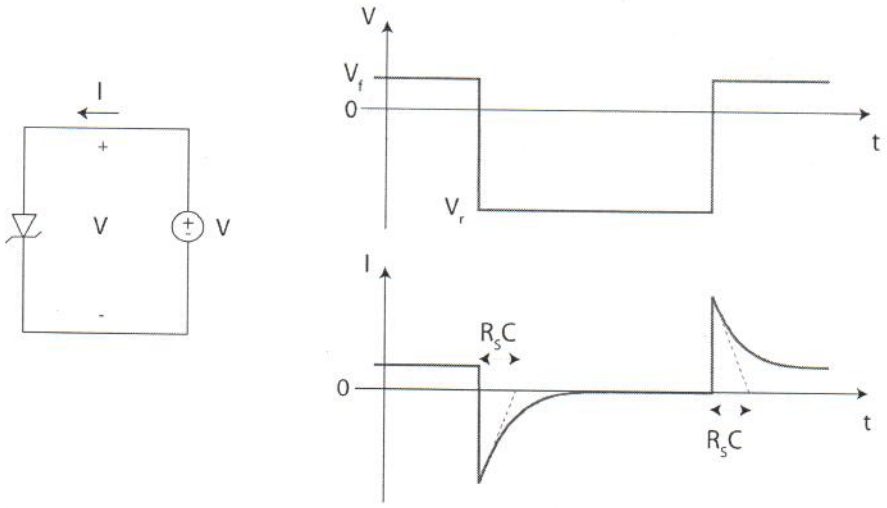



Figure 7.33: Switching transients in a Schottky diode. In response to a sudden switch in the voltage applied to the diode, the diode current shows spikes that decay with a characteristic time constant $R_sC$.

can be quite slow due to minority carrier storage. Since this is absent in a Schottky diode, these devices can switch much faster. This makes them preferable in many applications such as power management and ICs for automatic test equipment.

An appreciation for the limiting mechanisms can be obtained by considering the simple switching example depicted in Fig. 7.33. In this exercise, we have a Schottky diode switching from a forward voltage $V_f$ to a reverse voltage $V_r$, and back (note that the way the voltage is defined in Fig. 7.33, $V_r$ is a negative number). The current through the diode behaves as shown in the figure. Right after the switch-off transient (from $V_f$ to $V_r$), the current through the diode shows a prominent negative spike that decays away in an exponential way. Similarly, right after the swith-on transient (from $V_r$ to $V_f$), the diode current also exhibits a prominent positive spike that also decays away in a exponential fashion. In order to calculate the magnitude of the current spikes and, more importantly, the time constants of the exponential decays, we need to examine in detail the switching transients. This is best done by substituting the diode for its equivalent circuit model, as shown in Figs. 7.34 and 7.35 for the switch-off and switch-on transients, respectively.

Fig. 7.34 shows the details of the current paths during the switch-off transient. At $t = 0^-$, the diode is in forward bias with a voltage $V_f$ across. There is then a forward-bias current flowing through the diode of a magnitude $I_f$. This current produces an ohmic drop in the series resistance. As a result, the junction voltage is $V_f - I_f R_s$. As $V$ switches from $V_f$ to $V_r$, the junction voltage cannot abruptly change and remains with a value $V_f - I_f R_s$. Hence, at $t = 0^+$ the resistor has a voltage $V_f - I_f R_s - V_r$ across. The polarity of this voltage is reversed with respect to its sign at $t = 0^-$. Hence, the current flowing through the diode terminals at $t = 0^+$ is:

$$I(t = 0^+) \simeq \frac{V_r - V_f}{R_s} + I_f \tag{7.73}$$

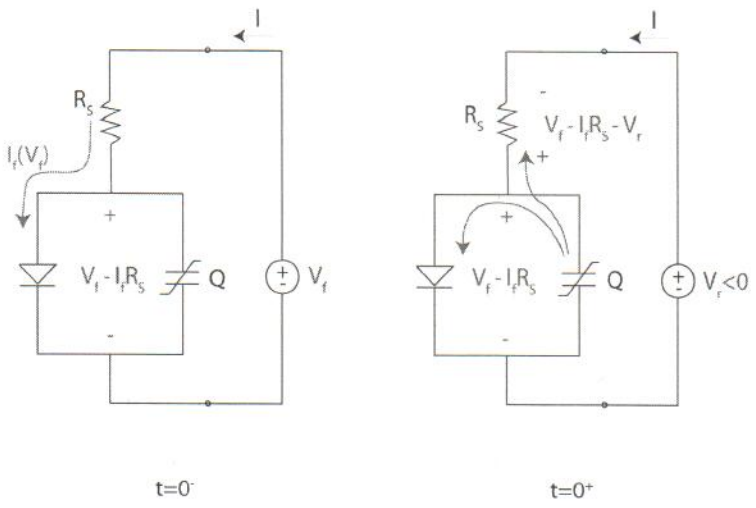



Figure 7.34: Switch-off transient in a Schottky diode. Left: for $t = 0^-$. Right: at $t = 0^+$.

Notice that this current is negative, *i.e.* it is a reverse current.

As $t > 0^+$, the charge storage element associated with the junction starts discharging through $R_s$ as well as through the diode. However, due to the exponential I-V characteristics of the diode, the diode current path gets shut off very quickly. After that, the junction charge element discharges only through the series resistance $R_s$ until it attains its final value $V_r$.

A rigorous analytical solution of this switching event is complicated because of two reasons. The first one is the initial fast discharge through the diode. Additionally, as we studied above, the junction charge in a Schottky diode depends on voltage. If $Q$ was voltage independent, past a short initial transitory, the time evolution of the diode current would be exponential with a characteristic time constant of a value $R_sC$, where $C$ is the associated junction capacitance. With $Q$ changing with voltage, the decay is not a pure exponential. Still, the characteristic time of the current decay is of the order of $R_s$ times the average $C$.

In the switch-on transient (Fig. 7.35), $V$ switches from $V_r < 0$ to $V_f$. At $t = 0^-$, the capacitor associated with the junction has a voltage $V_r$ across and the current going through the diode is the saturation current $-I_s$, a very small value. At $t = 0^+$, the voltage across the diode changes abruptly to $V_f$, but the junction voltage cannot change. Hence, a voltage $V_f - V_r$ appears across the series resistance of the diode with the polarity indicated in the figure. This produces a current of a magnitude:

$$I(0^+) \simeq \frac{V_f - V_r}{R_s} \tag{7.74}$$

Since the junction remains reverse biased, this current spike flows into the junction capacitance that as a result starts charging up. As it does, the voltage across the junction becomes less negative, the voltage across the resistor is accordingly reduced and the charging current is reduced

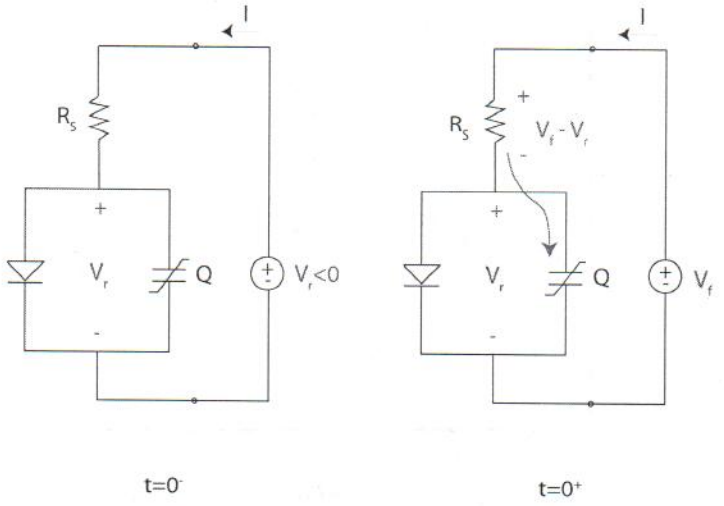



Figure 7.35: Switch-on transient in a Schottky diode. Left: for $t = 0^-$. Right: at $t = 0^+$.

as well. This is a typical $RC$-dominated transient characterized by an exponential evolution with a time constant equal to $R_sC$. Again, the fact that $C$ depends on the junction voltage results in a transient that is not a pure exponential.

The result of this detailed analysis is that the Schottky diode can respond to large-signal voltage switching in the scale of $R_sC$. Fast switching demands minimization of both $C$ and $R_s$. This contrasts with the switching times of the PN diode, which in addition to this $RC$ time constant, are delayed by the need to provide or eliminate the stored minority carriers in the quasi-neutral regions.

**Exercise 7.6**: *Simulate a switching transient of a Schottky diode in SPICE, as sketched in Fig. 7.33. The forward voltage is +0.45 V. The reverse voltage is −3 V. The SPICE parameters of the diode are:* **IS**= 5.5e − 13, **N**= 1.03, **EG**= 0.89, **RS**= 11, **CJO**= 3.24e − 13, **VJ**= 0.5, **M**= 0.339, **XTI**= 2, *and* **TT**= 0 *(units in Table 7.1). Extract the characteristic time constants of the transients and compare them with simple estimates.*

Fig. 7.36 shows the output obtained from HSPICE for a case in which there is a delay time of 2 *ps*, the reverse voltage is applied for 10 *ps* and the forward voltage is applied for 10 *ps*. Both turn-on and turn-off transients exhibit a near exponential dependence. For the turn-off transient, the extracted characteristic time constant is about 4 *ps*. The $R_sC$ time constant expected from the junction capacitance at $V = 0.45$ *V*, is about twice that, 7.8 *ps*. This suggests that the diode itself is effective in contributing to the discharge of the capacitor at the beginning of the transient.

The turn-on transient is characterized by a time constant of about 2 *ps*. The expected value of the $R_sC$ product corresponding to $V = -3$ *V*, is 1.9 *ps*. The excellent agreement reflects the fact that in reverse bias, the diode does not prevent the charge of the capacitor.

Right after the transients, the terminal currents are about 300 $mA$. This agrees with the expected value of $3.45/11 = 0.31$ *A*.

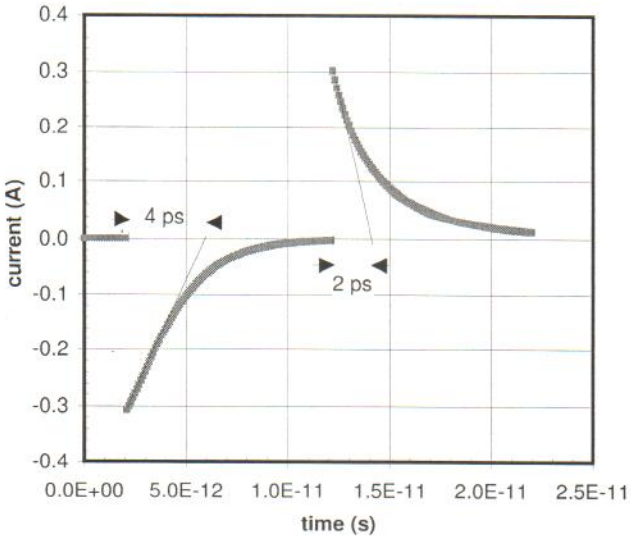
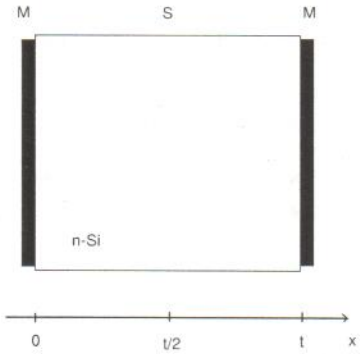



Figure 7.36: Switching transients in a Schottky diode as simulated by HSPICE (see exercise 7.6).

## Problems

**7.1** Consider a thin Si sample with two metallic contacts on each side, as shown in the figure below. The structure is at room temperature in thermal equilibrium and the doping is $N_D = 10^{15}\ cm^{-3}$ everywhere. The Schottky barrier height of both metal-semiconductor junctions is $q\varphi_{Bn} = 0.7\ eV$.




For the following three situations below, *i*) calculate the electron and hole concentrations and the electric field at the center of the sample $x = t/2$, *ii*) sketch the corresponding energy band diagram throughout the structure, and *iii*) compute the built-in potential of each metal-semiconductor junction.

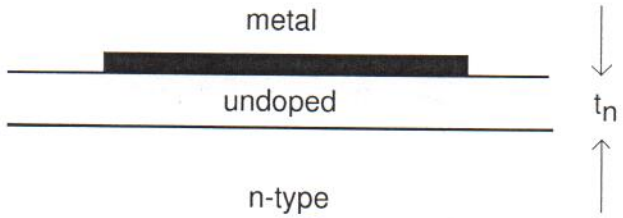
a) For a sample thickness of $t = 5\ \mu m$.

b) For a sample thickness of $t = 1\ \mu m$.

c) In the limit of a very thin sample.

**7.2** The **Mott diode** consists of a metal-semiconductor junction in which the semiconductor layer immediately adjacent to the interface is undoped, as indicated in the figure below. Consider

a case in which $W_M > W_s$, where $W_s$ refers to the work function of the n-type bulk.




**a)** Sketch the space charge, $\rho_o$, electric field, $\mathcal{E}_o$, electrostatic potential, $\phi_o$, and energy band diagram in thermal equilibrium.

**b)** Under the depletion approximation, calculate expressions for $\rho_o$, $\mathcal{E}_o$, and $\phi_o$ as a function of $x$ in the semiconductor. Leave everything in terms of $x_d$, the depletion region extension into the n-type region.

**c)** Compute an expression for $x_d$ in terms of material parameters.

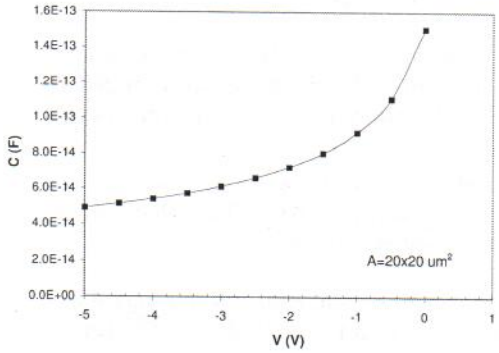
**d)** Modify the expressions for $\rho$, $\mathcal{E}$, $\phi$, and $x_d$ for forward and reverse bias.

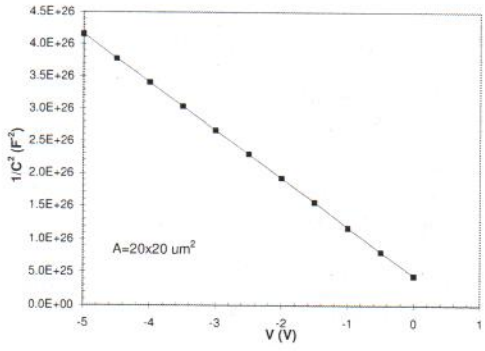
**e)** Derive an expression for the C-V characteristics. Sketch.

* **7.3** You have received a Si Schottky diode from your research supervisor with the instructions of measuring the Schottky barrier height. This is a rather novel device fabricated by some collaborators of your supervisor using an exotic metal. You are given a single sample.

You have a lot of experience in this procedure. Even though it was not asked of you, you start by carrying out C-V measurements at room temperature. Your plan is to proceed with temperature-dependent I-V characteristics from which you will extract the Schottky barrier height.

Tragically, as the C-V trace comes to its end, the device blows up. You panic. Another student that is working in your lab suggests that you can extract the Schottky barrier height from the C-V characteristics that you have just obtained. Through a microscope you measure the area of the metal as $20 \times 20\ \mu m^2$. You forgot to ask whether the substrate is n-type or p-type. Below are your measurements. $V$ represents the voltage of the metal with respect to the semiconductor.




Extract the Schottky barrier height from the available information.

• **7.4** Consider an ideal metal/n-type semiconductor junction. If without changing anything else, you change metals so that the Schottky barrier height $q\varphi_{Bn}$ goes up, how do the following device parameters change? Provide reasons for all your answers.

| | | | | |
|---|---|---|---|---|
| Current at a certain forward voltage, $I(V_f)$: | ↑ | ↓ | no effect | it depends |
| Capacitance at a certain forward voltage, $C(V_f)$: | ↑ | ↓ | no effect | it depends |
| Activation energy of saturation current, $E_a(I_S)$: | ↑ | ↓ | no effect | it depends |
| Charge in depletion region in equilibrium, $Q_d(V = 0)$: | ↑ | ↓ | no effect | it depends |
| Dynamic resistance at a certain forward current, $r_d(I_f)$: | ↑ | ↓ | no effect | it depends |

**7.5** A certain n-Si Schottky diode is characterized by the following set of SPICE parameters: **IS** $= 9 \times 10^{-14}$ $A$, **N** $= 1.00$, **EG** $= 0.95$, **RS** $= 20$ $\Omega$, **CJO** $= 3 \times 10^{-13}$ $F$, **VJ** $= 0.8$ $V$, **M** $= 0.5$, **XTI** $= 2$, **TT** $= 0$, **BV** $= 20$ $V$, all determined at 300 K.

a) Carry out SPICE simulations of the I-V characteristics of the device at room temperature and plot $|I|$ vs. $V$ in a semilog scale from $V = -5$ $V$ to $V = 0.5$ $V$.

b) Carry out I-V simulations at different temperatures ($T = 300$, 325, 350, 375, 400 $K$). For each temperature, extract $I_S$ by extrapolating the forward-bias characteristics to $V = 0$. Make an Arrhenius plot of $I_S/T^2$ vs. $1000/T$. Extract the Schottky barrier height from this plot. Compare with **EG**.

c) Perform SPICE simulations of the C-V characteristics at room temperature. Plot $C$ vs. $V$ in a linear scale from $V = -5$ $V$ to $V = 0.5$ $V$.

d) Perform SPICE simulations of a switching event as in Fig. 7.33 in the notes from $V = 0.45$ $V$ to $V = -3$ $V$ and back. Extract the characteristic time constants of both switching transitions and compare them with simple estimations carried out as in the notes. Also compare extract the peak currents at $t = 0^+$ and compare with simple calculations.

**7.6** [1] Consider an integrated Al/n-Si Schottky diode implemented in a bipolar process, as shown on the right of Fig. 7.22. The area of the metal/semiconductor interface is 40 $\mu m^2$. The total area of the n-tub is 800 $\mu m^2$. The doping level in the n-region is $1 \times 10^{17}$ $cm^{-3}$. The doping level in the substrate is $1 \times 10^{16}$ $cm^{-3}$. $q\varphi_{Bn}$ for Al on n-Si is 0.68 $eV$.

a) Derive values for standard model parameters for this diode (except **RS**). Include parameters to characterize the capacitance to the substrate. Model this as an asymmetric p-n diode with $\phi_{bi} = 0.9$ $V$, as discussed in Ch. 6. Refer to the new parameters as **CJOS**, **VJS**, and **MS**.

b) Compute and sketch the I-V and C-V characteristics of the diode for a wide range of voltages. Ignore parasitic effects. What is the maximum positive voltage that can be applied to the diode? What about the maximum negative voltage?

**7.7** [2] Consider an ideal metal/n-type semiconductor junction at room temperature. The metal has a work function $W_M = 4.05$ $eV$, the semiconductor is Si and is doped with $N_D = 10^{17}$ $cm^{-3}$.

[1] You will need to have studied Ch. 7 before attempting this problem.
[2] You will need to study the contents of Ch. 8 before attempting to solve part d) of this problem.

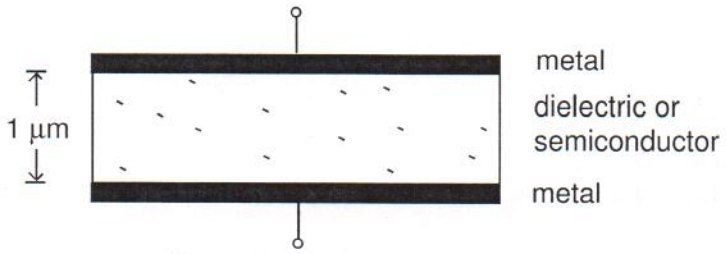
a) Sketch to scale the energy band diagram in equilibrium.

b) Compute the built-in potential.

c) Compute the electron and hole concentrations right at the metal/semiconductor interface.

d) Estimate the spatial extent of the space charge region in the semiconductor.

**7.8** Consider a parallel-plate capacitor with Al electrodes as sketched below. The separation between the plates is 1 $\mu m$.

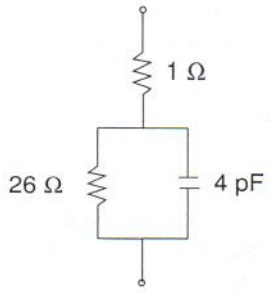



Calculate the capacitance per unit area of this device when the following materials are utilized as dielectric:

*i*) $SiO_2$
*ii*) intrinsic Si
*iii*) n-type Si with $N_D = 1 \times 10^{15}\ cm^{-3}$
*iv*) n-type Si with $N_D = 1 \times 10^{17}\ cm^{-3}$

In all cases, the contact between the Al plate and the dielectric material is intimate.

• **7.9** A Schottky diode biased at a forward voltage of $V = 0.3\ V$ has the small-signal equivalent circuit indicated below at room temperature. The Schottky barrier height of this diode is $\varphi_B = 0.9\ V$.




V=0.3 V

a) Estimate the current through the diode for $V = -1\ V$. State any assumptions you need to make.

b) Estimate the forward voltage across the diode for $I = 100\ mA$. State any assumptions you need to make.

c) Estimate the capacitance of the diode for $I = 100\ mA$. State any assumptions you need to make.

**7.10** Consider a Schottky-barrier diode on an n-type Si substrate. Suppose we *increase* the doping level in the semiconductor $N_D$. Nothing else is changed. Indicate the impact that this would have on the parameters and figures of merit listed below.

Circle one: *increase* $=\uparrow$, *decrease* $=\downarrow$, or *no effect*. For each item, give the reason for your choice.

| | |
|---|---|
| built-in potential, $\phi_{bi}$: | $\uparrow$ $\downarrow$ *no effect* |
| current at a certain forward voltage, $I(V_f)$: | $\uparrow$ $\downarrow$ *no effect* |
| capacitance at a certain forward voltage, $C(V_f)$: | $\uparrow$ $\downarrow$ *no effect* |
| switching time constant, $\tau$: | $\uparrow$ $\downarrow$ *no effect* |
| equilibrium electron concentration at the metal-semiconductor interface, $n_o(x = 0)$: | $\uparrow$ $\downarrow$ *no effect* |

**7.11** Consider a Schottky diode with a 10 $\mu m^2$ active metal-semiconductor junction area and a Schottky barrier height of 0.85 $eV$. This diode is forward biased with $V = 0.6$ $V$ at room temperature. Answer the following questions in the context of the thermionic-emission theory of transport in the metal-semiconductor junction.

a) Calculate the electron concentration at the semiconductor side of the metal-semiconductor interface.

b) Calculate the net electron velocity at the semiconductor side of the metal-semiconductor interface.

c) Calculate the diode current.

**7.12** A certain IC foundry offers a process that includes a "nominal" Schottky diode characterized by the following set of SPICE parameters at 300 K: **IS**= $1e-13$, **N**= 1.0, **EG**= 0.9, **RS**= 10, **CJO**= $1e-12$, **VJ**= 0.7, **M**= 0.5, **XTI**= 2, **TT**= 0, and **BV**= 10. This "nominal" Schottky diode has a junction area of 10 $\mu m^2$. Estimate the 3 $dB$ bandwidth, $f_{3dB} = \frac{\omega_{3dB}}{2\pi}$, of this "nominal" Schottky diode at a forward current of 1 $mA$ and at 300 K.

**7.13** Can the built-in potential of a metal-semiconductor junction in thermal equilibrium exceed the Schottky barrier height? Explain your answer. Draw relevant energy band diagrams as needed.

**7.14** This problem is about making some early design decisions for a process for a Schottky diode varactor (variable capacitor). The output of this exercise is a first-order sense of the Schottky barrier height of the metal and the doping level of the semiconductor.

The room-temperature specifications of this varactor are: i) a capacitance per unit area at 0 $V$ of $C_o = 1$ $fF/\mu m^2$, and ii) the capacitance must change by a factor of 2 between 0 and 2 $V$ (an 100% tuning range). To minimize power consumption, the Schottky diode must operate in reverse bias.

Assume a metal/n-Si structure. Provide values of $N_D$ and $q\varphi_{Bn}$ that meet the design specs.

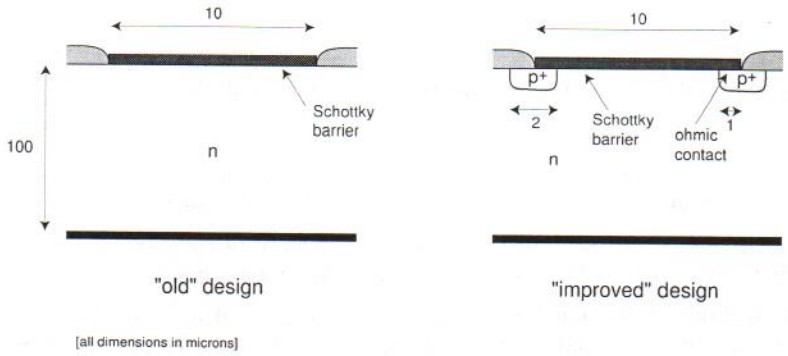
**7.15** [3]. In a metal-semiconductor junction that uses a metal with a large Schottky barrier height, an inversion layer of minority carriers might get formed at the metal-semiconductor interface in thermal equilibrium. Consider such a situation on n-type Si with $N_D = 10^{17}$ $cm^{-3}$.

[3] Attempt this problem after studying Ch. 8

a) Estimate the minimum Schottky barrier height for which this becomes a concern.

b) Sketch to scale the energy band diagram of such a metal-semiconductor junction in thermal equilibrium, i.e., label all key energy differences and horizontal dimensions.

**7.16** In order to improve the breakdown voltage of a Schottky barrier diode, a designer decides to introduce a $p^+$ guard ring, as sketched below. This problem is about evaluating the penalty in capacitance associated with this change.




In both designs, the Schottky barrier height is 0.8 $eV$, the doping of the semiconductor is $N_D = 10^{16}\ cm^{-3}$, and the wafer thickness is 100 $\mu m$. The ohmic contact to the body of the semiconductor is placed on the bottom surface of the wafer.

In the old design, the Schottky metal contacts the semiconductor over a $10 \times 10\ \mu m^2$ square window. In the new design, the Schottky metal is also $10 \times 10\ \mu m^2$. The $p^+$ guard ring is 2 $\mu m$ wide and is centered around the edge of the metal. The $p^+$ region has a doping level of $N_A = 10^{19}\ cm^{-3}$.

Assume that any minority carrier physics in the pn diode are dominated by the n-type substrate. Answer all the following questions at room temperature.

a) Estimate the capacitance of the old design at $V = 0\ V$.

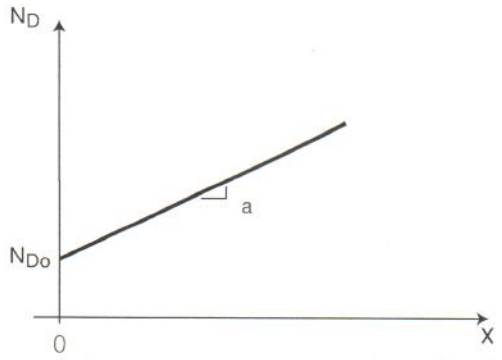
b) Estimate the capacitance of the new design design at $V = 0\ V$.

c) Estimate the capacitance of the old design at $V = 0.5\ V$.

d) Estimate the capacitance of the new design design at $V = 0.5\ V$.

**7.17** Consider a metal-semiconductor junction fabricated on a linearly graded n-type doped semiconductor. The doping distribution in the semiconductor is sketched in the figure below. The grading coefficient is $a$. $x = 0$ denotes the metal-semiconductor interface. The Schottky

barrier of this structure is $q\varphi_B$.




This problem is about formulating the electrostatics of this problem under the depletion approximation.

**a)** Derive an algebraic expression for the spatial dependence of the volume charge density across the structure. Leave expression in terms of the depletion region depth. Sketch.

**b)** Derive an algebraic expression for the spatial dependence of the electric field across the structure. Leave expression in terms of the depletion region depth. Sketch.

**c)** Derive an algebraic expression for the spatial dependence of the electrostatic potential across the structure. Explicitly state your choice of potential reference. Leave expression in terms of the depletion region depth. Sketch.

**d)** Derive an expression that allows for the solution of the depletion region width in terms of known parameters.

**7.18** Consider a metal-semiconductor junction made out of Al on n-type Si with $N_D = 10^{19}\ cm^{-3}$ at room temperature. For this system, the Schottky barrier height is $q\varphi_{Bn} = 0.68\ eV$.

**a)** Compute the built-in potential for this structure. Sketch the energy band diagram in thermal equilibrium indicating all key energies. Neglect carrier degeneracy.

Now consider inserting a thin p layer between the metal and the n-type semiconductor. The doping level is $N_A = 10^{17}\ cm^{-3}$ and the thickness is $x_j = 0.1\ \mu m$.

**b)** Qualitatively sketch the volume charge density of this structure in thermal equilibrium. Exploit the fact that $N_D \gg N_A$. Clearly label all instances of net electrical charge. Explain.

**c)** Qualitatively sketch the electric field across the structure in thermal equilibrium. Explain.

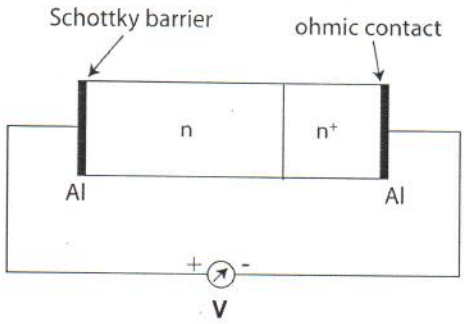
**d)** Qualitatively sketch the electrostatic potential across the structure in thermal equilibrium. Select as reference $\phi = 0$ in the bulk of the n-type region. Explain.

**e)** Qualitatively sketch the energy band diagram of this structure in thermal equilibrium. Explain.

**f)** Based on what you see in the energy band diagram, at a given forward voltage (metal positive with respect to n-type substrate), would this structure have more or less current than the structure in part a)? At a given forward voltage, would it have more or less capacitance than the structure in part a)? Explain.

**g)** Calculate the electric field at the metal-semiconductor interface in thermal equilibrium.

**7.19** Consider a Schottky diode built on n-type Silicon, as sketched in the figure below. The Schottky metal is Al which has a Schottky barrier height on n-type Si of $q\varphi_{Bn} = 0.68\ eV$. The doping level in the n region is $N_D = 10^{16}\ cm^{-3}$. In order to provide a good ohmic contact to this diode, there is an $n^+$ region with a doping level of $N_{D+} = 10^{19}\ cm^{-3}$. Al is also used as the ohmic metal.




Consider this structure in thermal equilibrium at room temperature. Use Boltzmann statistics to treat the $n^+$ region.

**a)** Sketch a complete energy band diagram across the entire structure (from metal to metal). Indicate the location of the Fermi level. Assume that both semiconductor regions are wide enough to fully accommodate the depletion regions associated with the metal-semiconductor junctions. Be neat in your sketch.

**b)** Compute the built-in potential of the metal-semiconductor junction on the left.

**c)** Compute the built-in potential of the n-$n^+$ junction.

**d)** Compute the built-in potential of the metal-semiconductor junction on the right.

**e)** Is there a voltage difference across the terminals of this device?

**7.20** The Bethe condition establishes the applicability of the thermionic emission model to the current-voltage characteristics of a Schottky diode. This problem explores the voltage range of applicability of the Bethe condition. Consider an $WSi_2$/n-Si Schottky diode at room temperature. The Schottky barrier height for this system is $q\varphi_{Bn} = 0.64\ V$. The doping level of the semiconductor is $N_D = 4 \times 10^{17}\ cm^{-3}$.

Follow this approach:

**a)** In what bias regime, forward or/and reverse, does the Bethe condition become problematic? Explain.

**b)** Estimate the mean free path of electrons in the semiconductor.

**c)** Estimate the maximum forward and/or reverse voltage for applicability of the Bethe condition in this diode.

**d)** At this(these) voltage(s), estimate the current density flowing through the diode.