Virtual Reality

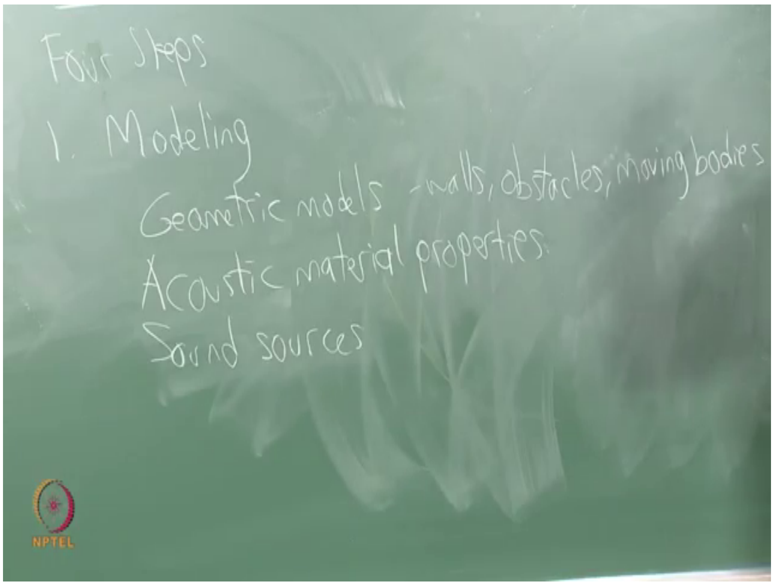
Prof. Steve Lavalle
Department of Multidisciplinary
Indian Institute of Technology, Madras

Lecture - 18
Audio (rendering)

Welcome back. I am continuing onward with audio for VR. So, we introduced that last lecture I talked about the human hearing system, and gave you some of the basic biology of how our hearing works also talked about propagation of sound waves, and made some comparisons to light and vision along the way. We talked about auditory localization and I finished up with that I talked about monaural cues and binaural cues.

And now I want to talk about the rendering part.

(Refer Slide Time: 00:52)



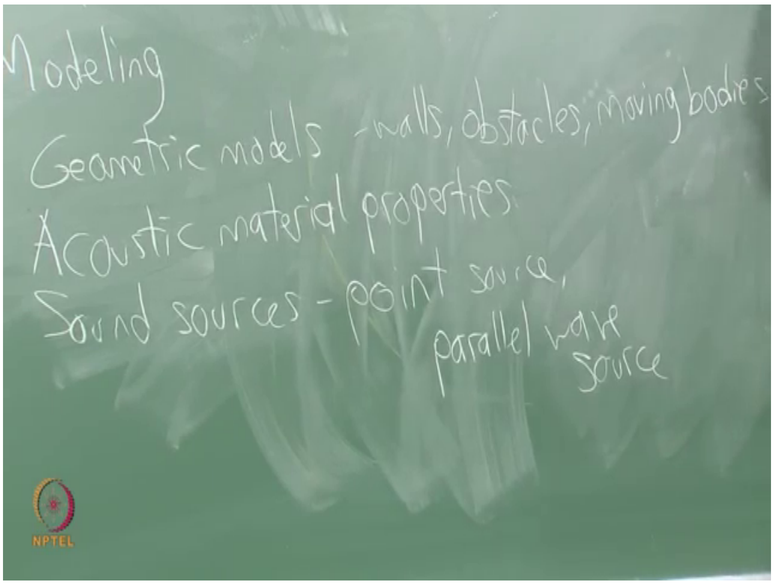
So, you have sound rendering, just like we had rendering for the visual part which we call computer graphics there is an audio part that has many similarities to the visual part. So, I will describe 4 steps and just give some comments about this as I go along. And it will be significantly more brief than it was for the visual case, so 4 steps, so one we have modeling.

So, for the visual case and even before we covered vision versus audio or anything we talked about geometric models right. So, we will again have geometric models, you can have stationary or movable models. You could have movable models you could have a moving object that generates sound or maybe sound will be bouncing off of a moving object.

So, walls, obstacles, moving bodies, perhaps you have a bird that is flying in VR and generating sounds. When we were modeling objects for computer graphics for rendering in that case we were concerned about the material properties in some cases right, we wanted to know how the light reflects off of the surface of our objects.

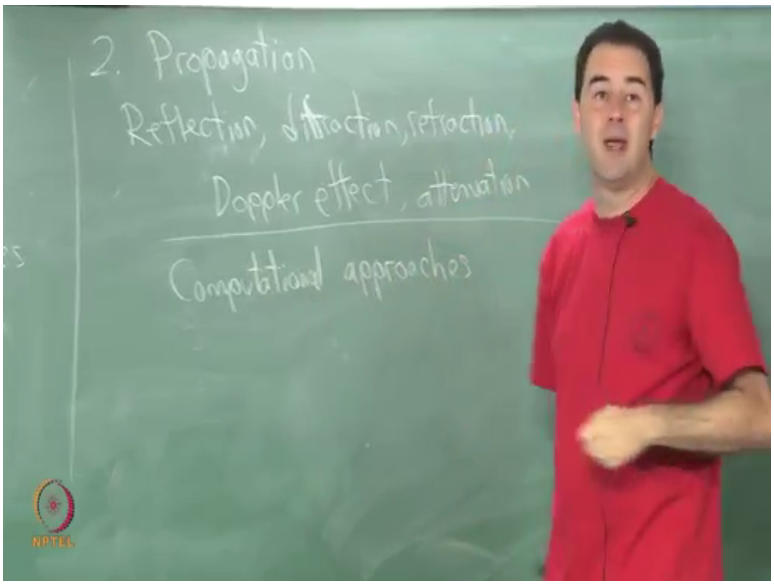
So, what do we want in this case we want to know how the sound reflects off or absorbs into the objects or maybe transmits through the objects or the diffraction that occurs. So, how does this happen? So, we have acoustic material properties, and just like we had light sources we have sound sources, when you can have a point sound source or you could have some kind of parallel wave source maybe an entire vibrating virtual wall that makes parallel waves. And so we could generate those directly or I mean in the case of a point sound source I guess we can make the same statement as we did for light that we never end up with truly parallel waves, right. So, we have these propagating spherical wave fronts, right. So, it could have point, source, or a parallel wave source, like a vibrating plate.

(Refer Slide Time: 03:34)



We can talk about the loudness of the source just like we talked about the brightness of a light source, all right. So, these are the modeling components that we have right just as we had for the visual case.

(Refer Slide Time: 04:10)

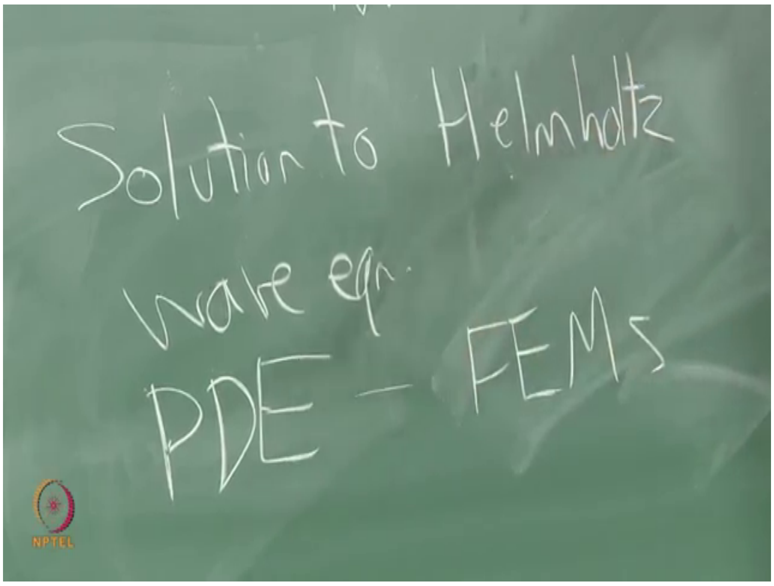


When we need to think about propagation through this virtual world that we are constructing, so using the acoustic material properties we want to figure out how is the sound going to propagate. And again we have all the things from before I will just write them here reflection, diffraction that should be two has there, refraction or transmission. If we are moving right if the avatar is moving through VR we may also get a Doppler effect right or if some object is moving towards us and making a sound.

So, that is something that we did not worry about that in the light case right. The red shift that occurs right in light is useful for tracking the motion of stars in the sky for example, right, but it does not end up being enough of a shift to be significant for VR, but here it ends up being important. So, the Doppler effect and overall amount of attenuation. So, waves get weaker as they propagate through the air at a more significant rate than light waves, so attenuation. So, these are all important types of propagation or factors that are involved in propagation.

And if we divide up computational approaches we want to think about doing a computational approach. Remember we are making an alternate world generator right. So, we talked about very early on in the course we are making a kind of simulator to propagate this audio that is being constructed in the virtual world. So, there are two big authories I would say at the highest level, we could do it in what I would call in a numerical approach versus a combinatorial approach.

(Refer Slide Time: 06:27)



In the combinatorial case we will use concepts like visibility or ray tracing that look familiar, right. So, this is the same kinds of concepts that we talked about and considered for graphical rendering for visual displays the same things can be used for audio rendering. So, we can talk about propagation of sound and we can look at rays of sound as the propagate and reverberate in an environment.

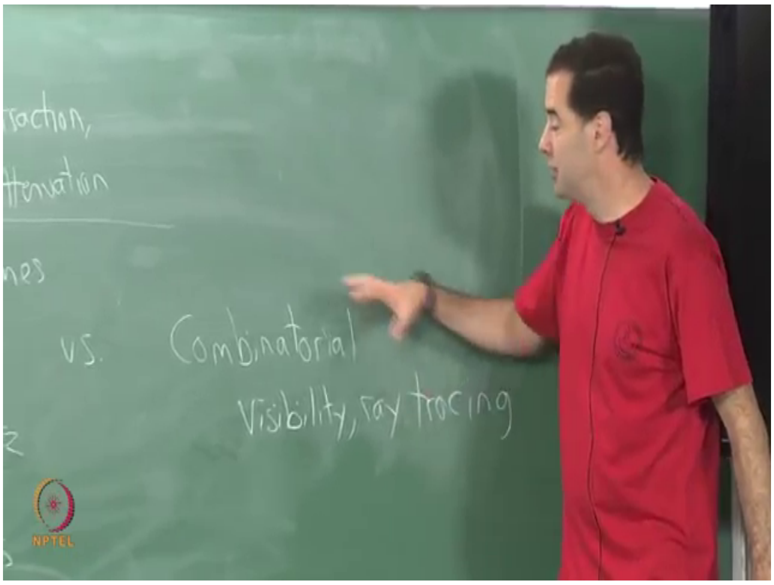
So, that is one possibility and you know this is fast and approximate in some kind of way and the question is going to be is it good enough. When we do this in computer graphics remember we looked at I am different shading models and we talked about how they were not physically realistic, but they seemed good enough because they look fine on a screen. Sometimes they look fine in VR sometimes they do not and we had these difficulties to deal with, right. People making computational simplifications by doing visibility based methods keeping them very simple not considering all of the propagations, simplifying the materials as much as possible making things like the Blinn Phong shading model which is not physically accurate, but is very fast computationally.

So, the same kinds of choices will exist on the audio side, but it may be even more difficult to tell whether or not you have done it right, right you may hear the sounds, but can you correctly localize them in the same way as you would in the real world. Very difficult to determine that you have to do more difficult kinds of experiments with human subjects determine whether or not you have got it right. Whereas, in the case of visual maybe you can just look at it and say that looks fine all right enough people look at it seems to be fine looks physically realistic enough and people are happy with it. So, you can go down this path, but again it becomes difficult to verify that these approximations are correct.

You can go the numerical way which is right out partial differential equations you can calculate what is called the solution to the Helmholtz wave equation which I will not write here. I will not make this a class on partial differential equations, but in other words you are doing numerical PDE calculations, using things like a finite element methods and so forth.

So, you end up in the regime of partial differential equations, you end up with lattice or grid representations over the entire three dimensional space, and you are doing some kind of numerical wave front calculations. So, the numerical computations propagate through the space. You can imagine that this is very complex. This is equivalent in terms of difficulty to the full global illumination problem that we had for lights right. If you want to do very real physically realistic rendering then you again want to resort back to numerical computations and maybe even do all of the pde calculations and people do that in computer graphics sometimes to make very highly photorealistic renderings.

(Refer Slide Time: 09:48)



So, you have the same kinds of choices for audio you can go this direction or you can try to find some kind of combinatorial quick visibility ray tracing kinds of hacks let us say that are hopefully good enough. So, these choices exist as well, this one being very expensive, but very physically accurate this one hopefully being good enough.

I think there is a lot of research and a lot of development to be done in this side here to make fast and efficient methods just as was done in computer graphics across the 1970s and 80s when people had very limited hardware and they still nevertheless generated. Very nice looking pictures and then eventually videos synthetically

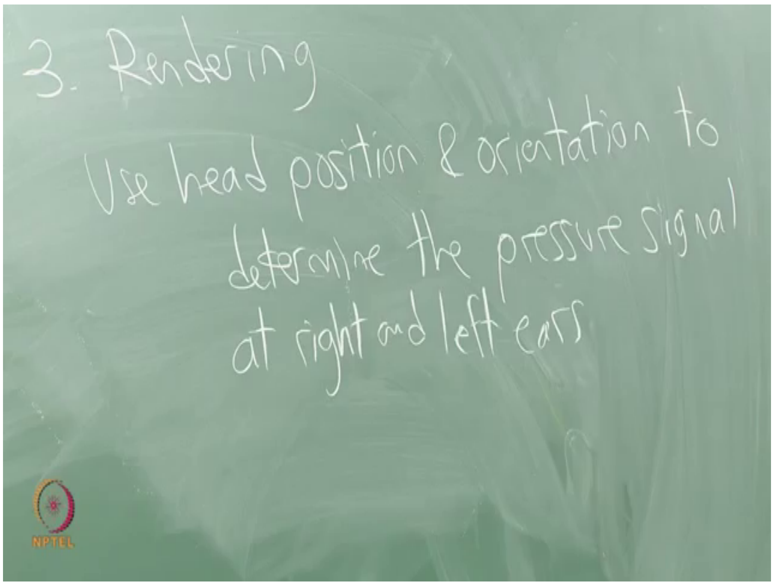
So, we are in the same kind of place hear with audio its known how much of these models and the acoustic material properties need to be fully specified, how much of this information is going to be critical for a good audio experience. You will hear something, but will it be reproduced in the in the same kind of way as it would in the physical world giving the abilities to localize the source for example, as you can in the real world. And that could be very important for the type of experience at your building. As I said in the beginning it depends on the task right what is the experience you are trying to make so.

If you are trying to make a horror experience way maybe some kind of scary experience perhaps where the sounds are coming from in some kind of not very well lit environment might be an important aspect of that right and so you might want to capture that very well. If you are just making an experience where you are just talking with your friends, maybe some amount of localization is important, but maybe not a large amount right.

So, depends on the application, but if you want to get these things right you may have to go further down this paths unless we develop unless the community develop some very good very fast methods here for generating synthetic sounds and propagating them in a way that is convincing.

Questions about that. So, that is one step is modeling another step is propagation. I want to talk about two more steps.

(Refer Slide Time: 11:42)



So, 3: Rendering, all right. Well, one of the things that we have to take into account for rendering and rendering is going to be to determine what, output needs to be given to the audio display which we may call speakers, right; so what needs to be presented there. Or one thing we have to do is we have to use the head position and orientation to determine the appropriate let us say air pressure signal at the right and left ears. Just like we have to give the right visual information to your virtual right and left eyes in VR you have to figure out where the wave should be hitting your virtual ears in right and left ears in virtual reality.

So that means, if you are moving your head around this has to be tracked and you have to adjust the sound that goes into your ears based on that. So, what goes into the audio display or the speaker should be dependent on your head position and orientation right. And if you grab on to a controller and move your character that will also change your position orientation in the virtual world, even though it is not changing in the physical world. The audio should be adjusted accordingly as well if you want to convince your brain that you are in fact, moving which may even further contribute to this vestibular mismatch or maybe it helps that problem. Maybe it overwhelms when you put audio and visual together helps overwhelm the vestibular signal. I do not know which one will happen you will have to do experiments to see.