Introduction to Semiconductor Devices

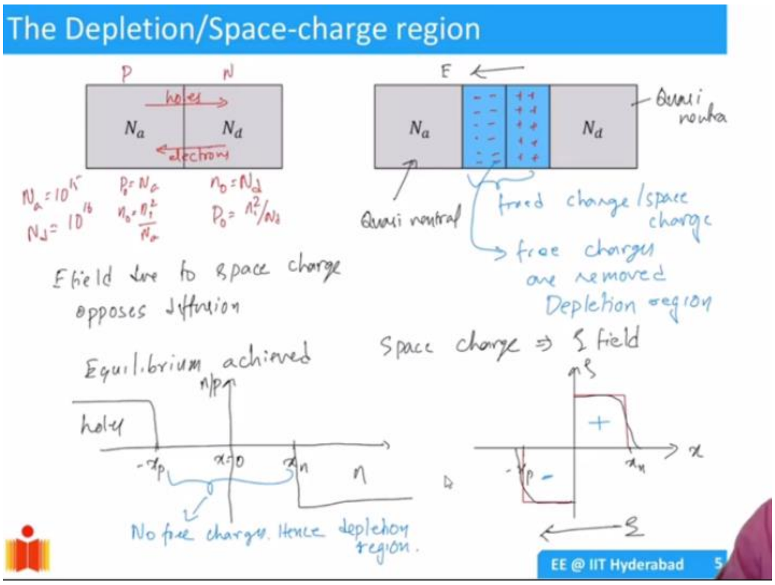
Dr Naresh Kumar Emani
Department of Electrical Engineering
Indian Institute of Technology – Hyderabad

Lecture - 4.2
PN Junction Electrostatics

This document is intended to accompany the lecture videos of the course "Introduction to Semiconductor Devices" offered by Dr. Naresh Emani on the NPTEL platform. It has been our effort to remove ambiguities and make the document readable. However, there may be some inadvertent errors. The reader is advised to refer to the original lecture video if he/she needs any clarification.

We discussed the physical aspects of a PN junction. Now, let us try to understand the electrical behaviour of a PN junction.

(Refer Slide Time: 00:18)



So, we saw that a PN junction is consisting of a P type region and an N type region. So, PN junction. Now, what happens? You know a P type region has excess of holes. So, the hole density(p_0) is basically N_a and electron density(n_0) in the equilibrium,

$$p_0 = N_a, \qquad n_0 = \frac{n_i^2}{N_a}$$

Similarly, for donors, the electron density(n_0) is going to be N_d and p_0 is going to be,

$$n_0 = N_d, \qquad p_0 = \frac{n_i^2}{N_d}$$

So, let us assume that we have doping concentrations in the range of 10^{15} or so. Let us say $N_a=10^{15}$ and $N_d=10^{16}$. Just for example. This is typical ranges. What would happen? What do you expect? So, the P region has excess of holes while the N region has excess of electrons. And whenever you have this concentration gradients, we have seen that there is this exchange there is a diffusion that happens.

So, basically holes will diffuse in this way. Holes will diffuse into the N side and electrons will diffuse into P side. So, when will this process stop? Will it go on forever? That is something to think about. So, what happens when holes diffuse into the N side? Holes are essentially positively charged particles. So, when they go what you are left with as an acceptor ion which is having a net charge of minus.

So, whenever your holes go into the N side, you leave behind and accept a charge so which I will represent by a negative sign here. So, it leaves some negative charges whenever this transfer happens. And correspondingly there is going to be some positive charges which are left behind because of the donor atoms. So, wherever there is a donor atom, electron is going to get away from that. So, the resultant charge would be positive.

And remember, these are fixed charges. These are fixed charge. And sometimes, we call them as space charge to distinguish it from the free charges. Electrons and holes are the free charges. They can keep moving around in the lattice whereas the fixed charge or space charge cannot move around in the lattice. And also, another aspect to think about it is since we have removed electrons and holes from here.

So, this region in the blue. Free charges are removed. Hence, we also call it a depletion region. The free charges are depleted. Depletion is reduced. So, since in this case you do not have any free charges. So, you have a depletion region without free charges whereas if you look at this region, we call this as a quasi-neutral region. And this also is called as quasi-neutral region.

So, essentially these are neutral regions because they have equal number of acceptor and holes and donors and electrons. So, these are not ionized. Oh sorry. I should not say that. Because when I say ionized you might confuse with the donor ionization. So, donor atoms are ionized but you have equivalently holes and the acceptor ions equal. So, that is why neutral. And similarly in the donor region you have donors and electrons.

So, quasi-neutral regions. So, what happens if you have a quasi-neutral region like this? Sorry. What happens if you have a depletion region like this? So, you have a region where you have a positive charge, and you have a region where you have a negative charge. So, there is going to be an electric field. Because of charge because of space charge implies electric field. What is the direction of this electric field?

This is very important to understand semiconductor devices. Whenever you look at any charges, look at what direction the fields are. So, now, the field is going to be in the direction opposing x like if I take x on left to right the field is going to be from right to left. Because, on the right, we have the donor atoms and those are positively charged. And on the left, you have acceptors which are negatively charged.

So, right to left. So, you have this charge. So, what if you have a charge? Well, this charge helps in opposing the diffusion. So, the E field due to space charge opposes diffusion. Why? Because diffusion is because of driven by the concentration gradient. So, there are excess of holes in the P side. So, they will try to get to the n side. But once this depletion region is formed, the electric field opposes further flow of holes.

Similarly, it also opposes the flow of electrons from the N side to the P side. So, because of that there is an equilibrium reached. So, the forward driving force is a diffusion the opposing force is a depletion field or depletion region that is formed and that is creating an electric field. So, because of that there is an equilibrium. And that results in what is known as a built-in potential equilibrium achieved.

So, basically this has a lot of implications. You know the size of the depletion region is going to have a lot of implications on our device function. So, what I will do is I will try to plot to get you get a idea what is happening. Let us try to plot first the charge densities. So, I will call x equal to 0 as my metallurgical junction. I will use x_n as an edge of the depletion region on the N side.

And I will use minus x_p as a edge of the depletion region on the p side. So, now, let me try to do this. We will take it like this just for convenience. So, now, let us plot what happens with the free charges? I want to plot n or p whatever. So, how is the charge distribution of n? The electrons are negative charge. Let us plot it here. So, the charge distribution is going to be like this n. Negative charge and they are extending from x_n to infinity.

So, this is how the charge distribution looks like. And, what about p. p is positive charged. So, edge of the x_p. So, remember you know these are I am drawing like a rectangle or it looks close to a rectangle not exactly a rectangle. But in the real case of course this is going to be some sort of a smooth curve. But you know we will take the approximate we will make an approximation that the edges of the depletion region are sharp.

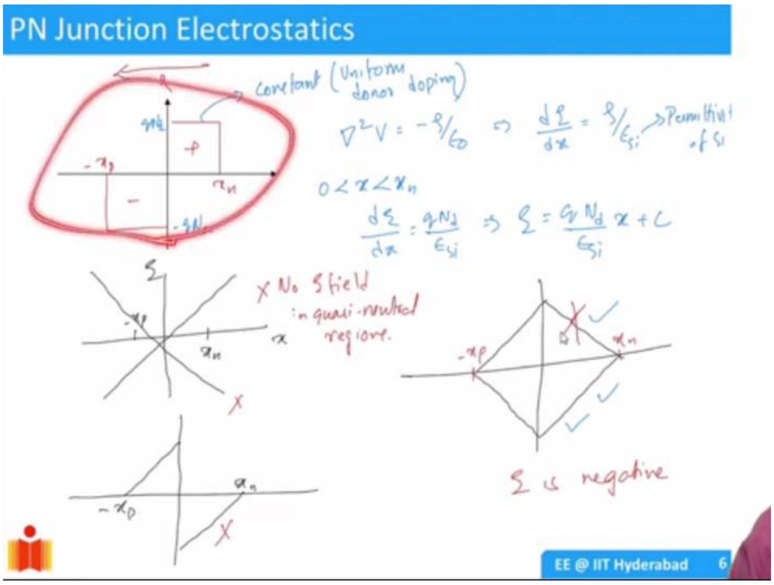
So, this is your free charges. And these are holes, and these are electrons. Remember no charges here no free charges. Hence, depletion region. The definition of depletion region it does not have any free charge. So, this is one part of it. Now, I also want to plot how does the space charge look like because that has a lot of implications for us. To do that, I will take axis. And then I will say I want to plot x versus ρ.

I will just give a common name ρ essentially the charge density space charge density. How would it look like? Take a moment and think about it. So, this is going to be your let us say x_n and this is going to be minus x_p and 0 is the junction. So, I will say that I will have acceptor ions in the p side. So, I will make it like this and I am going to have donor ions on the n side. So, this is going to be my positive charge.

This is going to be my negative charge. So, always remember. So, this is an approximation. This is like we call it a depletion approximation which is essentially meaning that it is going to be abrupt. But if you do the exact calculation it might turn out to be something like this close enough. Actually towards the end of this week, I will show you a small tutorial where we will actually show you how the calculation looks like with the exact simulations and then this depletion approximation.

So, we have these regions. So, what? So, we said that there is going to be an electric field in this direction. But then this will help us to calculate. We should be able to calculate the electric fields. Once we know the electric field, we can calculate currents and all that. We can also calculate the potential.

(Refer Slide Time: 10:54)



So, we would like to calculate the current. So, we have a situation where I will just repeat it for our completeness. So, there is a positive charge and a negative charge. So, this is let us say, ρ. What is the magnitude of the charge? This we already defined as x_n. This is minus x_p. What is the magnitude of the positive charge? Think about it carefully. So, if you are thinking about a diffusion process.

Because holes are diffusing and electrons are diffusing from the N side to P side. So, when they diffuse you might be tempted to think that there is going to be some sort of a gradual variation. But always remember when the electrons leave, the positive charge is coming because of the fixed charges and those are uniformly distributed because the doping is uniform. So, because of that your profile here this is constant because uniform donor doping.

So, it is a constant. So, what is the charge? This is going to be q times N_d. So, basically charge of electron times the donor atoms. And then similarly, on the acceptor side, it is going to be minus q times N_a. N_a is a density charge q minus because it is a negative charge. So, now, how do we solve this? To solve this, we need to solve Poisson equation. We might have been familiar with Poisson equation.

Essentially, it tells you that the Laplacian of potential V is going to be,

$$\nabla^2 V = \frac{-\rho}{\varepsilon_0}$$

ρ is your charge density. And we can rewrite this. Sometimes, we also call it Gauss law which essentially tells you that,

$$\frac{dE}{dx} = \frac{\rho}{\varepsilon}$$

And this epsilon is a permittivity of the material where we are solving it. So, in this case, we will solve it in silicon. So, we will say silicon.

So, this is permittivity of silicon or Si. If you solve it in Gallium arsenide, it is going to change lightly. So, we have Gauss law. How do we solve it? Well, to solve this, we can try. You know we have to do it in 2 regions because we have uniform doping in x greater than 0 and x less than 0. So, let us take the case x is basically between 0 and x_n. What should happen?

$$0 < x < x_n, \qquad \frac{dE}{dx} = \frac{qN_d}{\varepsilon_{si}}$$

This is the equation. So, this implies that electric field is going to be,

$$E = \frac{qN_d}{\varepsilon_{si}} x + C$$

I mean that makes sense. We have a constant doping. So, it is going to be linear. Before we solve it further let us try to think about what direction the field should be. What should be the quality to behaviour of the field? So, can I have a field which is looking it is going to be linear.

So, is it possible to have a field which looks like this, E versus x? Can it be this? This is a linear function. Would it work? Well, if you have something like this, it is not going to work. The reason is remember this is let us say x_n and this is x_p. Outside the regions outside the depletion region we said it is quasi-neutral regions having lot of free charges. Whenever, you have free charge, if you if there is let us say, by chance some electric field small electric field.

They are going to rearrange in such a way that the electric field becomes neutralized. So, because of that in the quasi-neutral regions you cannot have no electric field in the, no E field in quasi-neutral regions. So, this sort of function cannot work. Similarly, this function also will not work same arguments. So, you are reduced to, finding a function which has to be 0 on the outside regions only in the depletion region it has to be nonzero. So, is it possible to have something which is like this?

Well, it turns out that whenever you have sort of a jump like this. So, this is let us say this is minus x_p and this is x_n. If you have a jump in electric field like this, it has to be you know it has to have a charge at the interface which is not the case in this. So, this also will not work.

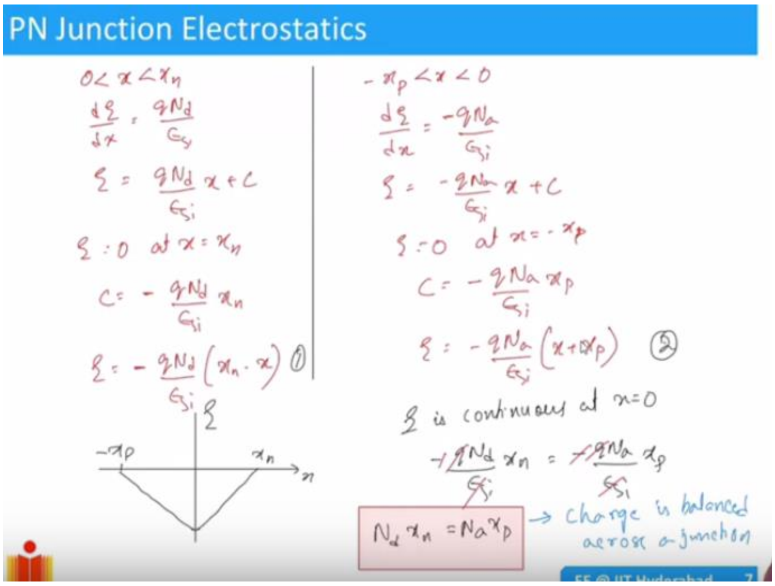
Because you might have studied in electromagnetics that whenever you have a jump in the electric field there is a charge density associated with the interface.

That will not work. So, what is left with us is basically which is looking something you know. Let me do it this way. So, what will be left with this? Something like this. This is one possibility. And this is another possibility. This will satisfy our equations. So, this or this. So, fields are 0 outside and it is continuous at x equal to 0. But, out of which only one is correct, other one is wrong. x_n, this is minus x_p.

And the correct answer is the bottom one. I am telling you the answer. We will show you mathematically. You can always get lost in the mathematics. But, physics cannot be wrong. Physics finally mathematics has to agree with the physical or qualitative picture qualitative understanding. So, the reason we say electric field is negative E field. E is negative. Because we saw in the space charge, we said that the field is pointing in the minus x direction.

So, the field is pointing in the minus x direction. As you know x is increasing, it should be the negative direction. So, because the field is opposing the in the opposite to x that is why it has to be negative. So, let us try to do it mathematically. So, let us just you know make sure that you have this in mind. So, let us take the equations and solve it. We will get to the answer that we just discussed.

(Refer Slide Time: 18:53)



So, let us say for x is in the range in the N region.

$$0 < x < x_n, \qquad \frac{dE}{dx} = \frac{qN_d}{\varepsilon_{si}}$$

This is your Gauss law. And so, E is going to be,

$$E = \frac{qN_d}{\varepsilon_{si}} x + C$$

So, remember what we said, at the edge of the depletion region the electric field has to go to 0 because beyond that it is just 0. I mean it is like a it is not a pure metal, but it has a lot of free charges.

So,

$$E = 0 \; at \; x = x_n$$

To the outside it, field has to be continuous remember in the you know quasi-neutral regions field has to be 0. Well, we will correct it but later on but right now, it will be 0. So, what happens? So, C is going to be what,

$$C = \frac{-qN_d}{\varepsilon_{si}} x_n$$

So, your field is going to be E,

$$E = \frac{-qN_d}{\varepsilon_{si}} (x_n - x)$$

So, this is the equation for the field. And then, what happens on the p side? Well something similar happens on the p side. Let us say I will consider between -x_p and 0.σ, remember is uniform.

$$-x_p < x < 0, \qquad \frac{dE}{dx} = -\frac{qN_a}{\varepsilon_{si}}$$

And so, that implies that your field is going to be,

$$E = \frac{-qN_a}{\varepsilon_{si}} x + C$$

And now,

$$E = 0 \; at \; x = -x_P$$

So, what does C turn out to be?

$$C = \frac{-qN_a}{\varepsilon_{si}} x_p$$

So, the E field is going to be,

$$E = \frac{-qN_a}{\varepsilon_{si}} (x + x_p)$$

So, now, is it making sense? Well, let us try to plot it. One of the ways to check that would be if you take a axis like this and then I take my x. This is your x and this is your E field. So, the way this is plotted this is a linear function and it is negative. So, this is going to be your E field. This is x_n. And on this side, you have minus x_p and the E field is going to be like this.

So, this is how electric field looks like. And one of the checks would be it has to be continuous. Whenever there is no additional charge at the interface it has to be continuous. So, E is equal to 0 at x equal to 0 sorry E not 0 I am sorry. E is continuous. x = 0. So, let us say this is 1 and let us say this is 2. E has to be continuous. So, I will put x = 0.

So, this is going to imply,

$$\frac{-qN_d}{\varepsilon_{si}} x_n = \frac{-qN_a}{\varepsilon_{si}} x_p$$

This is correct. Please make sure that you work out these things. These are not complicated math. It is simple algebra.

But you should work it out so that you are convinced. So, now, these things will go out. So, negative charge also goes out. So, essentially, we are getting a relation which is telling us that,

$$N_d x_n = N_a x_p$$

This is a fairly important relation for you to remember whenever you are analysing PN junctions. What this is telling you is that remember N_d is a doping density.

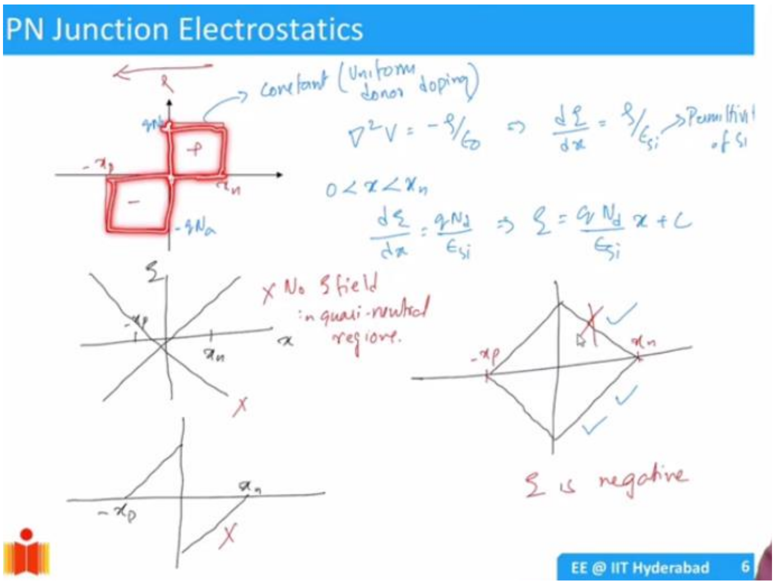
And N_d times sorry N_d is a donor density times the depletion width on the N side is going to be equal to Na times the depletion width on the p side. So, if you think about it is essentially saying that charge is balanced across the junction. You cannot have different amounts of charge on n side and n side. You can have scenarios where you have let us say I can have we will come back to it in a moment.

Basically, I can have a large N_d. If I have a large N_d my xn is going to be small and N_a can be small. So, x_p is going to be large. Eventually,

$$N_d x_n = N_a x_p$$

This is a criteria. So, this cannot be violated. The total amount of charge like in this picture what we are showing.

(Refer Slide Time: 25:32)

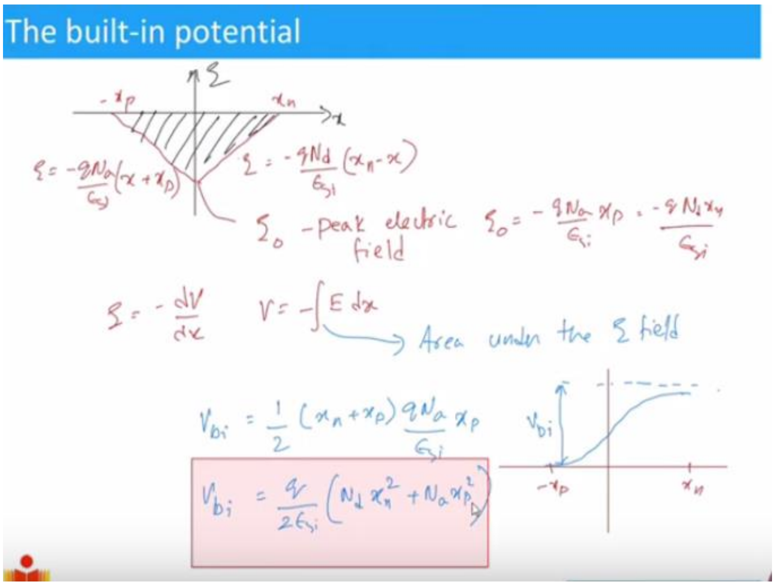


Here, this picture the area of the rectangle on both sides has to be equal. There is no other way because other regions are quasi-neutral. They are not contributing to any charge. So, the charged region should be in such a way that there is positive and negative, and the total amount of positive charge should be equal to the total amount of negative charge. So, this x_p is telling you the depletion width on the P side.

So, 0 to x_p is the depletion width on the p side multiplied by the magnitude of the charge N_a should give you the area of this rectangle. Similarly, x_n times N_d. So, this one should give the area of the rectangle on the positive side. So, the total charge should be equal to this. It has to be balanced. That is all this is telling you. So, do remember this equation. We will use it repeatedly.

So, we are able to get to the electric field. What then? So, once you have an electric field, we can compute what is known as the built-in potential of a semiconductor. So, that can be done fairly simply by considering let us say let us redraw the electric field for reference.

(Refer Slide Time: 26:55)



So, I am just doing it multiple times because I want you to be comfortable drawing these diagrams. So, these might look like sketches but then they have a lot of insight that you really benefit from. So, this is your x, and this is your electric field. So, this is going to be my x_n and this is minus x_p. So, what is my peak electric field? Remember, what is my peak electric field? On this side, the equation was let us try to guess it.

Equation was,

$$E = \frac{-qN_d}{\varepsilon_{si}}(x_n - x)$$

As x increases the electric field decreases. At $x=x_n$ it is going to be zero. So, if you try to memorize these equations it will be difficult. So, I would advise you to simply think about it this way. What will be the electric field on this side?

Electric field is going to be,

$$E = \frac{-qN_a}{\varepsilon_{si}}(x + x_p)$$

Now, at $x=-x_p$ it is going to be 0. So, these are the equations.

So, what is the peak electric field? Let us call this as E_0. E at zero is a peak electric field. How much is that? I mean it should be on either equation you can take,

$$E = \frac{-qN_a}{\varepsilon_{si}}x_p = \frac{-qN_d}{\varepsilon_{si}}x_n$$

So, this is going to be equal. We know we saw that $N_a x_p = N_d x_n$ .So, this is going to be equal. So, this is your electric field. And from here, we can calculate what is known as built in potential. Because whenever you have an electric field if you integrate it essentially you get the potential.

So, we know that electric field is going to be,

$$E = -\frac{dV}{dx}$$

So, V is simply going to be,

$$V = -\int E\, dx$$

So, I just need to integrate this. So, I mean there are 2 ways, a hard way, and an easy way to do it. The hard way would be to try to integrate it carefully and then solve for it. You can do it, but I will leave it to you as a homework or exercise. What I will do is I will take an easy approach.

I will say that this integral is simply area under the curve area under the E field line basically. So, the potential is simply going to be this area. That is the meaning of integration anyway. So, what is this going to be? I will say V built in is going to be E_0 which is half width into height. The width is x_n plus x_p into height is going to be E_0. E_0 is going to be let us take one of the expressions.

Let us take it to be minus q N_a well minus is not required because we are calculating area. It is not going to be minus q N_a divided by epsilon silicon into x_p.

$$V_{bi} = \frac{1}{2}\left(x_n + x_p\right)\frac{qN_a}{\varepsilon_{si}}x_p$$

If you try to rearrange it I can write it slightly differently. I can take q out and I will also take 2 epsilon. And then I will say $N_a x_p = N_d x_n$.

$$V_{bi} = \frac{q}{2\varepsilon_{si}}\left(N_d x_n^2 + N_a x_p^2\right)$$

This is an expression for built in potential of a device. Even though this is algebraically it seems interesting, we almost will never use this expression. It only is telling you it is this the relevance of this expression is only to tell you that the potential is quadratic. It is square root. It is not as linear function, but it is a square. So, the way the potential is going to vary across the junction is going to be this way.

So, you have minus x_p x_n. Of course, you know, we have to take a reference. So, we will take at x equal to x_p minus x_p as 0. And then you will see that the potential is going to change in a quadratic fashion like this. Outside these regions, electric field is 0. So, there is nothing to integrate. So, this quantity we call as V built in. So, right now, what I said was we argued it from the mathematical standpoint.

And said that there is going to be a electric if there is an electric field there is going to be a potential which is simply integral of the electric field. So, in the next lecture, we will talk about the physical how can we intuitively understand it based on the band diagrams and so on. So, next lecture will deal with that. Right now, I just want you to understand the basic electrostatics. What We have done is not very complex.

You should be able to draw the space charge region how you know if you have an abrupt junction. And then how the electric field looks like. And then you can also analyse how the potential is going to look like. So, please verify you know this expression that I have drawn here this is going to be the answer that you get if you try to integrate it. If you try to integrate the electric field, you will get this.

Please verify that you get that. And that will be a good homework for you. I will see you in the next lecture. Thank you so much.