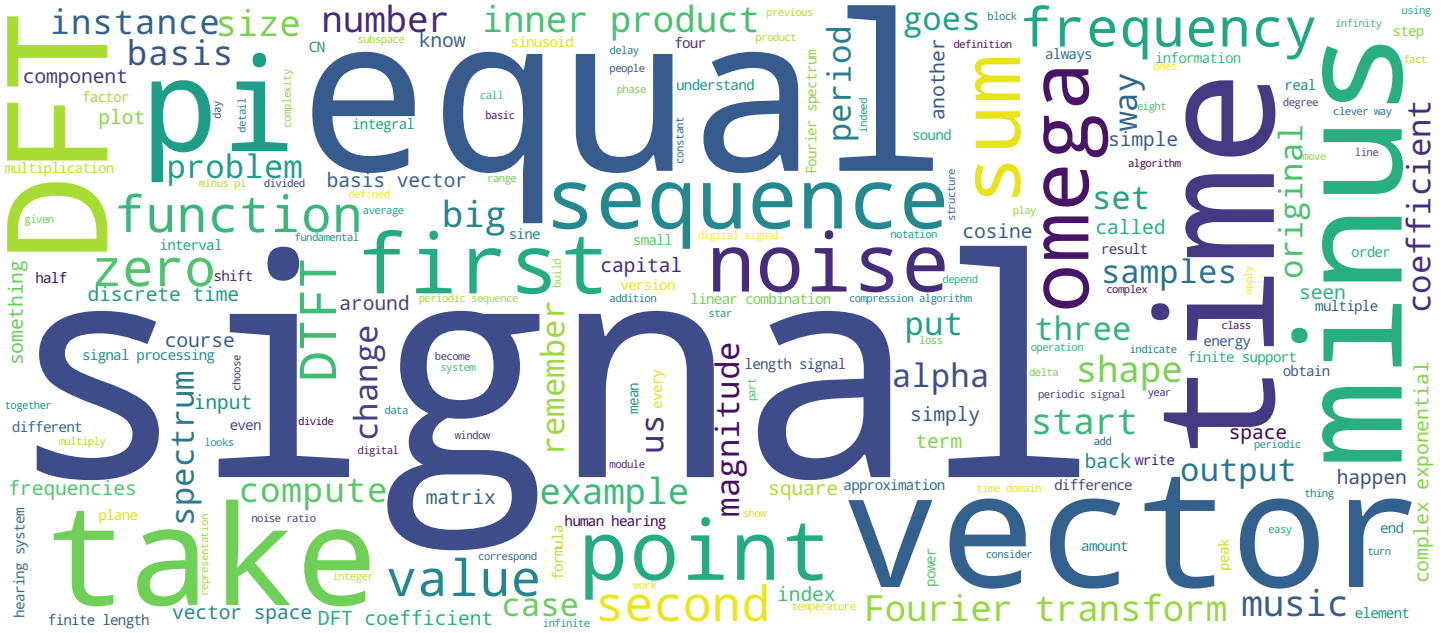
MP3 Compression trick

- How can an MP3 song sound so well, while being so compressed?
- Compression introduces noise.
- The trick is to shape the noise!

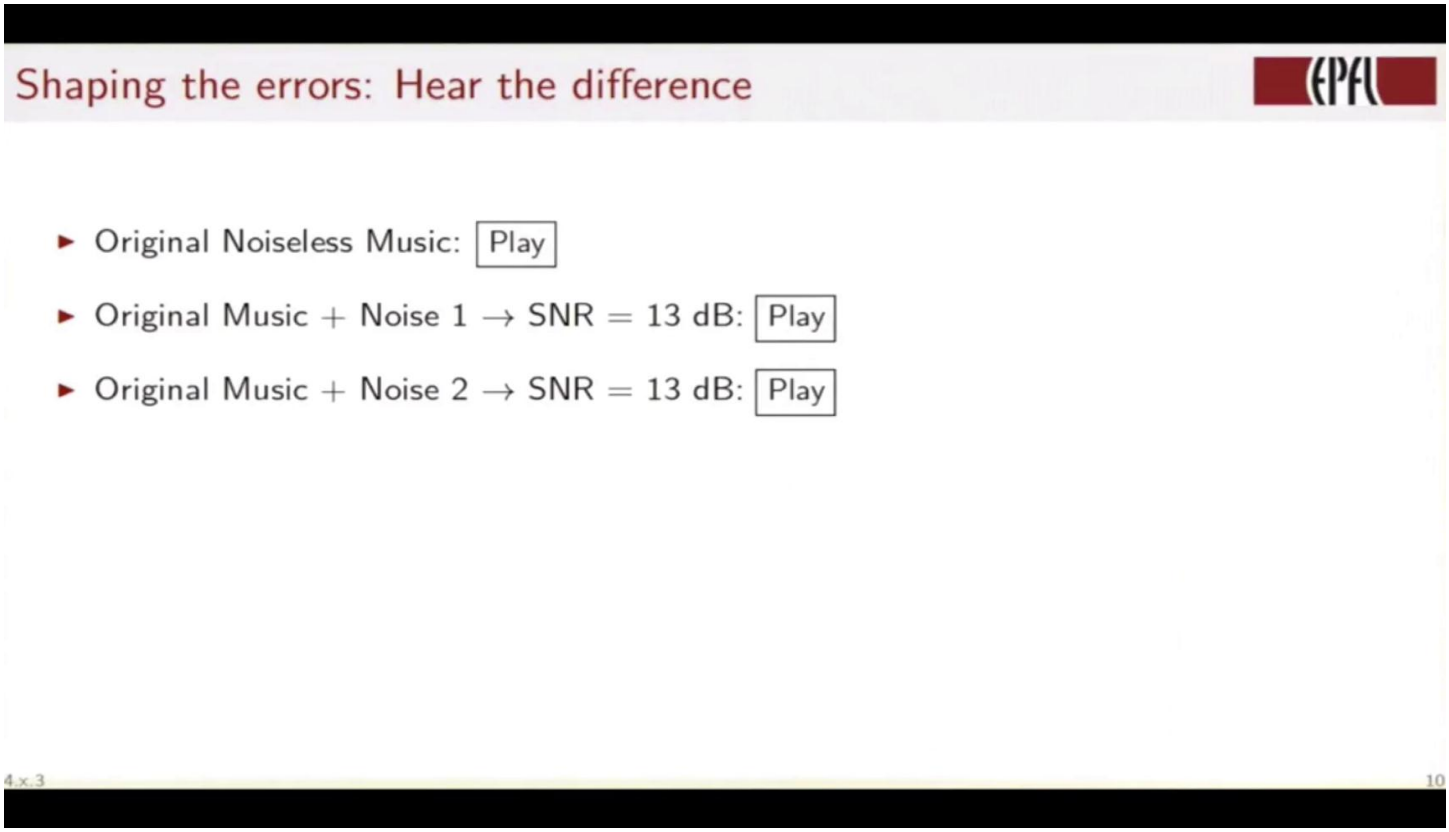
4 × 3



Search MOOC

Video

EPFL

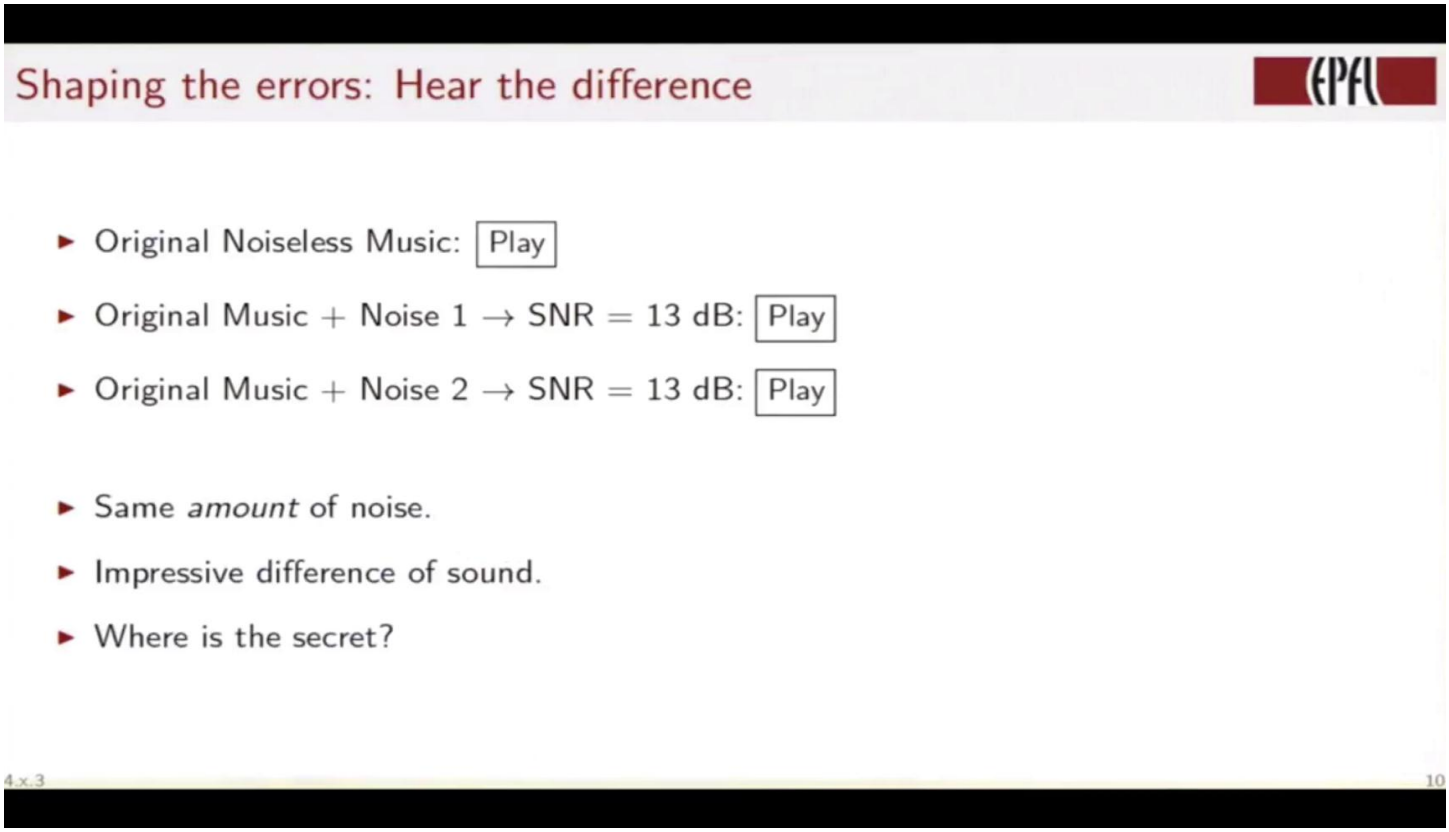


MP3 compressed music sounds pretty good, however, it has been compressed highly from its original format by a factor of 10, 20 or more. And therefore, some information has been lost, but the human hearing system has been fooled not to notice this loss. So compression, which will be discussed later in the class, does introduce some noise, some approximation error. And the trick is to shape the noise so that it's not picked out by the human hearing system. We are going to start by simply listening to music and two versions of adding noise. One which is a not so clever way to put noise, and the second one is a clever way, which is using MP3. Both noise levels are of the same magnitude, so they have the same energy of noise that is being added. So the signal to noise ratio, which is the ratio between the useful signal and the noise, is in both cases 13 dB. So let's start and play the original noiseless music. Now we play the original music plus what we call noise number one, which is simply white noise and has a 13 dB signal to noise ratio. I guess everybody would agree that this is pretty bad. It's pretty obvious that this is a noisy signal. Now we are going to add exactly the same amount of noise.

Notes

Summary

0m 00s



So SNR is also 13 dB. However, we shape the noise in a clever fashion. And it's probably pretty difficult to distinguish the original from this version, which is the MP3 version. So we have the same amount of noise. There is an impressive difference. Let's understand this secret.

Notes

Summary

1m 37s

And the secret, as you may guess, is found in the Fourier domain. So we plot the DFT spectrum of the music again on the log scale. So that's the top graph. And you see there is some spectral decay. And it has a very particular shape, which depends on this particular piece of music. Now the first noise, which is 13 dB signal to noise ratio noise, has an essentially flat spectrum across all the spectrum. Whereas the second noise, the clever noise, as you can see, has a shape that resembles the shape of the music. And so when the music has a lot of Fourier energy, we can add some noise, where the music has very little energy, we put very little noise. So that's called noise shaping.

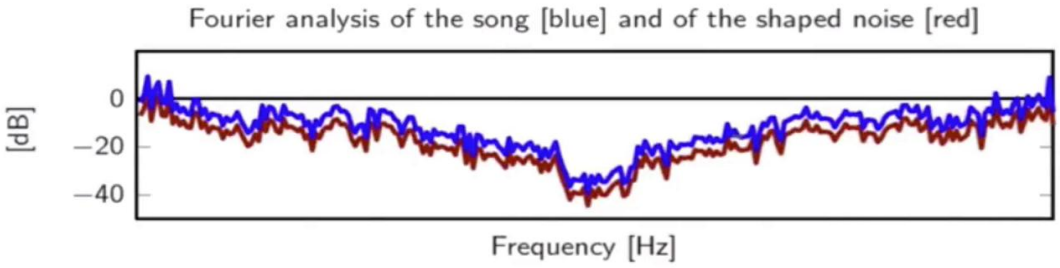
Notes

Summary

2m 05s

Conclusions

- MP3: complex compression algorithm that introduces errors.
- Errors shaped as the song in the Fourier domain → higher perceived quality.
- MP3 minimizes the perceived quality decay by shaping the compression errors.



4x:3 12

So we summarize this in this graph where the blue curve is the spectrum of the music. The red curve is a spectrum of the noise, exactly adapted to the music. And the compression algorithm used in MP3 is, of course, very complex. But one of the key elements is to shape the errors adaptively as the Fourier spectrum changes. Because, of course, music evolves over time. So the Fourier spectrum does not always look like this. It changes all the time. So you have to constantly adapt where you put the noise as the Fourier spectrum of the music changes. And so MP3 magically minimizes the loss of quality by shaping the compression errors in a clever manner. And this is done in something that is called the short-time Fourier transform domain, which we don't explain now in detail. But the idea is essentially a DFT that is local. So over small pieces of music you take a discrete Fourier transform, you look at the spectrum of the music and you put the noise exactly where it's not going to be heard. This is a magic method, and this is what has made MP3 such a successful music compression algorithm. By the way, this is called perceptual compression.

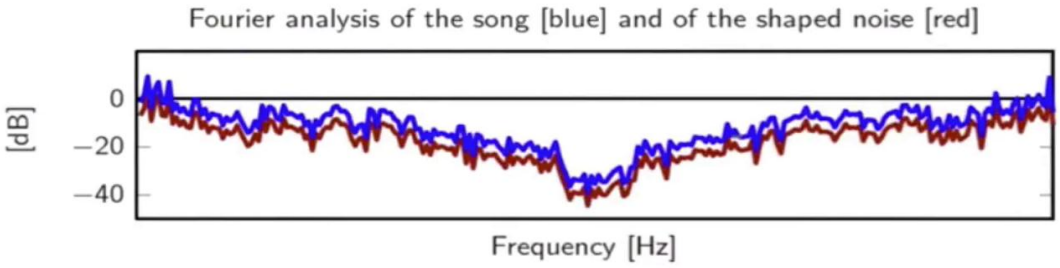
Notes

Summary

2m 56s

Conclusions

- MP3: complex compression algorithm that introduces errors.
- Errors shaped as the song in the Fourier domain → higher perceived quality.
- MP3 minimizes the perceived quality decay by shaping the compression errors.



4 x 3 12

And this is an art because you have to understand not only the mathematics, but all the perceptual effects of the human hearing system. And the people that derived the MP3 compression standard, and that's a lot of people are involved in this, have actually brought together all the qualities from mathematical analysis to human perceptual knowledge to derive a very powerful algorithm. By the way, MP3 is not the only audio compression algorithm out there, there other ones. But all the successful ones will have noise shaping in the same style as what we have just seen for MP3.

Notes

Summary

4m 12s