2025-10-17, 2D Fourier Xform

Ok, so here is what I UNDERSTOOD about 1D Fourier, I used sound analysis to help me understand.

1. The sound signal S(t) is a series of numbers sampled at different time, ie., 44140 times in each second. So I have 44,140 data points
2. The assumption (?) is that the sound signal is a mix of different waves at different frequencies. In other words, $\sin(t) = \sin(\omega_1 t) + \sin(\omega_2 t) + \cdots$, so if we know the underlying frequencies, we don't need all the 44,140 datapoints.
3. So we create different waves, multiply each frequency wave by the signal, point-to-point, sum them up, if the result is zero, is saying no such frequency, the nonzero result tells us the magnitude of such frequency component.

Throughout the experience, I never at once felt the need for 1) matrix; 2) nor 'complex number. I did realize that using only one sine, I am missing the phase information. Hence, I also add a cosine part, but everything done separately, never as complex numbers. Easy to understand in my opinion.

So 2D Fourier mechanically appear to be similar to many 1D analysis. Now the signal is no longer one series, but instead many series, if we use the arrangement of the pixels, naturally it has the shape of a matrix, hence now it's more natural to use matrix notation. But this is only the book-keeping part of matrix implementation, nothing fancy. Reading through it felt easy to run the 2D analysis in a multiple 1D way: run the same analysis for each row, now each row of pixels become one row of frequency component (each value signals the magnitude of such frequency component). You start with a box of pixels, you end up with a box of frequencies.

Ok, now this is new step, we run the same analysis along the vertical direction, so again we end up with a box frequencies, but now the value tells us the frequency component in this 'column' of pixels, rather than row of pixels.

GPT does not seem to be bothered, about something of conceptual leap. I think 1D is pretty easy to understand, but this 2D, the u,v COMPONENT of frequency is RADICAL! It means frequency has directions now, and you decompose it into two orthogonal directions. This is hard to understand! Then I asked for explanation, which reads more confusing. It mixes more jargons, because it does not know how WE experience waves.

If you have a source, sound, dynamics from a rock thrown into a lake creating ripples, you can measure wave from ANY direction! So wave is never a single line object heading in one direction! In sound analysis, we can omit that because we are recording at a fixed location, the wave obtained looks single along time axies. We wouldn't overthink, it is what it is.

But now in a picture, it's 2D, waves can be observed in any directions.

Unlike sound wave where you CAN'T hear waves, what's great in this image case is YOU CAN ACTUALLY SEE WAVES! Now imagine ripple, I think it's a single frequence wave, traveling from the originally at all directions, RADICALLY! But now if you were to collect the pixel data not in a radial direction, but as rows and columns, you will be observing waves at different frequencies!

2025-10-20

So when you were told that 'sound is made up of waves at different frequencies', what do you feel about it? Belief or disbelief? Similarly, when people told you that pictures are made up of waves at different frequencies, what is your take?

The way I make myself believe that sound is made up of different frequency waves is, I used Audacity, pick a sound wave at a single frequency, then play it. So I believed it. Then I was happy to move on using Fourier analysis to analyze the frequency components. But if you think about it, I've never experienced sound as oscillating wave! A sound which I feel constant is an oscillating wave, and I certainly did not hear the wiggling!

Now to the image. First of all, if image is made up of waves, what do those sine waves represent? In the case of sound waves, they represent the oscillating amplitude (air pressure), even though my ear cannot capture the change. As GPT explained, the values represent pixel value, in the case of black-n-while picture, they represent the brightness of that pixel. If the brightness values are oscillating, then we will SEE waves. I think that's a fundamental difference between sound and images:

1) In sound analysis, we were told there are waves, even though we play 'wave', we don't actually hear the jiggling.
2) But in image analysis, we were told there are waves, then we can see the oscillating brightness turned into wave that we can see.
   - But I still find it hard to believe that any images are made up of waves at different frequencies!

So I asked GPT for a step-by-step exercise showing us how by adding more waves, we can transform from simple wave image to other shapes. It gives me 20 pages in 1 min, impressive. But looking at them, I have questions.

1) Did I see wave in the first case, the single wave? But which way is it oscillating? Yes it's horizontal stripes, but does it mean the value is oscillating in x direction? Ok GPT confirms that I was right.

Then I asked for more interesting shape, because the image of some straight lines intersecting at an origin is not interesting. Then it proposes a mountain silhouette, which is boring too. But

then a idea hit me: the point of using waves to create image is not for animation! If you know the shape, you just draw it! There is no need to back-solve for waves and then reconstruct using waves! It's an overkill! The value of Fourier analysis in imaging is medical imaging where you CAN'T see the shape! You collect signals, you use those signals to analyze the waves to reconstruct the image! But then it leads to more questions…

If we know the image, meaning we have the pixel values at each point, we can run Fourier analysis, coming up with the waves, then recombing to recreate the images. But in the case of medical imaging, we are not analyzing each pixel, because we don't have a picture to start with. We have signals, but what do those signals mean so that will allow us to come up with the waves, then these waves are for what? In the first case, the waves are for pixel values directly, but here what do these waves represent then?