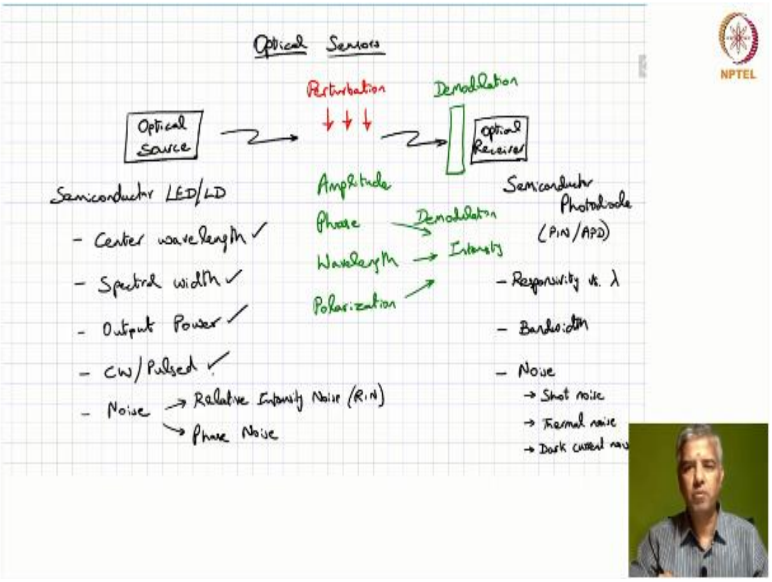
Optical Fiber Sensors
Professor Balaji Srinivasan
Department of Electrical Engineering
Indian Institute of Technology, Madras
Lecture 6
Optical Receivers - I

(Refer Slide Time: 0:16)



Optical Sensors

Perturbation

Demodulation

Optical Source

Optical Receiver

Semiconductor LED/LD
- Center wavelength ✓
- Spectral width ✓
- Output Power ✓
- CW/Pulsed ✓
- Noise → Relative Intensity Noise (RIN)
  → Phase Noise

Amplitude

Phase → Demodulation

Wavelength → Intensity

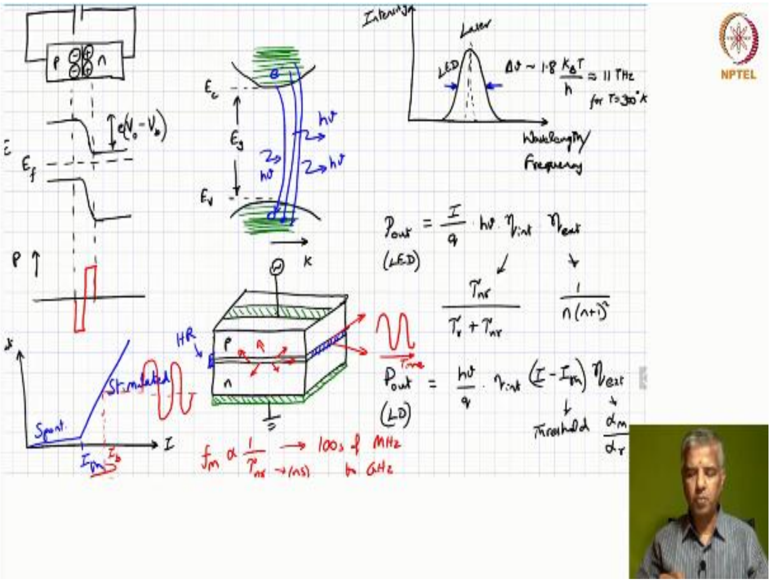
Polarization

Semiconductor Photodiode (PIN/APD)
- Responsivity vs. λ
- Bandwidth
- Noise
  → Shot noise
  → Thermal noise
  → Dark current nois

Hello friends, in the last lecture we started talking about the design of optical sensors and we realized that any optical sensor is going to have an optical source and the radiation from the optical source is going to be exposed to some perturbation and that perturbed light, the modulated light is going to be picked up by an optical receiver.

So, in our last lecture we were looking at typical optical sources which are the semiconductor light emitting diodes or laser diodes, we were looking at specific properties like center wavelength, spectral width, output power and how to directly modulate that light so that it could be continuous wave light, modulated light or pulsed light and of course, we have not talked so much about the noise and sources that we will come back to later.

(Refer Slide Time: 1:38)



But today what we would like to go into is talking about optical receivers, so before we go into that, let us just quickly recap, we said semiconductor, light emitting diodes or laser diodes is going to have, are going to be based on Pn junctions, forward by at Pn junction typically and it is typically a double heterostructure, so that all the carriers that are injected into this structure are going to be trapped or confined within this active region which has lower band gap energy compared to the surrounding regions and there we typically have this recombination of an electron with a hole and that is generating light.

And then we look at what is the difference between a light emitting diode, which is relatively in a larger spectral width, which is a laser diode whose spectral width is going to be typically much, much more narrow. So, this is for a laser, whereas this is for an LED.

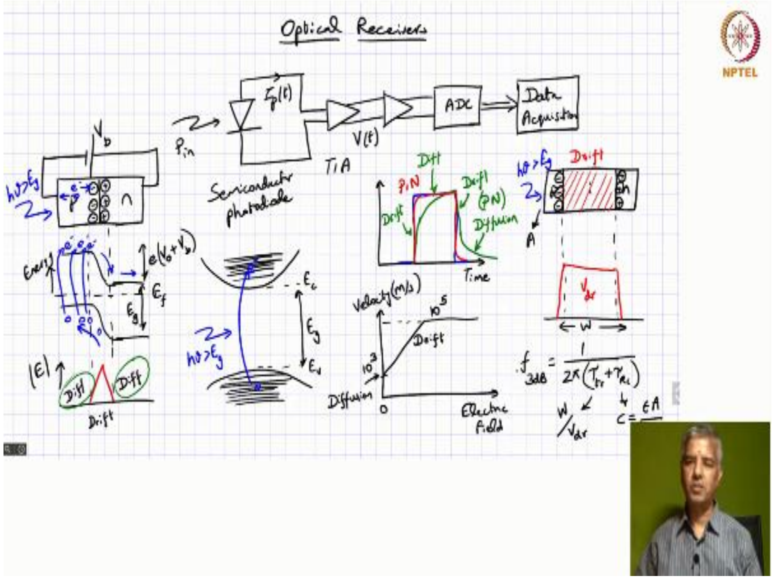
And then we were looking at the output power characteristics and we said the LED is going to have an output power given by this expression where eta int corresponds to the internal quantum efficiency and eta external corresponds to the fraction of photons that are escaping from this structure and the amount of light that you get to see.

Of course, I did make a small mistake here, this is actually not n square, rather this is this should be just n by itself, I had made a mistake yesterday or in the last lecture. So, that actually is the output power for an LED whereas as an output power for a laser diode is going to be given by this expression.

And this again happens to be the amount of power that is generated within the structure and if you are looking at the power that is extracted from one of the facets that is going to be actually having this other factor eta e this that corresponds to or you can you can just say eta external also, that is the power that you get to extract from this cavity. And this is typically given by eta m or rather I would put it as the mirror loss, the mirror loss corresponding to this facet over the total loss within that resonator.

So, that is the fraction of power that you get out of this light source and we looked at how the output power varies as a function of the injected current and we said that when you have a relatively low current levels, the same structure would act like a LED which is based on spontaneous emission of light, whereas if you go to a current level beyond the threshold current level, then you get to stimulated emission and we looked at how we could possibly modulate that light and all.

(Refer Slide Time: 5:23)



So, let us actually move on to the topic of interest for us today, which is Optical Receivers. What constitutes an optical receiver so what do we need to do in an optical receiver, we have a certain amount of light that is incident on the receiver and then you need to detect that and then convert it to a photo current, maybe using a semiconductor photodiode, typically.

And then from that photo current you need to extract voltage, you need to convert that to a voltage and then you need to possibly digitize that signal, that analog signal for most applications you would like to digitize that and so you might have a analog to digital converter and then the those digitized samples are finally acquired in a data equation unit.

So, typically the front end of your receiver. So, this is let us say is the input power that is falling on this photodiode, that photodiode generates a certain photocurrent which you may convert into a corresponding change in voltage. So, let us say this photocurrent, let us call that Ip of t and that is converted to a voltage, let us call this V of t, which you know which corresponds well, that photo current is converted to a voltage using what is called transimpedance amplifier, TIA in short.

So, this output voltage now needs to possibly be amplified further and then that amplified voltage is taken out and then you have what is called this analog to digital converter which generates your digital samples and then that, those digital samples will finally be fed into a data acquisition unit. So, this entire chain constitutes the optical receiver, so you are starting with a certain optical power and you finally have some electrical signal which you are acquiring as data.

So, this is a typical configuration, now the question is how do all these, all these components work and then and what are the typical issues that, well the typical issues as far as an optical receiver is concerned, we try to quantify that in the last lecture is what is the signal-to-noise ratio, that you have for this optical receiver, what so, to break that further you need to say for a given optical power what is the corresponding voltage that you are generating and then in the process of generating that voltage maybe you have a certain noise that you are accumulating in this receiver. So, you need to quantify that noise as well.

And of course, the other important aspect is how fast can your optical receiver respond to changes in the incoming radiation. So, that is characterized by the bandwidth of your receiver and we will start looking into some of those issues. To start with let us actually look into how this semiconductor photodiode, how that semiconductor photodiode works and how do we design a semiconductor photodiode for this optical receiver.

To look into that maybe we can look at a Pn junction as we did previously, so we start with let us say a P-type and an N-type semiconductor and if we draw the energy level diagram for this it will look something like this, this is your fermi energy and then for the N-type semiconductor the conduction band is going to be closer to the fermi energy with respect to the valence band and in the P-type semiconductor it is going to be the other way, the valence band is closer to the fermi energy compared to the conduction band.

So, when you bring these materials together of course, we said we are going to have the majority carriers which are the holes in the P-type semiconductor going to move over to the N-type and the electrons which are the majority carrier in the N-type are going to move over to the P-type and they leave behind some fixed charges, so these are negative charges on this side and positive charges on the other side and that is what we call as the depletion layer, it is depleted of mobile carriers.

Anyway so, when you look at the energy level diagram, so you have something that looks like this and because of the presence of this fixed ions over here, we can say that when you look at the field across this and maybe I can use a different color to show this clearly, the field, the electric field because of these charges are going to look like this, so this is actually the energy and here we are looking at the magnitude of the electric field.

So, it sort of there is a built-in electric field because of these fixed charges over here and of course, you can choose to enhance that field by providing an external bias, in this case you can enhance it by doing a reverse bias which means the negative of the battery is connected to the p side and the positive of the battery is connected to the n side.

So, by increasing that potential difference that is applied to the Pn junction you can enhance the barrier over here, so why is that important? We will come back and look at that in a minute, so now we will look at the mechanism for light absorption how it works like a detector.

Well, we know that this corresponds to Eg that is the band gap energy so if you have any light coming in correspond, corresponding to an energy h nu and if h nu is greater than Eg it is greater than the band gap energy then you have absorption happening across the structure.

So, when light actually comes across this region it will tend to take an electron from the valence band and take it to the conduction band and it leaves behind a hole over here. So, such absorption happens all along the material and it generates this electron hole pair.

To see this clearly maybe we can look at what we did previously for the semiconductor light sources, we can go to the Ek diagram so that is basically your conduction band, this is your valence band and this difference in energy corresponds to Eg. So, this is Ec, this is Ev and then of course, there are a lot of these different levels that are possible and as we go away from the bandage you are going to have higher density of energy levels and so on.

So, what we are talking about is if we have a photon h nu whose energy is greater than the band gap energy, then you are having these transitions wherein you have an electron at the conduction band and you leave behind a hole in the valence band. So that is the primary criteria that your input energy, the photon energy has to be greater than the band gap energy.

So, now, we need to then see how the process of converting this to a photo current happens. Now, you have these photons that are streaming in and falling on the material and then that actually generates this electron hole pairs and the electron hole pairs if it is generated over here.

Let us say this is an electron and a hole here, the holes will move this way, the electrons will move this way and so the electrons essentially that are generated over here are going to go out to the external circuit like this. Whereas, the holes that are generated in these regions are going to see the external circuit and that constitutes a photocurrent in the external circuit.

So, what is wrong with this? This seems to work very fine, just a regular Pn junction. Well, there is one fine aspect to this which is actually limiting the response of a Pn junction, so when you talk about response what we are talking about is, let us say you have a pulse of light come in, a pulse of light that is coming in, so here we are looking at this as a function of time, now how fast does this detector respond to a change in the power, over here, so that is key question.

Well, if you want to look at that, you have to look at what is the velocity with which these charges are swept out because that actually determines what is the time over which the photo current is registering in the external circuit. To look at that you need to understand something about the carrier transport, when you look at the carrier transport, the velocity of these carriers as a function of time, well sorry, instead of considering it with respect to time, let us just consider that as a function of the electric field that is present across the structure.

If you consider the electric field across the structure, the velocity is going to be typically, if the field is quite low the velocity here would correspond to a certain level and as you increase the electric field, you will see that the carriers are accelerated across the structures, basically the velocity increases and it attains a certain velocity which is called the saturation velocity and then further increasing the electric field does not change the velocity.

So, this increase in velocity as a function of electric field is what we call as the drift phenomena, whereas if you do not have an external electric field, the charges move under the influence of diffusion. The key point here is this diffusion velocity it tends to be in the order of 10 power 3, if you are representing this in terms of meters per second, whereas the drift velocity, this saturation drift velocity can be in the order of 10 power 5 meters per second. So, there is going to be a huge difference in the transport mechanism of the carriers, whether it is under the influence of the electric field or it is not.

So, you can clearly see in this picture that these regions where you have a certain electric field the carriers would move under the influence of this electric field and move at drift velocity, whereas outside of this region, outside of this depletion region it moves with the diffusion velocity.

And the key point here is, if you change your external bias, let us say this is Vb, it is a negative bias that you are providing or a reverse bias that we are providing, if you increase your reverse bias, then this electric field gets enhanced, essentially the fermi energies are pride open or they are going to be displaced either side and correspondingly this barrier here is going to get enhanced.

You remember, when we talked about the barrier in the case of a forward bias junction, we said this corresponds to e times v naught minus this external voltage is what we were seeing as the barrier, in that case the barriers are shrunk that in allows the transport of carriers on either side.

Whereas in this case the barrier is actually enhanced, so enhanced by that voltage that you are applying over there. So, how is that helpful? Well, when you have an enhanced barrier, then you have a very large slope over here, that means these charges are going to get swept across at high velocity and that is what we are talking about in terms of the drift velocity, when the field is enhanced over here then that corresponds to a fairly high drift velocity.

Now, when you look at the overall response, this is how it is going to look. So, the photo currents that is generated in the external circuit it will have a component that comes up relatively quickly but then it is going to respond a little slow and then once again of course, when the power is shut off it is going to respond relatively quickly and then it is going to be a slow component over here. So, what is happening?

Well, these regions over here and over here, these would correspond to the drift velocity because those carriers which are under the influence of drift they will actually go out to the external circuit first so that actually creates this initial signal, initial photo current and these regions which are relatively slower that is because of diffusion, all these events over here, wherever diffusion is happening, those will constitute a relatively slow component.

So, this is once again diffusion, this is drift. So, essentially we are seeing that in a Pn junction diode, the diffusion process is slowing down the response of the photodiode and if you wanted to actually have a faster response, you will go on to, you will have to modify the structure, how do you modify the structure to achieve a faster response? You can actually insert an intrinsic layer between the p and n regions, so you can essentially make what is called a PIN structure.

So, what is the advantage of going for a PIN structure? Well, the advantage is that just like the previous case on the positive, in the p layer you are going to have the majority carriers, the holes move out and then that is going to create this fixed charges that are negative charges and similarly the electrons are going to move out of this region and that is going to cause some positive charges over there.

So, when you look at the electric field because of this, what you will see is when you plot the electric field across this thing you are once again plotting the magnitude of the electric field, the magnitude of the electric field will be something like this, it will increase over here because of the presence of this negative charges and then will be almost uniform and finally this region it will, it will come back to 0.

So, you can see that most of your structure here is under the influence of the electric field, so in all these regions you essentially have the drift mechanism dominate and because of that when you, when you look at your overall response, your response is going to be something like this, it is going to be much faster compared to the response that you get from a PIN junction diode.

So, once again so what I m showing in green that corresponds to a Pn junction diode and whatever I am showing in red it corresponds to a PIN photodiode. So, you will see that most of the photodiodes that you have in the market are corresponding to a PIN structure and that is specifically for because of the reason that most of the structure especially around the i region, the electric field is relatively high and it is relatively uniform as well.

So, that means that most of your whatever absorption is happening just like what we are talking about there, you have a certain absorption that is happening across the structure but most of the absorption is happening within this, so this is h nu which is greater than Eg and most of the absorption is happening within this i region, that is the dominant aspect of the structure and there it is undergoing drift, transport and because of that you have relatively faster response in a PIN.

So, we can quantify all of this, we can quantify this in terms of the frequency response of this photodiode and that actually if you say the 3 db frequency which corresponds to the response of the photodiode where it actually goes down by 3 db or it goes down by half the response that was at lower frequencies, that will correspond to a factor that is what you call as the transit time and you have what is called the rc time constant for this structure.

So, what is the transit time? Well, the transit time is the longest time taken for any of these carriers to reach the external circuit. So, we know that in this region you have this drift velocity and so for any carrier that has to traverse across this region, let us say this entire region is of width w then tau tr would corresponding would correspond to w over the drift velocity. And tau rc corresponds to the rc time constant where r is the serial resistance that you have over here and c corresponds to the capacitance of this junction which is given by epsilon times A over w.

So, w once again is the width of this i region which almost is the width of the PIN and can be approximated as the width of the PIN junction diode because of the fact that the p and n regions are relatively small compared to the i region. So, that is what is in the denominator and A is the area over which this light is incident, so that is the cross sectional area that you have for this photodiode.

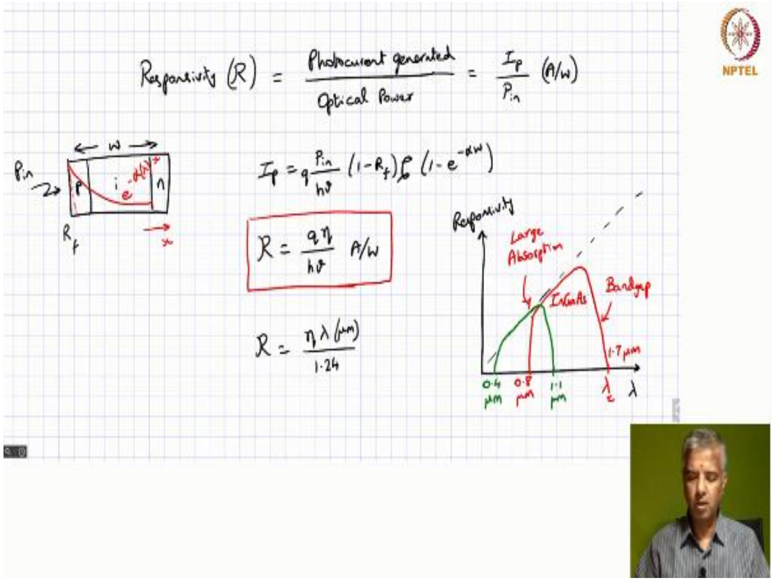
So, that seems to be contradictory, for one term w is in the numerator and other term w is in the denominator. So, how do you resolve that that aspect of it? Well you can resolve that by making this capacitance relatively low or the rc time constant, you can make it relatively low compared to the transit time and the way you can make it relatively low is by controlling the capacitance, you cannot do much about w because that will affect the transit time. So, what you typically do is you work with this area, if you can make the area relatively small, you can keep this capacitance value small and which means this rc time constant you can make it small with respect to the transit time.

So, you get to a point where you are typically working with w to have a control on the transit time and through that you can achieve the necessary 3 db frequency. So, just to give you an idea if you want to have a receiver with nanosecond response time, nanosecond rise time, why do you need a nanosecond rise time?

Well, it could be an application like we talked about in the last lecture, it could be an application in distributed sensing where you have a 10 nanosecond pulse that you need to capture at the receiver. So, if you want to have nanosecond type of response, you can say roughly something the order of gigahertz bandwidth you need, you at least need some hundreds of megahertz bandwidth to pick up a 10 nanosecond pulse.

And in that case what we are dealing with is, vdr is typically like the saturation velocity 10 power 5 meters per second, then what you need to work with is this w, so if you make w small enough you can actually achieve a relatively high bandwidth. So, that is the design philosophy that you work with. Now, of course, there is another aspect to this, which is, if you make w small you may actually compromise on the responsivity of the photodiode, so what is responsivity? So, responsivity is, just go to the next page.

(Refer Slide Time: 35:04)

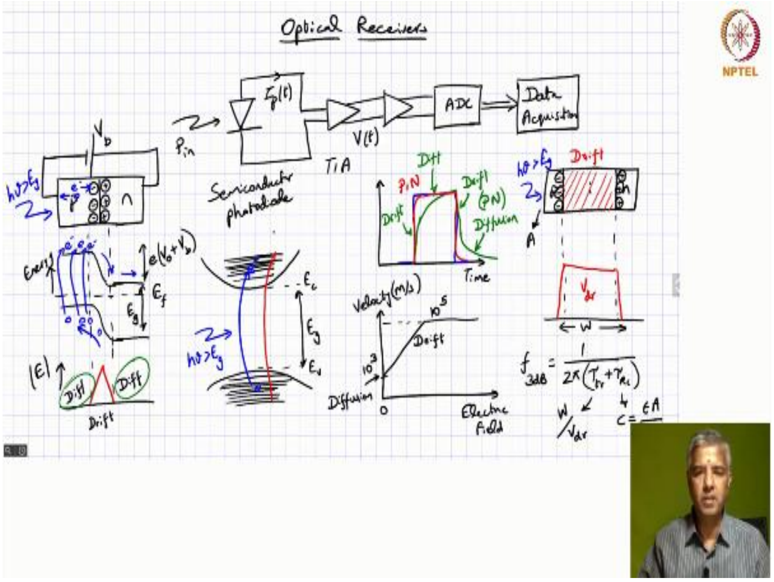


$$\text{Responsivity } (R) = \frac{\text{Photocurrent generated}}{\text{Optical Power}} = \frac{I_p}{P_{in}} \; (A/W)$$

$$I_p = q\frac{P_{in}}{h\nu}(1-R_f)\zeta\left(1-e^{-\alpha W}\right)$$

$$\boxed{R = \frac{q\eta}{h\nu} \; A/W}$$

$$R = \frac{\eta\lambda(\mu m)}{1.24}$$



Optical Receivers

The responsivity of the photodiode it is denoted by R is given by the photo current generated for a unit amount of optical power incident on the photodiode. So, you can just say this is Ip is the photocurrent and like we talked about in the previous figure Pn would be the power that is incident on the photodiode, so this is expressed in terms of amps per watt.

So, then you ask the question what is the photo current that is generated in this photodiode, how do you quantify that? So, Ip can be written as, first of all there is a certain rate at which you have these, you have these photons that are coming in, so what is that rate, that will be given by Pn over h nu, but if you look at a, let me just draw this PIN structure again.

If you look at light that is incident on this structure, not all the light that is coming in, so this corresponds to P and not all that light actually goes into the structure because this material is typically like indium gallium arsenide, let us say which has a refractive index greater than three and outside you have air which is refractive index of 1. So, there is a huge mismatch in the refractive index and that causes a certain reflectivity and that reflectivity you can say is what you call as the facet reflectivity Rf.

So, you would say that only the fraction which is not reflected, which is not reflected would be 1 minus Rf, Rf is actually you know a fraction between 0 and 1 and 1 minus Rf is the fraction of light that actually enters into the structure and then it has actually got to get converted into, it gets absorbed and it gets converted into a electron hole pair.

Now, it is not every photon that goes in gets converted to an electron hole pair, there is a certain photoelectron conversion efficiency, let us call that ceta and ceta typically when you are talking about direct semiconductor, direct band gap semiconductors it is in the order of 0.9, for an indirect band gap semiconductor it could be down to about 0.7 but it is around those values. So, that is actually the conversion efficiency.

So, now I am starting to talk about charges. So, I will have to multiply this by q and then that is the electric charge and then it is like, it corresponds to the number of photons that are absorbed within this structure, so when you look at absorption happening within that structure, that actually tends to go as an exponential, why is it an exponential? Because it conforms to what is called Beer’s Law.

So, you have a certain uniform absorption that is happening across the structure, so the or the power that actually, that gets absorbed it is expressed as e power minus alpha, let us say this direction is x, so e power minus alpha x. But within this entire structure whose width if you say corresponds to w, then the amount of light, so the value of this at this point corresponds to e power minus w. So, the amount of light that gets absorbed corresponds to 1 minus e power minus alpha times w, so this is an interesting point.

So, we previously said w, over here we said w has to be as small as possible, so that the transit time is small and correspondingly you can increase your bandwidth but what is the corollary to that? Well, if w is small then e power minus alpha times w is, so that becomes a relatively large number so, 1 minus so that tends to closer to 1.

So, 1 minus e power minus alpha w, this entire factor tends to be a relatively small number and if that is the case then the amount of photocurrent that you are generating is also small, so that is the trade-off, if you want high bandwidth from your photodiode, you typically go for smaller w but if you go for smaller w the photocurrent that you generate is relatively small.

And because of that if you looking at the responsivity which is Ip over Pn, so if you look at this then that is going to be given by q over h nu and this entire, these three factors here I can express it as eta, I will just, those three factors I will just combined them into one factor that I call as eta.

So, then eta is going to be low because of w being low and your overall responsivity which is expressed in amps per watt is going to be relatively lower. So, that is actually a trade-off you have as far as photodiodes are concerned, if you want faster response, you get lower responsivity.

And of course, in applications where you do not worry, you do not care too much about the fast response, you want to pick up signals that are just changing in the order of milliseconds or even seconds, then you do not worry about making w so small, so you can have a thicker width of this intrinsic region and thereby you can achieve higher responsivity as well.

So, it is interesting to see how responsivity works as a function of lambda. So, you see that q is a constant, h is a constant, nu is can be written as, nu is the frequency so it can be written as c over lambda so and c is a constant, so you will essentially see that R can be expressed as eta times, sorry eta times lambda, where lambda is expressed in terms of microns over 1.24, so if you just work out the math, you put this in terms of h is in Planck's constant, you put the value down for q and c you will come out with a round number of 1.24.

So, what does that tell you? That it tells that the responsivity is increasing as a function of lambda, so longer the wavelength more will be the responsivity. How far can you stretch it? If I am looking at the responsivity as a function of wavelength, we are saying that for a certain value of eta, the responsivity will be larger as a function of lambda, but that is not, you cannot keep increasing the responsivity, well that also brings up the question can responsivity be greater than 1?

Yes, surely responsivity can be greater than 1 because of the fact that it is scaling with respect to lambda, whereas eta if you look at it, eta cannot be greater than 1 because this is actually a fraction, this is a fraction is less than 1 and this is also a fraction that is less than 1.

So, this eta which is called the internal quantum efficiency, that quantum efficiency it can be a value that is less than 1 but R can be greater than 1, but like I said R cannot be arbitrarily large because of the fact that as you increase in wavelength, your frequency is decreasing and if your frequency is decreasing h nu that is the energy of the photon is decreasing and there comes a point where h nu is less than the band gap energy and if in that case you cannot have any absorption happening within the structure.

So, that is actually corresponding to this point over here, so you can call this as lambda c that is like the cutoff wavelength for a given material and for anything less than lambda c you have a responsivity so, yeah, it goes like this, but it cannot arbitrarily go down, it cannot have a responsivity over very short wavelengths because of the fact it tends to be limited like this.

Why is it limited like this? Well, you look at this picture and you say this is the kind of absorption that you have in the material, at one of the longer wavelengths, as you go shorter in wavelength.

Okay, what you are doing is, let us just go back to the previous page, as you go shorter and shorter wavelength you are actually going through a transition like this from one of these deep energy levels to one of the deep energy level, deep energy levels in the valence band to another deep energy levels in the conduction band.

And that means the probability of absorption is much higher which means that your alpha value, so alpha is actually a function of lambda, so the way I should write this is probably, I should write it as alpha which is a function of lambda multiplied by x, that denotes this variation over here.

So, if alpha is very large then you have a situation where all your absorption is happening within a very thin slice, as you go shorter and shorter wavelength, your absorption starts happening over thinner and thinner regions and you get to a point where all your absorption is happening right at the surface.

And that of course, generates an electron hole pair but then there is a very high density of electron hole pairs that are generated and then they start the electrons, for example, start colliding with each other and then they have to traverse this entire region where it can get reabsorbed and all that.

So, not a lot of that those charges are generating an external photo current and that is what is denoted over here. So, on this side we have a constraint because of the band gap, whereas on this side you have very large absorption and that actually limits the responsivity, large absorption corresponds to very small penetration depth and small penetration depth once again will lead to very poor responsivity.

So, if you do this for say indium gallium arsenide, for indium gallium arsenide the cutoff wavelength is around 1.7 micron wavelength and this sort of thing typically happens about 0.8, 0.8 microns so, this is for indium gallium arsenide. So, all these wavelengths are in the near infrared, so this means that it is for near-infrared sensors, indium gallium arsenide is a very good material and so it is for communication applications where the detection is typically happening around 1.5 microns indium gallium arsenide is very good.

But if you have a sensing applications where you are using visible radiation which is like 0.4, 0.7 nanometers and so on, you are not actually able to detect that using indium gallium arsenide detectors. So, what you use there is actually a silicon photodiode, so for a silicon photodiode the cutoff is about 1.1 micron, the band gap energy will correspond to a cut off wavelength of 1.1 micron.

So, if you add the same eta value it will tend to be like this, so this is typically good until about 0.4 micron, of course, we have applications where you need to go into the uv region, you tend to have what is called a uv enhanced silicon photodiode and in that case what you do is, the entire structure is made very thin and so it is very good responsivity over here.

But it is very poor responsivity for longer wavelengths because if you make this entire PIN structure to be relatively thin, then most of the longer wavelength light is going to just go straight through, so it is not going to register a lot of photo current within this structure. So, those are the typical characteristics of PIN photodiodes.