29. Where do Image Spectra come from?

In this section, we first discuss delta functions, and then analyze again the notion of a delta function sampled signal. This leads to a simple understanding of the physical origin of image spectra. Then we extend the interpretation to the amplitude modulated pulse train.

(a) Gaussian pulse

Consider this gaussian pulse and its gaussian spectrum computed from (1.1),

$$x(t) = \frac{1}{\sqrt{\pi}\alpha}\ \exp[-(t/\alpha)^2] \qquad X(\omega) = \exp\left[-\left(\frac{\omega}{2/\alpha}\right)^2\right]$$

As the width α of pulse x(t) pulse gets smaller, the width 2/α of the spectrum gets larger. Think of x(t) as being a summation of cosines. It takes higher and higher cosine frequencies to trim down the Gaussian as it gets narrower. For any α, the area under this gaussian x(t) is 1. As discussed in Appendix A, in the limit $\alpha \to 0$ this x(t) becomes exactly δ(t) and in this limit X(ω) = 1. We showed this in (8.3), but it is obvious for the explicit Gaussian pulse shape shown above.
ok to here

(b) Delta spike

So the spectrum now of δ(t) is X(ω) = 1. Now all frequencies up to infinity are required to force this spike to be zero on both its sides, which are now infinitely close together. For the gaussian, at least X(ω) tapers off after a while when you get around ω = 2/a. Here, you are stuck with a constant energy spectrum all the way out to a billion GHz and well beyond.

You might say that all this spectral energy is in the sharp needle-like edge of the delta function . If you had a voltage v(t) = δ(t) going into a 1K resistor, the pulse energy would be equal to δ(0)/1000 watt-seconds, which is infinite. It takes more than all the generators in the universe to create one delta function, and the reason is that its spectral energy must be supported all the way out.

(c) Delta Spike Sampled Signals and the Origin of Image Spectra

In Section 20 it was shown that if you "sample" a reasonable function y(t) with genuine delta function spikes, you end up with a Fourier integral spectrum that contains an infinite set of copies of the main spectrum of y(t). One tends ask oneself some questions:

- Where do all those image spectra really come from?
- How can they be there going out to infinite frequency?
- Are they really there or not?

A good example to consider is y(t) = a simple square box a little more than 3 times the delta function spacing. Then the sampled signal w(t) is a set of 3 delta spikes all by themselves, all of equal height, separated from each other by some Δt.

Since each delta has spectral energy out to infinity, and w(t) is a superposition of 3 of them, our second question is answered. No problem with infinite frequency here.

Where does the "main spectrum" come from? Well, you might argue that there is a ghost of that square box sitting there in the three delta spikes. It is the skeleton of the box, it ought to have a *little* of the main spectrum. But why *image* spectra?

In fact, from (20.5) and (9.2) we know what the answer must be for all the spectra, main + image:

$$W(\omega) = \sum_{m=-\infty}^{\infty} \tau \, \mathrm{sinc}(\tau/2[\omega - m\omega_1]) \qquad \omega_1 = 2\pi/\Delta t \qquad \tau = 2\Delta t$$

This looks like a rough sum to try and compute. It is the superposition of little sinc patterns, but they are close together and overlap a lot. The second zero of the main spectrum is smack on top of the center of the first image spectrum, and so on.

Now, in answer to the first question above -- where do the image spectra come from -- we give our simple answer. Recall that $X(\omega) = 1$ is the spectrum of $\delta(t)$, and recall the phasing rule for time shifts. Thus, if we add our three little delta's spectra we get this:

$$W(\omega) = e^{+i\omega\Delta t} + 1 + e^{-i\omega\Delta t}$$

In other words, the image spectra are caused by *interference* between the spectra of the three delta functions. This is also the "cause" of the main spectra as well, since it is really no different from the rest. The above sum is one we do in fact know how to do,

$$W(\omega) = 1 + 2\cos(\omega\Delta t)$$

Clearly this thing has a peak of 3 when $\omega = 0$, and that peak repeats whenever $\omega\Delta t = m\,\pi$, and these repeats are the image spectra.

You always get repeats when you have phase interference.

Suppose we close the spacing between the delta functions so now 101 of them fit in our box y(t). The signal is more complicated because now we have to add up 101 phases, and the sum looks like:

$$W(\omega) = 1 + 2\cos(\omega\Delta t) + 2\cos(2\omega\Delta t) + 2\cos(3\omega\Delta t) + \ldots 2\cos(50\omega\Delta t)$$

Now it's easier to see the general picture from the sinc(ω) sum above, since there is no longer any significant overlap. Basically, this set of 101 delta functions is acting like a phased array radar. There are interference maxima located at certain values of ω where the phases are all "in phase". This happens when the phase difference between adjacent delta "transmitters" is a multiple of 2π. Thus, we set $\omega\Delta t = m2\pi$ and find that the "radiation beams" are located at $\omega = m(2\pi/\Delta t) = m\omega_1$. The main spectrum is the central beam at m=0. The other beams are the image spectra.

So we now summarize our answers to the posed questions. First, the image spectra are generated by interference between the delta function spectra from simple superposition. Second, the interference is maximum whenever $\omega = m(2\pi/\Delta t)$, for m = any integer. Since integers don't stop anywhere, neither do the image spectra. They go to infinite frequency. Finally, yes they are really there, if you can really make a set of delta functions.

(d) The spectrum of an amplitude modulated pulse train.

We know the answer, it is given in the box in Section 25,

$$W(\omega) = (1/T_1)\, X_{\mathbf{pulse}}(\omega) \sum_{m=-\infty}^{\infty} Y(\omega - m\omega_1)$$

This situation is perhaps easier to handle than the delta function limit discussed above. $x_{\mathbf{pulse}}(t)$ can be a pulse you can really make. If this pulse is a narrow square wave, our signal w(t) looks like that shown in Figure 26.1. The curve there is y(t).

Now where do the image spectra come from? It's the same thing. Each pulse in the pulse train has a spectrum that is equal to $y_n\, X_{\mathbf{pulse}}(\omega)$ multiplied by a phase determined by where that pulse is sitting. Here is the interference sum if there are only three pulses of amplitudes y_{-1}, y_0, y_1:

$$W(\omega) = y_{-1}\, X_{\mathbf{pulse}}(\omega) \bullet e^{+i\omega\Delta t} + y_0\, X_{\mathbf{pulse}}(\omega) \bullet 1 + y_1 X_{\mathbf{pulse}}(\omega) \bullet e^{-i\omega\Delta t}$$

$$= X_{\mathbf{pulse}}(\omega) \left[y_{-1} \bullet e^{+i\omega\Delta t} + y_0 \bullet 1 + y_1 \bullet e^{-i\omega\Delta t} \right]$$

Although the magnitude of each term is now different, the phases are still going to add up at the same places as before, at places where the phase between terms is a multiple of 2π. So you *know* there are going to be an infinite number of image spectra regardless of how many terms y_n there are, and you know *where* they are going to be located: $\omega = m(2\pi/\Delta t)$. You see also how $X_{\mathbf{pulse}}(\omega)$ factors out to be a common factor in the final answer.

So why is each spectrum a copy of $Y(\omega)$? By definition, the sum shown above is $Y'(\omega)$, the Digital Fourier spectrum. And we know that $Y'(\omega)$ is always a sum of the $Y(\omega)$'s. But this does not quite answer the question just posed.

Instead of dwelling on this mystery, here is an observation that seems closely related and interesting. Suppose you only got the central spectrum $Y(\omega)$, and no image spectra. You would then have obtained $Y(\omega)$ from just the three sample points y_{-1}, y_0, y_1. But there are many functions y(t) that hit these three points, and they cannot all have the same spectrum $Y(\omega)$, see Figure 26.1.

It is the overlap between the image spectra and the main spectrum $Y(\omega)$ that resolves this puzzle. The sum of the main and all image spectra is exactly the equation above with the three terms. This sum only knows about the three numbers y_{-1}, y_0, y_1 and Δt. As you deform the curve y(t) such that it still hits

these three values, you alter Y(ω), so you alter the shape of the main and all image spectra. However, despite this alteration, they still add up to what they did before the alteration! You could thus say that the image spectra reflect the *lack of knowledge* of y(t). As you learn more about y(t) by having more sample points closer together, the image spectra recede away. When you have perfect knowledge of y(t), the image spectra are completely gone, and you have only the main spectrum Y(ω).

If y(t) is band-limited, then we have to adjust the above argument just slightly. As the sample points become closer, the image spectra do move out. But once the samples are close enough at spacing T_1 such that Y(ω) fits entirely below $\omega_1/2$, the spectra no longer overlap and Y(ω) is complete. Although only a finite number of points on y(t) are being sampled, Y(ω) is completely known because there is no longer wiggle room (below the $\omega_1/2$ limit) for an altered function y(t) to fit the points.

What about the energy in our three-pulse signal? How many generators are now needed to make it? The bracketed three terms above by themselves require infinite energy, just as in our delta spike example, but we are of course saved by the overall multiplying factor $X_{\mathtt{pulse}}(\omega)$ which keeps the lights from dimming in Madagascar. For the box shaped pulse, we know that the envelope of $X_{\mathtt{pulse}}(\omega)$ goes down as $1/\omega$, as shown in Figure 17.1. The area under $|W(\omega)|^2$ is quite integrable, and by (10.5) is equal to the area under $|w(t)|^2$ times 2π. With w(t) as shown in Figure 26.1, the result is easily computed and is quite finite.