Biomedical Ultrasound: Fundamentals of Imaging and Micromachined Transducers

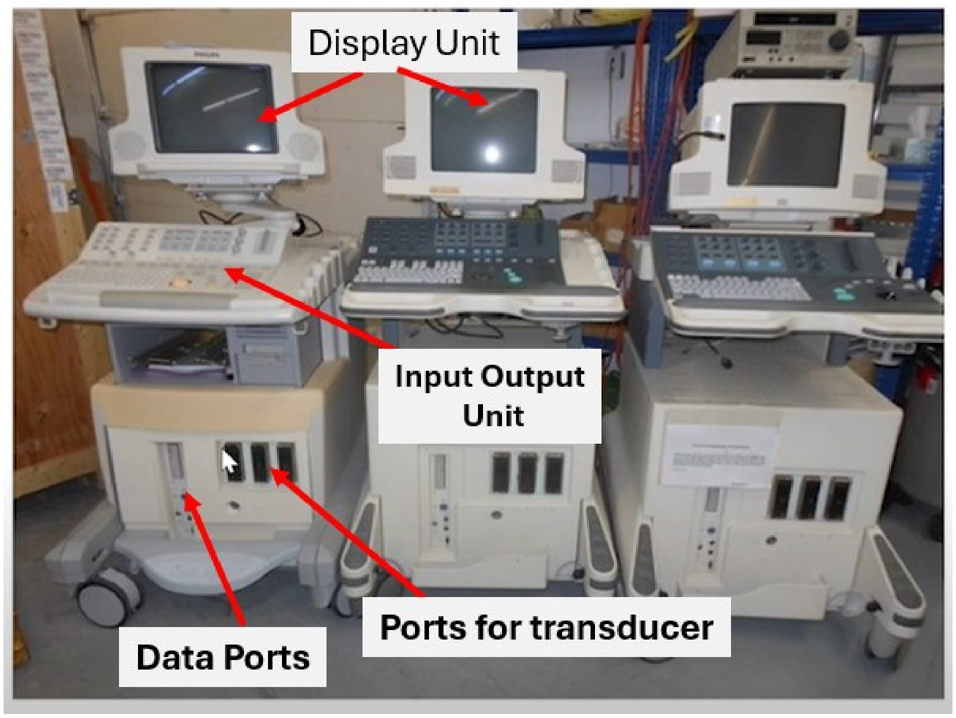
Prof. Karla P. Mercado-Shekhar, Prof. Himanshu Shekhar, Prof. Hardik Jeetendra Pandya

IIT Gandhinagar, IISc Bangalore

Lecture: 09

Ultrasound imaging systems

Hello and welcome to the lecture on ultrasound imaging systems. So as you will recall, ultrasound is a non-invasive, relatively inexpensive and non-ionizing medical imaging modality. Here you can take a look at some ultrasound imaging systems. These are slightly old fashioned systems, but the general components of the system remain the same. There's a display unit, input output unit and ports from which you can collect data. These are also, connectors to which you connect ultrasound transducers for imaging.



Recently, there's been a lot of development in portable ultrasound systems. You also have cell phones based mobile phone based systems. Here's an example of a ruggedized ultrasound imaging system, which is fairly portable tablet size. And, a laptop based Sonosight imaging system. Here in this picture, you can see that ultrasound is being performed in space. So as it turns out, ultrasound being portable is ideally suited for looking at the health of our astronauts. And the astronaut is being scanned using this portable laptop-based ultrasound unit. The systems are getting miniaturized.

Sistema d’Emergèncias Mèdiques (SEM), CC BY-SA IGO 3.0, CC BY-SA 3.0 IGO <https://creativecommons.org/licenses/by-sa/3.0/igo/deed.en>, via Wikimedia Commons

Ashashyou, CC BY-SA 3.0 <https://creativecommons.org/licenses/by-sa/3.0>, via Wikimedia Commons

https://upload.wikimedia.org/wikipedia/commons/e/ec/ISS-36_Luca_Parmitano_and_Chris_Cassidy_in_the_Columbus_lab_2.jpg

Typically, there is a relationship between the size of the system and the power of the system and the image quality that you can produce with the system. However, as the electronics is getting miniaturized, the systems are getting more and more powerful.

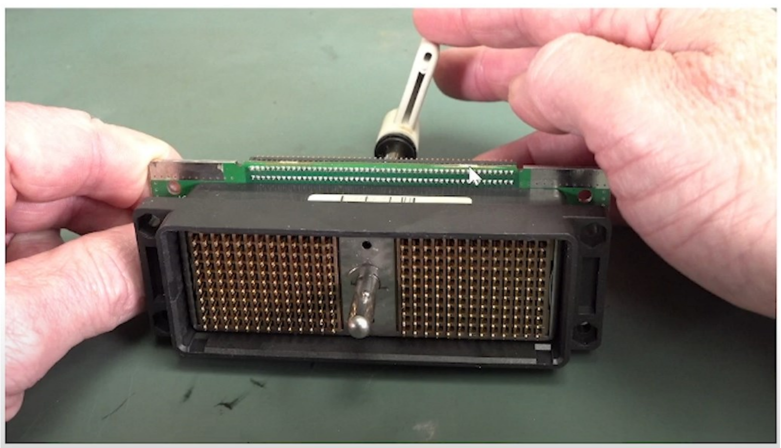
This is a typical ultrasound transducer. From its aperture, you can see that it's a curvilinear array. Typically these transducers have an operational bandwidth of 50 to 100%.

This is a fractional bandwidth. So the fractional bandwidth refers to the full width half maximum bandwidth or the 3 dB bandwidth divided by the center frequency.

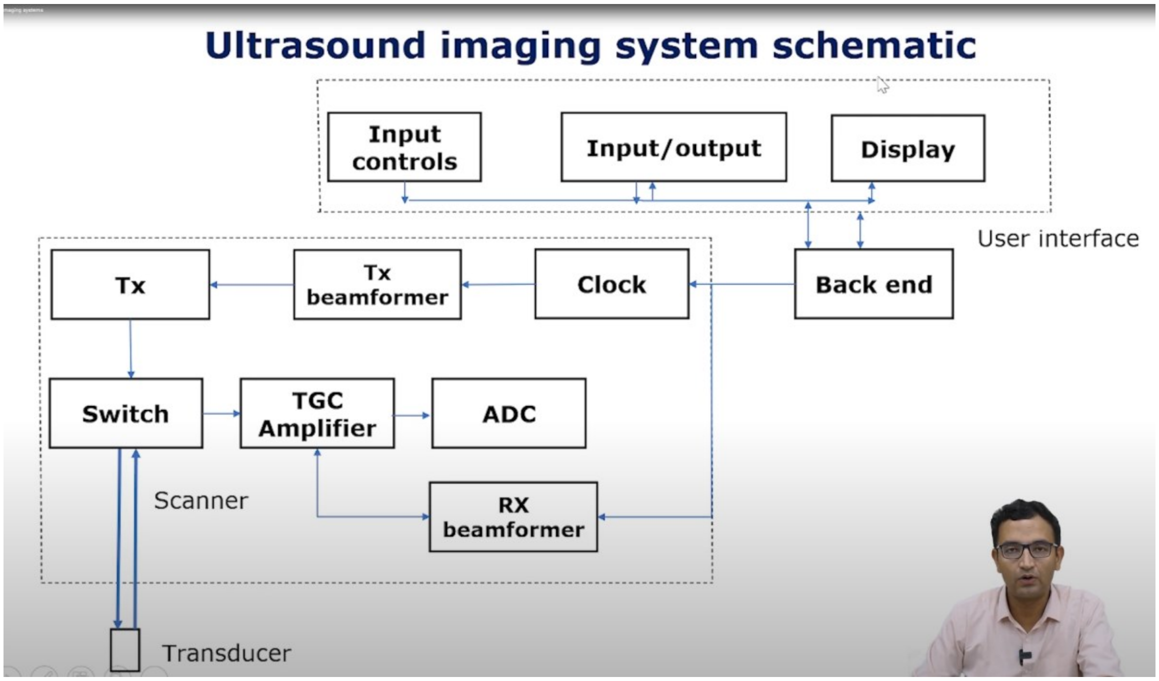
$$fractional\ bandwidth = \frac{3dB\ bandwidth}{center\ frequency}$$

So say if the center frequency is 4 MHz and the bandwidth is 2 MHz, then the fractional bandwidth would be 50%. Now, typically the name of the transducer has some information about its bandwidth. Like for example, consider the L11-5 transducer. This tells us that the operating frequency is between 11 MHz and 5 MHz and the center frequency will be somewhere midway between these points.

The transducer will be connected to the system and the connector is shown here. As you can see, there are a large number of connecting points. This is because the transducer has a large number of array elements. So this can be seamlessly connected to the transducer system and locked using the lock behind.



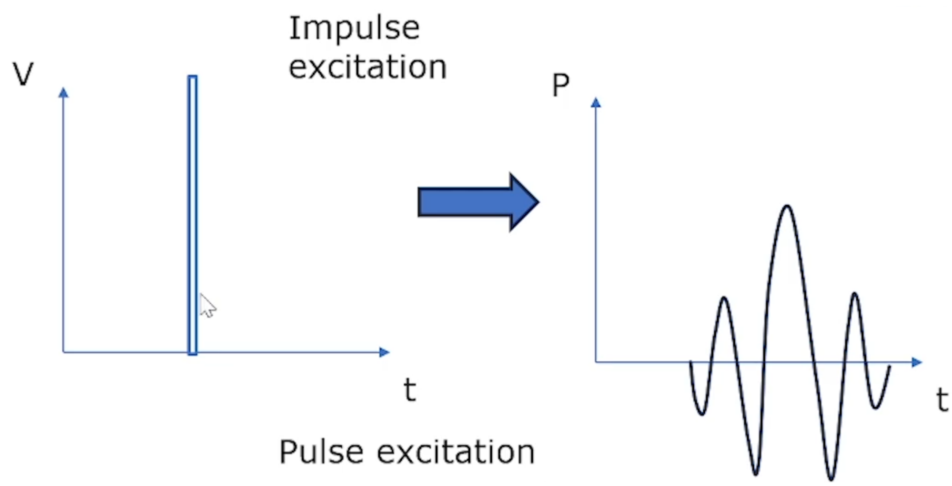
Here I'm showing a schematic of an ultrasound imaging system. There is a user interface which has input-output controls. There's a display. Then you can think of the main scanner unit with some important components such as the clock, the transmit beam former, the transmission system, the switch, the transducer itself, the time gain compensation amplifier, the analog to digital converter, and the receive beamformer.



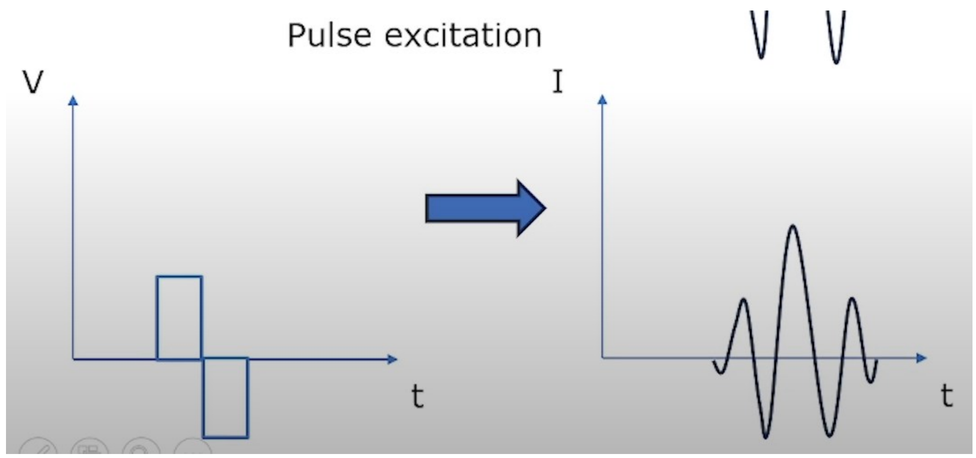
And then there's a backend for processing and display. So we'll just take a very brief overview of these different components of the ultrasound system. So when it comes to the user interface, it can be configured by the user. You have a system display with menus and control. You can pull up certain data, you can add information about the scans or the patient details and other details about the scan parameters. Now, typically there's a microprocessor or a computer which directs the operations of this ultrasound scanner and the excitation pulses which are sent to the transducer. These high voltage pulses are typically controlled by a microprocessor. Now, in terms of the receiver, the received signals go through a time gain compensation amplifier. We'll be discussing in detail on what this time gain compensation amplifier is.

The scanner unit contains beamforming and signal processing unit, filters for improving the signal to noise ratio, the demodulation unit components and specialized electronics for Doppler processing. All these will be discussed in more detail in future lectures. Then the backend of the scanner contains units for scan conversion of the data. For example, you may need a conversion from polar to Cartesian coordinates. Then there is color mapping, which assigns the color map and the dynamic range of the image.

And of course, there might be post-processing based on signal processing approaches to improve the sharpness or the quality of the image. So all these are contained in the backend of the ultrasound system. So typically the transmitter contains hardware for exciting the transducer. And we can generate an impulse for generating the signal from the transducer. So when the transducer is excited by an impulse, it will actually result in the impulse response, which can typically be modeled as a Gaussian modulated sinusoid. So this is a typical pulse.

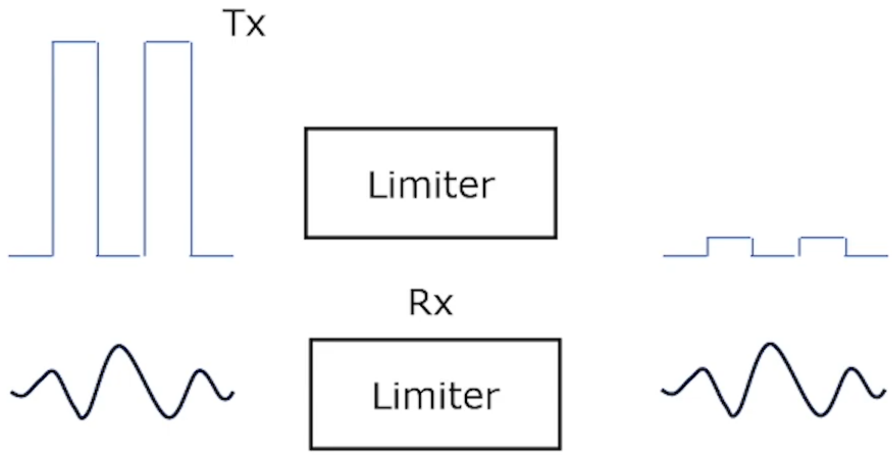


Now it turns out rather than impulses, generally we use short pulses or we have pulses which have different phases. So this is a biphasic pulse here. So you can use these different kind of excitations to generate the transmitting pulse.



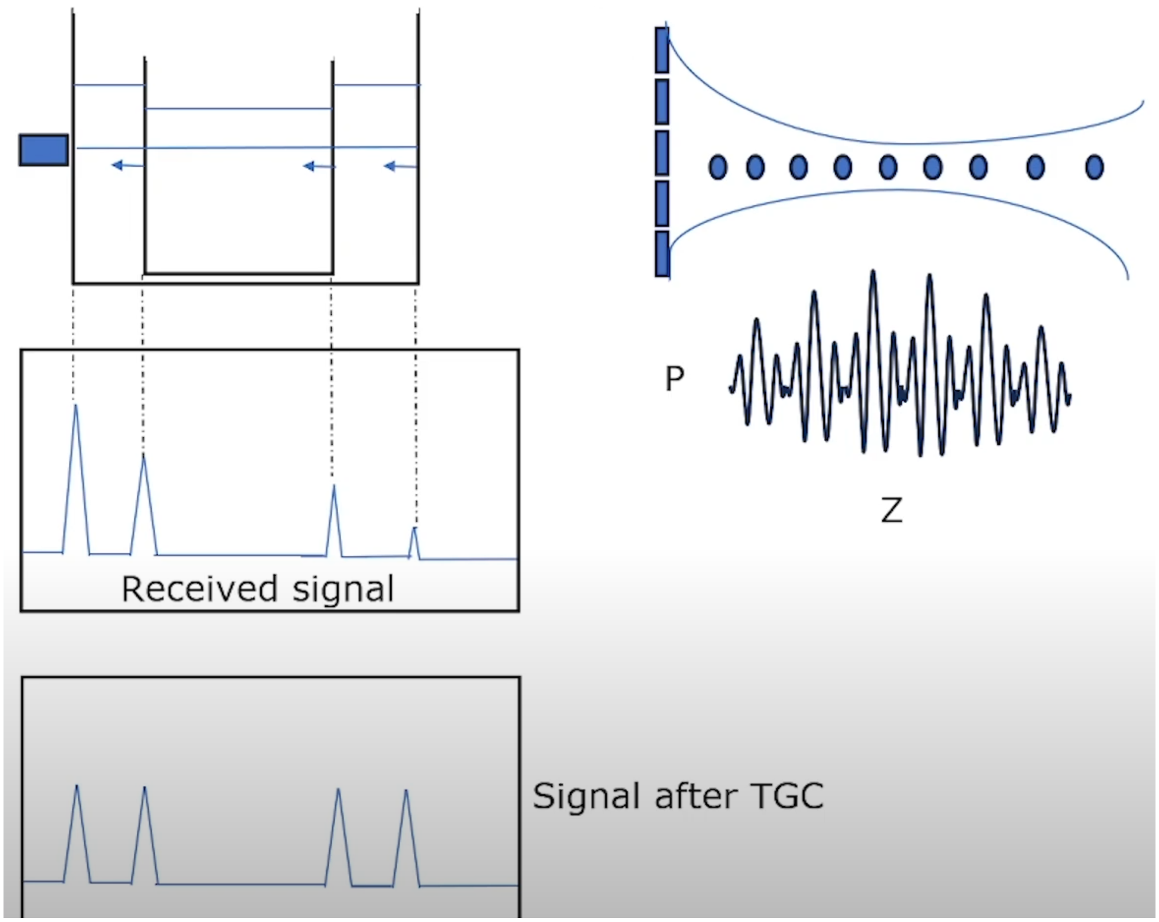
Now there's a hardware which is known as the expander and limiter. Generally, the same transducer is being used to send and receive the signal. When you send the signal, you send a high voltage to the transducer. And this high voltage can be something on the order of hundreds of volts. And when you are receiving the signal, this is a very weak scattering from tissue typically. You also know that only a very small percentage of the incident energy is actually scattered back. So the signal can be of millivolts. So now there's a huge difference between the transmitted signal as well as the received signal. So even if a small component of the transmitted signal leaks into the received signal you will completely spoil the SNR.

So there is a component called limiter, which essentially separates this transmit energy from the receive energy. It does not allow the transmit energy to appear in the received signal chain, but smaller signals, which are scattered signals, are not blocked by the limiter.



And then there is a unit called the expander. So if you typically send these two pulses to the transducer, the transducer will ring a lot. And the amount of ring down will depend on the quality factor of the transducer, which is a measure of how sharp the resonance is, what the backing material is, et cetera. But this ring down is not desirable. So, you do not want this ring down to reach the receiving component or the receiving signal chain. So, the expander blocks this ring down signal. Also the limiter prevents the high voltage signals produced by the transmitter from reaching the receiver as we just discussed.

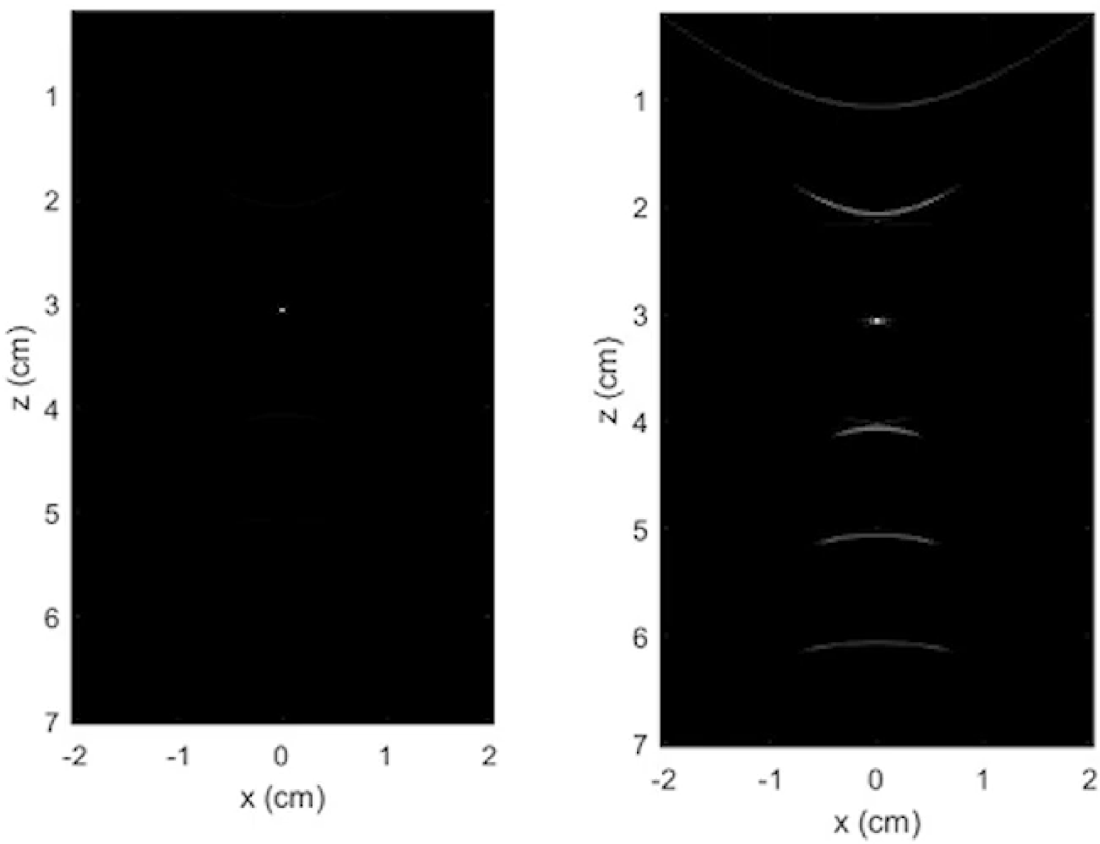
Now consider an ultrasound transducer here on the left side of the figure below, and there are a couple of interfaces in the medium and let's assume that these interfaces have uniform reflectivity that is, the reflections from each of these interfaces will be equally strong.



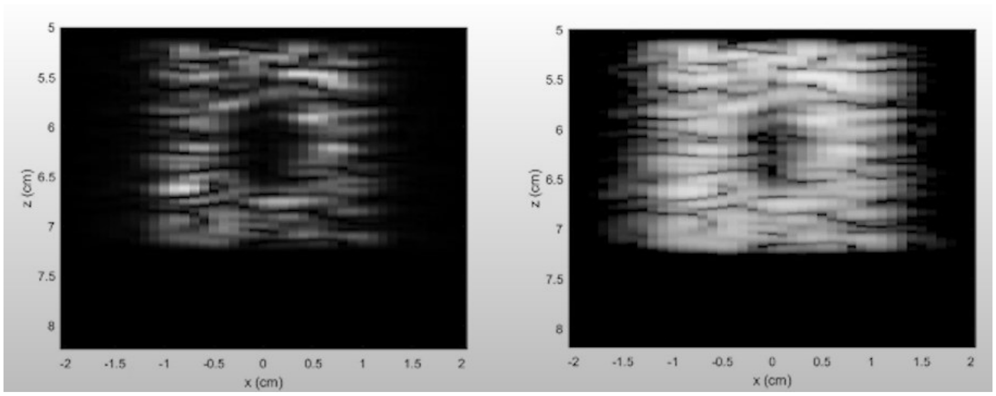
So in this case, what you would expect is that as you go deeper the reflections coming from deeper targets would be weaker because of attenuation. And not only that, there is a diffraction component associated with it. There is a beam shape and because of the beam shape, the beam has the narrowest width at the focus and then it diverges. So, as the beam diverges, there is a redistribution of energy. So, if you have a scatterer which is in the wider portion of the beam, the energy reflects back will be lower. So we would like to equalize for this because all these scatterers are of the same amplitude. Essentially, they are of equal scatterer strength. Therefore, what we do is called time gain compensation or TGC or also known as depth gain compensation.

Essentially as you go deeper you are losing signal so you would like to amplify the deeper signals more than the shallower signals coming from shallower targets or targets located close to the transducer. So this is called time gain compensation or depth gain compensation and a specialized amplifier called the TGC amplifier performs this task.

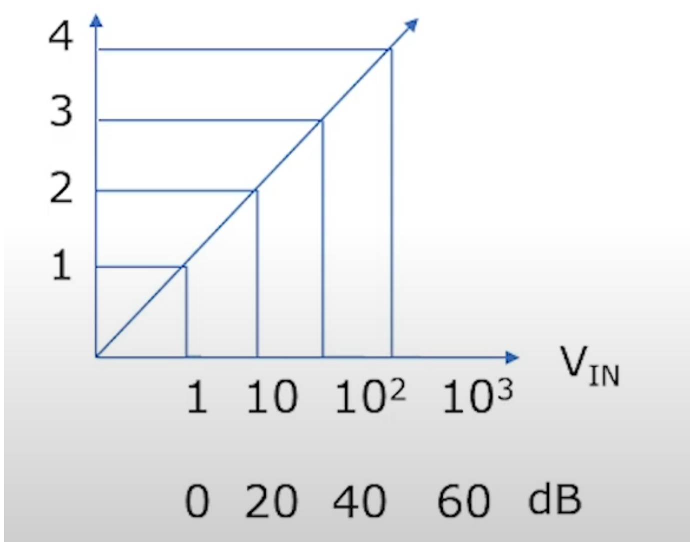
Now, typically ultrasound transducers have a very high dynamic range. They can easily have 100 or 120 dBs of dynamic range, which makes it difficult to see the signals which are very weak. So the idea here is that we can employ a log compression or some other kind of dynamic range compression which will make the weaker scatterers appear visible in images. So in simulation image below, there are a large number of point scatterers in the beam and only one scatterer is visible. But once you perform log compression you can see that those weaker scatterers also become visible. Here they appear as streaks and that is because of the beam properties. In the near field the beam is wider, and it is the narrowest at the focus and then again it diverges again after the focus. That is why you are getting this point spread functions which are a bit extended. The main point here is that you are able to see these targets which were otherwise not visible without this log compression.



Now here is a simulation where a log compression was performed on the image and the dynamic range was set. So you can see some of the weaker targets appear more echogenic here because of this log compression.



So how is this done? One way would be to use a linear amplifier, then get the signal, and then using some kind of DSP processor or a computer, perform this image dynamic range compression. However, there is a better way of doing this in hardware itself. There are amplifiers which are called logarithmic amplifiers where they straight away take the voltage values and amplify them in such a manner that the resulting signal will be log compressed. So, here is an example where the input signal is changing in powers of 10, but the output signal is changing linearly. So, the output is changing as 1, 2, 3, 4, but the input signal is changing as 1, 10, 100, 1000, etc.

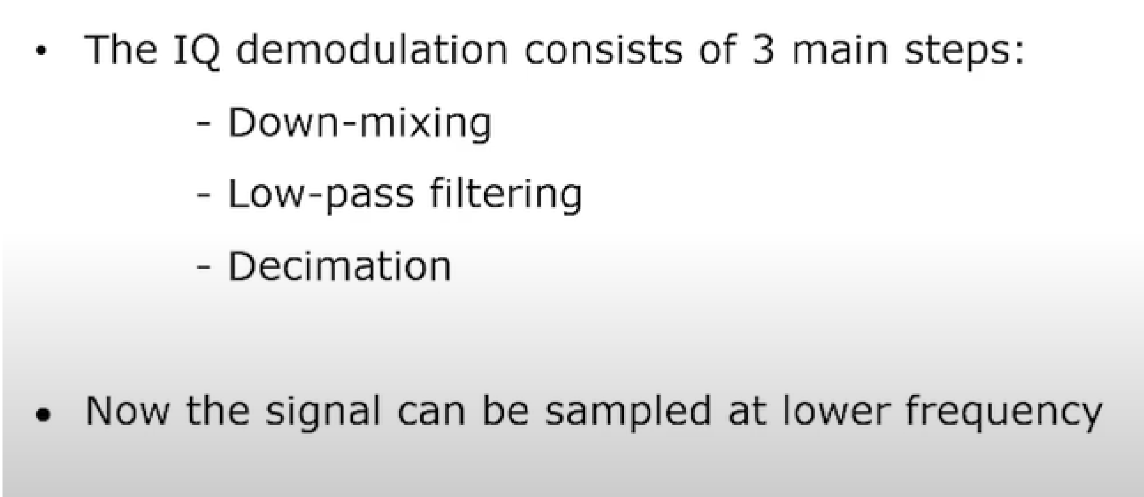


And these are the decibel units. So if the dynamic range is 60 dB, that corresponds to scatterers ranging from 1 unit strength to 1000 unit strength. So this kind of amplifier will be used so that log compression can be performed. Now we are not going to discuss the details of the amplifier or the hardware implementation of the amplifier. But this is something for you to look up.

Next, let's discuss base banding or IQ demodulation. As we are aware, ultrasound generates large volumes of data. Consider a 10 MHz ultrasound transducer, and let's say you have an array with 128 elements. So now all these 128 elements will receive data at 10 MHz. Now to sample this data effectively, we have to follow the Nyquist sampling theorem. And also we have to consider that 10 MHz is the center frequency, not the highest frequency component in the signal. We have to consider the presence of high frequency noise, et cetera. And therefore, typically the thumb rule is that the sampling frequency needed will be 5 to 10 times the center frequency.

So let's say if the sampling frequency is 10 times the center frequency, then it is 100 MHz. So 100 million points per second. And now if you have 128 channels, you have to multiply by 128. So 128 times 100 million points per second is what you are acquiring. And this is a very large volume of data. And not only that, we need electronics which are very high end to be capturing data at such high frame rates. So we need some method by which we can get away with sampling at lower rates. So this is where we use a trick which is called base banding or in-phase quadrature (IQ) demodulation.

You may have heard of IQ data. This is also used in communications and this technique is essentially borrowed into ultrasound from communications theory. So, IQ refers to in-phase and quadrature data and I will explain what this does. So this process of IQ demodulation consists of three steps.



You take the signal and downmix it. Downmix means bringing down the frequency. Then you perform low pass filtering to get the spectrum of interest. And then you decimate the data or you bring down the sampling frequency of the data so that the data volumes can be reduced. And now the signal can be sampled at lower frequency after this base banding or IQ demodulation has been performed.

So how would we perform this base banding mathematically? Let's recall that phase shift in time domain for any given signal is equivalent to a frequency shift in frequency domain. So to recall this, you need to revise the property of Fourier transforms

We take any signal which is of high frequency and either increase its mean frequency or bring down the frequency. And how would we do that? To realize this, let us recall the Euler's theorem.

$$e^{j\omega t} = \cos \omega t + j \sin \omega t$$

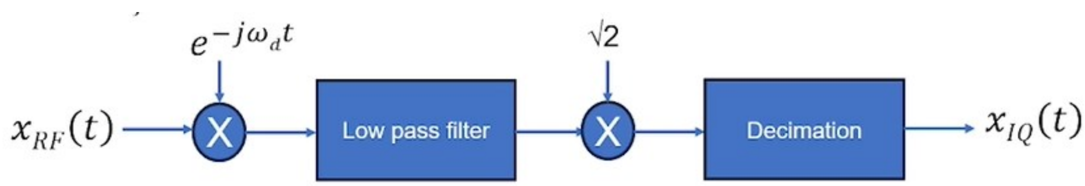
Where,

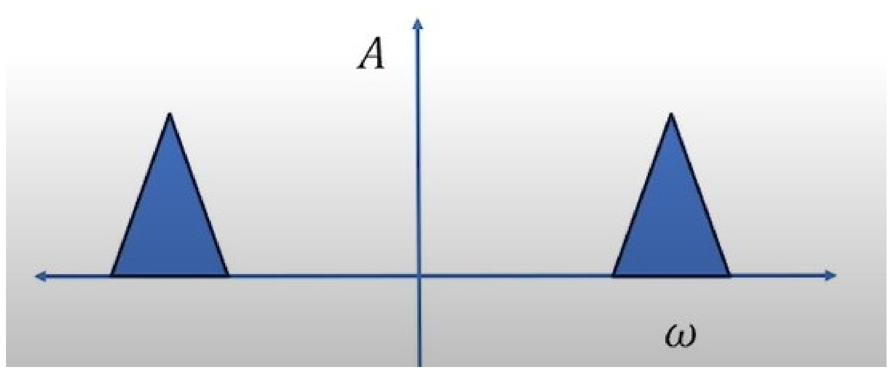
$$j = \sqrt{-1}$$

Typically we denote it with i, but in electrical engineering i is used often for current and that is why we use the notation of j here. So, similarly,

$$e^{-j\omega t} = \cos \omega t - j \sin \omega t$$

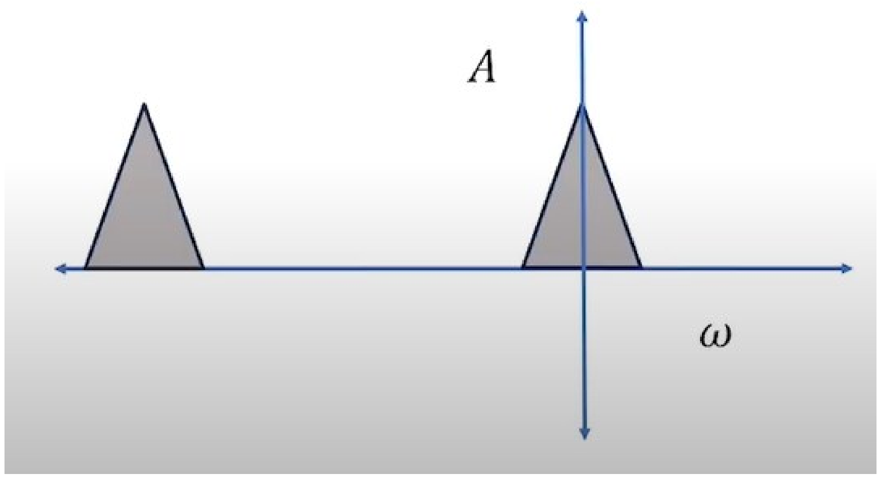
Now recalling this Euler's theorem let's see if you have a RF signal $X_{RF}(t)$ and you multiply it with $e^{-j\omega_d t}$,here, d indicates a downshift, and ω_d is the frequency by which you are going to downshift the signal. Once you do this, there are other steps, we will discuss them step by step, but you will need to low pass filter the signal, multiply by square root of two, and then you can decimate it to get the IQ version of the data.





So let us imagine that this spectrum in the above figure is of a real signal And since the signal is real, it will also have a replica at the negative frequencies. So this is how the signal looks initially before we give the phase shift. Now, once we give the phase shift in the time domain, it will lead to a frequency shift in the frequency domain.

So now we have shifted the frequency down by ω_d. Now what happens is that by suitable choice of this ω_d, we have brought down this mean frequency of this spectrum to zero as in the figure below. So what this means is that the signal is now a baseband signal. However, you have an asymmetry in the spectrum and you have an other negative frequency term which does not seem very helpful to us. So we are going to use a low pass filter to extract only the baseband signal, which is of our interest and we are going to reject this other negative frequency signal.



Now if you're going to reject the negative frequency signal, and only accept the baseband signal, it is important to remember that you will have loss in energy because both spectra has some contribution to the energy.

So to get rid of that loss, we are going to scale up the amplitude by a factor of $\sqrt{2}$. After this, we have the baseband representation of the pass band signal. So what we have done essentially is base banding. We have brought the pass band signal, which was at a frequency of ω_d, to a center frequency of zero. And that is where we are right now.

Once you have brought the frequency down, Now let's talk about sampling the signal. So since the maximum frequency is also reduced considerably, we need not go for very high sampling frequency. We can now decimate the signal by a certain factor of interest and after that we can perform the sampling. So once we decimate the signal, the signal represents the IQ data. So now let's discuss how we are going to implement this IQ demodulation. So it can be done either in the analog domain using hardware or using DSP processor chips.

So these DSP processing units can handle complex numbers easily. So if you are actually doing this with analog hardware, then instead of complex numbers, you can take sine and cosine. These signals are phase shifted by π/2. And we know that multiplication by j also shifts the phase of a signal by π/2 in the complex plane. So you can perform an analogous operation by using sine and cosine terms instead of using complex numbers.

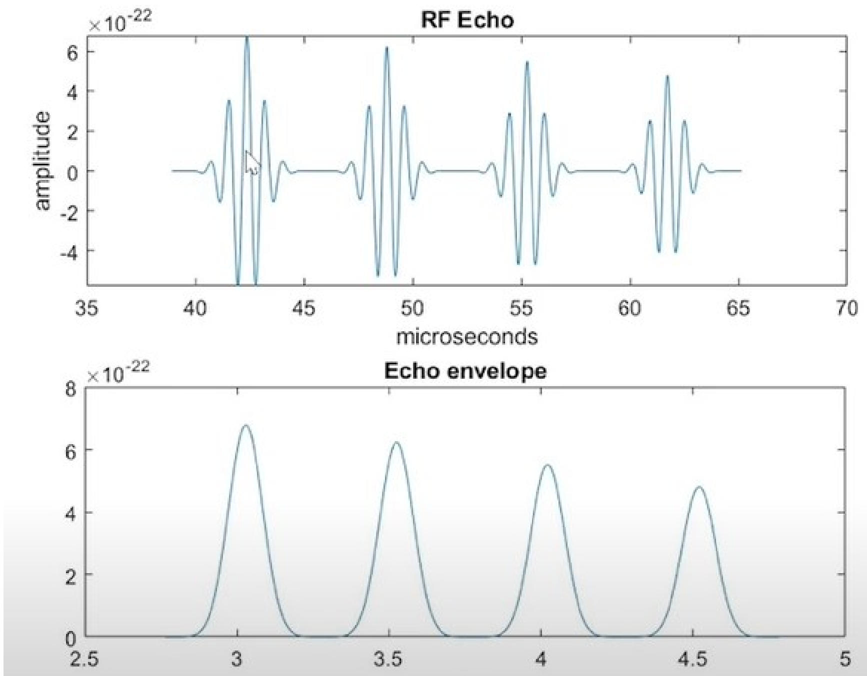
Just like you can convert from RF data to IQ data, the vice versa can also be performed wherein you can get IQ data and convert it to RF data. How you do that is by a few steps. First you need to interpolate the data to increase the number of points and this is done by adding zeros between your samples. Now, if you look in the frequency domain, the effect of zero padding is that some replica spectra get added to your spectra. So we only want the spectrum of interest. That's where we do low-pass filtering. Once we are done with low pass filtering and we have the spectra with the right number of points and which is isolated in terms of the spectra of interest, we can perform up mixing. Just like in the last process we performed down shifting or down mixing by multiplying by $e^{-j\omega_d t}$, here we will perform up mixing by multiplying by $e^{j\omega_d t}$. And then what you get is still a complex valued signal. So now you can take the real value of this complex signal and this will give you the RF data.

IQ to RF data conversion

- IQ data can be converted back to RF data
- Interpolation (zero padding and low pass filtering)
- Up-mixing
- Taking the real value of the complex signal

Typically for processing the data, particularly if you're doing some research, you might need to convert from IQ to RF data. And this is the general process.

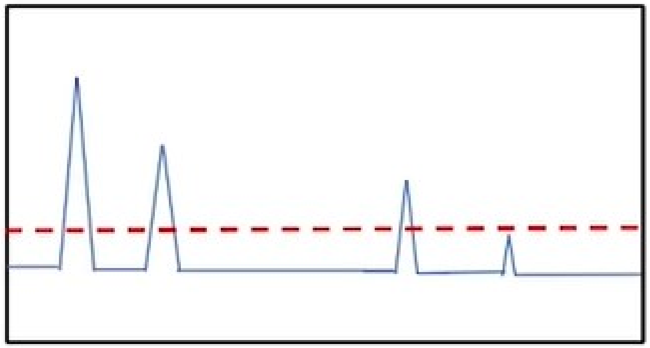
Now, when we get the signal back, the signal has these high frequency oscillations, which were present in the pulse. So let us assume that you have four scatterers and the figure below is the RF echo from those four scatterers. So we would like to see a high intensity at those four spatial locations, but you have some high frequency oscillations which you would like to get rid of.



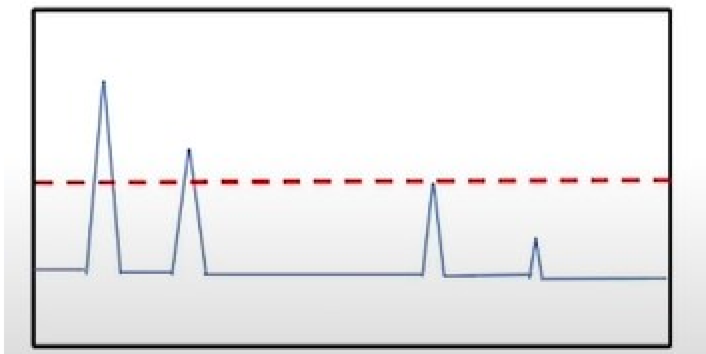
So you would actually like to see a signal like the “Echo envelope” signal in figure above, instead of the “RF Echo” signal, and this can be performed by just looking at the envelope of the signal and not the high frequency modulations. So now if you recall this envelope detection process, you have seen it somewhere. If you remember amplitude modulation and demodulation, where we take a high frequency carrier and we multiply it with a message signal. And then our message signal actually appears in the envelope of the modulated waveform. And when we want to demodulate, we simply look at the envelope and we get rid of the carrier signal. So it is kind of similar here as well. We are interested in the envelope and we would like to get rid of the high frequency carrier signal.

So this can be considered analogous to a type of demodulation. And we need to get rid of these oscillations to prevent ripple artifacts. So there is a mathematical operation called Hilbert transform, which can do a great job at extracting this envelope. There are some other approximate transforms and other approximate mathematical operations as well, which require less computational overhead and can do a similar job. But after doing the envelope detection, what you can see is the peaks present at the location of the four scatterers.

Now typically your signal will have some noise as well, and one way of denoising is a threshold based noise rejection which is present in lot of ultrasound systems. Here you assume a certain threshold and you say that any signal below this threshold is actually noise and you are not going to consider it. You can vary this threshold to eliminate different levels of noise.



Low rejection threshold



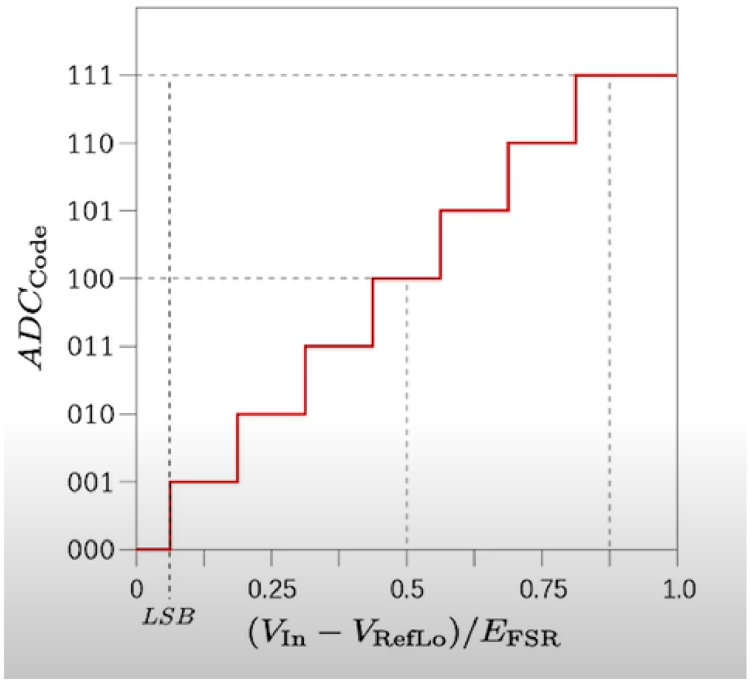
High rejection threshold

So as you can imagine there will be a trade off. You can decide this threshold based on a percentage value relative to the maximum. For example, you can say that any signal which is less than 5% of the peak amplitude, will be rejected because it is assuming to be noise. Or you may say that any signal which is less than 10% of the peak amplitude, is going to be considered as noise. But there is a trade off because as you increase the threshold, you are going to lose some useful signal as well. So there's a trade off between sensitivity to weak scatterers (because the weak scatterers will have low amplitudes) and signal to noise ratio. So this trade off should be considered judiciously and it will help improve image quality.

So far in brief we are discussing some of the steps required in getting the ultrasound image information and this is a very brief discussion we are only scratching the surface, but nonetheless I believe this will be a starting point for you to get into in depth study of these topics.

Next, let us discuss the digitization step. When you have a signal, you are able to convert it to a digital signal by two processes. One is sampling and the second is quantization. So sampling will make the signal discrete along the time domain and quantization will make the signal discrete along the amplitude axis.

So here, let's assume that you have a three bit analog to digital converter. So 3 bits will give you 2^3 =8 levels. So you can say 0 to 7 levels. And the amplitude of your signal will be quantized in 7 levels. Now clearly 7 levels is not sufficient. Typically in ultrasound scanners, you have at least 12 bits or even higher number of bits for digitization.



There is also a certain quantization in amplitude here. If a signal is in between two levels, say between 011 and 100,but closer to 011, then it will be brought down to the floor. If the signal is slightly higher than the mean value (more closer to 100), then it will be taken to the ceiling. But in either case, there is a certain approximation, and that approximation is modeled as a noise. So you have something called digitization noise, which gets added as a result of this analog to digital conversion. So now this digitization noise can be reduced if there are higher number of bits, which means there are more levels.

The more the levels, the better the approximated the digital signal will be to the analog signal and it will give you better amplitude resolution. So once this digitization is performed the echo is available to you in a digital format. Once you get the echo in the digital format, you can perform processing for display purposes, or do color mapping, and other post-processing methods , for example, various filters, denoising filters, et cetera, and then you can display the image. So I would like to summarize this lecture on ultrasound imaging systems.

We discussed about a very basic architecture of the ultrasound system. We discussed the process of time gain compensation or depth gain compensation. There's also beam gain compensation, which is done along with time gain compensation. We discussed the use of logarithmic amplifiers for dynamic range compression, which helps us visualize weaker scatterers with more fidelity. We discussed the process of in-phase and quadrature demodulation. This is a technique borrowed from communications theory, which helps us to reduce the computational overhead and the data overhead involved in ultrasound imaging.

We discussed techniques for noise rejection. We discussed analog to digital conversion and the final processing and display of the image. So I believe you found this lecture useful. I'll see you in the next lecture. Thank you.