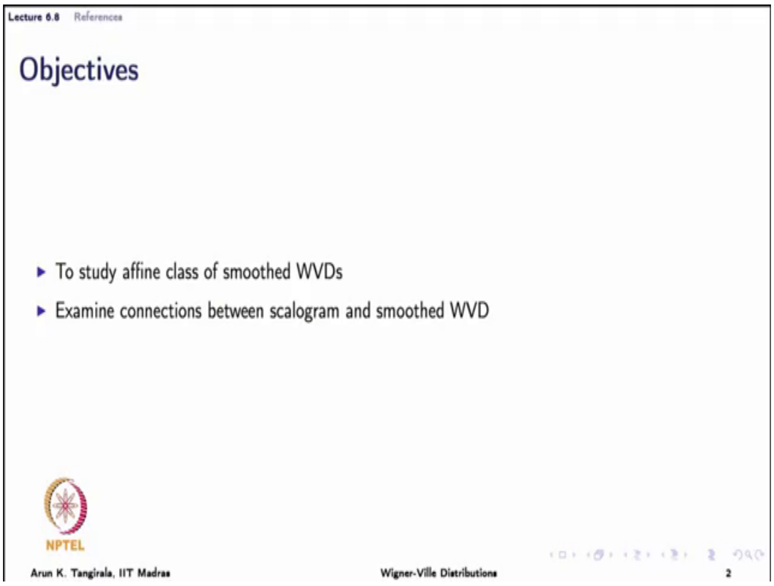
Introduction to Time-Frequency Analysis and Wavelet Transforms
Prof. Arun K. Tangirala
Department of Chemical Engineering
Indian Institute of Technology, Madras

Lecture - 6.8
Affine Class and Closing Remarks

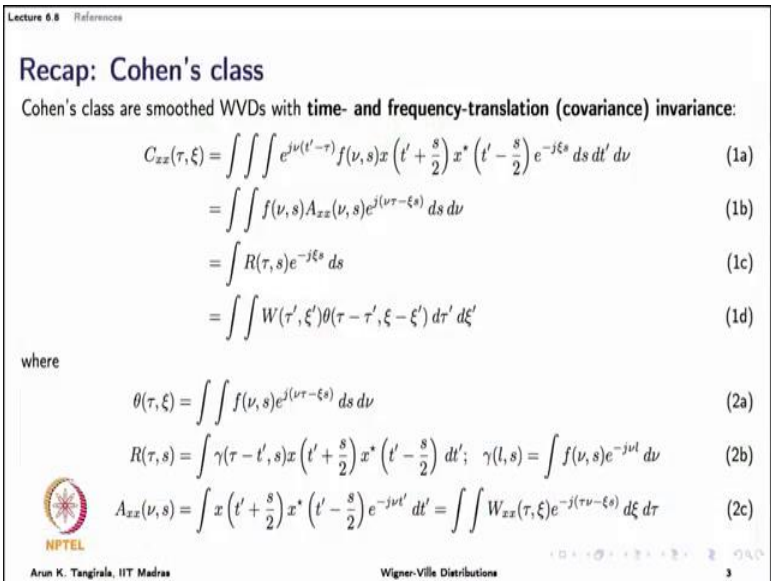
Welcome to lecture 6.8, which is the last lecture on the topic of Wigner-Ville distributions. In this lecture, we are essentially going to look at very briefly look at this affine class of transformations, which are essentially affined invariant class of distributions. In the previous lecture and the lecture 6.6, we had looked at time and frequency invariant class of distributions, which are known as the Cohen's class. And we essentially extend the ideas of the Cohen's class, now to a different class of distributions. The idea is that we will require the new class of distributions to be affine invariant. And I will tell you what is affine invariant, essentially got to do with scaling. And since this is a closing lecture on this topic, we will make a few remarks that will hopefully give you an overview of this entire topic of Wigner-Ville and why we went through this class of distributions corresponding to Wigner-Ville and the smoothed ones.

(Refer Slide Time: 01:26)



So, the objectives of this lecture are primarily to study or in fact to briefly study affine class of smoothed Wigner-Ville distributions. And in that process, we will examine the connection between the smoothed Wigner-Ville and the scalogram. One of the best ways to look at this is these affine class of transformations, and the content of this lecture is to simply go back to lecture 6.6 and 6.7, and replace the time and frequency translation invariance with what is known as a time and scale or dilation invariance. This was one of the properties that we listed as a requirement of a joint energy density. So, that is the first step. And the second step is to replace this spectrogram with this scalogram. And the third one is to replace the window in the spectrogram with the wavelet. And what I mean by this will become clear shortly.

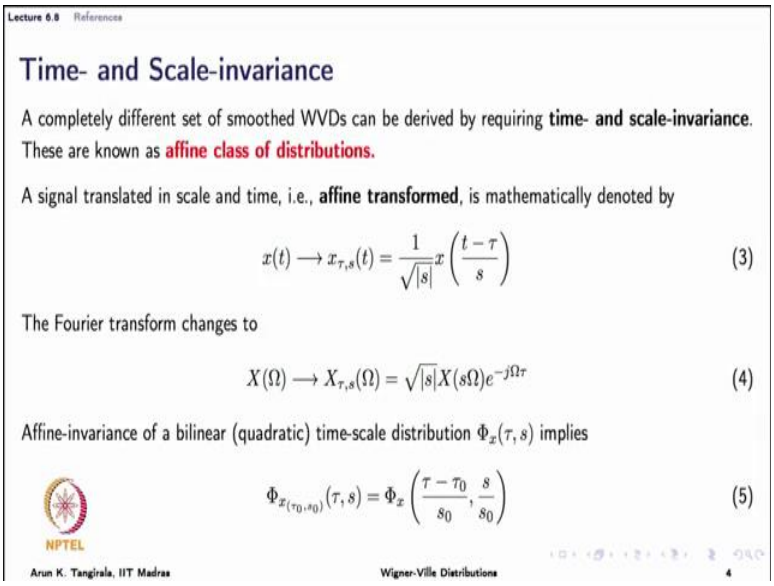
(Refer Slide Time: 02:28)



So, just to recap, because we are going to use lot of analogies here; let us recap the Cohen's class of distributions, I know this is the lot of math in a single slide, but this is nothing new, it is a collection of all the definitions of Cohen's class that we have seen in the previous two lectures. The top one is the one that is given by Cohen. And it also deserves to be on the top on a lighter note, because it has three integrals. So, it ((Refer Time: 02:54)) highest. And then we have the other forms in terms of the time-lag kernel in terms of the ambiguity function. For example, the second one is in terms of the ambiguity function. The third one is in terms of the generalized auto-correlation function that we talked about in a previous lecture. And the last one is the one actually that we began from or the way we introduced Cohen's class as a smoothed Wigner-Ville distribution with convolution being the specific smoothing operation that we are performing.

So, theta is the smoothing kernel there. And the reason for this convolution is because we wanted time and frequency translation or shift invariance. And convolution operators will give you the desired property. The remaining three equations are essentially telling you what is theta – the relation between theta and f; that is, the smoothing kernel and the weighting kernel; and the expressions for the generalized auto-correlation as well as the ambiguity function. So, this is in a nutshell, mathematical summary of the Cohen's class of distributions.

(Refer Slide Time: 04:11)



Now, what we want is time- and scale-invariance instead of time- and frequency-invariance. In many operations, in many applications, I would like to have invariance to scaling of features rather than frequency shifts. Is scaling different from frequency shift? Yes, of course, because we know qualitatively that scale and frequency are inversely related. So, when I shift a feature and frequency, does not mean that there is a scaling of the signal in time. Therefore, time- and scale-invariance is different from time- and frequency-invariance. Of course, there are a few distributions that have both properties. And in fact, Wigner-Ville has all – time, scale, frequency shift invariance, everything.

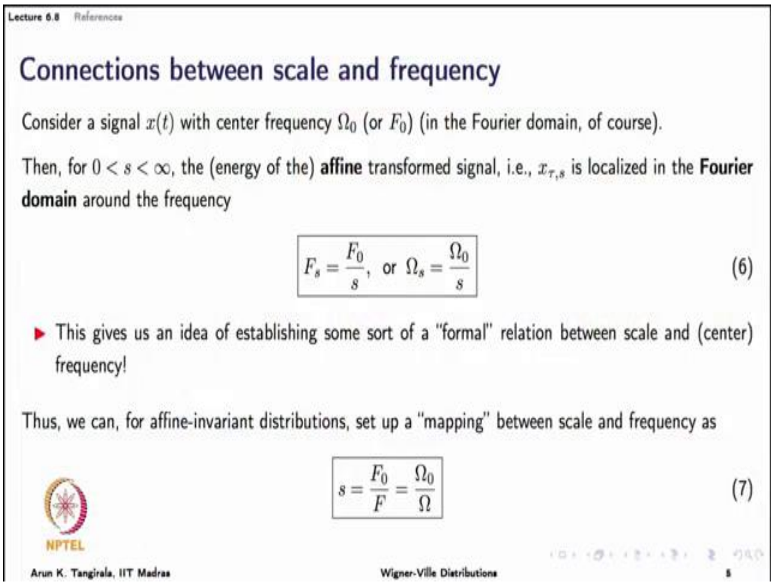
Let us understand what is meant by this affine class of distributions; in fact, the term affine itself. Affine essentially refers to a time – translationing time and scaling of the signal. And I have given the mathematical definition of this affine transformation of a signal in equation 1. When a signal is translated by tau units and time and scale by a factor of s, then it is said to be affine transform. And this 1 over root mod s is to ensure that the scale or the affine transform signal – scaled and translated signal has the same energy as x of t. We have seen this before. And we have mod s here to take care of negative values of scaling. In practice, we do not do that; but mathematically, you should be equipped to handle that. And therefore, we have a modulus of s. When a signal is affine transform; then the Fourier transform changes in this manner, that is, a Fourier transform of this new signal is given by this expression. It is related to the Fourier transform of the original signal in this manner. And this is fairly easy to derive using the properties of the Fourier transform.

Now, note that I am using the big omega just to emphasize that we are in continuous time. And the important thing to notice is that now, I have X of s omega; we will use this observation later on to derive a relation between the frequency and the scale. Now, coming to affine invariance of a bilinear time-scale distribution; until now, we have been talking of bilinear or quadratic time frequency distributions. Now, we are going to talk of distributions in the time-scale plane. You should not get intimidated here; all we are doing is instead of using the time-frequency plane, we are now analyzing the signal in the time-scale plane. So, my basis functions in other words are no longer characterized by time and frequency; but they are rather characterized by this time and scaling parameter s. However, we know the relation between – qualitative relation between s and frequency. That will make the understanding a bit easier.

So, what we mean by affine invariance of a quadratic… We will leave the term quadratic or bilinear; it is understood; we are only looking at such distributions. So, what we mean by affine invariance of a time-scale distribution denoted by phi is that the time-scale distribution of the affine transformed signal, in other words, translated and scale signal should be related to the time-scale distribution of the original signal in the same manner as you have actually scaled the signal itself and translated.

So, if I evaluate the distribution of this new signal – affine transformed signal at any point tau and s in the time-scale plane, it should be the value of the distribution of the original signal translated by tau naught and s naught. So, what are these tau naught and s naught? These are the translation and scaling parameters that I have used to transform the signal. So, the distribution of the new signal should be distribution of the original signal or the old signal translated by tau naught and scaled by s naught. So, that is what we mean by affine invariance. And compare this with the time- and frequency-shift invariance. There we had a similar requirement; we said if I shift the signal in time and in frequency. Again, shift in frequency means modulation. Then, the distribution also should reflect the same thing. It should follow the same shift in time and frequency.

(Refer Slide Time: 09:06)



Now, before we move on to talking about this smoothed Wigner-Ville distributions that give me affine invariance, it is useful to understand a connection between scale and frequency, because now you can think of this lecture – this entire lecture as a stepping stone for wavelet transform. This offers a very nice transition from Wigner-Ville to smoothed Wigner-Ville to scalograms or wavelet analysis. Therefore, it is important to understand the connection between scale and frequency. Many a times, until now, we have noted the connection between scale and frequency. We have said that the share and inverse relationship whether it is an introductory lecture or when we were discussing the properties of Fourier transform. We observed that low scales correspond to high frequencies; and high scales correspond to low frequencies.

Now, we will mathematically understand, why this is so. We have already looked at it semi mathematically before. Consider a signal x with center frequency omega naught. We have now a clear understand – understanding of what we mean by center frequency; center frequency is a center around which the energy is concentrated. And we have also given mathematical definitions of the center frequency in unit 2 or unit 3. So, assume that the signal has energy spread around omega naught. Then, the energy spread of the affine transformed signal – now we will restrict this scale values to anywhere between 0 to infinity. We will exclude 0. The energy of the affine transformed signal is now localized in the Fourier domain. Although I am using the scaling parameter, I am asking suppose I scale a signal; then how does its energy change? And we have noticed this before even mathematically and qualitatively.

So, let us now go back to equation 4, which will help us understand how the energy or what is the center frequency of the energy of the transform – affine transformed signal. Look at this relation. It says the Fourier transform of the affine transformed signal is related to that of the original signal in this fashion – X of s omega. Therefore, if x, that is, the original signal has a center frequency of omega naught; then the center frequency of the affine transformed signal should be omega naught by s. It should not… It is not omega naught times s. Remember X of s omega is centered at omega naught. Therefore, the transformed signal will have a center frequency at omega naught by s. So, that should be fairly clear.

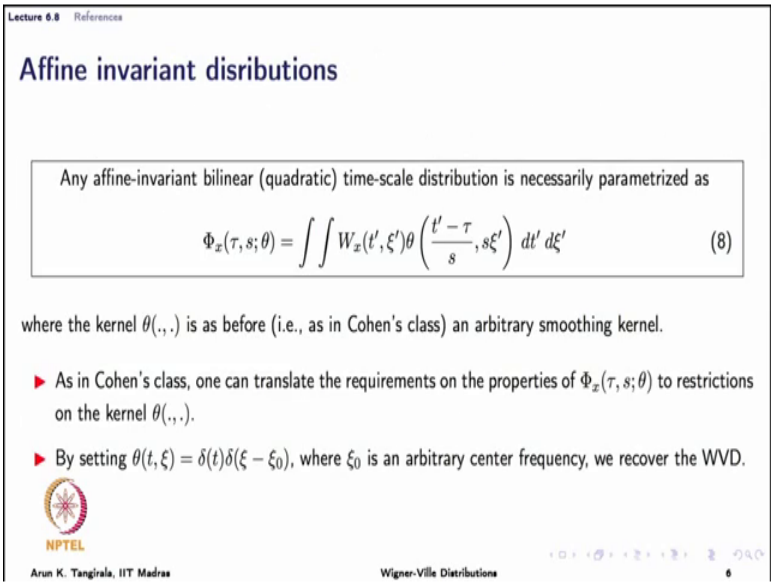
What this means is if I am using a value of s equals 2 let us say; then the center frequency of the new signal is omega naught by 2; which means it shifts to the left; exactly what we discussed when we talked about the scaling property of the Fourier transform. You should go back and refer to that lecture. So, now this is clear. Whenever I scale a signal by a factor of s, its center frequency shifts by factor of 1 over s; all right?

Or, you can say if I am scaling the signal by factor of 1 over s depending on how you call it; but if I am scaling a signal in this fashion as I given in equation 3, then it center frequency shifts to omega naught by s. This now gives me a nice relation between scale and frequency of any signal.

Now, here we cannot achieve an absolute relation between scale and frequency. There is always this reference center frequency; which means I can only strike a relation between scale and frequency with respect to that of the original or the unscaled signal. So, I always need that reference point; there is no absolute relation. And we will observe this even more when we talk of wavelet transforms and so on. So, the point is the scaling factor s and the frequency. So, s is the scaling value that I have used to scale a signal. And this big F is now the center frequency of the transformed or the scaled signal; they are related in an inverse manner. So, if I ask you what is the frequency or the center frequency of a scaled signal; then I need to give you also the center frequency of the unscaled signal.

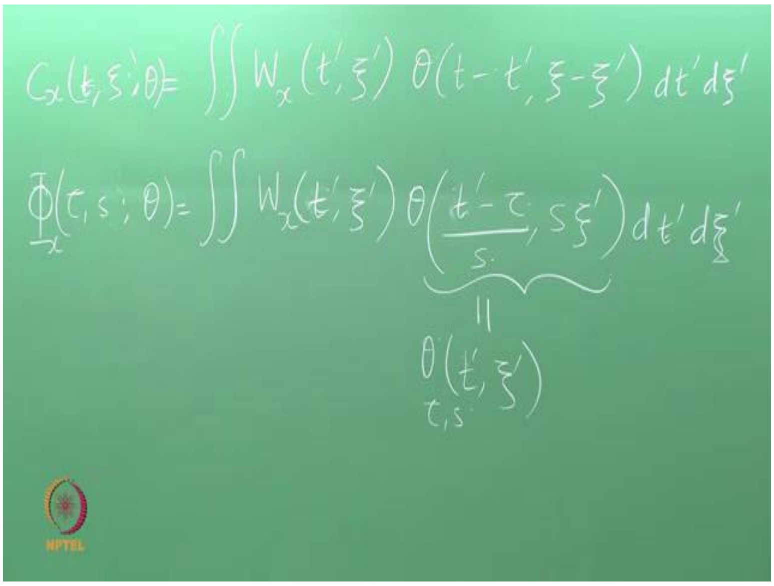
Then you can calculate the center frequency of the scaled signal. I hope that is clear. Whether you express this in cyclic frequency or angular frequency, the relationship is the same. It is just at the center frequency, then gets expressed accordingly. So, the summary is there is no absolute relation between scale and frequency; there is only a relation between scale and center frequency. We will get a different interpretation of this relation when we discuss wavelet transforms. So, you can say s is inversely propositional to F; and that proportionality constant is the center frequency. That is another way of looking at it.

(Refer Slide Time: 14:30)



Now, we return to the topic of affine invariant distributions. Any affine invariant time-scale distribution is necessarily parameterized as given in equation 8. Now, I know that the integral is the big intimidating in the first place; but let us relate this to what we saw for Cohen's class. So, let us go back to the recap and look at equation 1d, because that is the nicest analogy that we have. Here we said that we want a time and frequency invariance. And therefore, we said that convolution operators will produce the time and frequency invariance. So, now theta is a kernel that… Here theta is a kernel that we used for achieving that property for this smoothed Wigner-Ville. Here also, we use the same notation, but we say now that I am going to smooth the Wigner-Ville. Once again, I am smoothing the Wigner-Ville with a different kind of operation. I can take a particular any kernel theta 2-dimensional function; but I cannot smooth with, I cannot attain affine invariance by smoothing this Wigner-Ville with an arbitrary theta. I have to take an arbitrary theta and scale it accordingly. Why is that necessary? Let me just quickly explain that to you.

(Refer Slide Time: 16:06)



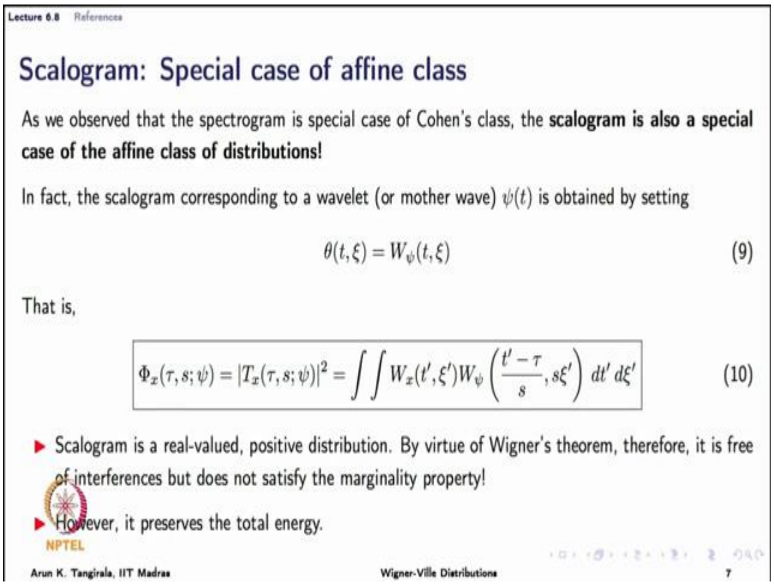
So, equation 8 essentially tells me that this smoothing is performed in this way – W x – I have now used only a single subscript; we have been using double subscript, but we will switch over to the single one. W x of tau prime times xi prime times – it is a theta of tau prime minus tau by s times s xi prime. And then we are integrating t; so this should be t as per the notation used in the equation. So, this is the integral that we are evaluating. If you quickly recognize, this is nothing but a scaled kernel itself. So, I pick a kernel theta and I scale it essentially in a same way or using the same values of tau and s at which I am evaluating this smooth Wigner-Ville. That is the main point here.

Compare this with what we did in Cohen's class. In Cohen's class, it is a same Wigner-Ville that I start with. However, what I do is I have here a kernel that is shifted in time and frequency. Now, you notice the parallelism. Whenever I want time and shift invariance, then I pick a kernel and shift it in time and frequency and then smooth the Wigner-Ville with that. Now, I want time and scale invariance. So, I pick a kernel, that is, I pick a 2-dimensional function and then scale it and then use the scaled version to smooth the Wigner-Ville. Here I shifted by the same amount; you can even write it as t prime minus t and xi prime minus xi; things would not change. So, I am shifting it by the same amount as at the point at which I am evaluating the smooth Wigner-Ville.

Here as well, I am scaling it by the same values of tau and s at which I am evaluating the smooth Wigner-Ville. That is the parallel. If you understand that analogy, then the rest of the development is more or less straightforward; however, if you look at historically, the development of affine invariant transforms came about in 1980s; whereas, Cohen's class existed as early as 1960s. So, there is a huge time gap in terms of the years that it has taken to come up with the affine transforms. And by then wavelets were also becoming quite popular.

And now, we shall see the connection between this affine invariant class of distributions and scalogram; pretty much like we saw the connection between the time and shift invariant and spectrogram. And before we see the connections, again the same story; we said earlier in Cohen's class that if I want a certain property on the smooth Wigner-Ville distribution, I impose the restriction on the kernel; likewise also, I can impose the restriction on theta to achieve a desired property on the new smooth Wigner-Ville distribution, which is a represented by phi. And you can also check that by setting, by choosing this particular kernel, you can recover the Wigner-Ville. Again, this xi naught is the center frequency reference that I use to establish the connection between scale and frequency.

(Refer Slide Time: 20:06)

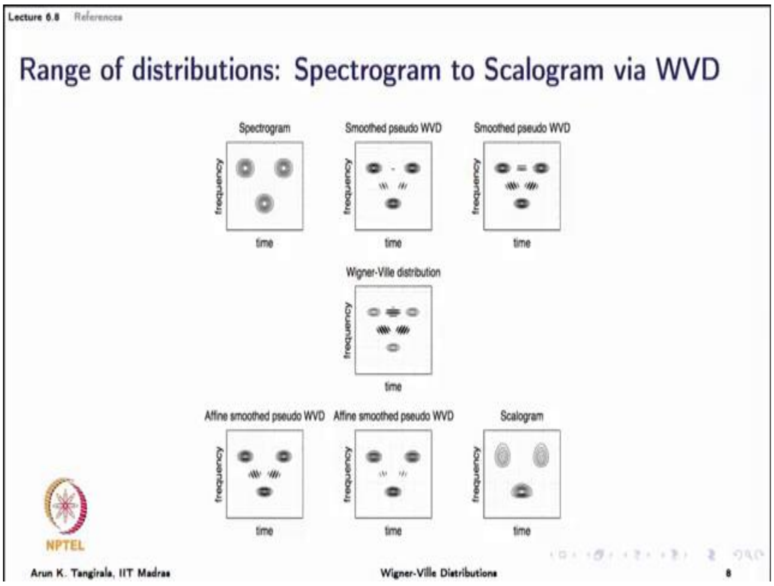


But more importantly, what we can learn from this lecture is that the scalogram is a special class of the affine class or the affine invariant class of distributions in the exactly the same manner as a spectrogram was. There we said, if I set the kernel, that is, the time-shift invariant – the convolution kernel to the Wigner-Ville of the window; then I essentially achieve spectrogram. Here as well, if I set the kernel that I am going to scale, that is, the scaling kernel or the smoothing kernel to the Wigner-Ville of the wavelet, then the resulting smoothed Wigner-Ville distribution is nothing but the scalogram. And that is what is given in equation 10 for you. So, here you have the notation that we have been using in this lecture phi evaluated at tau comma s, that is, the distribution evaluated at this point in the time-scale plane using wavelet. So, now, we are saying the kernel is a wavelet itself. Strictly speaking, the kernel is actually Wigner-Ville of the wavelet; but we do not really write W of c.

And if I had performed a wavelet transform, which is a subject of the remaining two topics in this course; then I would have obtained this transform and then taken the squared magnitude of that. That is exactly what you will get by smoothing the Wigner-Ville with the scaled Wigner-Ville of the wavelet. So, in other words, I choose the kernel itself as a Wigner-Ville of the wavelet; scale it and then smooth the Wigner-Ville; I get the scalogram. And is this the way I am going to evaluate the scalogram? No, this is more of theoretical interest; this is to show that I can recover the scalogram from smoothed Wigner-Ville as well; exactly the way we pointed out for spectrogram.

Now, obviously, the scalogram by definition is a real valued positive distribution. Therefore, by virtue of Wigner's theorem, it is free of interferences; but unfortunately, does not satisfy the marginality requirement; that is, it does not have the marginality property. However, it preserves the total energy. And we will talk about this scalogram exclusively when we talk of wavelet transforms. The purpose here is to show you that by smoothing the Wigner-Ville in a particular manner, you can recover the scalogram.

(Refer Slide Time: 22:42)



I will conclude this lecture again with a figure that I have taken from the time-frequency toolbox tutorial. This is a beautiful picture; I really like it, because it conveys essentially the purpose of the entire topic of Wigner-Ville distributions in this course. Anybody who is working with spectrogram and scalogram not just casually, but more seriously should be aware that the spectrogram – between the spectrogram and the scalogram, there are many – many grams. So, we begin… In this lecture, we began with the Wigner-Ville distribution; I will leave it to you to determine now, having gone through Wigner-Ville, you should able to guess what kind of signal would have this kind of a Wigner-Ville distribution. You should get used to these kinds of analysis, because then only, you have really understood at least the essence of the distributions here. So, I have interference terms as usual appearing in Wigner-Ville.

If I go in a particular direction; that is, if I smooth in a particular manner, I get the spectrogram. If I go in a particular other direction, I get the smoothed pseudo Wigner-Ville… We call that the smoothed pseudo Wigner-Ville is not guaranteed to be free of interferences; but I can obtain interference-free smoothed pseudo Wigner-Ville with the choice of the windows in the time – smoothing windows in the time and frequency; recall the previous two lectures. And with another choice of windows, I get even more interferences; but what you should notice is as I go from spectrogram to one variant of smoothed pseudo – another version of smoothed pseudo Wigner-Ville, I am improving on the time-frequency localization of the energy. But then I have to sacrifice on these interferences.

Now, though the ones that you have at the top are time- and frequency-shift invariant class of distributions. Just now, in this lecture, we have studied affine class of distributions or affine invariant distributions. And those are at the bottom here. Again, if I choose a particular kernel – the smoothing kernel in the affine class, then I get the scalogram. Now, you should observe, what scalogram does has to what spectrogram does. This is something that we have discussed time and again. The spectrogram has uniform time-frequency localization, that is, uniform smearing in time and frequency in the entire time frequency plane. Although this is there are no numbers given on the plot for you, qualitatively, it exactly represents what each method does.

Now, coming to scalogram, which we have talked about it, it reveals the properties of wavelets, which we have talked about many times. At low frequencies, it has very good frequency localization relatively compared to the high frequencies. And at high frequencies, it has poor frequency localization. And that is because, to extract the high frequency content of the signal, wavelet transforms uses short-lived high frequency wavelets. And therefore, the time localization is good; but by virtue of the duration bandwidth principle, you lose out on the frequency localization. And now, you can apply this same analogy that you have seen on the top to the bottom as well; you have what are known as affine smoothed pseudo Wigner-Ville of different kinds depends on how you choose your smoothing in the time and scale plane. There we said how the nature of the localization and interferences presence or absence of interferences depends on how I choose my windows in time and frequency. Here it is all about choosing the windows in time and scale of course, because scale and frequency are related; you can once again think of frequency if I…

The main point is this the affine smoothed pseudo Wigner-Ville like the smoothed pseudo Wigner-Ville is actually based on the concept of separable smoothing. Whereas, the spectrogram and scalogram are based on couple smoothing; that is, I do not have independent smoothing happening in time and frequency; I would like to ideally have that and I would like to do that in the best possible manner, but the duration bandwidth principle limits that; it is actually coupling in implicitly; it is acting behind the scenes and preventing me from choosing very fine smoothing in time and frequency. And scalogram and spectrogram respect that very nicely, but in a different manner; and therefore, generate different classes of distributions. So, this figure really summarizes also the entire course itself, which tells you the… Basically, the purpose of which is to tell you that there are all these distributions; and of course, to give you the math behind that. So, hopefully, with this slide, you have kind of understood essence of the topic of Wigner-Ville distributions including the smoothed ones; how smoothing can be done in timing and frequency; whether they have to be coupled or whether they can be separable; if they are separable, then what is the price that you pay; if they are coupled, then what is the benefit that you get and what is the price that you pay. And more importantly, that you go from spectrogram to scalogram via Wigner-Ville distribution.

Although we have shown that the spectrogram and scalogram are special cases of Wigner-Ville, it is possible to start from one of these distributions that are shown here and arrive at the remaining other. So, it is not that as if Wigner-Ville is a very sacrosine distribution and everything originates from Wigner-Ville; it is not that. Mathematically, we have shown that one – all the remaining ones that are surrounding Wigner-Ville are special cases of smoothing the Wigner-Ville. But you can always start from spectrogram for instance and keep improving. You keep making improvements in one direction, you get smoothed pseudo Wigner-Ville; in another direction, you get Wigner-Ville; and you go further, you can reach the scalogram itself. The important thing to notice is that the scalogram has been designed for time and scale invariance; and the spectrogram has been designed for time and frequency-shift invariance. So, with this these words and with this discussion, we will draw close on the topic of Wigner-Ville distributions. And we are all set now to discuss the topic of to learn wavelet transforms, continuous wavelet transforms, discrete wavelet transforms. It is an exciting world, because there are a number of other applications.

Wigner-Ville distribution did not allow us to look at filtering, because we know that it is hard to recover the signal uniquely from the WVD. Whereas, it is possible to do that with the short-time Fourier transform or the wavelet transform. So, we will talk about the applications of the wavelet transforms to signal estimation apart from the time-frequency analysis as well. So, see you in the next lecture, which will begin with the introduction to the continuous wavelet transform.

Thank you.