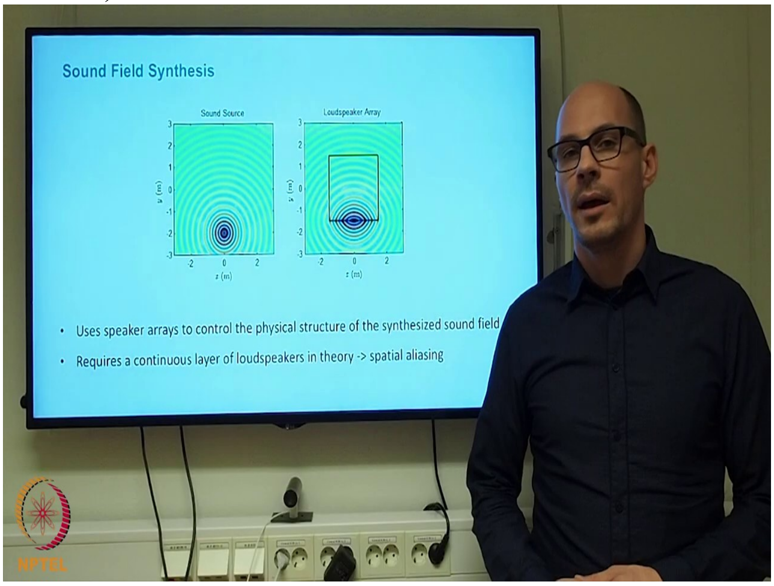
Audio for Virtual Reality
Professor Jens Ahrens
Division of Applied Acoustics
Chalmers Institute of Technology
Sound Field Synthesis

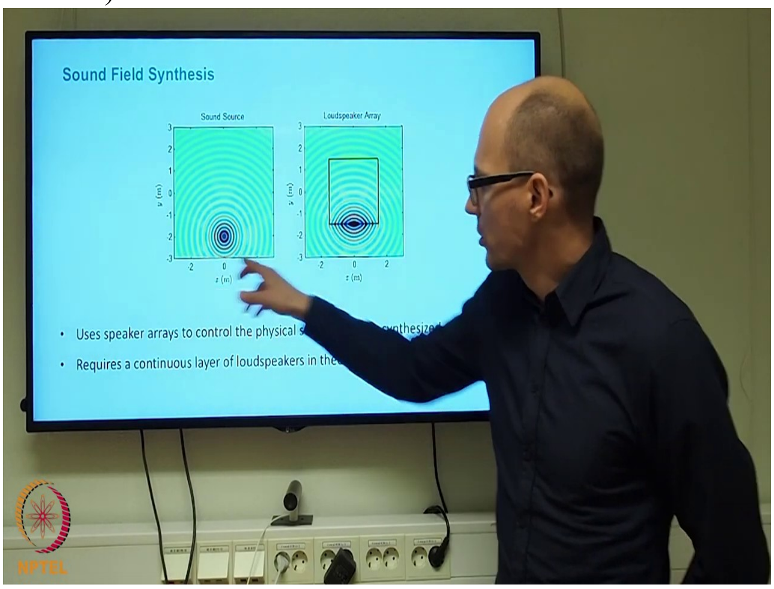
(Refer Slide Time: 00:15)



One method or one approach that overcomes basically all limitations of the previously mentioned approaches may be termed sound field synthesis.

And the concept is illustrated in these two simulations. So we are looking at

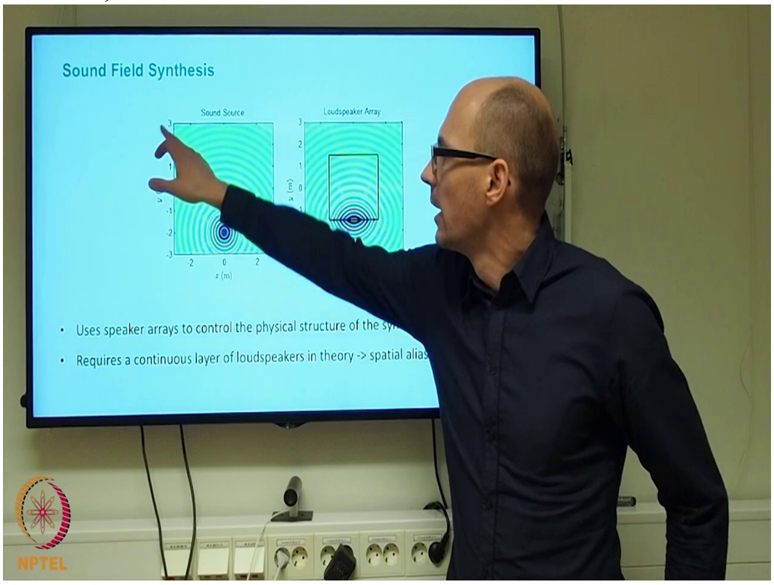
(Refer Slide Time: 00:32)



a portion of space. This is more, so could be for example top-down view on the horizontal plane.

Here we have one spatial dimension

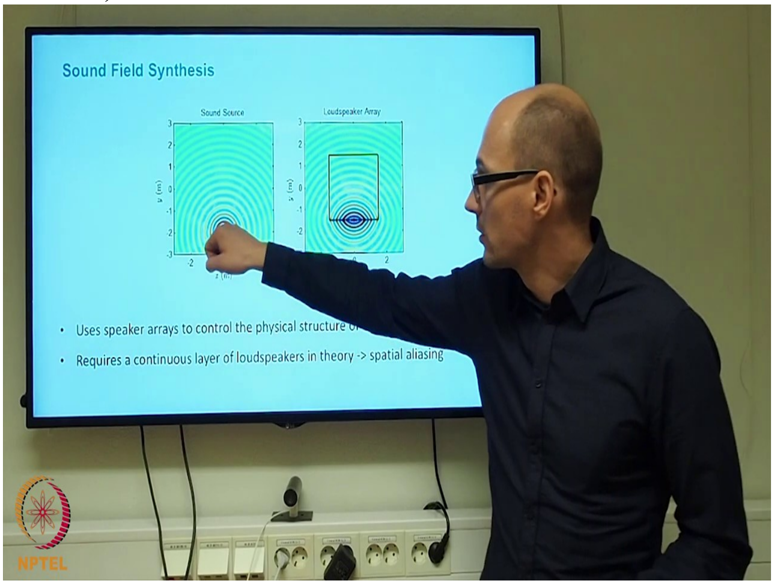
(Refer Slide Time: 00:40)



and another spatial dimension and we are looking at the sound field that is radiated by virtual sound source.

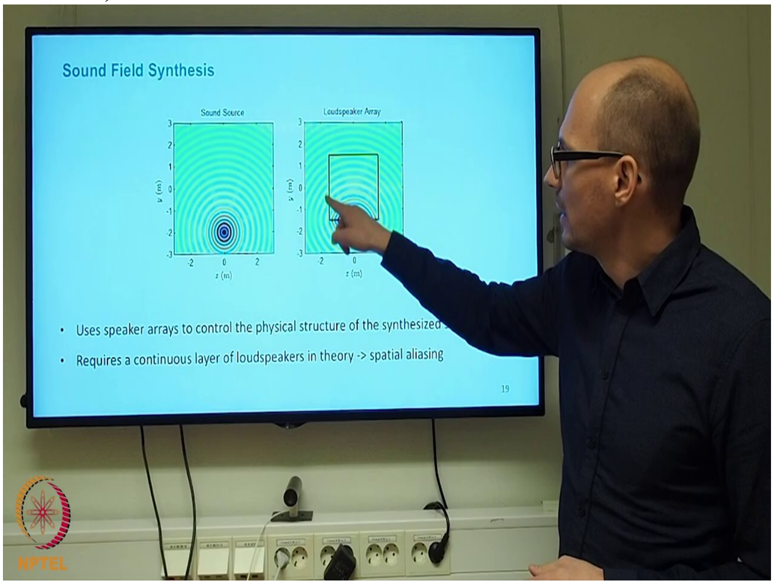
In this case, it is a spherical wave that is radiated by a

(Refer Slide Time: 00:51)



source that is located at this position and the idea behind sound field synthesis is that it needs to get many, many loudspeakers

(Refer Slide Time: 00:58)

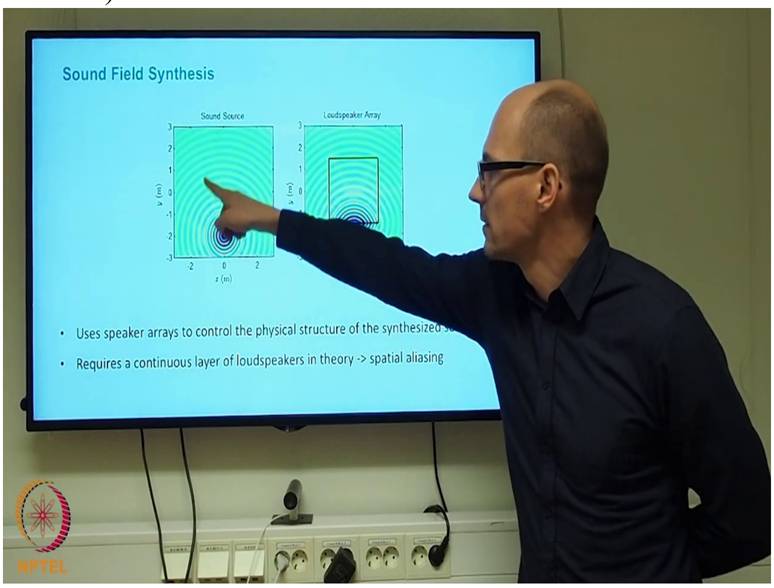


around the receiver, around the listener area

And you drive these individual loudspeakers of this larger sample such that the sound fields that they emit, they superpose and produce a copy of the desired sound field.

So if you compare the sound field inside this square which is the assumed listener area, this receiver area in this case and the sound field

(Refer Slide Time: 01:22)



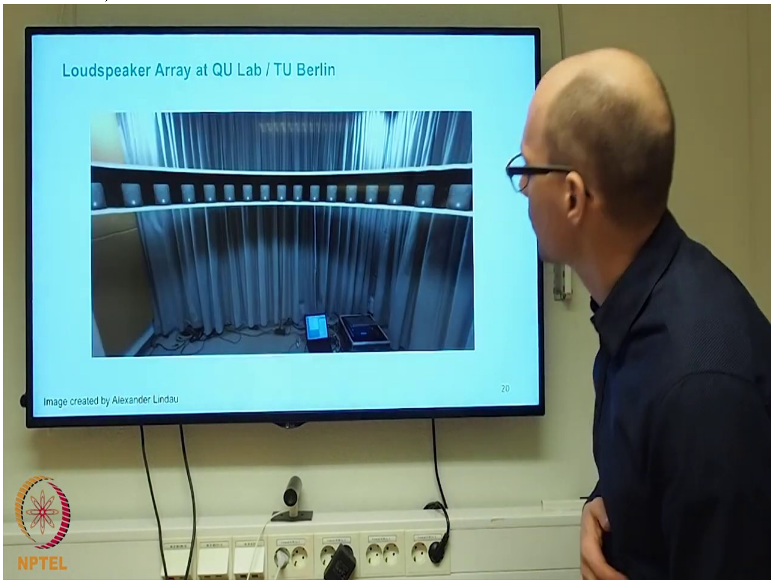
the corresponding area on the left block, you will see that they coincide very much.

There are tiny differences but conceptually they are the same. The limitation of this is that theoretically this is all possible if we have a continuous distribution of loudspeakers meaning we have a continuous layer of an infinite amount of infinitesimally small loudspeakers.

If that is the case and our target volume is enclosed with the surface of loudspeakers then we can perfectly reproduce any sound field that is possible in real life. That would be perfect reproduction.

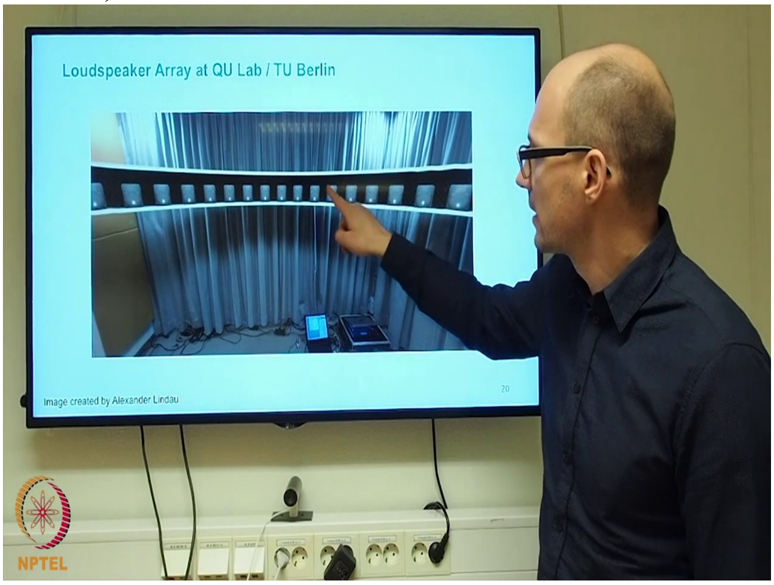
Unfortunately real world systems, they look more

(Refer Slide Time: 02:12)



like this so they are composed of a very finite set of

(Refer Slide Time: 02:15)

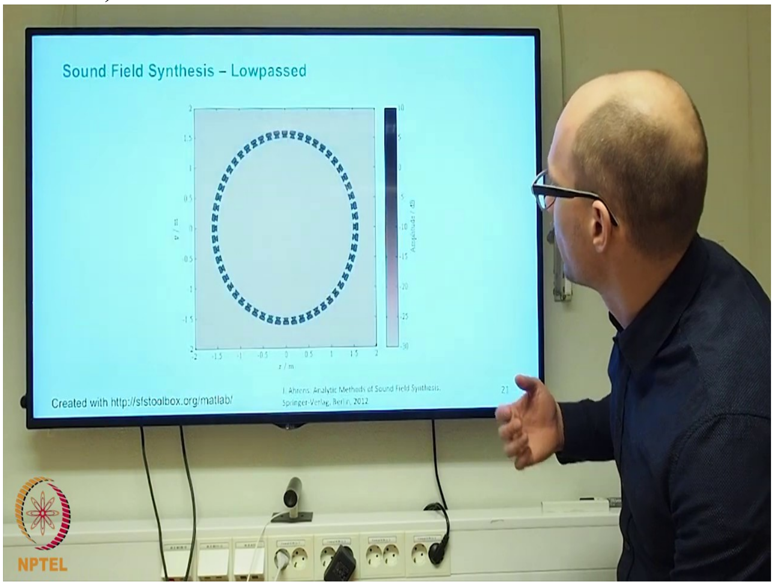


discrete loudspeakers at certain spacing. And which has the consequence that this sound field synthesis is only correct below a certain frequency which is called the spatial aliasing frequency.

So this array that you see here is the one I worked with when I was at the University of Technology at Berlin. It is composed of 56 loudspeakers arranged on a circle with the diameter of 3 meters.

And you see over time,

(Refer Slide Time: 02:48)

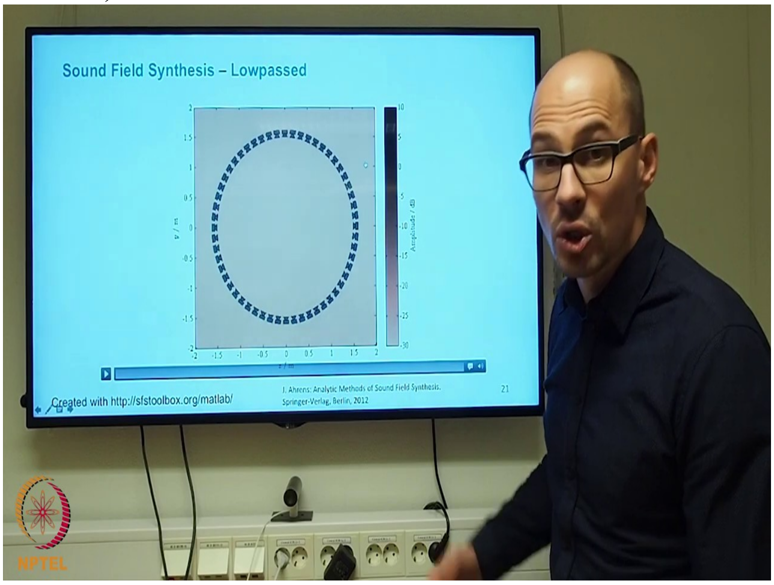


the same loudspeaker this time in a simulation. So let us compute and let us solve all these required integral equations that describe the, the situation.

And to compute the loudspeaker signals that would be required to produce a straight wavefront that propagates upwards in the plot.

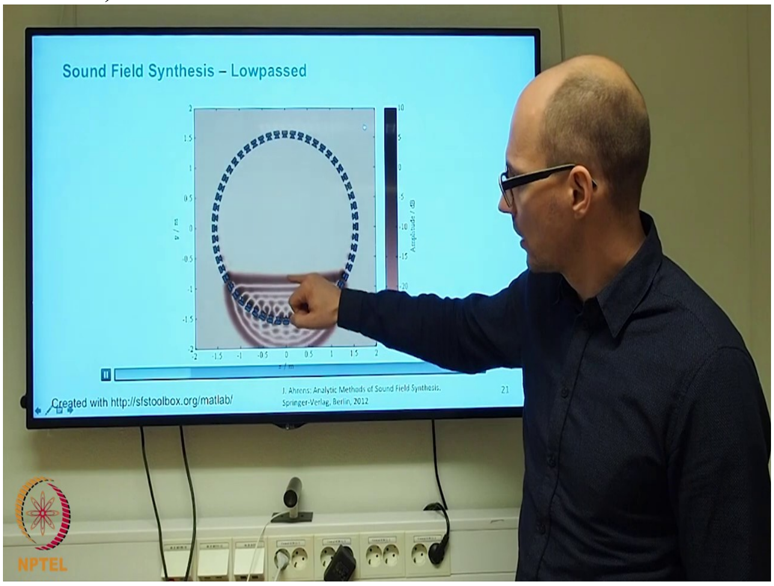
If we are limiting the frequency content of that plane wavefront to

(Refer Slide Time: 03:14)



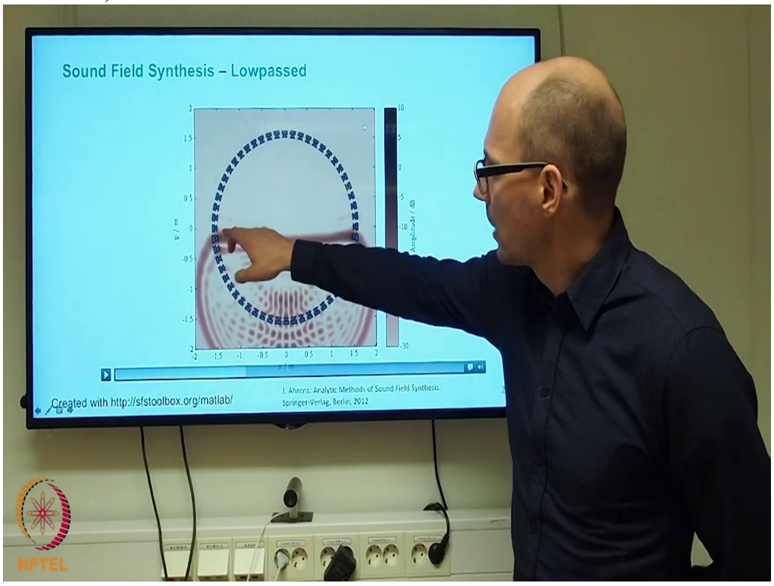
frequencies below the areas in frequency which is somewhere around 1700 Hertz in this case then it looks as follows.

(Refer Slide Time: 03:26)



This is where the wave propagates. Let us stop here. We can see that

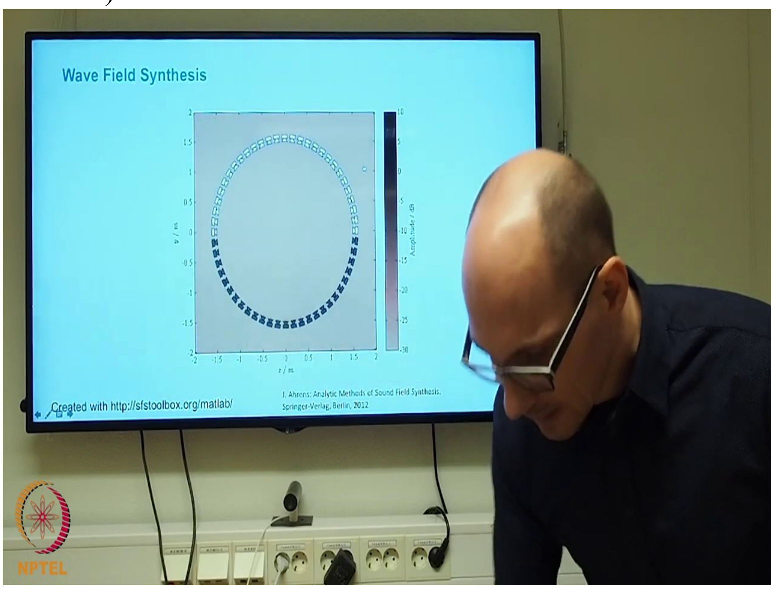
(Refer Slide Time: 03:30)



the wavefront indeed is very straight. There is a little bit of stuff behind the wavefront but this is not an artifact of the method. This is the consequence of the fact that we are using only lower frequencies or that we allow the waves to carry only lower frequencies and so please ignore this.

We can see a plane wavefront, a straight wavefront that does indeed propagate upwards in the plot. So that looks very successful but

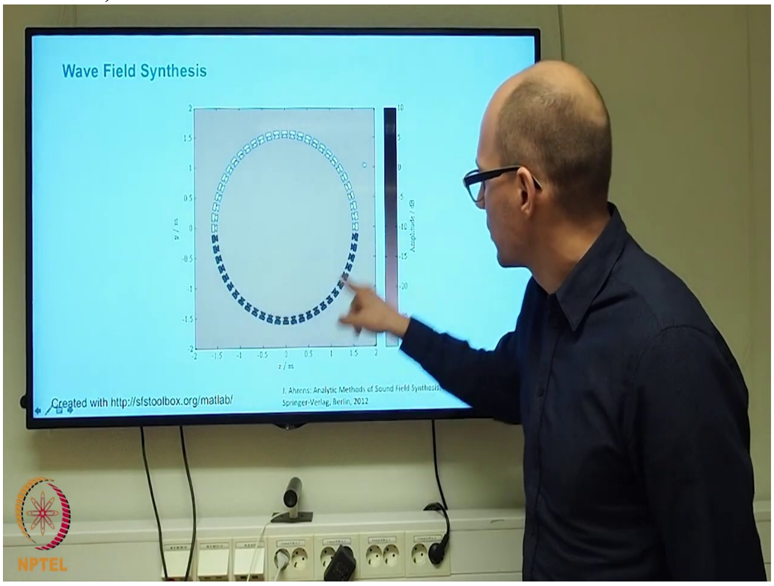
(Refer Slide Time: 03:58)



if we are allowing also higher frequencies which we will need for high quality audio reproduction we need frequencies up to 16, 17, 18 kilo Hertz or even higher.

If we are using one particular solution which is called wave field synthesis and which happens to use only the lower

(Refer Slide Time: 04:18)



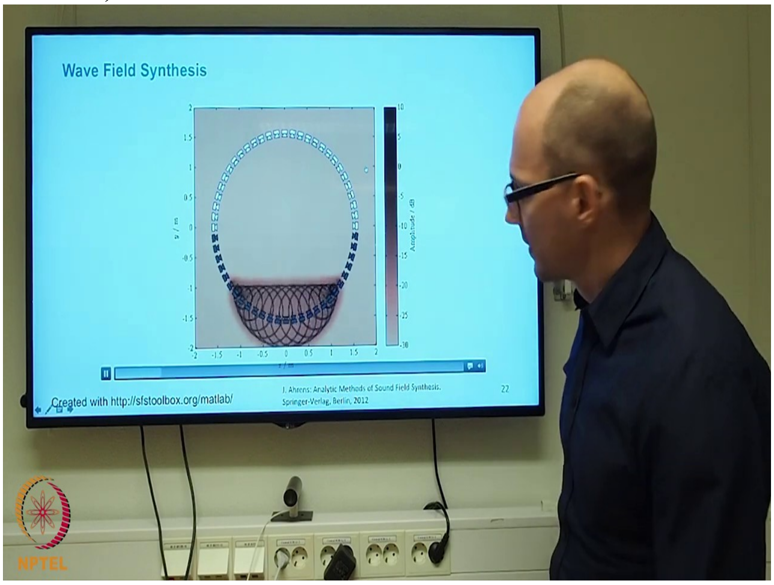
loudspeakers in order to produce the wave that propagates upwards, so

(Refer Slide Time: 04:22)

white filling indicates that these loudspeakers are not active in this particular scenario.

And now we allow the entire audible frequency range in terms of signal content, and again we drive the system

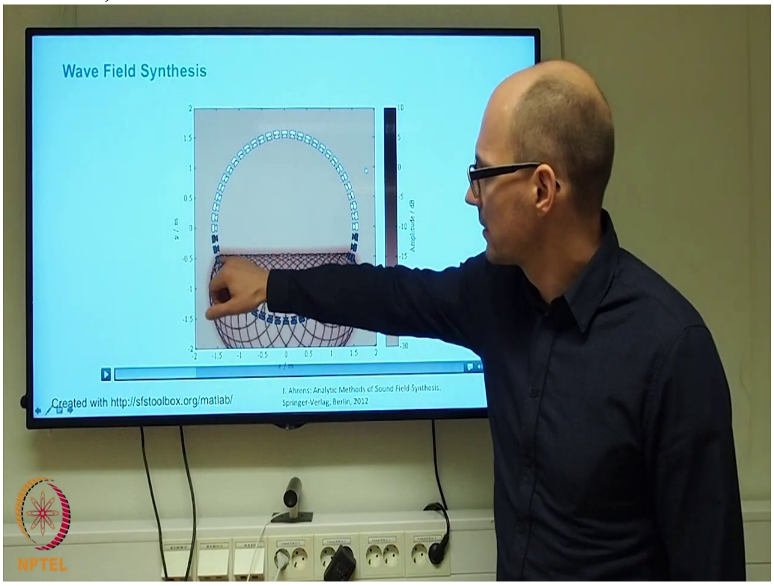
(Refer Slide Time: 04:36)



to produce a straight wave, plane wave that propagates upwards. You can see that.

Yes,

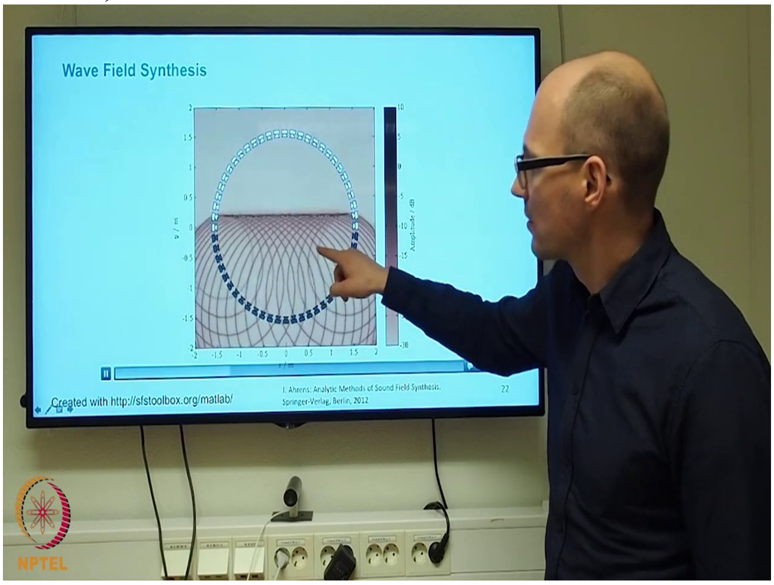
(Refer Slide Time: 04:41)



we have again a very straight wavefront. It is much sharper than in the previous simulation, that is because simply we also have high frequencies and which allow for this wavefront to be more pronounced.

But what we see is that behind the desired wavefront we have additional wavefronts

(Refer Slide Time: 05:02)



that propagate in different directions you can see clearly. And these are called spatial aliasing.

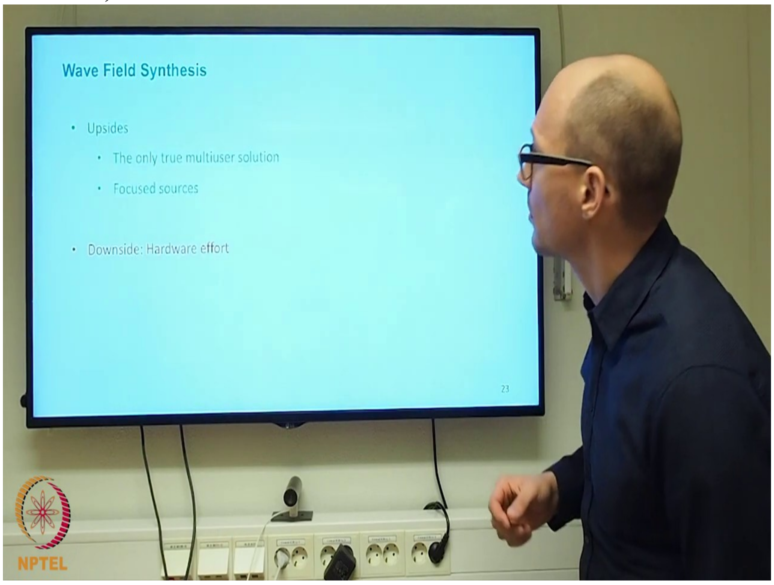
And they are the consequence of the fact that we are not using a continuous layer of loudspeakers but a very discrete one with a finite number of loudspeakers.

And there is unfortunately nothing we can do about this. But interestingly when doing this in practice and listening to this, it turns out that this impairment due to spatial aliasing, the perceptual impairment meaning how much worse the sound is, is much lower than we would think.

So these systems they can sound really excellent when designed and driven properly and typically these systems that use spacing between the loudspeakers which is in the order of 10 to 15 centimeters and that causes spatial aliasing frequency somewhere between 1500 and 100 Hertz and 1500 Hertz and 2000 Hertz.

So

(Refer Slide Time: 06:10)



to summarize the upsides are it is multiuser and the downsides are it is incredible amount of hard work that you need. So if you want to equip a cinema with such a loudspeaker system that is only horizontal then you are already need something four or five hundred loudspeakers.

If you are thinking 3 D, if you want to enclose the cinema with a surface of loudspeakers with a 15 centimeters spacing, you will need thousands and thousands and thousands.

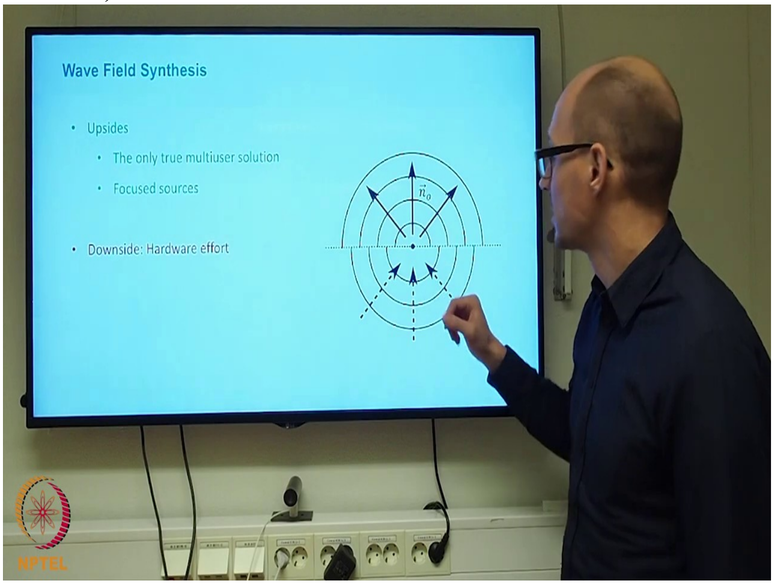
The largest system that exists nowadays in the world is installed also at The University of Technology in Berlin and it uses 800 or even more than 800 independent channels but it is also horizontal only because otherwise the hardware, the required hard work would be prohibitive.

Wave field or sound field synthesis in general there is also other methods, besides wave field synthesis but this one is the most popular, the one and the only one that can be properly implemented in practice.

The upside of sound field synthesis is the ability to render what is called focus sources. These constitute the following.

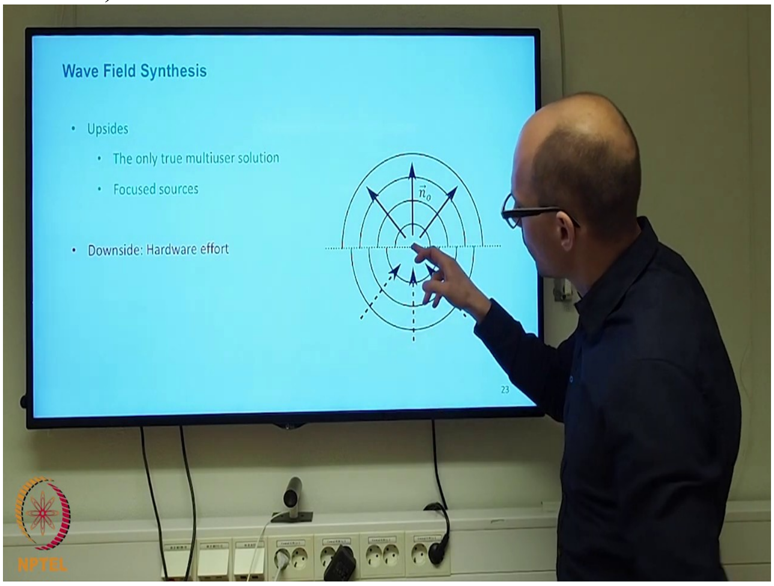
A focused source

(Refer Slide Time: 07:30)



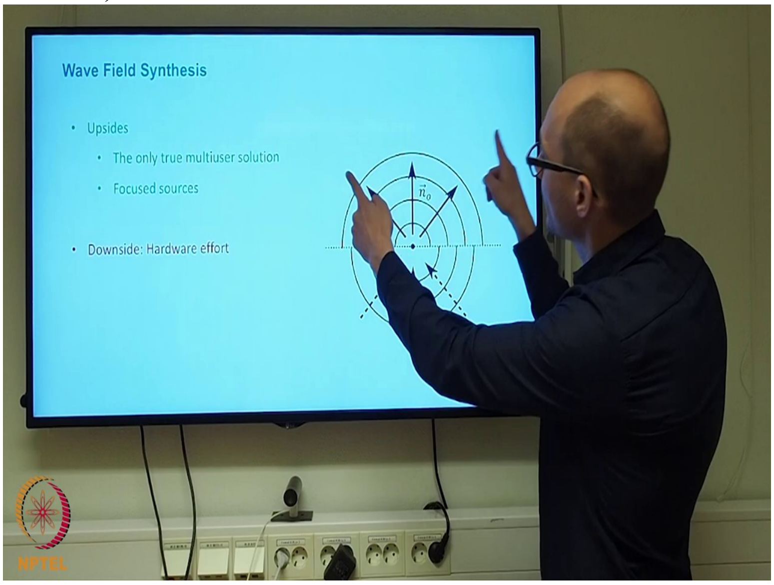
is not really a sound source, or not virtual sound source, it is a sound field that converges

(Refer Slide Time: 07:36)



towards a point, then it passes the point and the sound

(Refer Slide Time: 07:40)



field diverges on the other side of that focused point.

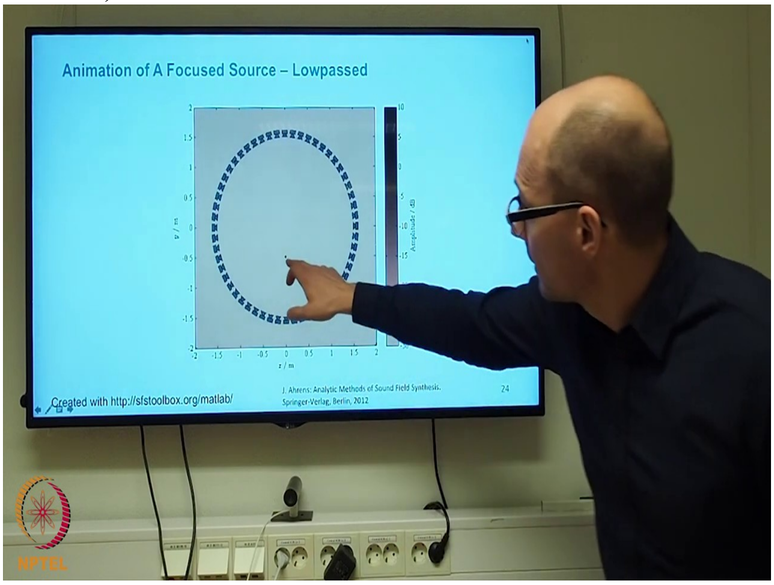
And the diverging part of the sound field resembles very much the wavefronts that would, that would be apparent if they were a sound source at the location of the focused point radiating sound.

So a user located in the diverging part of that sound field will, will localize a sound source at the location of the focus point, so in mid air in front of the loudspeakers. And that can be very impressive.

So you can, for example put rain drops in rain scene inside the audience area with other methods like stereophony or ambisonics or so, simple ambisonics then if you would, you would be feeling like sitting in a dry bubble and then rain is taking place somewhere at the distance from the loudspeakers or even further, this is something that can be mitigated with a focus, focus source in sound field synthesis.

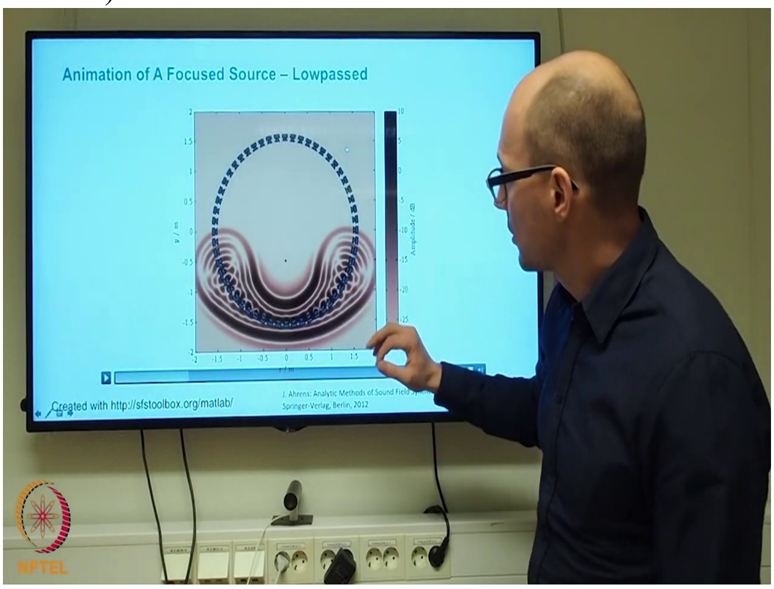
Let us also look at the time domain simulation of this. This dot

(Refer Slide Time: 08:44)



represents the focus point and now we are looking only at lower frequencies. So this is without spatial

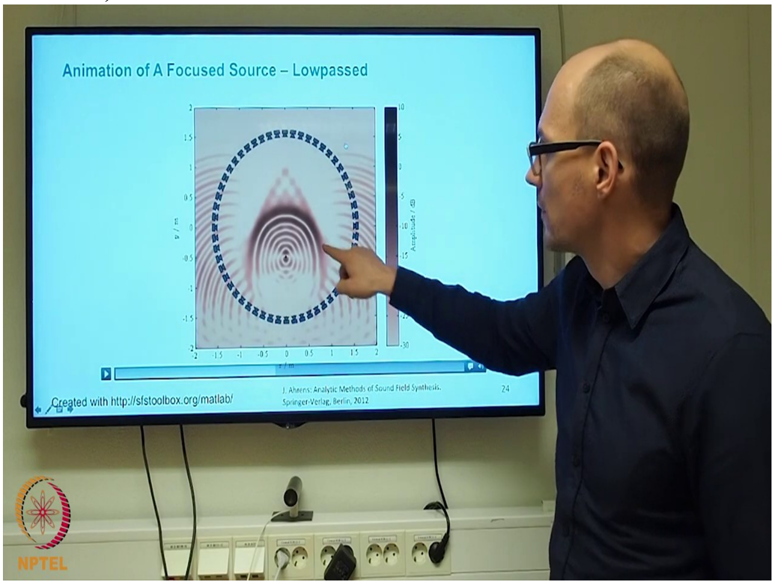
(Refer Slide Time: 08:53)



aliasing.

And we can see that indeed the sound, the wavefront converges to the focus point. It passes the focus point and now the diverging

(Refer Slide Time: 09:02)



part very much resembles the wavefronts that would be apparent if this focus point were radiating sound.

So clearly person located, excuse me, in the diverging part of the sound field we will localize a sound source in front of loudspeakers in mid air. A person located in the converging part of the sound field will localize whatever, it depends on the head orientation and many other situations so this will be irritating, and should be avoided.