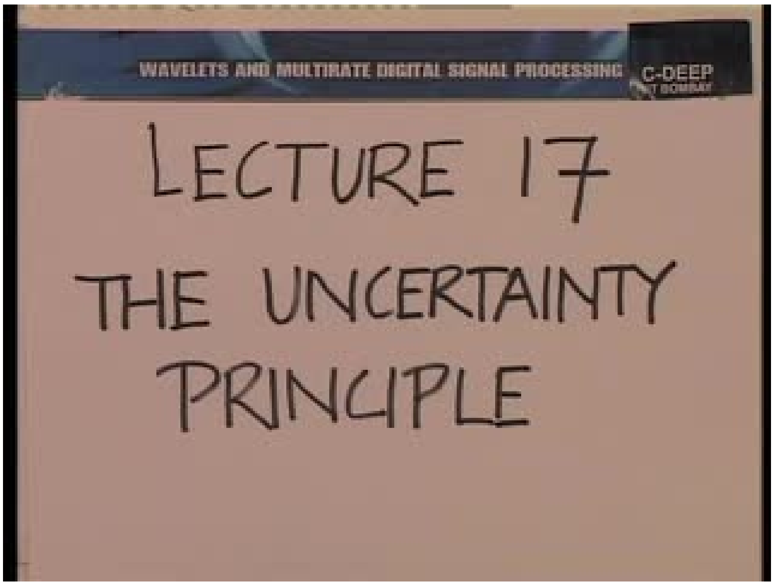
Advanced Digital Signal Processing -Wavelets and Multirate
Prof. V. M. Gadre
Department of Electrical Engineering
Indian Institute of Technology, Bombay

Module No: # 01
Lecture No: # 17
The Uncertainty Principle

A warm welcome to 17th lecture on the subject of wavelets and multirate digital signal processing; we build in this lecture a very important principle. In fact, in some senses the principle that lies at the heart of the subject of wavelets and time frequency methods namely, the uncertainty principle. Therefore, as you note today, we shall devote the whole lecture to a discussion of the uncertainty principle; laying the foundation of what uncertainty means first and then proceeding to obtain certain numerical bounds on confinement in 2 domains simultaneously.

(Refer Slide Time: 00:56)



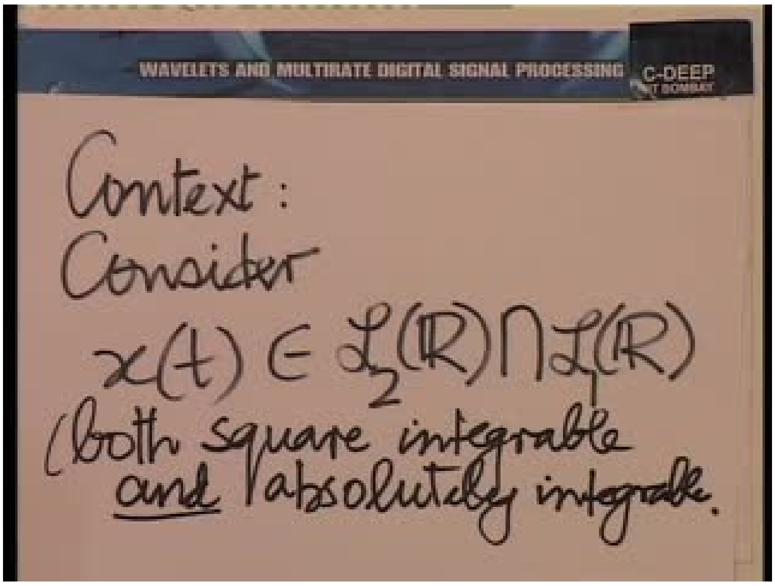
Let me first give an informal or a diffused non formal introduction to the idea of containment. Well, we did a little bit of that yesterday, in the previous lecture. But, what I intend to do now is to say a little more in terms of formality and then proceed to write down the mathematical relationships or definitions. Recall, we said that there is a very tight or a very strong kind of a notion of containment that would ask that you have compact support in both domains time and frequency. The function must be non-zero strictly over a finite interval of the real axis and must be non-zero strictly over a finite interval of the real axis in the frequency domain as well.

So, both in time and in frequency, you demand that the function be non-zero only over a finite part of the independent variable or the real axis. This is a very strong demand and yesterday we mentioned that it cannot be met ever. In fact, I also hinted at the idea behind the proof. It related to the fact that if you noted that the function was finitely supported, compactly supported on the real axis, there was certain properties of that function specifically the existence of an infinite number of derivatives which made it impossible for the function to be compactly supported or non-zero only on a finite interval of the independent variable in the natural domain. Natural domain can mean time, can mean space, whatever. Anyway, this was what we called the strong version of containment and we said that this was not possible. But, we had asked whether a weaker notion of containment could be admitted. Namely, we do not insist that the function be strictly non-zero over a finite interval but that most of its energy most of its content so to speak in some sense be on a finite interval of the independent variable which indexes it. Simultaneously, in the transform domain in the frequency domain we insist that most of the content be in a finite interval of the frequency axis this seems like a more reasonable requirement and to a certain extent this requirement can be met and as I said to give a diffused or a non-formal presentation of how it can be met, I shall begin this whole discussion by saying that we are finally going to come out with certain bounds on how much you can contain in the 2 domain simultaneously.

So there are several steps to reach this destination. The first step is to put down in a non-diffused, in a formal way what you mean by containment; what you mean by most of the content being in a certain finite range. We are also hinted at the approach we would take to do this briefly in the previous lecture.

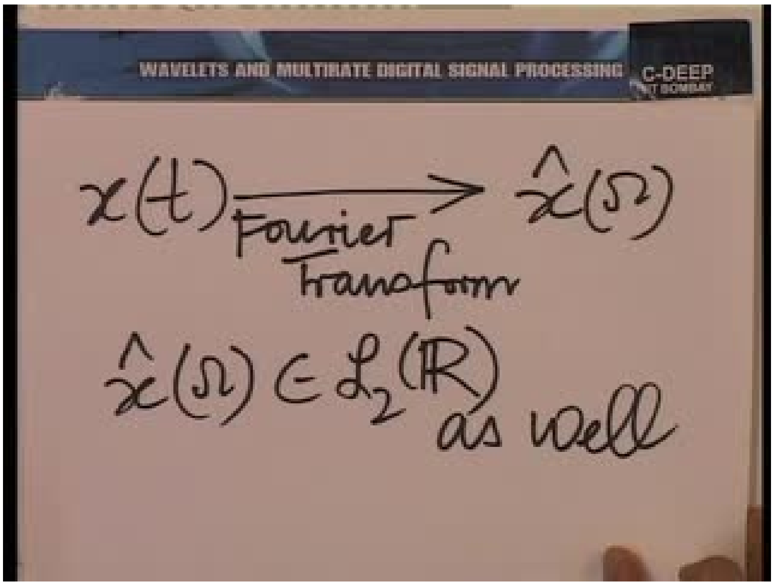
We had said that there are 2 ways of looking at it. You could think of the magnitude squared of the function and the magnitude squared of the Fourier transform as a 1 dimensional object and then you could talk about the center of that object, center of mass if you like. You could talk about the spread of the object around the center of mass by using the notion of radius of gyration or if you prefer to speak in the language of probability densities, then you could employ the idea of the density built from the squared magnitude of the function and another density built from the squared magnitude of the Fourier transform. You could then look at the mean of these densities and the variance of these densities. The variances are indicative of the spread; so this was a non-formal introduction. Now, we need to formalize it and that is what we shall do precisely. To begin with, put down formal definition, a formal explanation of the idea of spread.

(Refer Slide Time: 07:36)

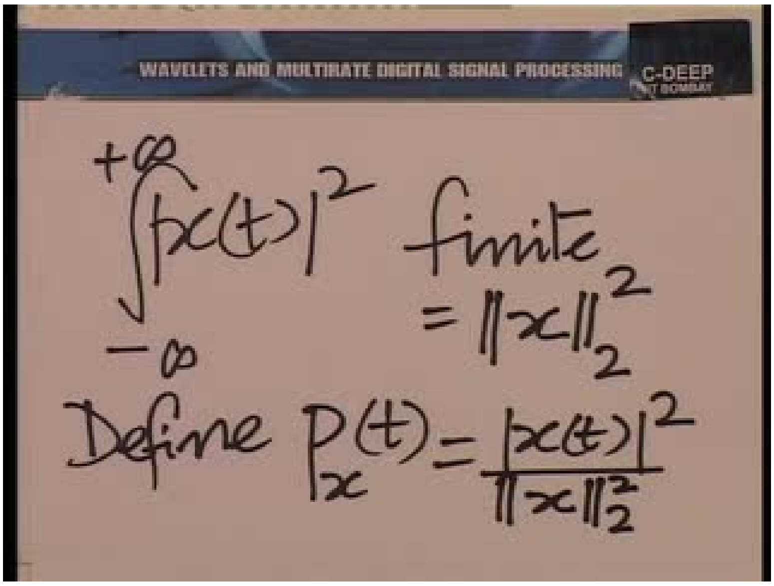


Now, we have to define the domain in which we are going to work. We are going to work in L 2 r we had agreed to that. It is always going to be the space of square integral functions. In fact, I must mention that sometimes we are actually going to work in the intersection of the space of square integrable functions and absolutely integrable functions. To be on the safe side, let us put down that requirement right now and let us put down the tie test of the requirements; namely, that the function belongs to the intersection of these 2. So, the context: consider, a function let say x of t which belongs to the intersection of L 2 R and L 1 R which means it is both square integrable and absolutely integrable; I think we should note that.

(Refer Slide Time: 08:54)

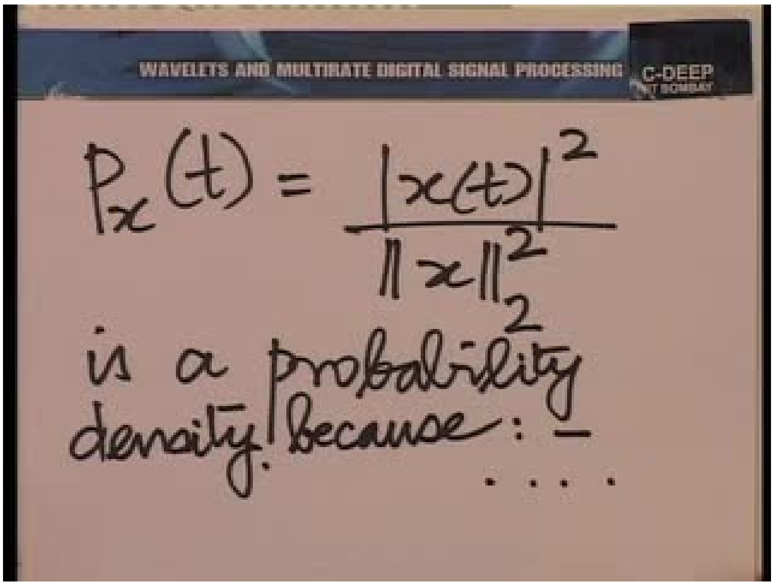


(Refer Slide Time: 10:15)

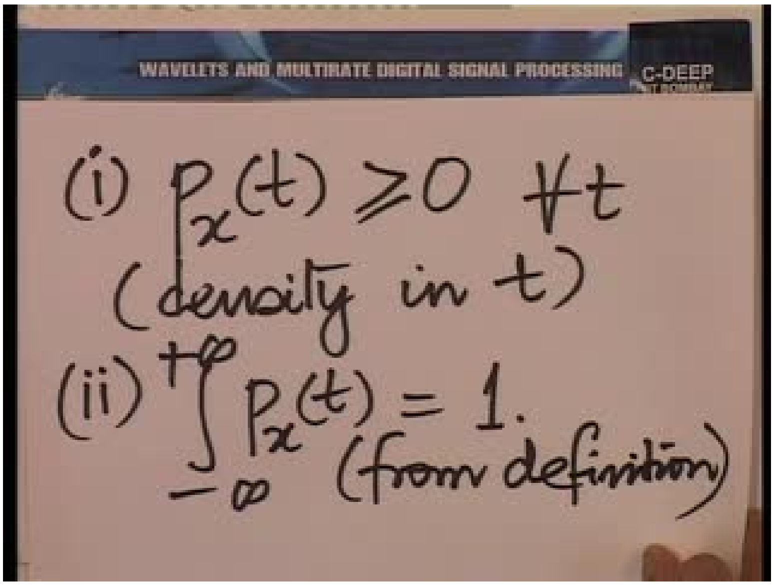


Alright, now because the function belongs to L 2 R, we are sure that its Fourier transform also belongs to L 2 R. so let us x t have the Fourier transform x cap - capital omega and we know that capital x, I mean x cap of capital omega belongs to L 2 R as well. So, we first define a density or a 1 dimensional mass if you would like to call it that. We know that both x t and x cap omega are square integrable and therefore, if we take the magnitude squared of x t and the magnitude squared of x cap omega, they would enclose a finite area under. In fact, the 2 areas would be essentially the same but for the factor of 2 pi. Again, if we chose to do away with angular frequency and used hertz frequency, that 2 pi factor would also go away. Anyway, what we are saying is, mod x t squared integrated from minus to plus infinity is finite let us in fact use the standard notation for this the norm of x in L 2 R the whole squared and therefore define a density P x (Refer Slide Time: 11:10) given by mod x t squared divided by the norm again squared.

(Refer Slide Time: 11:23)

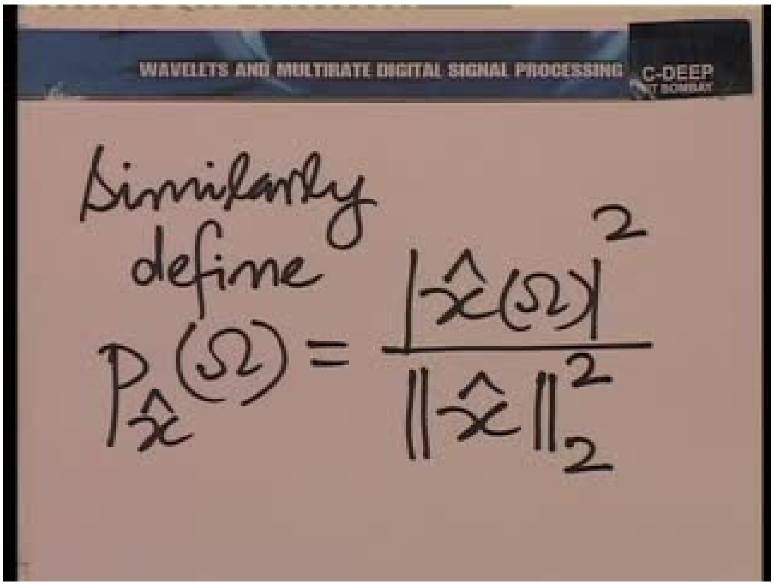


(Refer Slide Time: 11:59)



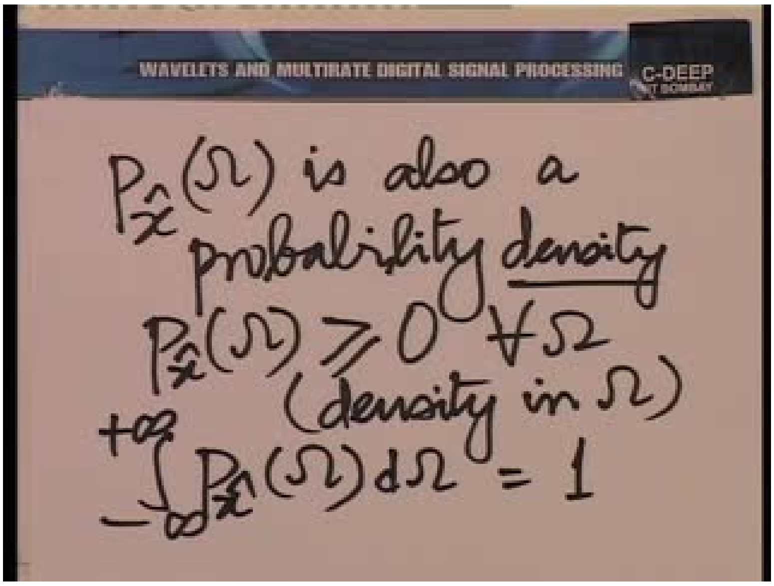
Now, a few remarks and in fact, we should write them down one by one. P x t as we defined it namely, mod x t squared by the norm in L 2 R of x the whole squared is a probability density. Why do we say this? Well, because of the following reasons: let us list them one by one. Number 1: P x t is greater than equal to 0 for all t it is a density in t. Of course, see you may think of t as a random variable and this is the density on that the integral over all t of P x t is easily seen to be 1 from the definition. Essentially, the integral of P x t over all t would in the numerator again have the L 2 norm of the function x and of course, the denominator is indeed the L 2 norm of the function x both squared the numerator and denominator and therefore, they would cancel out to give 1.

(Refer Slide Time: 13:11)

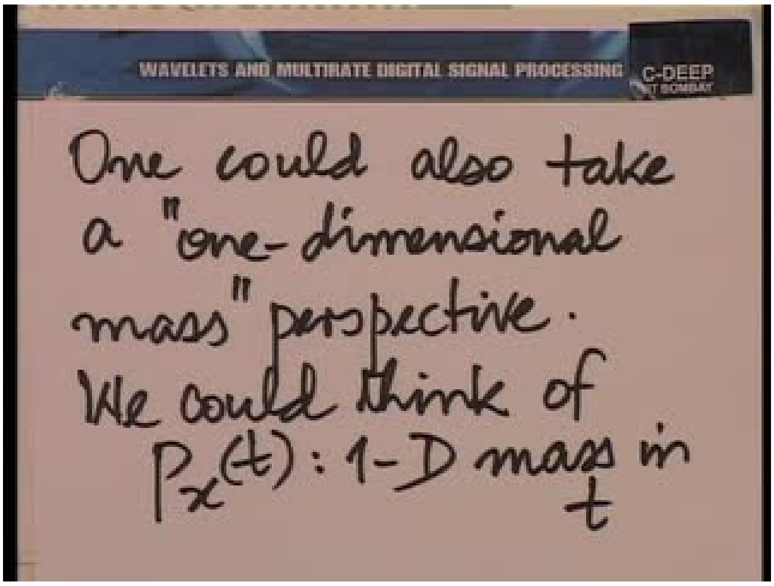


Similarly, let us define a density in the Fourier domain in the angular frequency domain. There we shall write P x cap as a function of omega to be mod x cap omega squared divided by the norm of x cap squared. Here again, we are assured of the denominator being finite because of the L 2 R business.

(Refer Slide Time: 14:08)

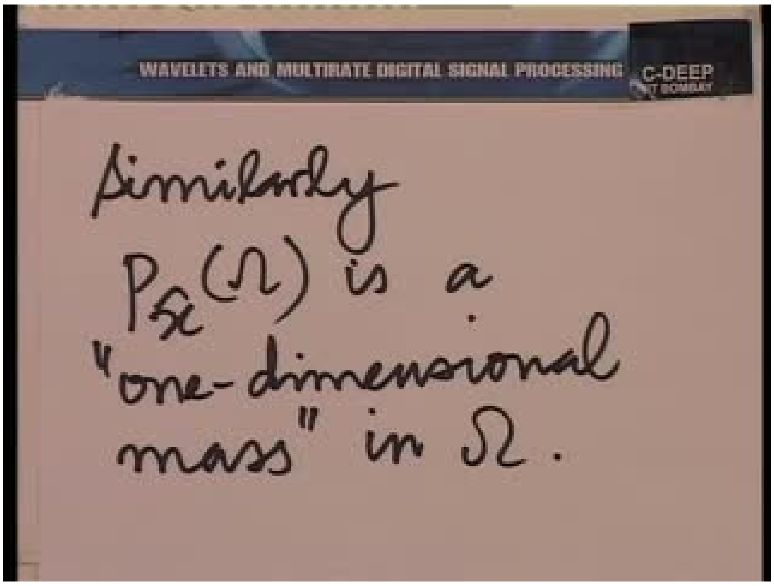


(Refer Slide Time: 15:27)



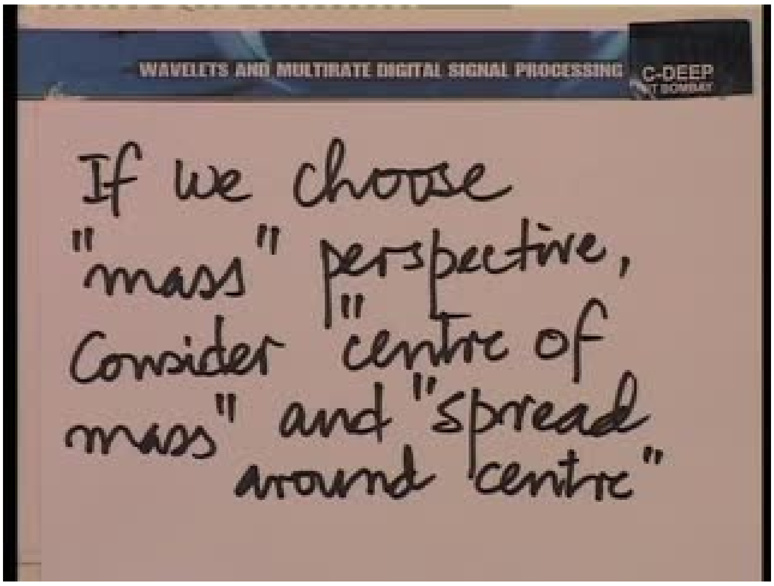
So, again we shall for completeness and formalism, note that this is the probability density. P x cap omega is also a probability density. Indeed, P x cap omega is greater than equal to 0 for all omega; it is the density in omega. The integral over all omega from minus to plus infinity of P x cap is 1. That is also easy to see by very definition.

(Refer Slide Time: 16:47)

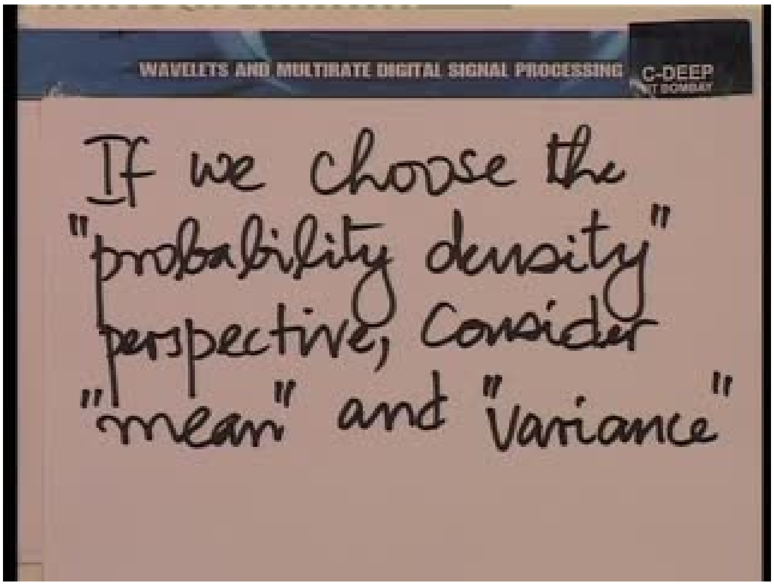


Now, we have taken the probability density perspective but, we could as well take the so called 1 dimensional mass perspective and let me also note that. One could also take a 1 dimensional mass perspective. That is, we could think of P x t as a 1-D mass in t and similarly, you could think of P x cap omega as a 1 dimensional mass in omega. So, what I am saying is, all of us or all the objects around us are masses in 3 dimensional space so here we have a simplified situation you have a mass in 1 dimensional space; that 1 dimensional space can be the space of t or the space of capital omega. Similarly, as I said p x cap omega is a one dimensional mass.

(Refer Slide Time: 17:37)

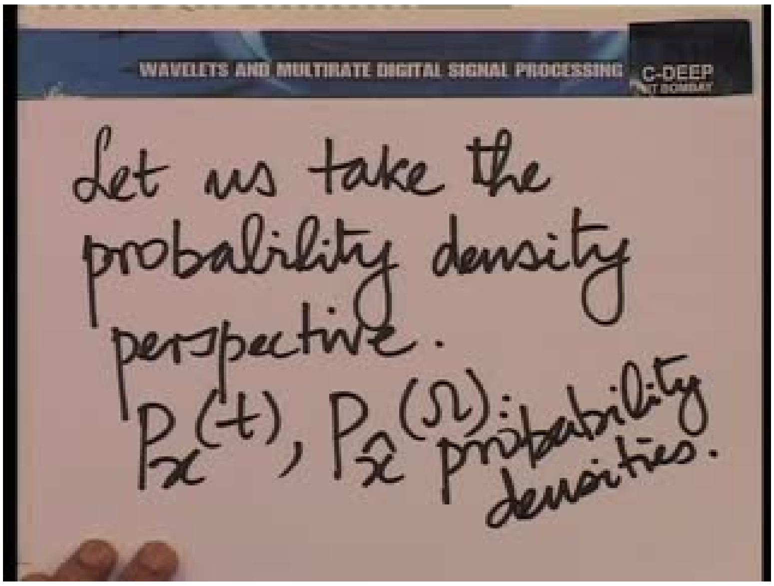


(Refer Slide Time: 18:36)

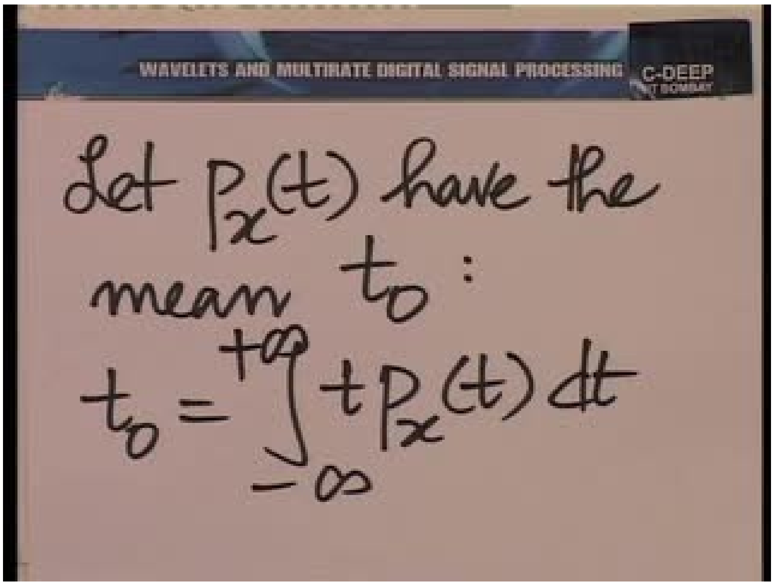


Now, once you take the mass perspective then immediately, you have the notion of centre of mass, centre of gravity if you like to call it that and when you take a probability density perspective you have the notion of mean. Either of them is they are equivalent; so let us make a note of that. If we choose the mass perspective, consider the centre of mass and spread around the center. Now, incidentally as I said the spread around the centre in mechanics is often measured by quantity called the radius of gyration. If we happen to take the density perspective, consider the mean and the variance. Now, we must assume that these quantities can be calculated and we shall do that. It is possible that the variance be infinity; that is a subtle point. So, we are not always guaranteed of finite variance and in fact, that is not a contradiction to what we have been saying so far. We are trying to find a lower limit to where the quantities go in the 2 domains simultaneously. So, if the variance happens to be infinite which it will actually in some situations, we shall simply say that is the worst possible case that we can encounter.

(Refer Slide Time: 20:24)

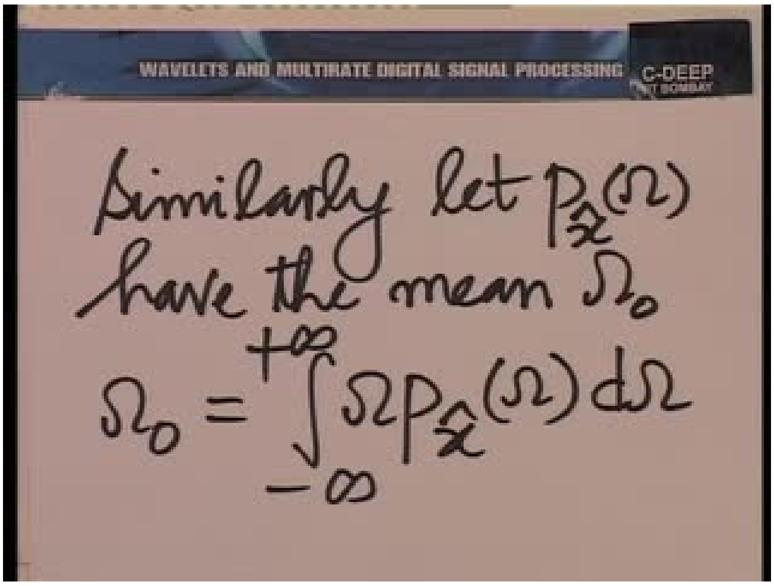


(Refer Slide Time: 21:11)

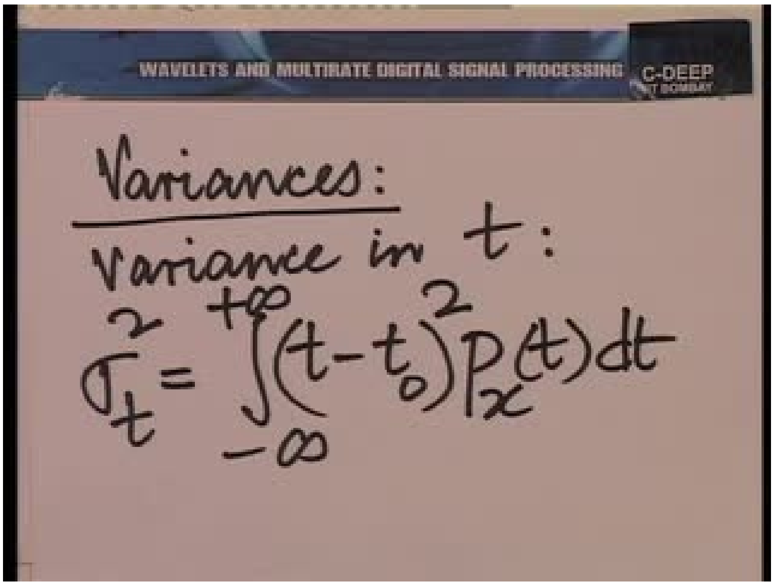


Anyway, let us consider given this function x t we will prefer to take the probability density perspective. So, we will think of P x of t and P x cap of omega as probability densities and we will then write down the mean. So, indeed let P x t have the mean t 0, what would that mean? t 0 would then be integral t times P x t over all t from minus to plus infinity. Simple, the definition of mean; you recognize the same definition, to hold good for the centre of mass here. Essentially, you are calculating the moment by choosing the fulcrum to be 0 and therefore, getting a different fulcrum or a point at which the moments are all balanced.

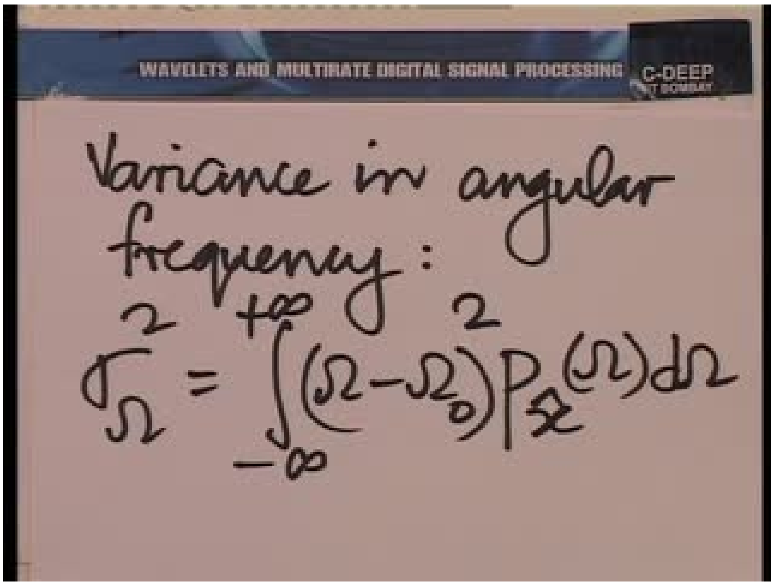
(Refer Slide Time: 22:10)



(Refer Slide Time: 23:12)

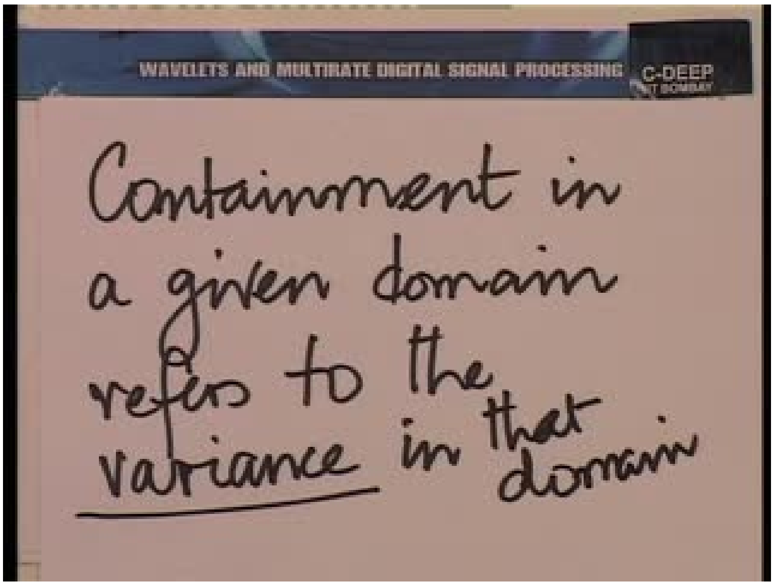


(Refer Slide Time: 24:03)



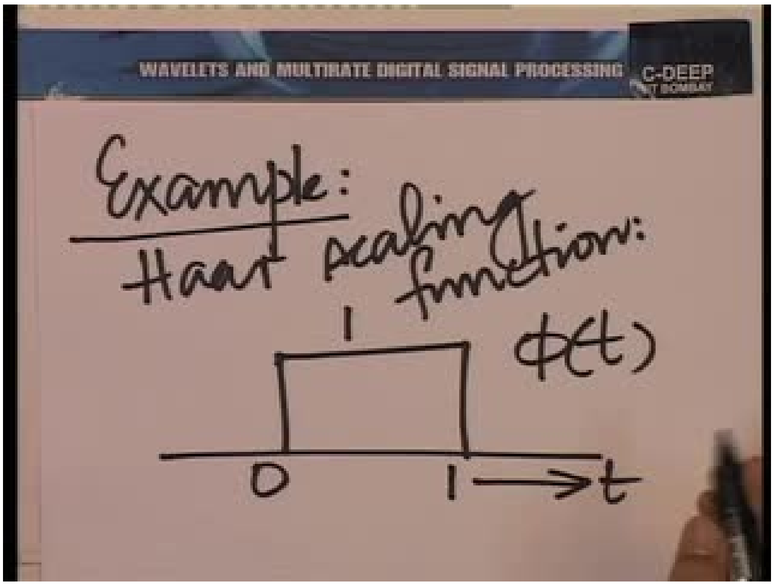
Similarly, let P x cap omega have the mean omega naught, whereupon omega naught would be integral from minus to plus infinity omega P x cap omega d omega; again, the centre of mass, if you like to look it at that way, in the frequency domain. Now, once we have the mean and just as a tongue-in-cheek statement, we assume the means are finite. Normally, they should be in some pathological situations we may have a problem we are not looking at those pathological situations. So, assuming these means are finite, let us look at the variance, so the variance in t would then be given by and we will define it to be sigma t squared by definition; this should be t minus t 0 the whole squared times P x t integrated over all t. Similarly we could talk about the variance in frequency. So, variance in angular frequency; sigma omega squared is integral from minus to plus infinity, omega minus omega naught the whole squared P x cap omega d omega. Once again, tongue-in-cheek, we are assuming these variances to be finite. In any case, here we do not have such a problem even if the variances are infinite we will accept it. We will say that it is the extreme. In the worst case, whatever they be, finite or infinite, we accept.

(Refer Slide Time: 26:16)

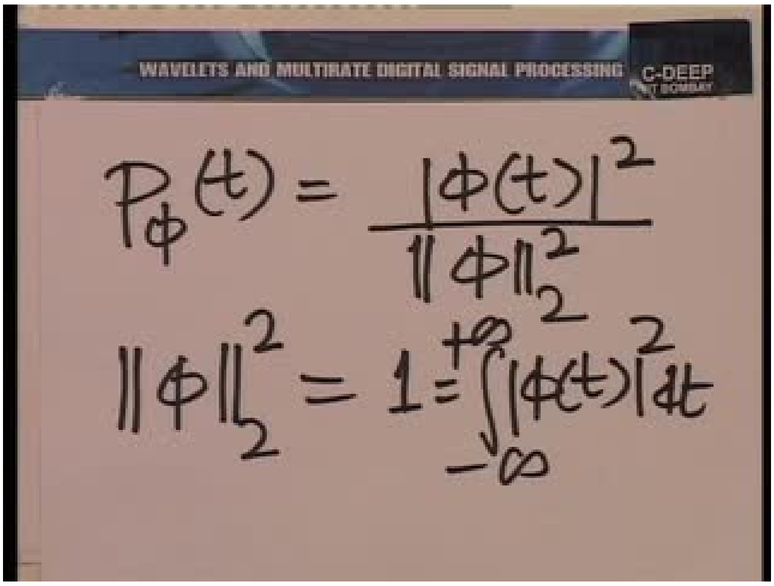


Now, it is very clear; you see, if you look at a probability density or perhaps, if you to choose to think of this as one dimensional masses, it is very clear that the variance is an indication of the spread. So, the larger the variance the more the density is set to have spread around the mean. The smaller the variance, the more the density on the mass is said to be concentrated. Now, we have a formal way to define containment. In fact, we shall now make a very simple definition. We will say containment in that particular domain refers to the variance or if you like the square root of the variance, positive square root. So, let us put down this statement formally. We will say containment in a given domain refers to the variance in that domain. So, containment in time is essentially the sigma t squared quantity and containment in the angular frequency is essentially the sigma omega squared quantity. Now, we ask ourselves how small can we make any one of these quantities for a function, for a valid function. In a few minutes we will be convinced there is really there is no limit. In fact, let us take the haar scaling function as an example; let us calculate the variance.

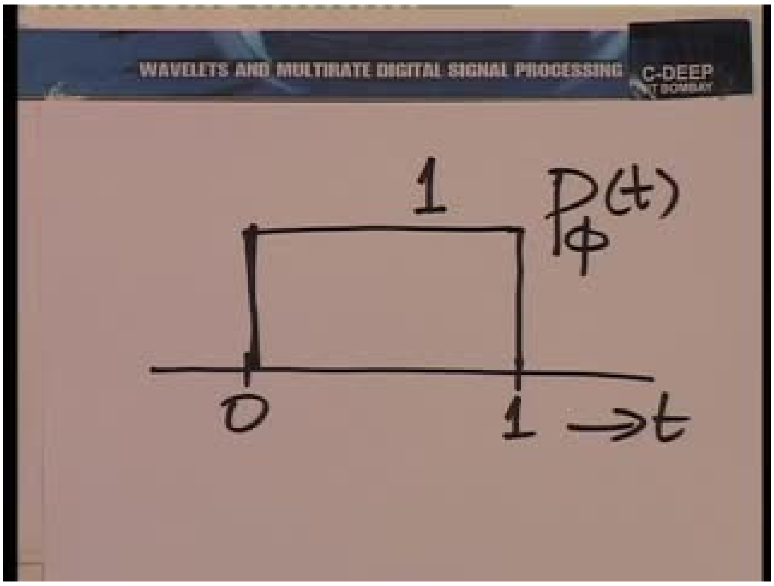
(Refer Slide Time: 27:26)



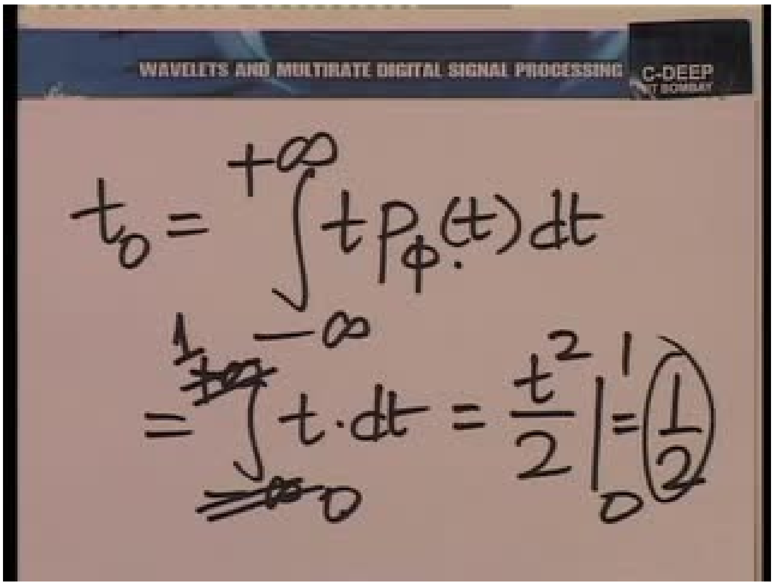
(Refer Slide Time: 28:10)



(Refer Slide Time: 28:50)

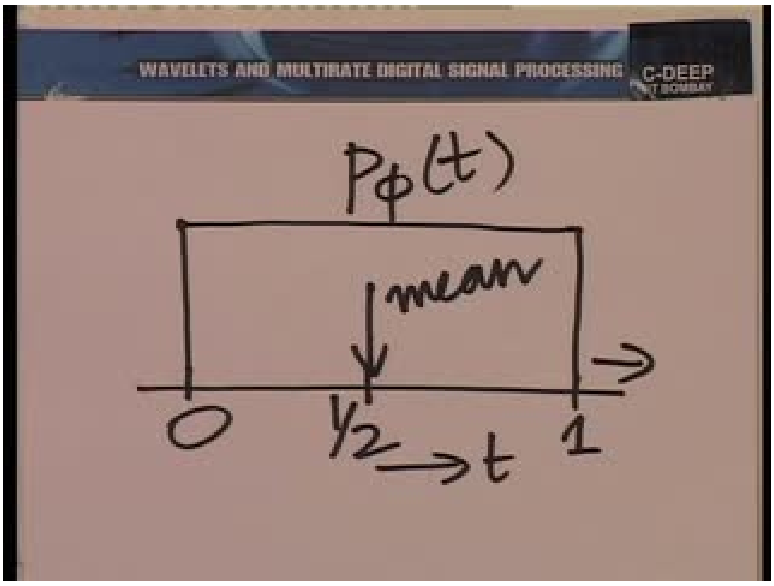


(Refer Slide Time: 29:30)

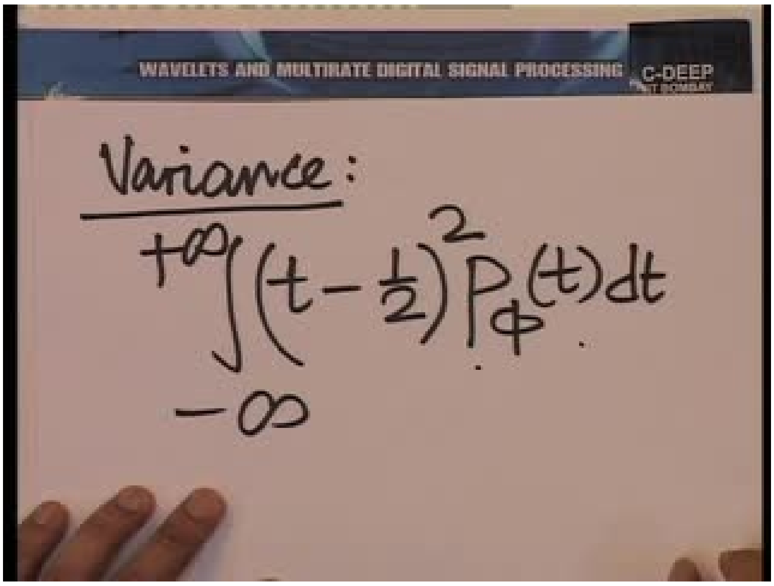


You see the haar scaling function phi of t is 1 between 0 and 1 and 0 else and then of course, it is very easy to write down the density here. It is very easy to see that the squared norm in L 2 P of phi is 1. It is essentially the integral phi t mod squared dt over all t; easily seem to be 1 and therefore, very luckily P phi t looks very much like phi t. This is how P phi t looks. Our job is easy; let us find the mean. In fact, even before I formally set out to find the mean I can estimate the mean graphically. The mean is going to be in the centre at half; that is obvious. But, let us do it formally. So, t 0 would be integral t P phi t dt over all t and this essentially mounts to integral t dt from 0 to 1 so I have replaced P phi t by 1 and I have replaced the limits by 0 to 1. This is obviously t squared by 2 evaluated from 0 to 1 which indeed is nothing but half, as we expected; so the mean is indeed half.

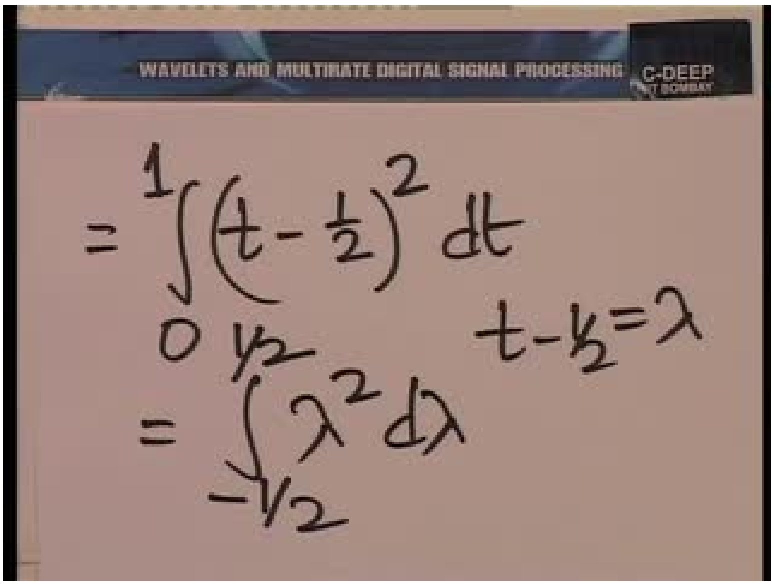
(Refer Slide Time: 30:17)



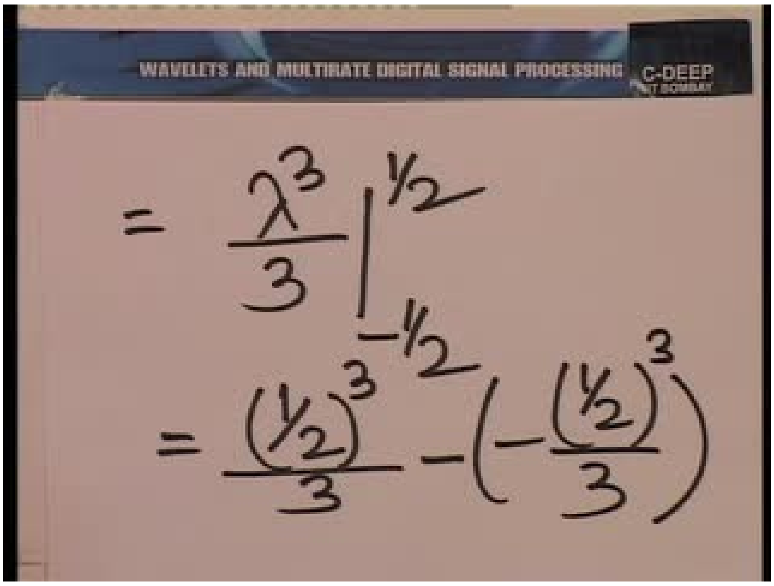
(Refer Slide Time: 30:50)



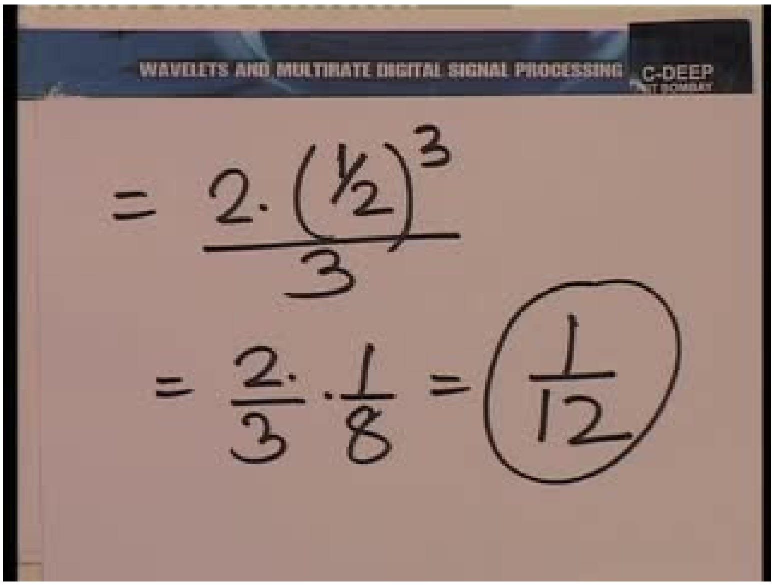
(Refer Slide Time: 31:27)



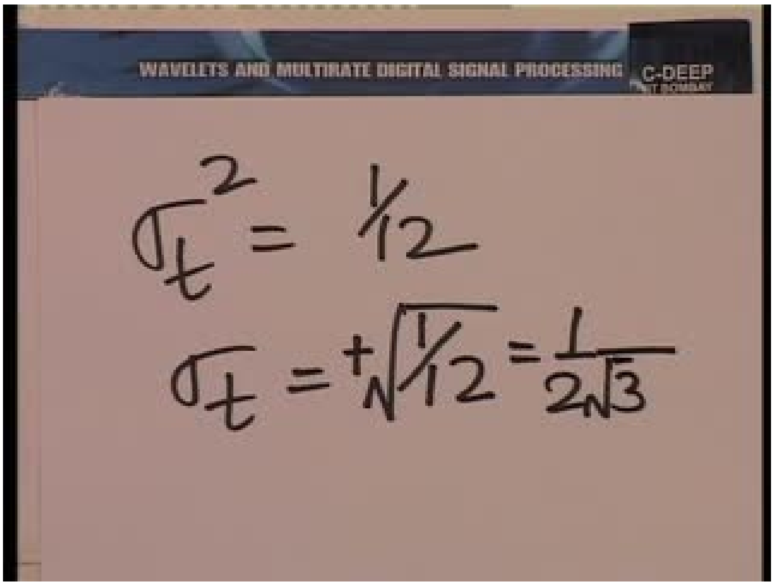
(Refer Slide Time: 32:10)



(Refer Slide Time: 32:40)

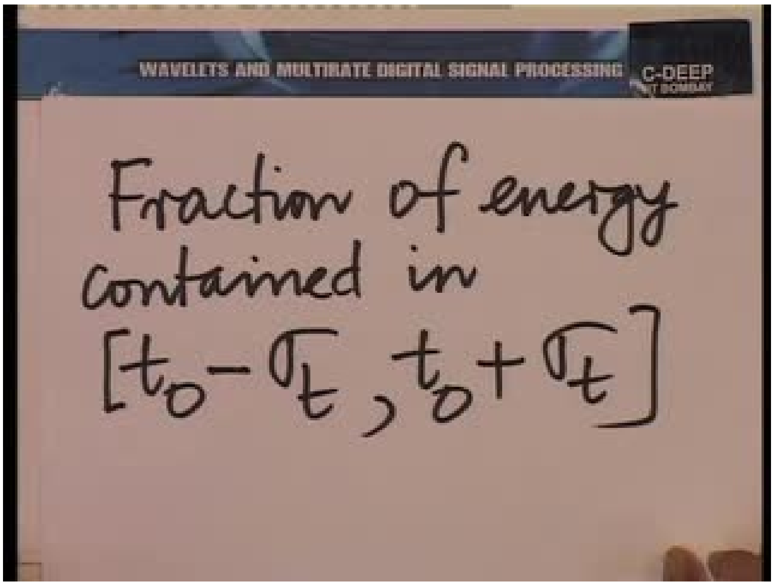


(Refer Slide Time: 33:05)

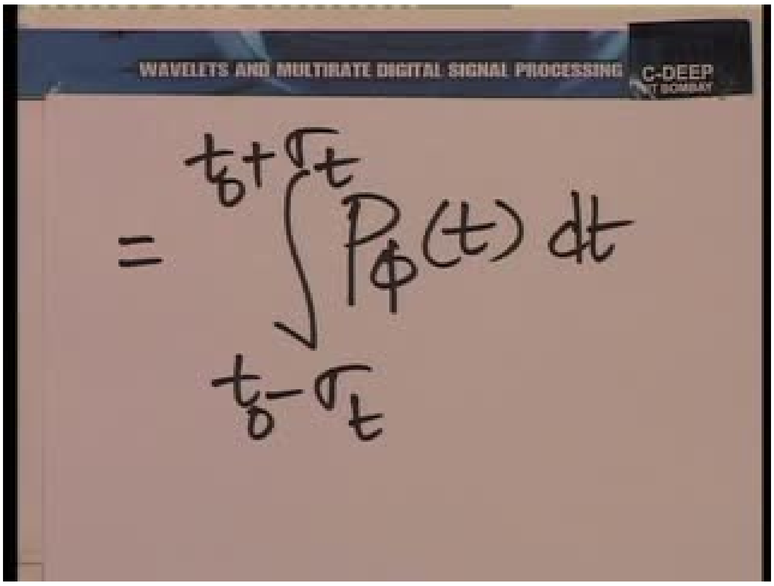


Now, we need to calculate the variance that is little more work but, not too much variance. Indeed, the variance would be given as t minus half the whole squared times P phi t dt over all t. Once again, noting that P phi t is 1 only between 0 and 1 and 0 else, we can rewrite this to get t minus half the whole squared dt and if I care just to replace t minus half by another variable lambda, I would get this to be well when t is 0 lambda would take the value minus half; when t is 1 lambda would take the value half. This would be lambda squared t lambda to be integrated and then we have an easy expression. So, let us half cubed by 3 minus of minus half cubed by 3 and therefore, you have 2 times half cubed by 3 which is 2 by 3 into 1 by 8 or 1 by 12. So, this is the variance now sigma t squared is 1 by 12 and therefore, you may take sigma t to be square root the positive square root of 1 by 12 which is 1 by 2 square root 3; as you can see it sigma t is less than half.

(Refer Slide Time: 34:15)

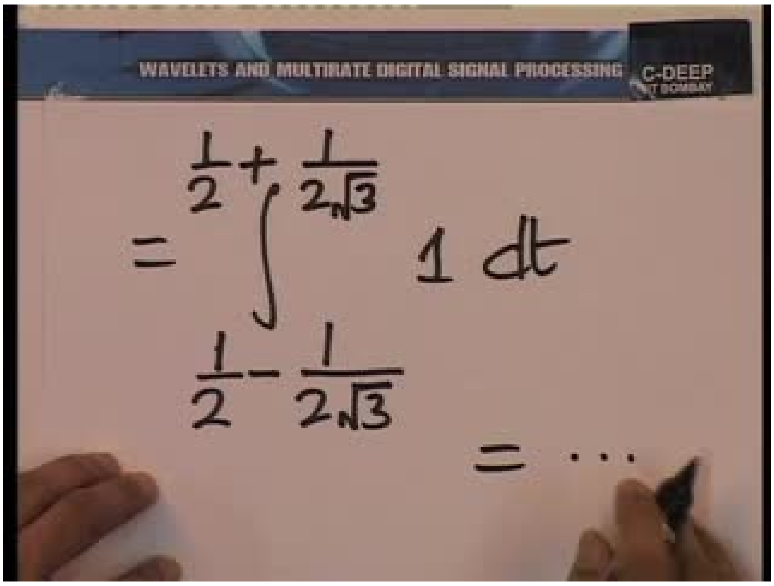


(Refer Slide Time: 34:51)

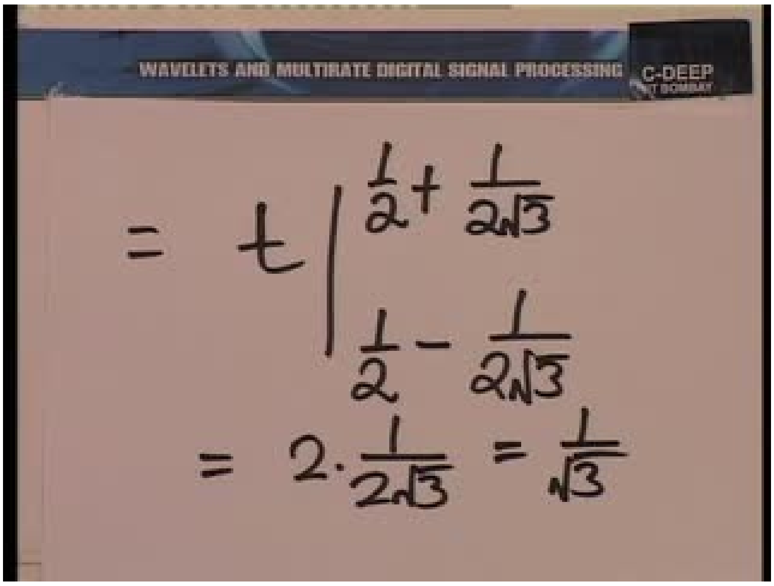


So, in a certain sense we do not really use the number half to denote the spread of phi t around its mean. The variance does not say it whole all the way to half; it says the spread is a number slightly less than half. Most of the energy is contained in that region around the mean captured by the variance. In fact, if you wish to be very specific the fraction of the energy contained here would be, it is essentially, would be the integral of the density between t 0 minus sigma t to t 0 plus sigma t. So, that would essentially be now… I do not really intend to calculate this quantity for this case; it is a very simple calculation of course integrated with respect to t but, what I am trying to emphasize is, we are asking for 100 percent.

(Refer Slide Time: 35:43)

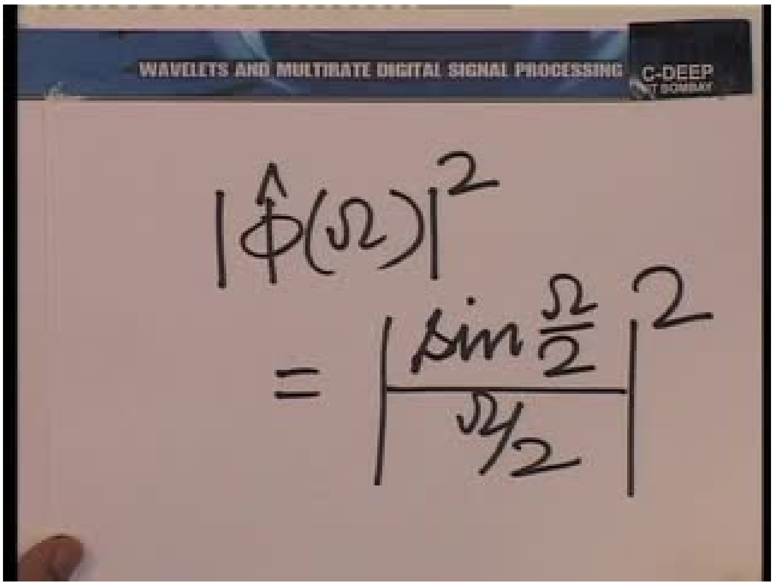


(Refer Slide Time: 36:04)

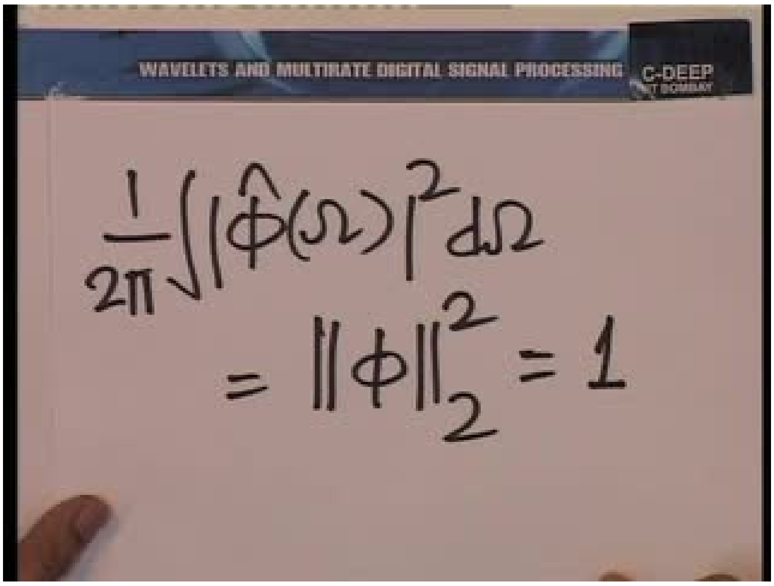


We are saying, tell me the region over which 100 percent of the energy lies that region could very well be the whole real axis. We are saying well, at least a significant part of it. Now, in this particular case may be it is a good idea to actually calculate it, how much is this really. So, it is actually and this is essentially so; it is t evaluated from half minus, so this is easily seen to be, 2 times 1 by 2 under root 3 which is 1 by square root of 3. Now, certainly not a very large fraction like 90 percent; what is about 1 by 1 point 7 more than 50 percent anyway. Incidentally, this fraction is not going to be the same for the different function. It depends on the density but, what we are trying to say is that the variance is one accepted measure of spread and very often the variance actually tells us where most of the function is concentrated. Even in the case of this function, if you look at it carefully, what we are saying is quite a bit of the function is contained between half minus 1 by 2 square root 3 and half plus 1 by 2 square root 3. So, it is not an unreasonable range that we choose.

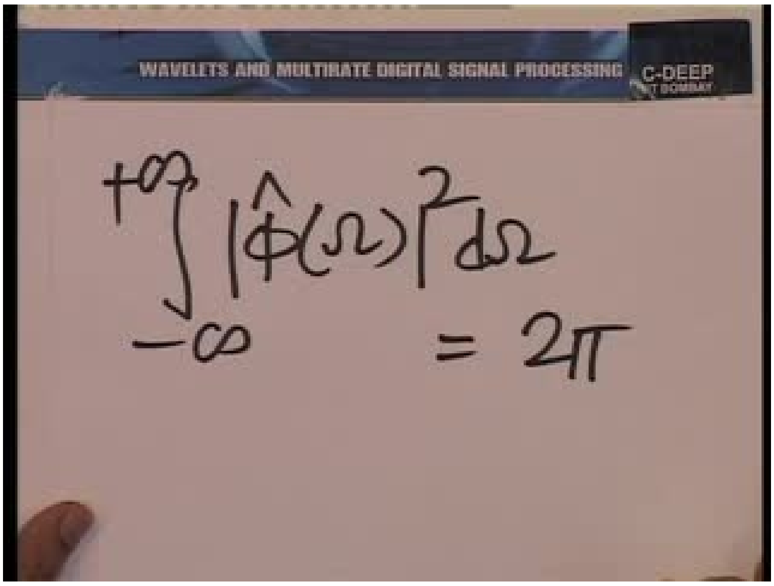
(Refer Slide Time: 38:03)



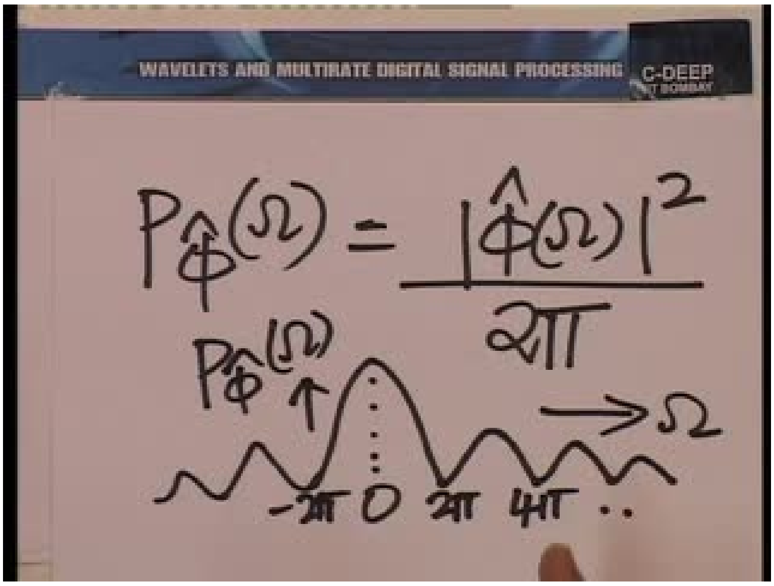
(Refer Slide Time: 38:38)



(Refer Slide Time: 39:01)

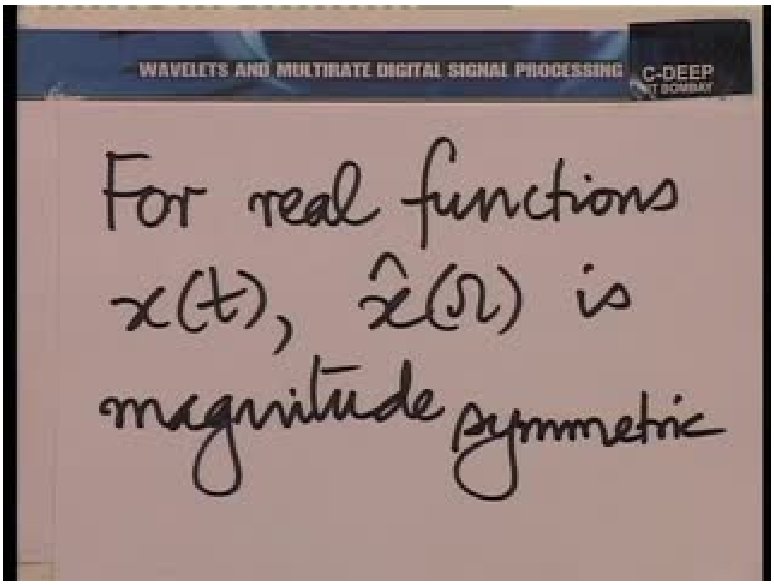


(Refer Slide Time: 39:22)

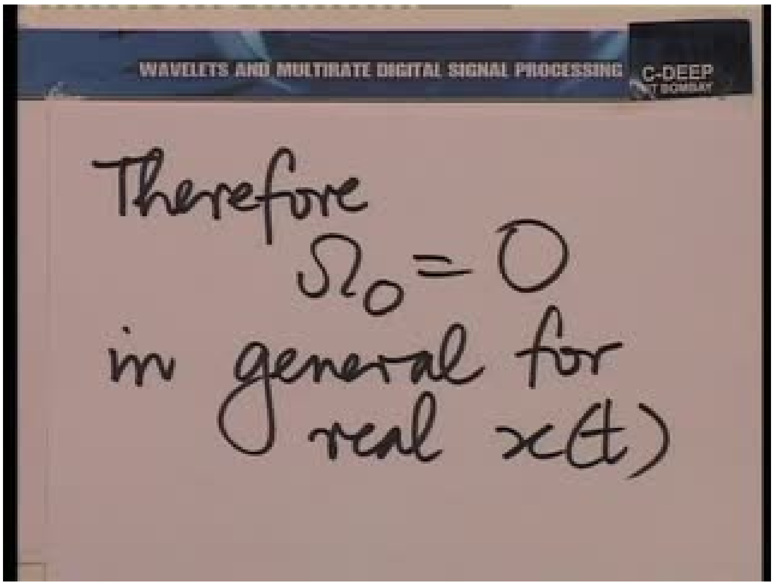


Now, we ask about the variance in frequency of the same function and there we are going to have a very pleasant or unpleasant surprise. So, let us look at phi cap omega. In fact, we are so interested in phi cap omega; we are interested in phi cap omega mod squared and that has the form. Essentially, sin omega by 2 by omega by 2 the whole squared mod and you could integrate this you know, indeed as you know the integral of phi cap omega mod squared d omega divided by 2 pi would essentially, be the norm of phi L 2 R norm of phi the whole squared, which is easily seems to be 1. Therefore, integral phi cap omega the whole squared d omega over all omega is essentially 2 pi. Therefore, we essentially look at the quantity mod phi cap p phi cap rather which is essentially of the form phi cap omega the whole squared divided by 2 pi. Let me sketch this; in fact, we are familiar with it, would have an appearance like this is 0, this is 2 pi, 4 pi and so on. So far, we know this; we have been doing this more than once.

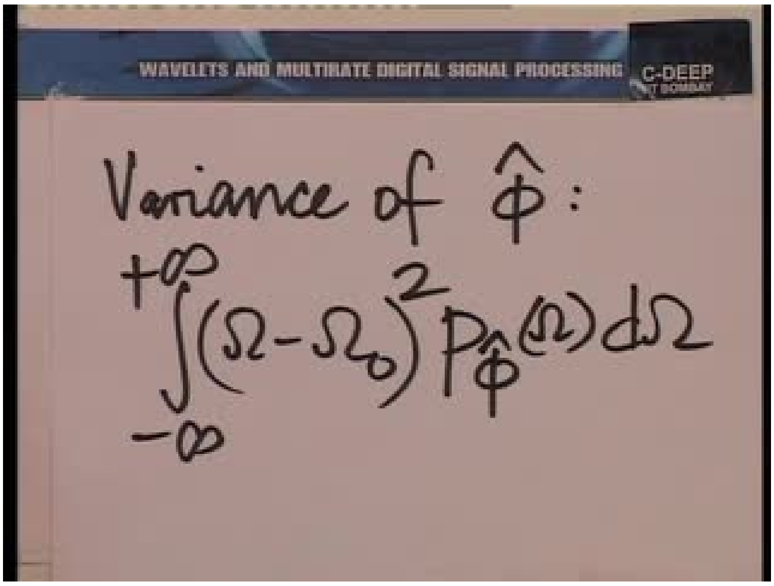
(Refer Slide Time: 41:07)



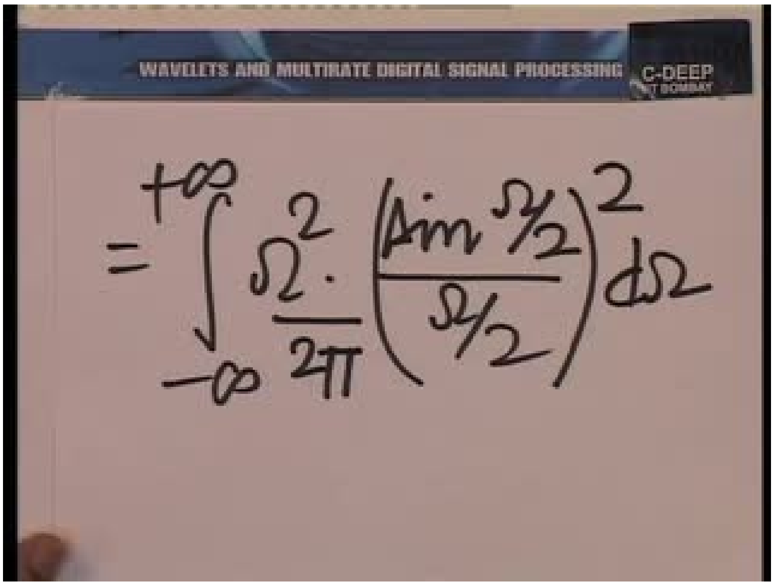
(Refer Slide Time: 41:35)



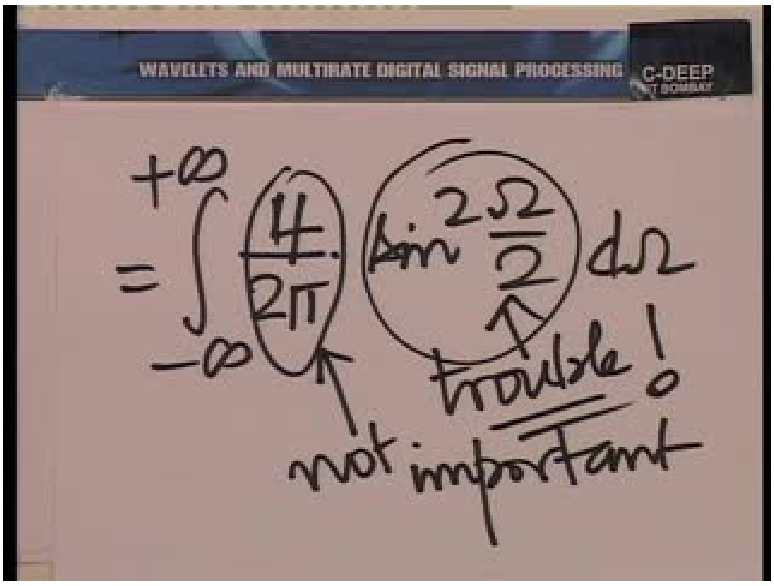
(Refer Slide Time: 42:12)



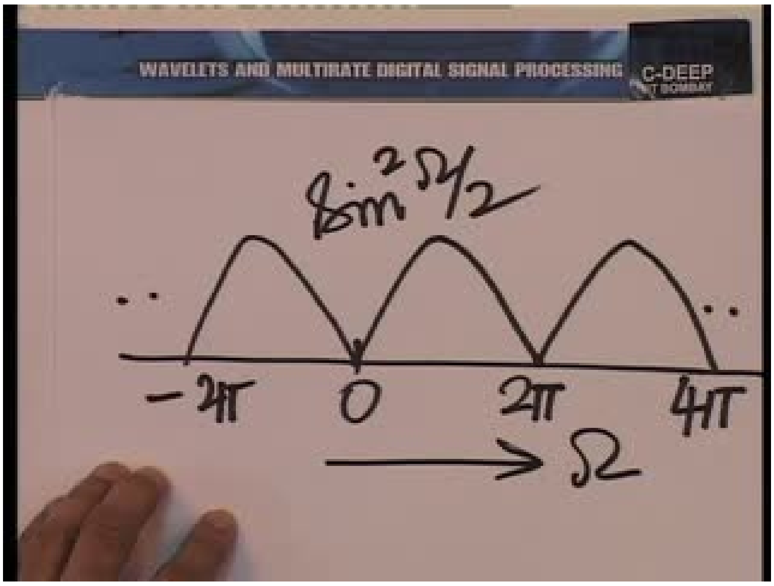
(Refer Slide Time: 42:42)



(Refer Slide Time: 43:09)

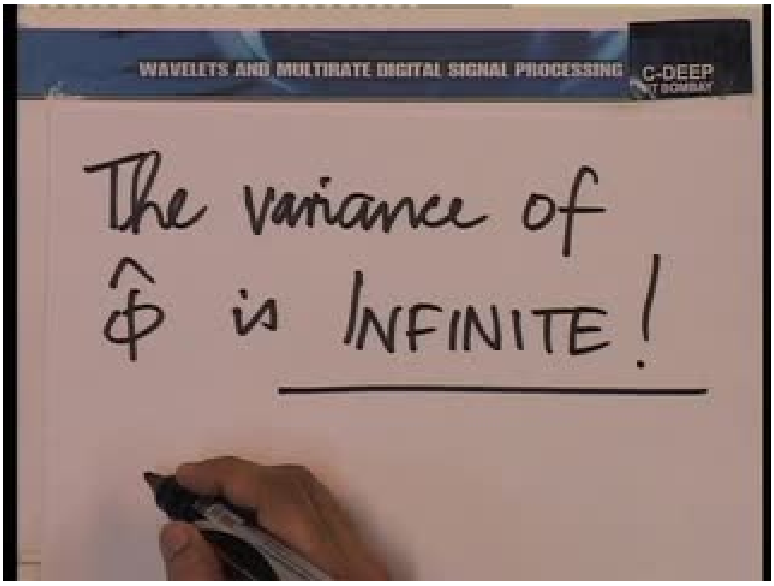


(Refer Slide Time: 43:47)

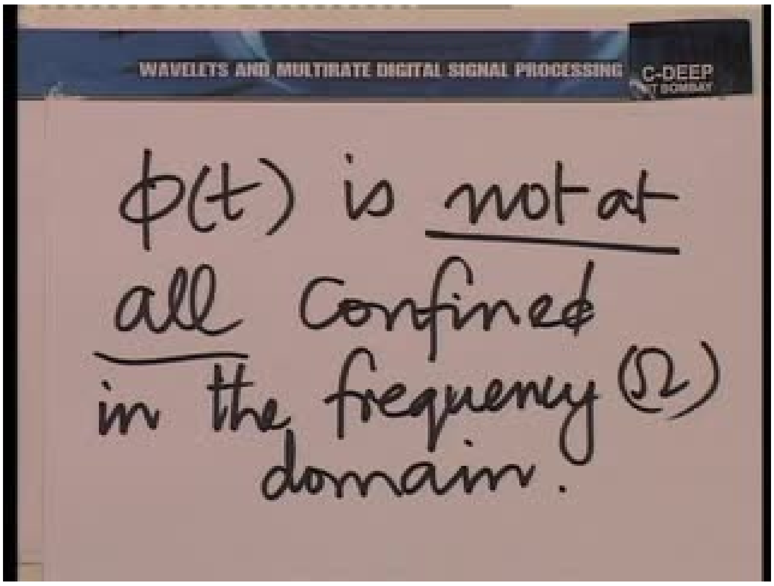


Now, it is very easy to see what the mean of this function is. The function is symmetric around omega equal to 0 and therefore, the mean is 0. By the way, this is not a surprise for many real functions. We would find the mean for all real functions; the mean of the density on the frequency axis is going to be 0. The Fourier transform for real function is magnitude symmetric and therefore, it is not surprising that for a real function the mean as understood in this sense, is always going to be 0 on the frequency axis. Let us make a note of that; it is a very important conclusion. For real functions, x cap omega is magnitude symmetric; therefore, the mean is 0. Now, comes the variance and here we have a very unpleasant surprise waiting for us. I say unpleasant because, maybe we should have something better. So, the variance of phi cap would be calculated as follows: integral over all omega minus omega naught the whole squared P phi cap omega d omega and if we make the required substitutions we have this is essentially omega squared times sin omega by 2 divided by omega by 2 the whole squared divided by 2 pi here d omega and here we are in serious trouble this is integral minus to plus infinity 1 by 2 pi sin squared omega by 2 times 4 so 4 goes up there d omega. This is not important, just a constant; but, this is trouble, we are in serious trouble here. In fact, let me sketch what we are trying to integrate. This function, a periodic function with a period of 2 pi sin squared omega by 2 serious trouble as I said.

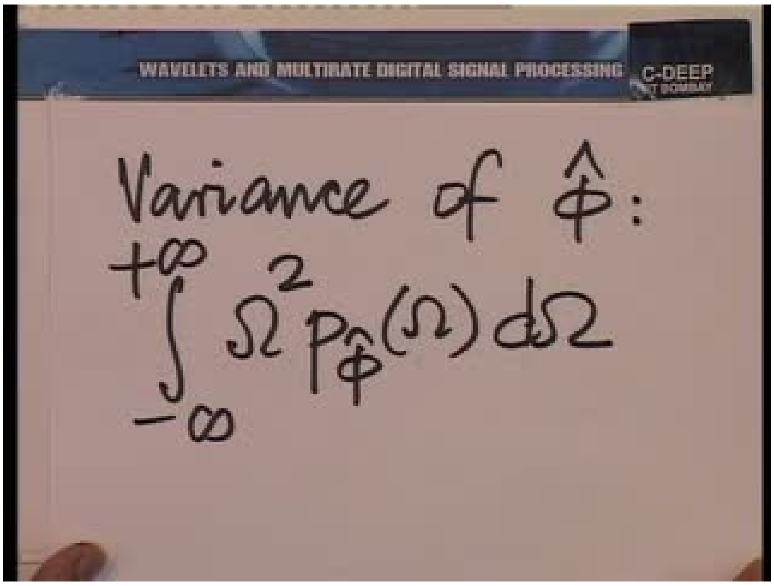
(Refer Slide Time: 44:45)



(Refer Slide Time: 45:06)

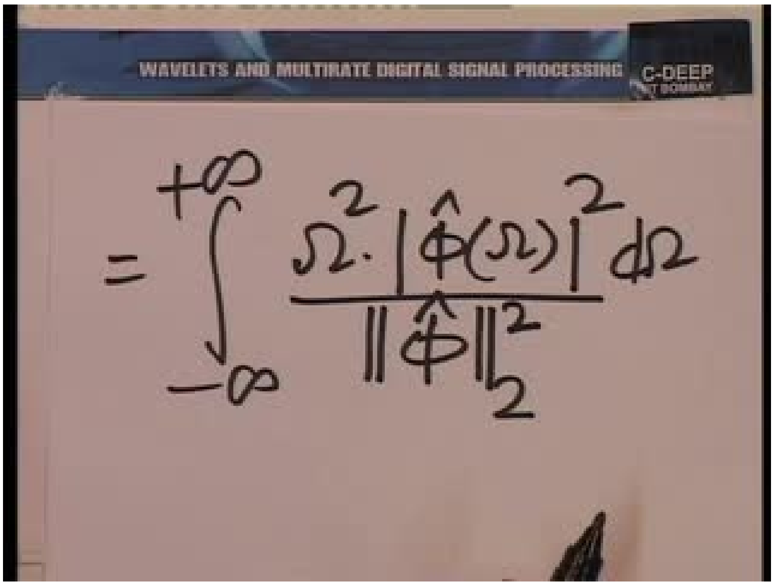


(Refer Slide Time: 47:35)

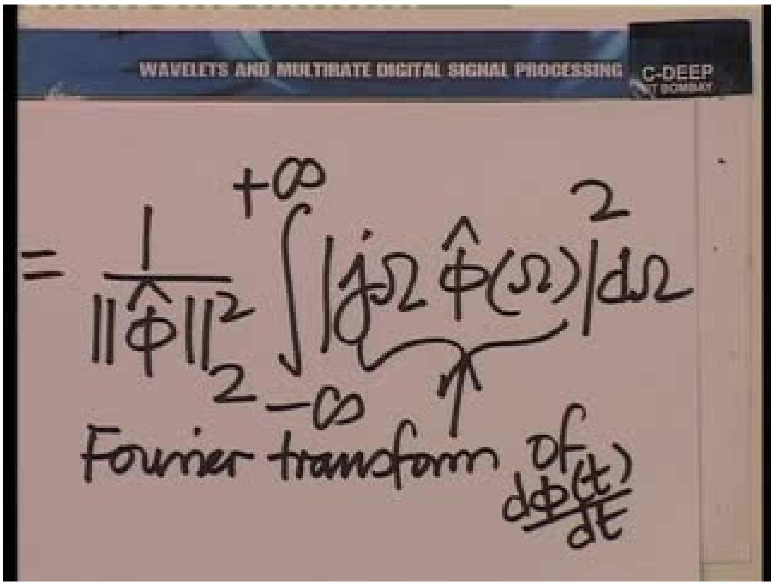


We are trying to integrate a periodic function from minus to plus infinity and obviously, that integrate is going to diverge. So, the fear that we had when we started with discussion with variance comes out to be true, right in the very simplest case of a scaling function that we know. The variance of phi cap is infinite; in other words, phi t is not at all confined in the frequency domain, at least in this sense. Now, all this… why in our discussion when we talked about time and frequency together and so on? The previous lecture we had been worried about the side lobes, as we call them. We said well, it is alright to look at the main lobe and talk about presence in the main lobe. But, then we have the side lobes and the side lobes are falling off only by the factor 1 by omega in magnitude. As you can see, the side lobes have created a problem after multiplication by omega squared in the calculation of variance. The side lobes create a periodic function to be integrated; a periodic non negative function and we are in trouble. So, this tells us again why we have to go much beyond the haar. We have been asking again and again, why we cannot be content with the haar multi resolution analysis. Now, we have one more formal answer. If we look at the scaling function for the haar multi resolution analysis, it is variance. In the frequency domain is infinite it is not at all confined in the frequency domain in the sense.

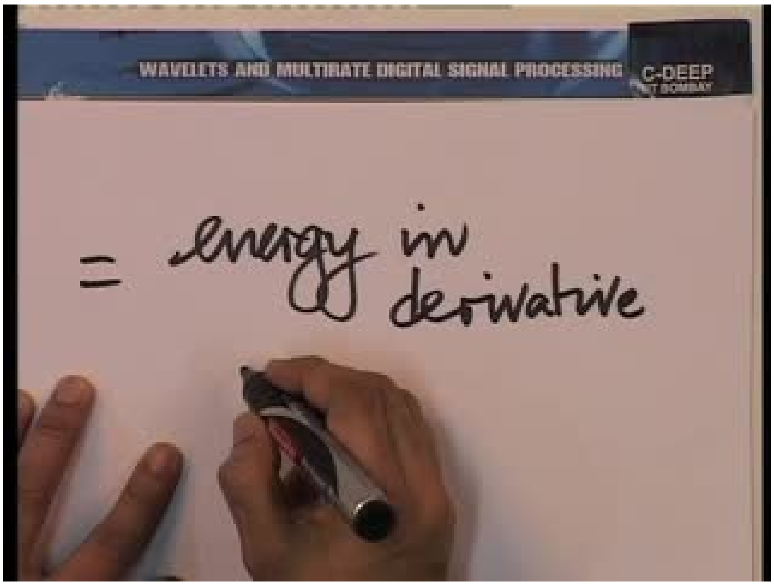
(Refer Slide Time: 48:03)



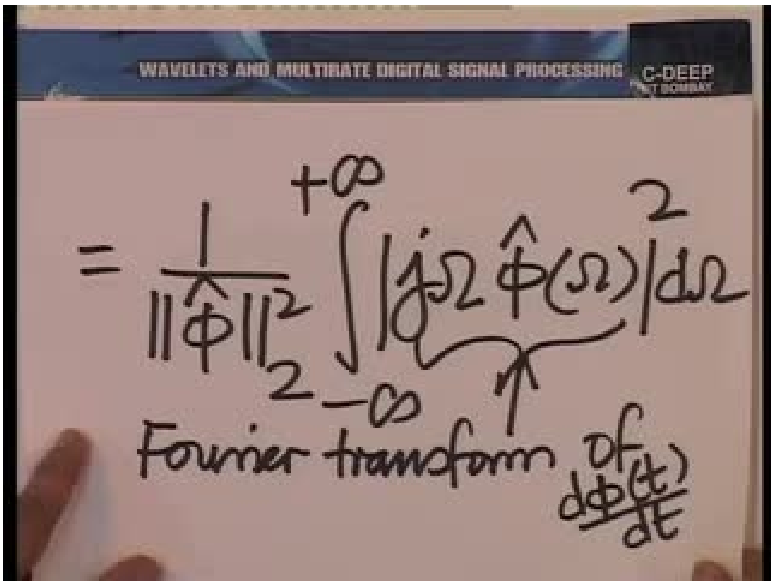
(Refer Slide Time: 48:37)



(Refer Slide Time: 49:53)

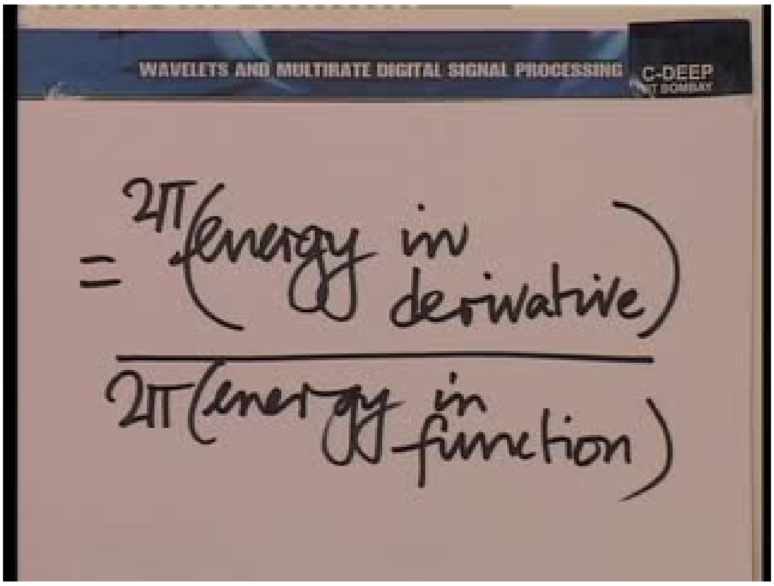


(Refer Slide Time: 50:11)

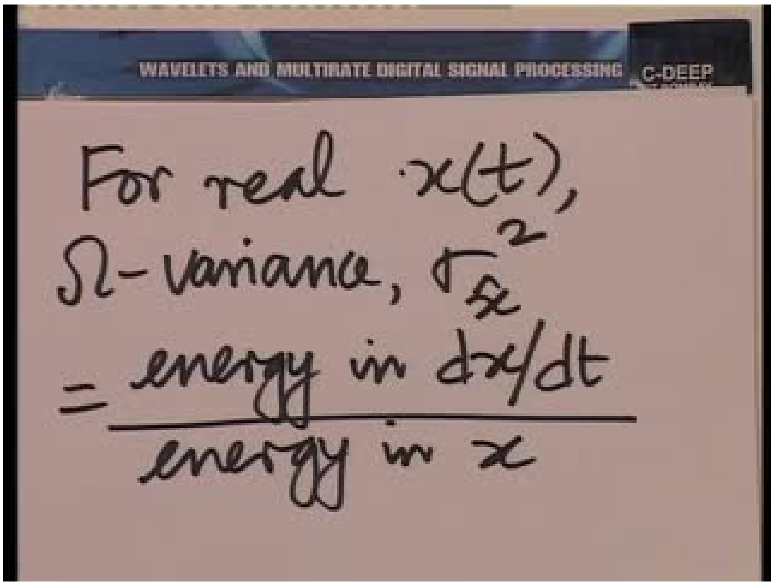


Now, it is a natural question to ask what is it that made this variance infinite? Why did we have a divergent variance here? In fact, we can answer that question if you only care to make a slight adjustment of the expression of variance. The variance of phi cap is finally, as you can see, given by omega squared P phi cap omega d omega and this can be written as omega squared times phi cap omega mod squared d omega divided by the norm of phi cap in L 2 R the whole squared. Now, this norm is a number; it can be brought out of the integral. So, I can rewrite this as 1 by norm of phi cap squared integral and I also do a little bit of rearrangement in the integral. I will write the integral as j omega phi cap omega the mod whole squared d omega. Notice, that if I take the modulus of j omega phi cap omega, it is essentially modulus omega squared times modulus phi cap omega squared. Modulus omega squared and omega squared are the same for the omega real, but then when we write it like this has a meaning. It is essentially the Fourier transform of d phi t dt, Fourier transform of the derivative of phi. So, essentially what we are saying is this variance is actually the energy in the derivative.

(Refer Slide Time: 50:24)



(Refer Slide Time: 50:57)



Remember, you would have a factor of 2 pi there because, this is the energy in the derivative. But, for a factor of 2 pi… So, this would be 2 pi times the energy in the derivative divided by 2 pi times the energy in the function. Please note that this inference that we have made is independent of what function we consider as long as the function is real. The variance in frequency is going to be this ratio: the energy in the derivative divided by the energy in the function. Let us make that remark for real functions, for real x t; the frequency variance omega variance or sigma x cap squared is essentially energy in dx t dt divided by the energy in x or the L 2 norm squared of the derivative of x divided by the L 2 norm of x. Now, we have the answer: why we ran into a problem for phi t? As you can see phi t is discontinuous so, when its derivative is considered, there are impulses in the derivative and impulse is not square integrable. Therefore, the numerator of this quantity diverges off. You look at it from that perspective; the moment we have a discontinuous function, we have an infinite frequency variance and there we are with this note, then we realize that if we want to get some meaningful uncertainty, some meaning bound, we must at least consider continuous function and we shall proceed to build on this concept further in the next lecture.

Thank you