

Topic 11: PN Junction Analysis 1

Preface:

This set marks the beginning of our analytical study of the PN junction. The following notes involve many derivations with concepts including charge conservation, Gauss's law, continuity of the electric field, and many more. Alongside these concepts lies a sea of mathematics and equations. I highly encourage reading through this set of notes several times. I have broken down the derivations step-by-step to provide a thorough derivation. As mentioned in the PN junction introduction notes, these equations will be used extensively throughout the remainder of the course.

Gauss's Law

Gauss's law will be prevalent throughout the PN junction analysis. This section provides a brief review of Gauss's law before presenting Poisson's equation. Formally, Gauss's law states that the gradient of the electric field is proportional to the charge density within some space.

$$\nabla \cdot \vec{\mathcal{E}} = \frac{\rho}{\epsilon} \quad (1)$$

$$\frac{d\mathcal{E}_x}{dx} + \frac{d\mathcal{E}_y}{dy} + \frac{d\mathcal{E}_z}{dz} = \frac{\rho}{\epsilon} \quad (2)$$

$$\Rightarrow \frac{d\mathcal{E}}{dx} = \frac{\rho}{\epsilon} \quad (3)$$

Where,

- ρ : The charge density in $\frac{C}{cm^3}$
- ϵ : The dielectric permittivity in $\frac{F}{cm}$. $\epsilon = \epsilon_r \epsilon_0$.
- ϵ_r : The relative permittivity of a semiconductor; unitless.
- ϵ_0 : The permittivity of free space. $\epsilon_0 = 8.85 \cdot 10^{-14} \frac{F}{cm}$
- ∇ : The gradient operator in three-dimensions.
- $\mathcal{E}_x, \mathcal{E}_y, \mathcal{E}_z$: The electric field component in the x, y , and z directions, respectively.

Three forms of Gauss's law have been provided above. Eq. 1 is the most general form utilizing the gradient operator. Eq. 2 expands the gradient operator into spatial derivatives acting on the three Cartesian coordinates. Lastly, Eq. 3 assumes a one-dimensional electric field. We will use Eq. 3 since our study is focused on one-dimensional analysis. Also, note the use of ϵ without a subscript. Different semiconductors have different dielectric properties. Hence we will use ϵ to encapsulate the relative permittivity of the material. For reference, ϵ_r for air, water, silicon, silicon Oxide (SiO_2) and Germanium are: $\approx 1, 81, 11.7, 3.9$, and 16 , respectively. The relative permittivity matters because it will decrease the strength, or effectiveness, of the electric field to permeate through a medium.

Poisson's Equation:

Recall that the units for electric field are $[\mathcal{E}] = \frac{V}{cm}$. Remembering the units is a convenient way to remember that $\mathcal{E} = -\frac{dV}{dx}$. If we plug in this expression into Gauss's law we arrive at,

$$-\frac{d^2V}{dx^2} = \frac{\rho}{\epsilon} \quad (4)$$

Eq. 4 is famously called Poisson's equation in one-dimension, and will serve as the basis for analyzing the PN junction. If we look at the units for ρ , we do not have any quantities with units of $\frac{C}{cm^3}$. However, we know that a coulomb is the fundamental charge multiplied by some number of charges. We know the number of charges per volume within a semiconductor is given the electron and hole concentration. Hence we can rewrite Eq. 4 as,

$$\frac{d^2V}{dx^2} = -\frac{q}{\epsilon} \left(\frac{\# \text{ of charged particles}}{cm^3} \right) \quad (5)$$

Eq. 5 does not help us much since we need to find a way to relate the number of charges on each side of the junction. We could attempt to apply Poisson's to the entire device but this would result in unnecessarily complicated expression. We could try and write the net charge across every region in terms of the four charge types: holes, electrons, positive dopant ions, and negative dopant ions. Within each region, we can then describe the behavior in a piece-wise expression. The piece-wise approach will only be valid if we can ensure some quantity is continuous across the regions of the device, such as the electric field and potential. Regardless of the region the net number of charges is characterized by,

$$\frac{\# \text{ of charged particles}}{cm^3} = p - n + N_D - N_A \quad (6)$$

The p and N_D terms are positive because both holes and the space charge due to donor atoms, such as phosphorous, are positive. Likewise, n and N_A are negative because both electrons and the space charge due to acceptor atoms, such as boron, carry a negative charge. Eq. 6 is the most general form of charge concentration and can be applied to any region. For example, in the p-type bulk, the number of donor atoms is zero; $N_D = 0$. We assume that the doping concentration is significantly greater than the intrinsic concentration implying that $p \gg n$. The large doping concentration implies that $p = N_A$. Therefore, the net number of charges is $p - \text{negligible} - 0 - N_A = 0$. A zero net number of charges implies the region is neutral, which confirms our initial assumption that the bulk is neutral.

Whenever we refer to Poisson's equation, we will use the following expression:

$$\frac{d^2V}{dx^2} = \frac{q}{\epsilon} (p - n + N_D - N_A) \quad (7)$$

Setup:

From Poisson's equation, we can solve explicitly for the built-in voltage and the built-in electric field. However, we need to place some restrictions or assumptions on the charge density for each region. For notation, I have defined several key points within the semiconductor. In the figure below the junction occurs at $x = 0$, the depletion region in the p-type ends at $x = -x_p$, and the depletion region in the n-type ends at $x = x_n$. These three points will provide initial conditions for the following analysis.

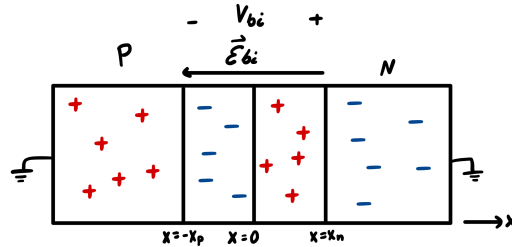


Figure 1: PN Junction at thermal equilibrium

To derive the built-in voltage across the junction we need to find the voltage at x_n and $-x_p$ with respect to some reference point. The difference of the two will give the built-in potential; $V_{bi} = V_n - V_p$. We could use ground as the reference, but this then assumes our equations are only valid when the terminals of the device are grounded. Luckily, there is a better choice - the Fermi level. We established at the end of the last notes that the Fermi level is constant in thermal equilibrium. Using circuit ground does not necessarily imply that ground is zero volts, so the Fermi level is a better choice. In the bulk, the

difference between the Fermi and intrinsic level provides us with the concentration. Hence, the Fermi level reference provides access to many key quantities.

$$p = n_i e^{\frac{E_i - E_F}{kT}} \quad (8)$$

$$n = n_i e^{\frac{E_F - E_i}{kT}} \quad (9)$$

Assuming the doping in both the n and p-type regions are significantly larger than the intrinsic we can say $p = N_A$ for $x \leq -x_p$ and $n = N_D$ for $x \geq x_n$. Rearranging Eq. 8 and 9 within the bulk, we can solve for the difference in Fermi and intrinsic levels.

$$E_i - E_F = kT \ln \left(\frac{N_A}{n_i} \right) \text{ for } x \leq -x_p \quad (10)$$

$$E_F - E_i = kT \ln \left(\frac{N_D}{n_i} \right) \text{ for } x \geq x_n \quad (11)$$

Taking advantage of the fact that $V = -\frac{E_i}{q}$, where E_i is an arbitrary energy level. We can divide the left side of Eq. 10 and 11 by q to convert the energy levels into potential.

$$\frac{kT}{q} \ln \left(\frac{N_A}{n_i} \right) = \frac{E_i - E_F}{q}, \text{ for } x \leq -x_p \quad (12)$$

$$\Rightarrow \phi_p = -\frac{kT}{q} \ln \left(\frac{N_A}{n_i} \right), \text{ for } x \leq -x_p \quad (13)$$

$$\frac{kT}{q} \ln \left(\frac{N_D}{n_i} \right) = \frac{E_F - E_i}{q}, \text{ for } x \geq x_n \quad (14)$$

$$\Rightarrow \phi_n = \frac{kT}{q} \ln \left(\frac{N_D}{n_i} \right), \text{ for } x \geq x_n \quad (15)$$

The negative associated with Eq. 13 is due to the charge of the hole being positive. The $V = -\frac{E_i}{q}$ does not assume the charge is positive or negative. If an electron is being considered then the additional negative cancels leaving Eq. 15 positive. Taking the difference of ϕ_n and ϕ_p gives the built-in potential.

$$V_{bi} = \phi_n - \phi_p = \frac{kT}{q} \ln \left(\frac{N_A N_D}{n_i^2} \right) \quad (16)$$

The figure below visualizes ϕ_p , ϕ_n , and V_{bi} in the energy band diagram. Since the conduction, valence, and intrinsic band all bend equally the bend in the intrinsic level is the same for any other band. ϕ_p accounts for the bend in the p-type depletion region and ϕ_n accounts for the bend in the depletion region within the n-type. Summing the total bend across both depletion regions gives the built-in potential as the total height of bending. To reiterate, we only take the difference $\phi_p - \phi_n$ because ϕ_n is dependent on $E_F - E_i$. Multiplying by the negative ensures both ϕ_p and ϕ_n are referenced with respect to $E_F - E_i$.

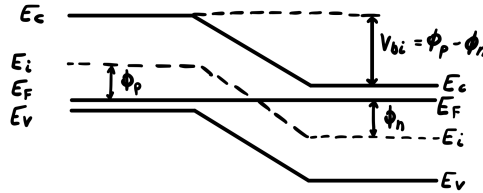


Figure 2: Energy band diagram potentials for a PN junction in thermal equilibrium.

It is quite incredible that we can determine the built-in potential for a PN junction by knowing only the doping concentration. For example, a junction with $N_A = 10^{15} \text{ cm}^{-3}$, $N_D = 10^{16} \text{ cm}^{-3}$ and $n_i = 1.5 \cdot 10^{10} \text{ cm}^{-3}$ results in a built-in potential of 637 mV. The 0.637 V should be reminiscent of the diode forward voltage typically assume to equal 0.7 V. In addition, the built-in voltage increases logarithmically. We could increase the doping by a factor of 1000 for both N_A and N_D and the built-in voltage becomes approximately 1 V. The built-in voltage is a weak function of doping, and this observation can help determine if you get nonsense from calculations. It is incredibly unlikely the built-in potential for a pure silicon PN junction to be above 1.5 V. The doping level required to even reach 1.5 V is even a stretch since a large number of free carriers will make the material behave like a metal.

Space Charge Regions

Again, SCR will be used as a shorthand for the space charge region of the PN junction. The space charge region is anywhere within a device space charge, or a depletion region exists. We know in the PN junction that the built-in potential is a result of the space charge. However, there are different methods of doping the junction to alter the behavior and properties of the region. The way a junction is doped is commonly called the doping profile and we will discuss the abrupt and linearly graded junctions. Note that other doping profiles do exist.

Abrupt Junction:

The abrupt junction implies an abrupt, or sudden, change in doping at the junction. It is practically impossible to have a sharp discontinuity in doping, but it provides a simple case to analyze. The horizontal axis will represent the distance along with the device and the vertical axis represents the number of charge carriers. Note that positive concentrations will be above the x-axis and negative concentrations below it. We will place the junction at $x = 0$ and assume the n and p-type regions are uniformly doped with N_D and N_A ions, respectively. The p-type region is to the left of the junction and the n-type to the right. The terms $-x_p$ and x_n denote the boundaries of the space charge region.

The visualization for the abrupt junction is provided below,

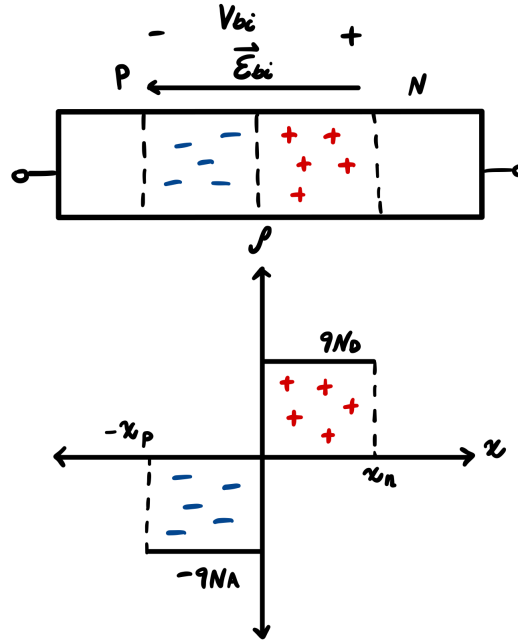


Figure 3: Abrupt junction space charge doping profile.

At $x = 0$ the charge density, ρ , changes from $-qN_A$ to qN_D . The space charge region in the p-type extends from $-x_p \leq x \leq 0$. The net charge of this region is negative due to the holes being stripped from the acceptor ions. Similarly, the n-type space charge is positive due to the electrons being stripped from the donor ions.

Our analysis will rely on applying Poisson's equation to both space charge regions. Beginning with the p-type, some simplifications can be made to the charge density. N_A is known since we know the doping concentration. N_D must equal zero since no donor ions are present in the p-type. We assume, by the depletion approximation, that the free carrier concentrations, n and p , are negligible.

Notice the built-in electric field across the device. The field will result in excess electrons drifting toward the n-type and holes toward the p-type. Since these charges are constantly ejected from the depletion region we assume their charge contribution to be small compared to the doping density. Hence, the charge density $\rho = N_A$ in the p-type material. Applying this to Poisson's we greatly simplify the equation.

$$\frac{d^2V}{dx^2} = -\frac{q}{\epsilon}N_A \quad (17)$$

Assuming uniform doping implies that the doping concentration is independent of distance, implying the right-hand side of Eq. 17 is a constant. Before we integrate, we can simplify further by substituting the derivative definition of the electric field.

$$\mathcal{E} = -\frac{dV}{dx} \quad (18)$$

$$\Rightarrow \frac{d^2V}{dx^2} = -\frac{d\mathcal{E}}{dx} \quad (19)$$

$$d\mathcal{E} = \frac{qN_A}{\epsilon}dx \quad (20)$$

Note we cannot simply integrate both sides with respect to x . Using Poisson's equation and separating the differentials requires us to integrate with respect to each differential. For example, on the $d\mathcal{E}$ side we will integrate from $\mathcal{E}(-x_p)$ to $\mathcal{E}(x)$. This may seem arbitrary but it is necessary to keep track of our references. We know the neutral p-region has no electric field. Hence, we can rely on the boundary condition of $\mathcal{E}(-x_p) = 0$. This step is minor but explicitly stating the bounds on the differentials will help with organization.

$$\int_{\mathcal{E}(-x_p)}^{\mathcal{E}(x)} d\mathcal{E} = \int_{-x_p}^x \frac{qN_A}{\epsilon} dx \quad (21)$$

$$\mathcal{E}(x) - \mathcal{E}(-x_p) = \frac{qN_A}{\epsilon}(x + x_p) \quad (22)$$

$$\Rightarrow \mathcal{E}(x) = -\frac{qN_A}{\epsilon}(x + x_p), \text{ for } -x_p \leq x \leq 0 \quad (23)$$

From Eq. 22 to 23 $\mathcal{E}(-x_p) = 0$ was substituted. To reiterate, it does not matter if you integrate from x to $-x_p$ or from $-x_p$ to x . However, you must be consistent with your bounds of integration throughout the derivation.

From Eq. 23 we notice 3 things: 1. The electric field is negative 2. The electric field approaches zero as x nears $-x_p$ and 3. The electric field is a maximum at $x = 0$. We must ask ourselves if these observations are feasible. The negative is expected since the electric field points in the negative x direction, and we integrated in the positive x direction. The zero condition at $x = -x_p$ was assumed and should be true. We will need to analyze the n-type space charge to further our observation regarding the maximum electric field.

Moving onto the n-type space charge, the derivation is nearly identical. We assume $\mathcal{E}(x_n) = 0$ and will integrate from x to x_n to match the direction from the p-type derivation; $\mathcal{E}(x)$ to $\mathcal{E}(x_n)$ for the left-hand side. The charge density is now positive with a value of qN_D . We only have donor ions, implying zero acceptors, and the concentrations n and p are assumed negligible by the depletion approximation. Applying to Poisson's equation,

$$\frac{d^2V}{dx^2} = -\frac{qN_D}{\epsilon} \quad (24)$$

$$\frac{d\mathcal{E}}{dx} = \frac{qN_D}{\epsilon} \quad (25)$$

$$d\mathcal{E} = \frac{qN_D}{\epsilon}dx \quad (26)$$

$$\int_{\mathcal{E}(x)}^{\mathcal{E}(x_n)} d\mathcal{E} = \int_x^{x_n} \frac{qN_D}{\epsilon} dx \quad (27)$$

$$\mathcal{E}(x_n) - \mathcal{E}(x) = \frac{qN_D}{\epsilon}(x_n - x) \quad (28)$$

$$\Rightarrow \mathcal{E}(x) = -\frac{qN_D}{\epsilon}(x_n - x), \text{ for } 0 \leq x \leq x_n \quad (29)$$

Again, we can make several observations to verify the equation makes sense. The electric field still points in the negative x-direction due to $x \leq x_n$. This is a good way to verify you have derived the equation correctly. If at the junction, $x = 0$, the direction of the electric field switched then an error has been made. The electric field is a continuous quantity, implying a discontinuity is not possible. We can also verify that at $x = x_n$ the electric field equals zero. Similarly, the magnitude of the electric field experiences a maximum at $x = 0$. With this information, we can plot the electric field over the junction in terms of the equations from the p-type and n-type regions.

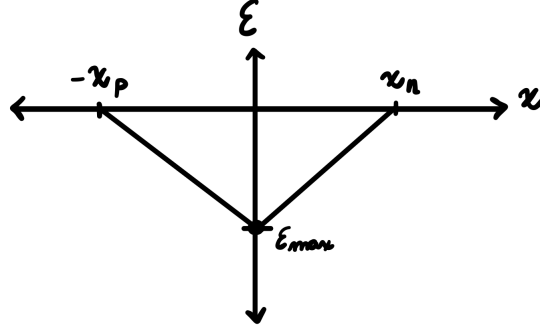


Figure 4: Abrupt junction electric field.

From equations 23 and 29, we know that the electric field varies linearly within each region. Hence, they are drawn as straight lines within each depletion region. We also know that the maximum electric field, $x = 0$, is the point that combines the two curves. Implying that the two curves must intersect with a maximum value at $x = 0$. We can find the value of the maximum electric field by applying the conditions to the above equations.

$$\begin{aligned} \mathcal{E}_{max,p} &= -\frac{q}{\epsilon} N_A x_p \\ \mathcal{E}_{max,n} &= -\frac{q}{\epsilon} N_D x_n \\ \Rightarrow N_A x_p &= N_D x_n \end{aligned} \quad (30)$$

Equation 30 provides an equation relating the doping density to the width of the depletion region. **Note** that the electric field curve tells us the direction. Repeating the derivation with the n-type to the left of the p-type would give a positive electric field.

$$w_d = x_n + x_p \quad (31)$$

Another question we may ask ourselves is how can we relate Eq. 30 to the charge density? Multiplying both sides of the equation by q converts the doping density into charge density.

$$q N_A x_p = q N_D x_n \quad (32)$$

Eq. 32 is simply a fancier way of expressing charge conservation. The charge density plot in Figure 3 must be conserved across the junction. Hence, **the area under each region must be equal** or the charger per unit area is conserved. This observation is incredibly important and will help us analyze other junction types and devices. If the charge must be conserved, we may ask several questions regarding the width of each region. The charge density and electric field plot are nearly symmetric but they do not have to be. Say the doping density in the p-type region is significantly larger than the n-type. Charge conservation implies the depletion region in the p-type material will be smaller than the n-type region. The excess charge in the p-type means less region is needed to hold a given charge. Inversely, the depletion region in the n-type will be large. A visualization of this is provided on the following page.

Charge conservation is mentioned as a way of determining the width of the junction and a way to classify special junctions. Take a moment and imagine how you would try and determine the width of the PN junction? We cannot simply cut the device in half and see where it starts and stops. We have to rely on the equations above to relate the material parameters we know, such as doping density, to the width. As for the special junctions, they will be discussed at the end of this section.

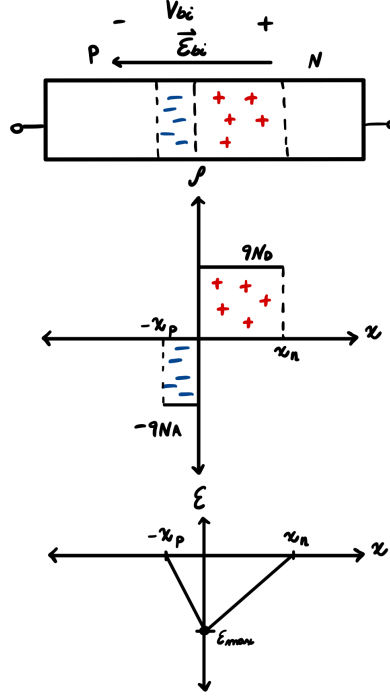


Figure 5: Semiconductor, doping profile, and electric field plot comparison.

Potential Derivation:

So far, we have only found the electric field. We can integrate equations 23 and 29 again to derive a closed-form expression of the voltage across the junction. Beginning with the p-type region:

$$\begin{aligned}
 \mathcal{E} &= -\frac{d\phi}{dx} \\
 \Rightarrow \frac{d\phi}{dx} &= \frac{qN_A}{\epsilon}(x + x_p) \\
 d\phi &= \frac{qN_A}{\epsilon}(x + x_p)dx \\
 \int_{\phi(-x_p)}^{\phi(x)} d\phi &= \int_{-x_p}^x \frac{qN_A}{\epsilon}(x + x_p)dx \\
 \phi(x) - \phi(-x_p) &= \frac{qN_A}{2\epsilon}(x + x_p)^2 \Big|_{-x_p}^x
 \end{aligned} \tag{33}$$

At this point, some assumptions need to be made. We will assume that the potential is taken with reference to the neutral p-type region. This assumption forces the $\phi(-x_p)$ term to become zero. The right-hand side of equation 33 is simplified by evaluating at the given points, where the $-x_p$ component will become zero. Hence,

$$\phi_p(x) = \frac{qN_A}{2\epsilon}(x + x_p)^2, \text{ for } -x_p \leq x \leq 0 \tag{34}$$

The subscript p has been included to denote the potential throughout the p-type region. Repeating the same process for the n-type region.

$$\begin{aligned}
 \frac{d\phi}{dx} &= \frac{qN_D}{\epsilon}(x_n - x) \\
 d\phi &= \frac{qN_D}{\epsilon}(x_n - x)dx \\
 \int_x^{x_n} d\phi &= \int_x^{x_n} \frac{qN_D}{\epsilon}(x_n - x)dx \\
 \phi(x_n) - \phi(x) &= \frac{qN_D}{2\epsilon}(x_n - x)^2 \Big|_x^{x_n}
 \end{aligned}$$

$$\phi_n(x_n) - \phi_x(x) = \frac{qN_D}{2\epsilon}(0 - (x_n - x)^2) \quad (35)$$

Now we seem to be stuck. We have already assumed the reference voltage on the p-side of the junction. Hence, we need to utilize the p-type equation to give us an initial condition for the n-type region. We know that potential is also a continuous quantity, implying the potential at the junction must be equal regardless of where we start the derivation. Evaluating Eq. 34 at $x = 0$, we can calculate the potential at the junction with respect to the p-type region.

$$\phi_p(0) = \frac{qN_A}{2\epsilon}x_p^2 \quad (36)$$

Equating both equations evaluated at zero, $\phi_n(0) = \phi_p(0)$,

$$\begin{aligned} \phi_n(x_n) &= \frac{qN_D}{2\epsilon}(x_n - 0)^2 + \frac{qN_A}{2\epsilon} \\ V_{bi} &= \frac{q}{2\epsilon}(N_Dx_n^2 + N_Ax_p^2) \end{aligned} \quad (37)$$

First, $\phi_x(x_n)$ has been replaced with V_{bi} because they are equal by definition. The built-in voltage is the voltage across the junction and the n-type potential evaluated at the boundary is also the voltage across the junction. If we wanted to plot the potential over the junction we would note that the two curves would be quadratic. The potential increases from the p to n side, implying the quadratic would be inverted across junctions. Outside of the depletion region, the potential is constant. We cannot say the potential is zero because we assume the voltage in the neutral region is constant. Another way to view this equation is the area under the electric field plot. The total area is just the area of two triangles with height \mathcal{E}_{max} and bases x_n and x_p , respectively. As an equation,

$$\begin{aligned} V_{bi} &= \frac{\mathcal{E}_{max}x_p}{2} + \frac{\mathcal{E}_{max}x_n}{2} \\ &= \frac{\mathcal{E}_m a x}{2}(x_p + x_n) \end{aligned} \quad (38)$$

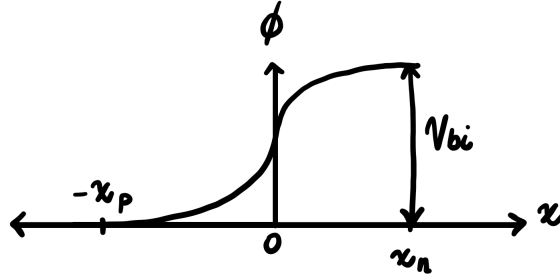


Figure 6: Abrupt junction $\phi_p(x)$ and $\phi_n(x)$ curves.

We defined earlier that the depletion region has a width. With the numerous set of equations above we can find a closed form of the depletion width. Relying on charge conservation we may eliminate either the x_n or x_p term.

$$\begin{aligned} N_Ax_p &= N_Dx_n \\ \Rightarrow x_n &= x_p \frac{N_A}{N_D} \\ w &= x_n + x_p \\ &= x_p \left(1 + \frac{N_A}{N_D}\right) \\ \Rightarrow x_p &= \frac{w}{1 + \frac{N_A}{N_D}} \end{aligned} \quad (39)$$

The above derivation could be found in terms of x_n , which will not be included. The derivation is identical regardless if you use x_n or x_p . To continue the derivation we need utilize the simplified V_{bi} equation to reduce terms.

$$\begin{aligned}
V_{bi} &= \frac{\mathcal{E}_{max}}{2}(x_n + x_p) \\
&= \frac{\mathcal{E}_{max}w}{2s} \\
&= \frac{qN_Ax_p}{\epsilon} \cdot \frac{w}{2} \\
&= \frac{qN_Aw}{\epsilon(1 + \frac{N_A}{N_D})} \cdot \frac{w}{2} \\
&= \frac{qN_Aw^2}{2\epsilon\left(\frac{N_A+N_D}{N_D}\right)} \\
\Rightarrow V_{bi} &= \frac{qN_A N_D w^2}{2\epsilon(N_A + N_D)} \tag{40}
\end{aligned}$$

$$w = \sqrt{\frac{2\epsilon(N_A + N_D)V_{bi}}{qN_A N_D}} \tag{41}$$

Equations 40 and 41 have abstracted key parameters of the PN junction purely in terms of doping constants. We know that $V_{bi} = \frac{KT}{q} \ln\left(\frac{N_A N_D}{n_i^2}\right)$. Hence, the width of the depletion region is purely in terms of well-known constants and the doping density. For the sake of file length, the derivation will not be done for the x_n , the formula and derivation are identical.

One-sided Abrupt Junction:

The depletion width and built-in voltage equations form the basis of many questions which could be asked in quizzes or exams. For example, does a large built-in potential have a small or large depletion region, and why? Eq. 41 tells us that increases to V_{bi} will increase w . Assuming the temperature is constant, the only way to modify V_{bi} is to change the doping. More charge is needed to satisfy a given V_{bi} , therefore, the depletion region must increase. Other questions we can ask is how does the ratio of N_A to N_D vary the depletion region width. Within the square root term, approximations can be made to $(N_A + N_D)$. Say we have a PN junction which has very large doping in the p-type and light doping on the n-type. To keep things abstract, let's say a 1000x ratio between N_A and N_D . The depletion width equation then simplifies to

$$\begin{aligned}
w &= \sqrt{\frac{2\epsilon N_A V_{bi}}{qN_A N_D}} \\
&= \sqrt{\frac{2\epsilon V_{bi}}{qN_D}} \tag{42}
\end{aligned}$$

Interestingly, when a junction is doped heavily on one side the depletion region is dependent on the lightly doped region. Alternatively, if take the equation $w = x_n(1 + \frac{N_D}{N_A})$, applying $N_A \gg N_D$ implies $w = x_n$. That is the PN junction is one-sided, extending into the region with lower concentration. This junction is the special junction mentioned earlier in the device. For context, one-sided junctions are incredibly common when connecting metal to semiconductors. Metals have their sea of electrons which would interact with the semiconductor forming a junction. These junctions form Schottky barriers and compose the basis of Schottky diodes. Placing a one-sided junction between the metal contact and the semiconductor helps smooth the transition between the free charge in the metal and the semiconductor's doping level. The charge density and electric field plot are provided below. The voltage curve is not included because it is quite similar to the stand PN junction, only with different curvature strength.

The equations developed for the PN junction may be simplified respectively. Going through the derivation again is a good exercise but not one that will be provided here. The relevant formulas are provided below

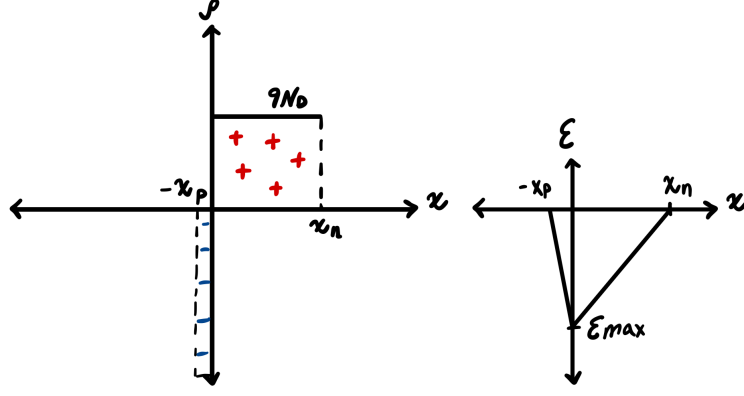


Figure 7: One-sided abrupt junction charge density and electric field plots.

for a P⁺N junction. Note that the (+) denotes which type is heavily doped.

$$w = \sqrt{\frac{2\epsilon(V_{bi} - V_{bias})}{qN_D}} \quad (43)$$

$$\begin{aligned} \mathcal{E}_n(x) &= -\frac{qN_D}{\epsilon}(x_n - x) \\ &= -\frac{qN_D}{\epsilon}(w - x) \end{aligned} \quad (44)$$

$$\mathcal{E}_{max} = -\frac{qN_D w}{\epsilon} \quad (45)$$

$$V_{bi} = -\frac{qN_D w^2}{2\epsilon} \quad (46)$$

Where V_{bi} can be found either by integrating the electric field curve, assuming no field lies within the p-type depletion region. Alternatively, you can derive the equations as done above but observe that little built-in voltage is present over the p-type depletion region. The above equations can be helpful if you know the junction is either P⁺N or PN⁺, however it is often safer in an exam to utilize the standard PN junction equations. If you have a PN⁺ junction it is as simple as swapping N_A and N_D because the dependence on the x_n and x_p simplifies to w in either case.

Bias Voltage:

The above equations are nice but apply only to thermal equilibrium. If anyone wanted to use the device, there is no way for them to predict the depletion region width when a bias voltage is present. Luckily, the inclusion of a bias voltage is simple. The fundamental assumption that the n and p-type bulk are neutral implies that the voltage across the terminal of the PN junction is applied directly across the junction. If we forward bias the device, the depletion region reduces ($w \downarrow$). Inversely, if the device is reverse biased the depletion region must increase due to the voltages being aligned ($w \uparrow$). To reiterate, forward bias decreases the voltage across the junction, whereas a reverse bias increases the voltage across the junction. Replacing V_{bi} with $V_{bi} - V_{bias}$ will implement the above relations.

$$w = \sqrt{\frac{2\epsilon(N_A + N_D)(V_{bi} - V_{bias})}{qN_A N_D}} \quad (47)$$

Note that $V_{bias} = V_{bi}$ reduces the depletion width to zero. It is an incredibly common mistake to forget the negative sign on the bias voltage. A negative term within the square root can serve as a sign that a mistake has occurred. The zero-width depletion region also serves as the point where the device is truly "on" and carriers can diffuse across the junction.

Conclusion:

This subsection concludes our initial discussion on the analysis of PN junctions. We have introduced the general process used to derive expression for the PN junction. Concepts such as Gauss's law, Poisson's equation, charge conservation, and many others were used to characterize the space charge region. Some calculus was necessary to convert Gauss's law and Poisson's equation into the built-in electric field and potential. We only focused on the abrupt junction in this set of notes, and will continue with the linearly graded junction in the next set. The process for a linearly graded junction is nearly identical to the abrupt, but presents some interesting simplifications.