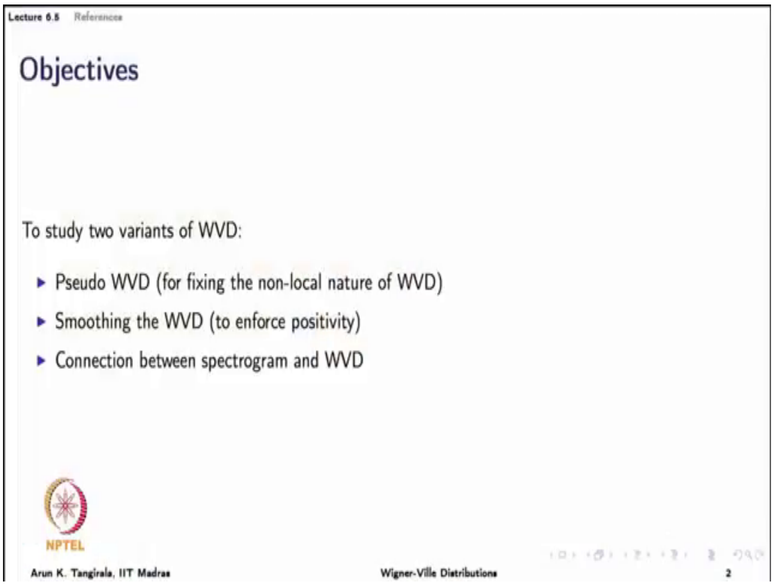
Introduction to Time-Frequency Analysis and Wavelet Transforms
Prof. Arun K. Tangirala
Department of Chemical Engineering
Indian Institute of Technology, Madras

Lecture - 6.5
Pseudo and smoothed –WVD

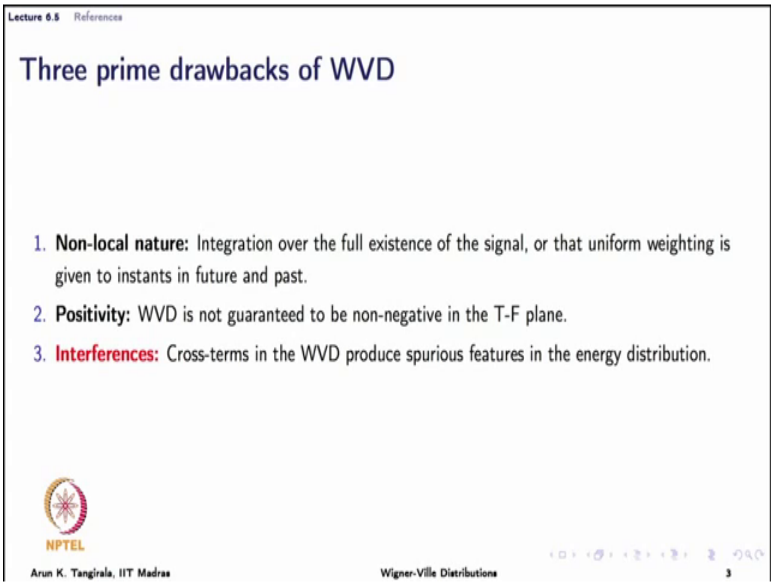
Hello friends, welcome to lecture 6.5 on the topic of Wigner-Ville distributions. In this lecture, we are going to start our journey on smooth Wigner-ville distributions. In particular we look at 2 variants of Wigner-Ville, one is called is pseudo Wigner-Ville and other is called a smooth Wigner-Ville. But we are going to look at an elementary just learn the elementary basics of smoothing. In the next lecture, we are going to take this up in more generalized fashion where we look at Cohen's class of distributions.

(Refer Slide Time: 00:45)



The idea here is to show how smoothing can bring about positivity. And more importantly talk about the connection between spectrogram and Wigner-Ville. We will revisit this connection between spectrogram and Wigner-Ville more formally in the next lecture.

(Refer Slide Time: 01:05)

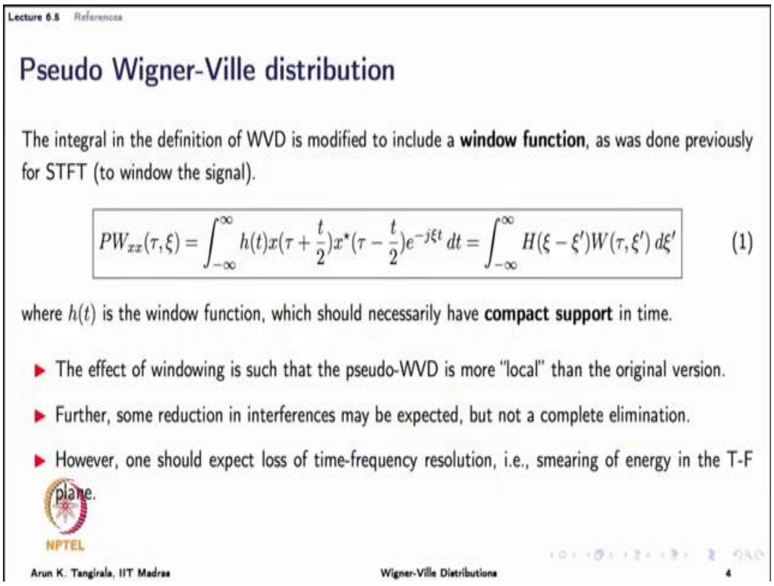


So, to recall we have 3 prime draw backs of Wigner-Ville; one is that it has a non local nature in the sense that it gives equal importance to the past and future of the time location of ((refer time 01:21)). So, if I am standing at a certain time and that is the point in time, and I am looking at the Wigner-Ville of this signal. If you recall the definition of Wigner-Ville it it looks at the entire extent of the signal from minus infinity to infinity that is not a problem. But it gives the same weight age to the local correlation that it computes at all values to the left and right of tau.

But that is not really desirable we would to like get the local properties and therefore, we would like to fix that issue and that is what we are going to do with this pseudo Wigner-Ville. The second and probably I more prominent drawback is the lack of positivity or non-negativity in Wigner-Ville. It may be positive it may not be so, therefore, there is no guarantee that these values of Wigner-Ville distribution will be non-negative in the time entire time frequency plane. And the third which is equally disturbing and prominent is the presence of the interferences whenever the signal has more than 1 frequency.

Well, in this is not applicable to chirps, for chirps Wigner-Ville is the best. But we are talking of multi component signals when I have more than one frequency at a time or 2 frequencies separated in time not necessarily in a chirp fashion then I have interferences so, how do we fix that as well. Of course, what we are going to discuss is only a part of how to address positivity and interferences. But essentially the idea's will be laid down and then will take that up in Cohen's class in a more generalized way.

(Refer Slide Time: 03:05)



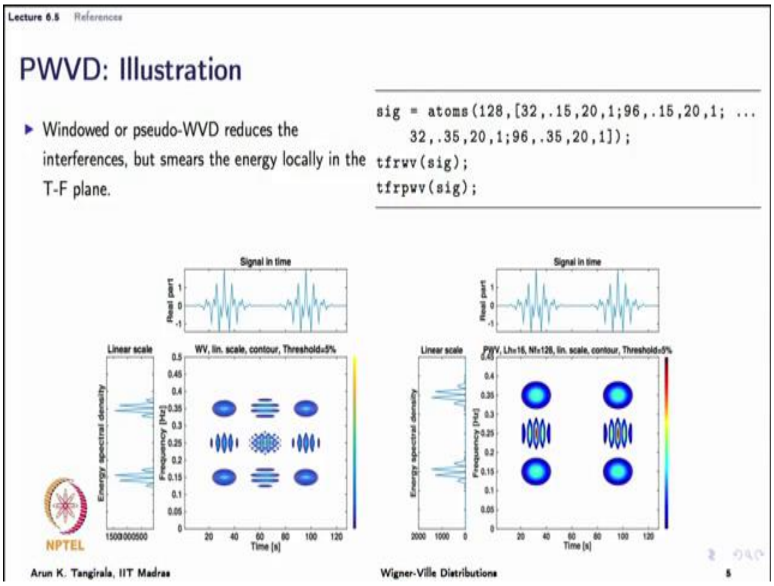
So, let us begin with pseudo Wigner-Ville, we said just now that the one of the key problems is the non local nature. And I would like to give more importance to the values of the signal with in the vicinity of my analysis point in time and give less importance to far of values whether it is in past or future. Well, by with our experience in short time Fourier transform we know that the remedy is to use a window function and that is what this h does for us it brings in a windowing nature. Obviously, this is the one of the prime requirement is that is h should have compacts support in time what do we mean by compact support?

It is a technical word to denote finite duration functions. So, obviously, I want this window to have finite length otherwise what is the point in performing the pseudo Wigner-Ville distribution at all. That is only difference between the Wigner-Ville and a pseudo Wigner-Ville. On the right hand side that is extreme right of this same equation one1I have rewritten this pseudo Wigner-Ville in the frequency domain to show you that the windowing in time is essentially resulting in a convolution of the Wigner-Ville along the frequency access.

So, what is happening is that the Wigner-Ville distribution is computed and then convolved with the window in frequency. And we will shortly learn that this convolution is nothing but smoothing in the context of enforcing positivity and so on. So, at the moment you can think that what is pseudo Wigner-Ville is doing is essentially it is smoothing the Wigner-Ville along one dimension which is a frequency. And we will learn later that we need to smooth in both dimensions if you want positivity that is a nice interpretation to begin with.

Now obviously, because I have windowed the Wigner- Ville the effect of windowing is such that it will bring down the time frequency localization or the nice localization that Wigner-Ville had for all these signals. Just like the way we saw in short time Fourier transforms. But, because we are smoothing in the frequency access there is the possibility that I will reduce the interferences. I may not be able to reduce interferences completely, because the smoothing is only happening along the frequency access.

(Refer Slide Time: 05:32)



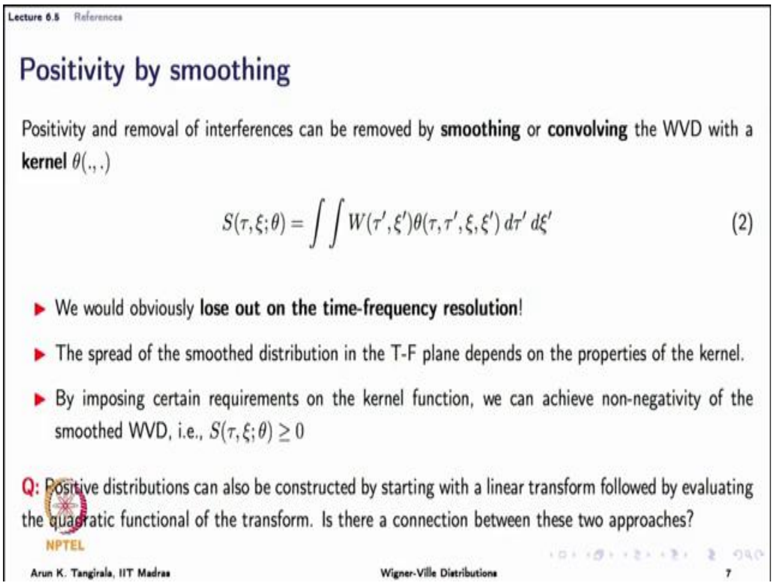
So, let us understand the consequence of using pseudo Wigner-Ville on an example. Again I have taken this example from the time frequency tool box tutorial. I have generated 128 samples of this signal by now; you should familiar with this routine called, atoms. And I have in fact, 4 atoms, but they are all centered at the same location here, two of them are centered at the same location. And I have a pair of frequency is here 0.15 there are pair of atoms with frequency center frequency 0.15. Another pair of atoms with frequency 0.35 that is what you see they are separated in time. When I evaluate the

Wigner-Ville of this signal I have of course, now I am using analytic representation that goes without saying the atoms routine also gives me analytic representation of the signal.

So, I have this 4 terms correctly being identified, which belong to the signal. But then these 5 terms which are the center the geometrical center of this 4 atoms their interferences. Well, one of the main drawbacks also is that it is going their going to be non-negative, but the main concern here is interferences how does the pseudo Wigner-Ville perform? Well, what it has done is it has gotten and rid of these interferences as you can see it has gotten rid of the 3 interfering terms. But is still retains these 2 interferences this is not managed to completely alleviate as we had argued before.

But at what expense did we eliminate these three interfering terms? The expensive is the time frequency resolution. So, we have now, a poorer time frequency localization compared to the original Wigner-Ville. There is the smearing of the energy in both time and frequency dimensions. This is exactly the compromise that we talked about and you should start expecting this as we work with smooth Wigner-Ville distributions you will improve the distribution in terms of removing interferences, but we lose out on the time frequency localization. That is the trade of that we have on the spectrogram and scaleogram being one such possibilities.

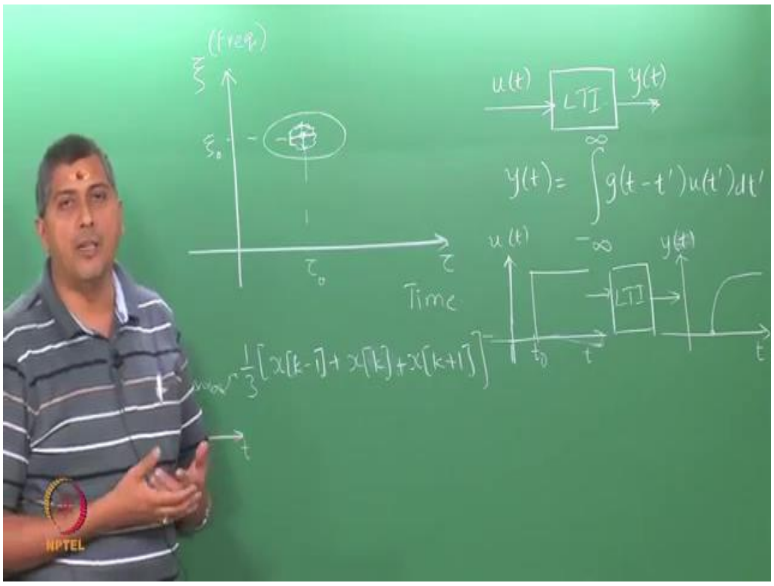
(Refer Slide Time: 08:02)



Now, having discussed the pseudo Wigner-Ville and seeing that, it actually perform some kind of smoothing in the frequency domain. We are now ready to talk about smoothing in both time and frequency and the purpose here is to enforce positivity. Earlier the purpose was to remove to bring about a local nature to Wigner-Ville and side effect was positive side effect was reduction and interferences, but not complete elimination. And there was no guarantee that the pseudo Wigner-Ville is going to be non-negative value. But now, I want non-negative valued distribution and what I am going to do now is start with the idea of smoothing the Wigner-Ville itself or convolving the Wigner-Ville. Why should smoothing of Wigner-Ville produce any positive numbers? Well, think of this elementary example where I have a sequence of positive and negative numbers.

If I perform some kind of averaging in a specific manner then, it is possible to always guarantee a positive outcome, that is the positive value number as a outcome of such a smoothing operation that is a exactly the idea here. Now, since the Wigner-Ville is a 2 dimensional function I need to perform smoothing in the 2 dimensional plane which means I have to walk along the time and frequency dimensions. And this theta is the 1 that will help me smooth the Wigner-Ville at every point with in the vicinity of tau comma g tau comma g is where and standing. And I am going to walk in the neighborhood of tau comma g and this tau prime and g prime will help me march around the tau g.

(Refer Slide Time: 09:43)



So, let me just draw that on the board for you. This is the tau axis or the time axis and this is the chi axis it is the frequency and time. Assume that your Wigner-Ville's centered around some point here. Let us call this is tau and chi are you can say g naught although we do not have tau naught and g naught in integral, but think of replace momentarily tau g in integral with tau naught and g naught. And by smoothing what we are going to do is we are going to walk along with in walk around this point within this vicinity may be in this region and go and add up numbers in a particular way. Now, how far do I go within in this vicinity and how add up is what theta will tell me.

Well, the first thing that we should know when we are talking about smoothing here is, if I am standing here, I am walking to the right to the left to the bottom and to the top in all directions surrounding this point and looking to the left right and so, on. And this in estimation theory is also clear when I want to construct the smooth estimate of a signal. So, let us say I have a signal x which has this kind of profile and let us say this is I do not have the x itself. But I have a noisy version of the signal and standing at some point in time t I would like to construct a smooth estimate. Because I have a noisy version of x I would like to get an estimate of the underline x signal by the smoothing operation.

One of the ways of doing it is I could look to the left right and the center and take an average a uniform average or a weighted average whichever way. So, a simple average would be this ideally, I should use the sample data for averaging. So, as you can see to construct this will give me this smooth estimate at this point assume that this is a k sampling instant. At this point this would give me a smooth estimate of course, there are number of possibilities here millions of possibilities. It is a very simple way of estimating the x this smoothing essentially will remove this noisy terms that present in x k and also surrounding points. And that is exactly the idea here also, I have some fluctuations here when I go from go around this point and by smoothing I am killing those fluctuations that are causing negative values to appear and also interferences to appear.

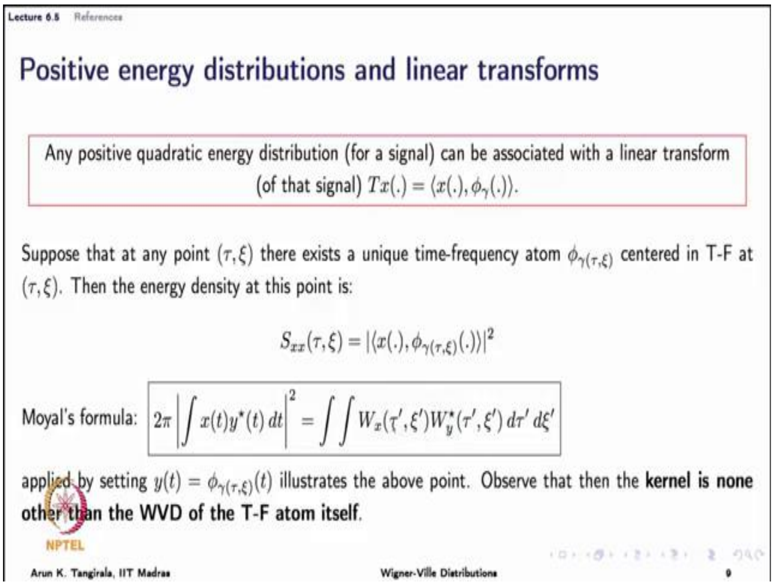
It is a very, what do you say schematic way of illustrating things, but the more formal operation is to per evaluate that double integral that is a first point. The second point is, we use the term convolving synonymously with smoothing and the reason comes from the systems theory, linear systems theory perspective. To understand that consider a standard linear time invariant system or a general linear time invariant system continuous time 1 whose input is u and output is y. And we know from linear systems theory that the output of this LTI system is model by this convolution equation what does this convolution do? Let us take a step input here so, if I have a step and the input and let us say the step is actually starting at t naught in time.

This is my input to the system and when it is goes through the LTI system. Assume that this LTI system has damping characteristics, the output comes out as a distorted step is what we normally see. But in fact, what we can also say is that the sharp corner in the step has now been smoothened out or flattened out. So, we can use the term convolution or smoothing or distorting equivalently depending on the context. Of course, the gain here and the rate at which it will reach the steady state all that depends on the filtering characteristics of the LTI characteristics. But the main messages is convolution and smoothing or identical or synonymous, but always smoothing does not have to been implemented as convolution. Whenever I am performing convolution it means smoothing, but not necessarily vice versa I could implements smoothing in a number of other different ways as well

So, now having understood the relation between smoothing of the idea of smoothing itself, and having understood the relation between smoothing and convolution. We should recall from our experience with pseudo Wigner-Ville that we are going to lose out on the time frequency resolution as the result of this smoothing. And we can bring about positivity by enforcing certain conditions on theta. This is what we are going to learn in the next lecture as to how by imposing certain requirements on the kernel not only can I bring about positivity, but also certain other desirable features of the distribution that is a modified Wigner-Ville distribution. Now, all the class of distribution that I get through this operation are called smooth Wigner-Ville and remember that I am performing are smoothing in the time frequency plane. And if I decide to perform convolution then this kernel would be rewritten as theta of tau minus t prime comma g minus g prime.That is that is what will be the consequence of assuming convolution as a way of smoothing.

So, the question that we should ask at this point is I can also arrive it positive distributions by starting from a linear transform and taking this squared magnitude like a do and short time Fourier transform or otherwise spectrogram or even scaleogram. That is what exactly I do I take a Fourier transform or a short time Fourier transform or a wavelet transform and take this squared magnitude. Then why should I break my head on performing this smooth Wigner-Ville? And the other question that is more important and interesting to ask is there a if there is a connection between these two approaches, that is the smooth Wigner-Ville and the squared magnitude of linear transforms will answer the second question first and talk about the first question second.

(Refer Slide Time: 16:57)



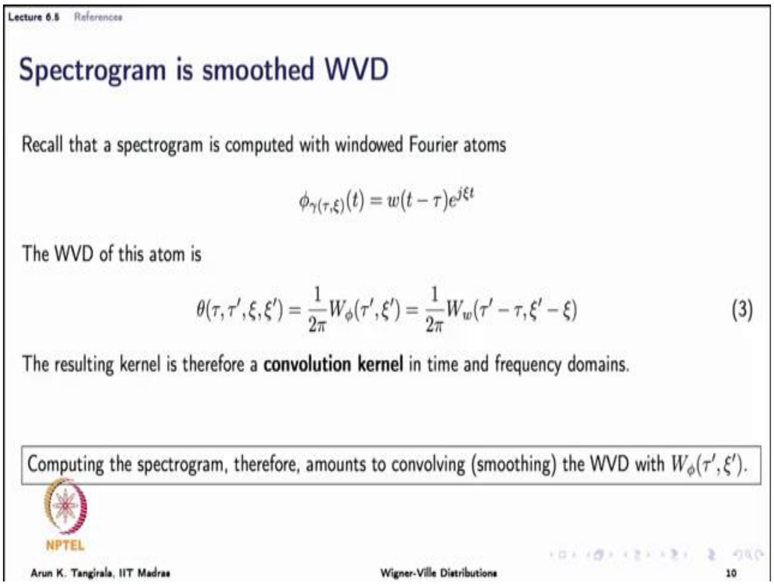
The main result that we have is, which is now established in the literature is that any positive quadratic energy distribution for a signal can be always associated with a linear transform of that signal where, we denote the linear transform by this operator T. Essentially that linear transform amounts take an taking an inner product between the signal x and a time frequency atom. Because we are talking in the context of time frequency plane, but always linear transforms can be associated with inner products. Now, this is very nice, because it says that instead of constructing smooth Wigner-Ville I might as well start with linear transforms and then take a square magnitude will soon discuss why going through the theory of smooth Wigner-Ville's very useful. To establish the proof we only established the one way proof for this we will use Moyal's formula.

Let us assume that at any point in a time frequency plane again tau comma g there exists a unique time frequency atom centered exactly that point and that I am going to take a transform of the signal with this time frequency atom. And then construct the energy density as a squared magnitude of the inner product coming out between the signal and the time frequency atom itself. The gamma that you see here is an additional parameter like a scaling parameter that you would be using for the time frequency plane, time frequency atom. Now, I invoke Moyal’s formula to establish the connection between smooth Winger-Ville’s and the energy density that I construct from linear transforms what does Moyal’s formula tell me?

It is says essentially that Wigner-Ville's reserved the square inner products, on the left I have the inner product square in a product between x and y and in the right I have the double integral of product of Wigner-Ville’s of x and y. Now, if I set y to the time frequency atom I get the desired result how does it come about? Well, when y is the time frequency atom this integral is nothing but the inner product between signals x on the time frequency atom itself. And the magnitude square is nothing but energy density well there is a factor of 2 pi that is a spoil the proof in any way.

On the right hand side as I am substituting y with the time frequency atom what do I have? I have double integral of the Wigner-Ville times the Wigner-Ville of the time frequency atom, compare this with the expression that we had in equation 2 here. So, what is happening here is that the smoothing kernel the, that is the term that is used for theta. The smoothing kernel when it comes to constructing an energy density from linear transform is nothing but the Wigner-Ville of the time frequency atom itself. In other words, if I start with a Wigner-Ville for a signal and then I decide to also perform a linear transform with certain time frequency atom what I can do is, I can do two things and I will get the same answer what are the two things? I take the Wigner-Ville of the time frequency atom with which I am going to transform the signal and smooth the Wigner-Ville of the signal that is one result. The other result is, simply linearly transform are transform the signal with this time frequency atom and construct the square magnitude I will get the same result. So, which means I can show that probably spectrogram and scaleogram which are constructed from linear transform are nothing but special cases of smooth Wigner-Ville’s all I have to do this identify the associated kernel.

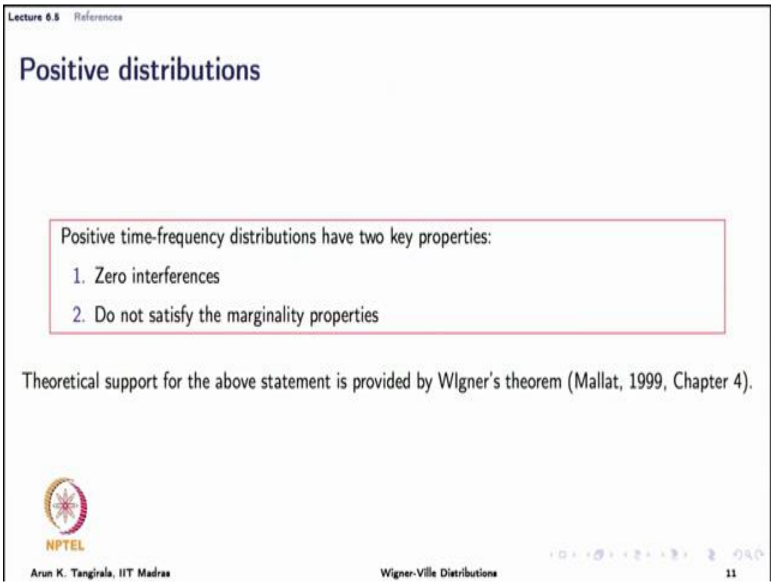
(Refer Slide Time: 20:55)



Let us look at the spectrogram as an example I know the time frequency atom associated with this construction of spectrogram is this, W of t minus tau e to the j g tau. Here there is no gamma perse I could bring in a gamma if I wish. This W is a window function it essentially the clip sinusoid that I have. Now, the Wigner-Ville of this time frequency atom is nothing but the Wigner-Ville of the window itself shifted in time and frequency why? Because the Wigner-Ville satisfies or respects this translation and frequency shifts. All I have to do now, therefore, to construct the spectrogram is I perform the Wigner-Ville of the signal compute Wigner-Ville of the signal and smooth it with the Wigner-Ville of the time frequency atom itself.

So, now I have to learnt another way of computing the spectrogram well does not the purpose here. The purpose here is to show the formal connections between spectrogram and smooth Wigner-Ville. And the fact that I can take from here by looking at the Wigner-Ville of the window is that, this kernel is a convolution kernel. Because I have a tau prime minus tau and chi prime minus g and when I convolve the Wigner-Ville of the signal with this kernel. It is going to be a when I smooth the Wigner-Ville with this kernel is going to the convolution in both time and frequency.

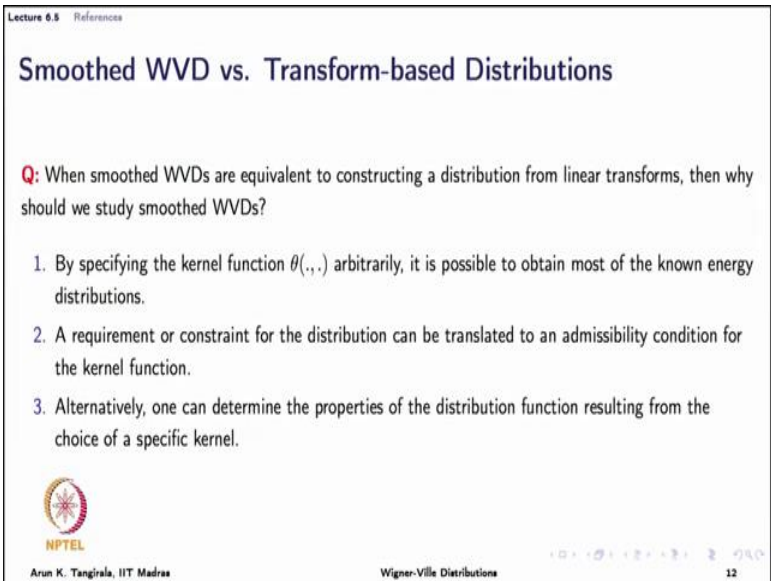
(Refer Slide Time: 22:28)



Now, although I have learnt how to achieve positivity of the Wigner-Ville of the smooth Wigner-Ville I should keep in mind that the moment I work with positive time frequency distribution two things are going to happen. One that the interferences are going to go to the 0 that has to happen that is by virtue of the Winger's theorem itself. And that this positivity time frequency distribution will not satisfy the marginality property this is again by the virtue of Winger's distribution theorem itself. And the proof of both this points although I have learnt how to arrive at positive quadratic time frequency distribution by way of smoothing the Wigner-Ville I should keep in mind 2 points. One that this positive quadratic time frequency distributions will not satisfy the marginality property.

We have studied this earlier in the form of Winger's theorem, two which is the nice thing that there will not be any interferences in this time frequency distribution or the positive time frequency distribution. Although the slide says positive time frequency it should be strictly positive quadratic time frequency distribution. Of course, only exception being the chirp which violates this, but strictly speaking for modulated chirps the distribution is not really quadratic. So in fact, in essence you are not violating this result itself both this points are proved nicely in Winger's theorem. And for a proof I refer you to Mallat's with the proof is fairly straight forward to understand.

(Refer Slide Time: 24:00)



We conclude with the answer to the first question that we raised where we said if I can construct the positive quadratic time frequency distribution starting from a linear transform such as short time Fourier transform or wavelet transform, then why should I work on smooth Wigner-Ville's? Well, there are number of advantages; one by specifying the kernel function I can actually arrive at a number of desirable properties that I want for the distribution. Because a particular distribution may suit only some class of applications for some other application some other properties may be desirable.

In which case I may not know how to modify the spectrogram or scaleogram they may not be suitable for the all applications we know the limitations. So, in view of that it is nice to know that there exist this freedom to exercise where we can arrive at the desired frequency distribution. Of course, there are going to be constraints, but there is considerable freedom. And on the other hand, I can also find out if there is requirement on the distribution, what kind of kernel will produce. They may not be just one kernel there may be a number of kernel that give me this.

So, these are the prime reasons why the smooth Wigner-Ville distribution calls for some formal attention and careful study. The next 2 lectures is going to be building on what we have learnt here. In the next lecture, we are going to look at Cohen's class of representations or distributions. And in the final lecture, we are going to look at what are known as affine invariant distributions, which are again extension of what we learn in

Cohen’s class. So, hopefully you enjoyed the lecture and if you want more details or formal details please refer to these books and papers that I listed here. We will meet again in the next lecture.

Thank you.