Understanding Avian Sound-It's not what you think!

Neil J Boucher

Audio Wings

July 2012

Synopsis

Here it is shown that the conventional interpretation of avian and other calls using the FFT is seriously flawed and that using higher precision spectrum analysis such as the LPC sheds new light on the nature and structure of those calls. What we found is astounding and could mean that we need to re-write the text books. We reveal here, for the first time some of the things we have discovered.

We have been using the LPC (Linear Predictive Coding) Transform for some years now in our SoundID recognition system. We found quite early in our studies, that even though the LPC is computationally more expensive than the FFT it gave better recognition results. This led us to explore why the FFT could not deliver equivalent results and it soon became apparent that the problem was that the FFT did not tell the whole story. In particular the LPC allows simultaneous zooming in both the time and frequency domain, whereas the FFT can only zoom in one of these domains at a time and it does so at the expense of the resolution of the un-zoomed domain.

Recently we have applied our LPC spectrogram to study some calls in detail. And because we can see the call structure in much finer detail than the FFT we expected to find some interesting things. We found much more than we expected.

Definitions

Before we go any further we need to define a few terms.

First, a *harmonic* is an integer multiple of a fundamental frequency. It usually is caused by non-linear behaviour of a system which is often alternatively called distortion.

Next we need to define *inter-modulation*. This is the beat frequency produced when two different signals are mixed in a non-linear medium. It should be noted that inter-modulation will not occur ordinarily in a linear medium (like air), but will occur in a non-linear medium like a vocal tract or ears. Harmonics and inter-modulation are different and mostly have different causes.

And finally a *wobulator* is a generator that deliberately wobbles its frequency (sometimes pseudo-randomly) around a centre point. These are widely employed in electronics and by birds, frogs and bats. The verb is to *wobulate*.

The Pied Butcherbird Call

The Pied Butcherbird is arguably the most accomplished singer in the bird family. It invents a new song every year and uses very complex sounds in that song. Even so it was surprising to find that in one little snippet of a song the Pied can demonstrate so many different call structures.

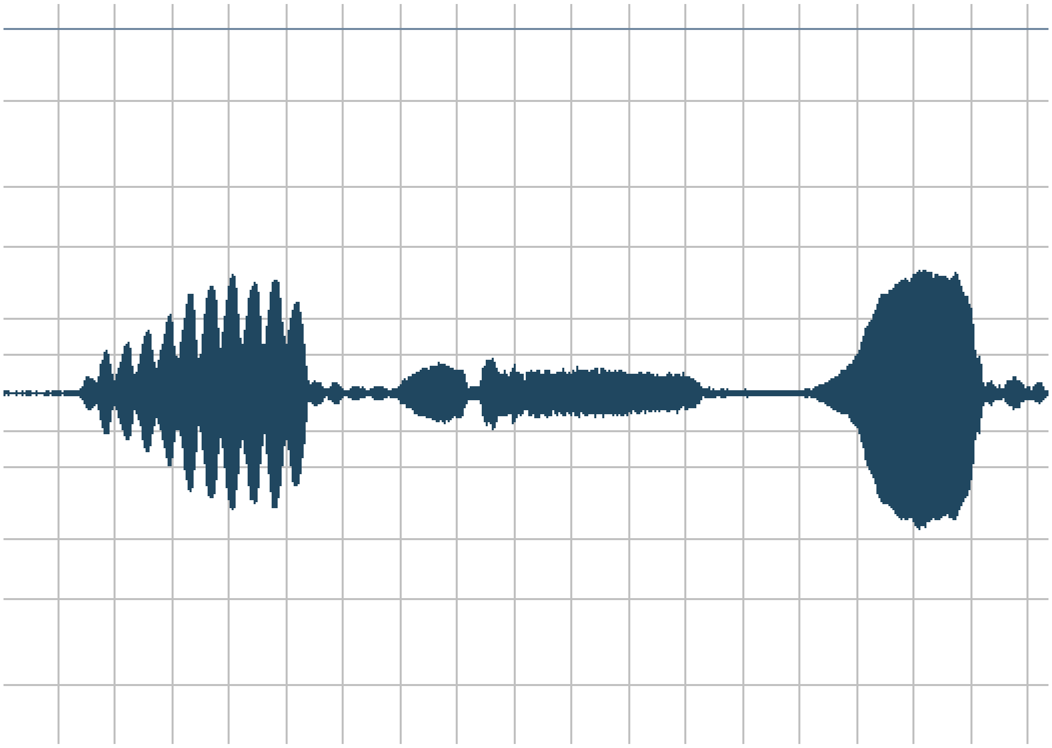
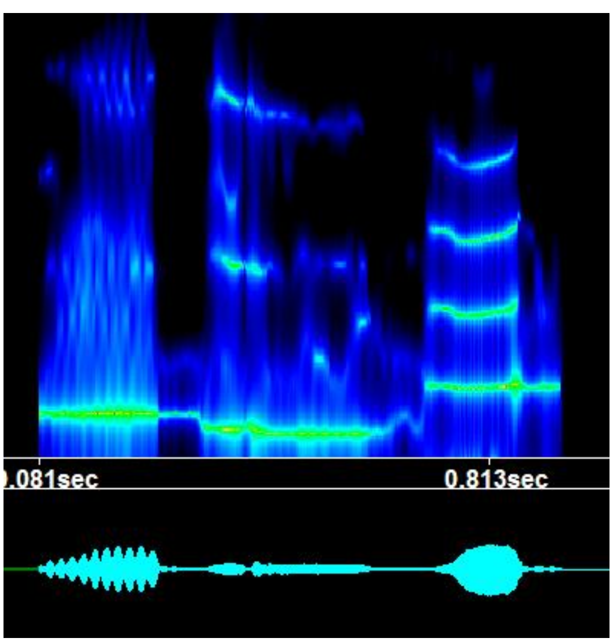


Figure 1: The waveform view of a Pied Butcherbird call.

Consider the Pied Butcherbird call in Figure 1 and its LPC transform in Figure 2. It is obvious from the LPC transform that three very different things are going on here. The FFT view as usual just shows a heap of fuzzy “harmonics”. The wave view in Figure 1 suggests that three different things are happening here and we will see that this is the case and that the wave view clearly identifies the same distinct sections as the LPC. In the first section of Figure 1 we see a much modulated envelope and that suggest that something interesting is happening. In the LPC spectrum of Figure 2, looking at the first section again we see no sign of anything that looks like harmonics, but rather a messy collection of miscellaneous high frequency components.



In the middle section there is what appears to be a high frequency signal with a harmonic and also a disconnected low frequency dominant signal. This is unlike the other spectrums.

In the last section of the call we see a simple envelope and if that is to produce higher frequency products (that look like harmonics, but don’t jump to conclusions), such an envelope suggests an entirely different mechanism from the earlier parts of the call.

Figure 2: The LPC view of the Pied Butcherbird call from Figure 1.

Note carefully, the WAV view of the first section. The shape of the envelope there suggests some interesting underlying mechanism generates it. We will see that this envelope could be a marker of two different types of modulation.

For comparative purposes, there is a FFT version of Figure 2 shown in Figure 3. The two are mostly in agreement for the end part of the call, but are telling different stories elsewhere. We will see that even where they do agree, the real situation is more complex than the FFT can reveal.

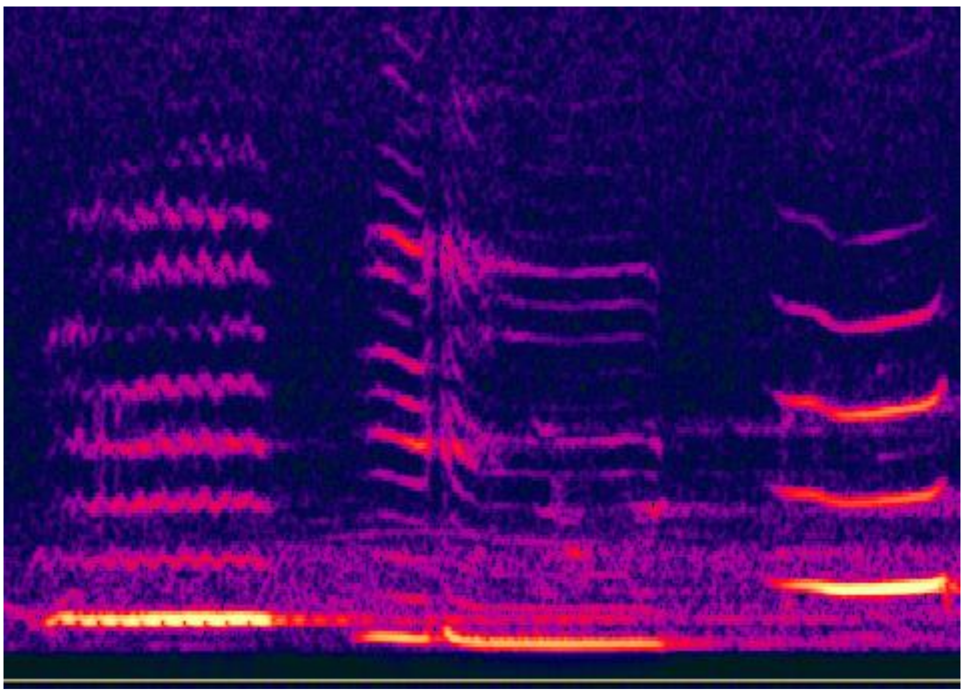


Figure 3: The fuzzy FFT view of the same part of a Pied Butcherbird call.

The LPC enables one to zoom in on small time slices of the signal without loss of frequency resolution (something that the FFT just cannot do), so we zoom in to see how the first part of the call is generated and we see something very interesting in Figure 4. We see that the frequency is being ramped up and down by about 1580 +/- 40 Hz (we call this wobulation). This will generate not nice clean harmonics, but a whole zoo of inter-modulation frequencies and so instead of "harmonics" (that the FFT shows in Figure 3) you really have lots of individual high frequency inter-modulation components as seen in the first section of Figure 2.

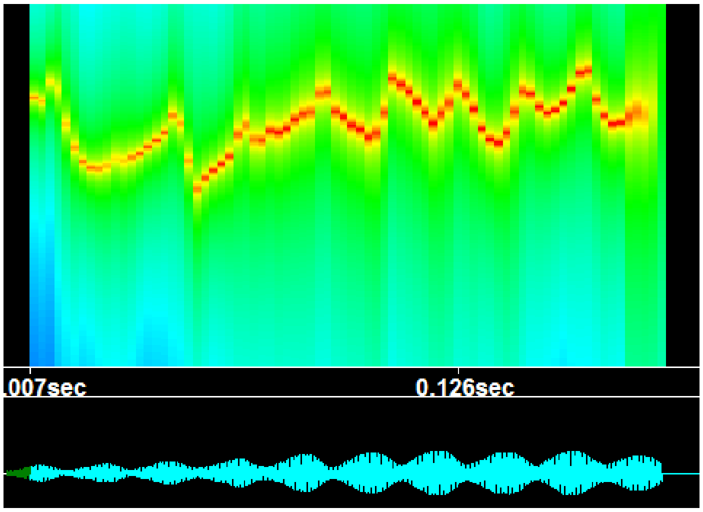


Figure 4: The structure of the first part of the call seen in the LPC view (peak frequency 1568 Hz).

An important point about wobulation is that the inter-modulation has not occurred in the vocal tract of the bird, and so it occurs in the ears of the listener. This means that different listeners will experience the sound differently.

If we zoom the frequency a little further on the same call and at higher resolution we see that both syrinx's are at work here. There seems always to be a dominant syrinx and the second one is significantly subservient (it could not even be seen in the lower resolution Figure 4.)

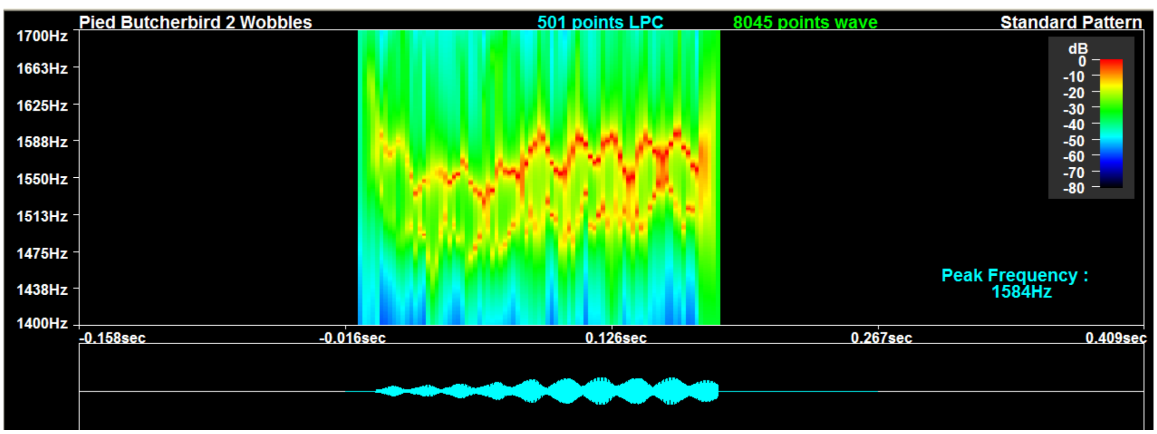
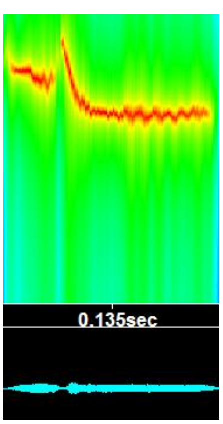


Figure 5: The higher resolution LPC view of the call.

Second Part of the Butcherbird Call



Now we come to the middle section and we see something else altogether different is happening. If we zoom in on the low frequency signal (seen in Figure 6 at about 932 Hz), we see that it consists of two more or less fixed frequency sections with a ramp to the second part. There is nothing much here to generate harmonics of any kind (except during the ramp). However the ramp is very brief and most unlikely to generate the "harmonics" seen in the FFT.

Figure 6: The low frequency portion of the middle part of the signal at 932 Hz.

Now let's get a closer look at the so-called "harmonic" in Figure 7. The first thing to notice is that it is only present in the first pulse in this middle section. Note particularly that it does NOT line up with the ramp in frequency and that its frequency of 6,273 Hz is far removed from the frequency of the lower frequency portion. The signal at approximately 12,000 Hz is almost certainly a harmonic of the 6,273 Hz fundamental which it can be seen to line up with in time. This signal is also significantly lower in volume than the other two parts of the call. So where does the 6 kHz signal come from; it can only come from another bird that has intruded into the recording. I went back and listened carefully to the original recording and indeed another bird was in the background.

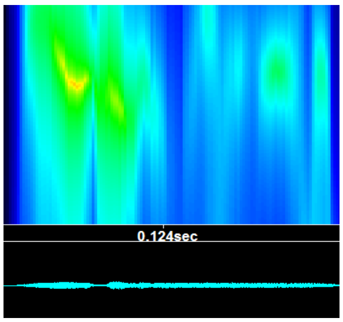
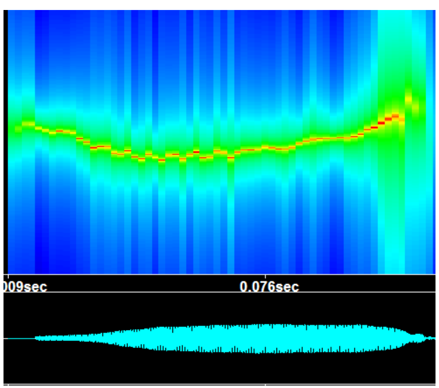


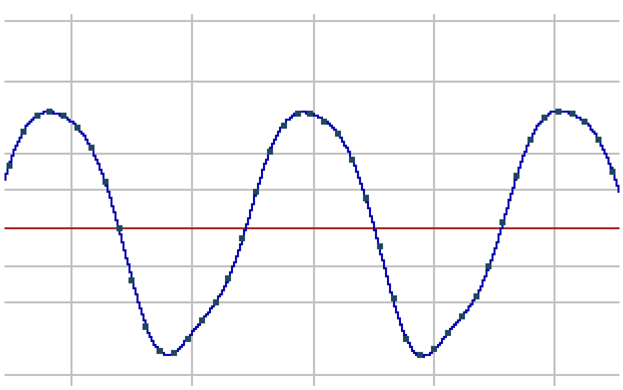
Figure 7: The 6.2 kHz region of the middle of the call.

The End Part of the Butcherbird Call



Now we come to the end part of the call. In this section at least, both the FFT and the initial LPC views are in reasonable agreement. Both are saying that we have good strong harmonics here. But if we look at the fundamental region seen in Figure 8 we notice that it is a rather clean signal varying slow in frequency. So we have to look somewhere else to see how these harmonics are being generated.

Figure 8: The 2476 Hz (peak) signal of the end portion of the call.



To see where the harmonics are coming from we should look where most bio-acousticians rarely tread. We look at the wave-view in Figure 9. When I first saw this view it looked very familiar. All cheap (and some not so cheap) signal generators from the 1960s to about 1980s produced asymmetric waveforms like this in Figure 8.

Figure 9: The wave-view of the signal in Figure 7.

Notice the "ugly" portion of the wave at the bottom (this ugliness persists throughout this section of the call). In the old signal generators this was caused by non-linearity's in the generator and did these things ever produce heaps of harmonics! It is probably just a co-incidence but the signal generators also most showed the distortion in the negative going part of the signal. Here the Butcherbird is deliberately producing harmonics by arranging a non-linearity in the vocal tract. Clever little beast!

Just in case you think I might be cheating a bit here, let's look at the first part of the same call (Figure 4), where the frequency was wobulating as seen in Figure 10. You can see that this part of the call is highly symmetrical.

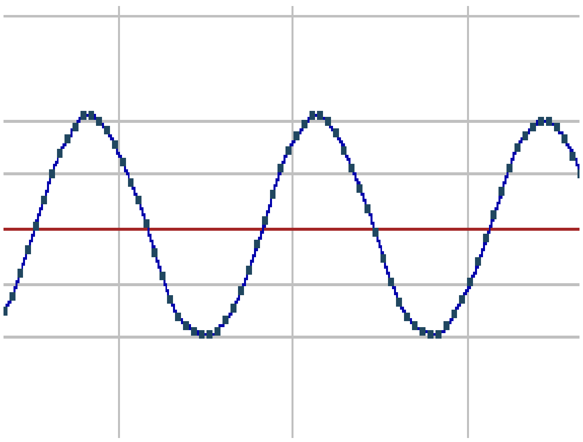
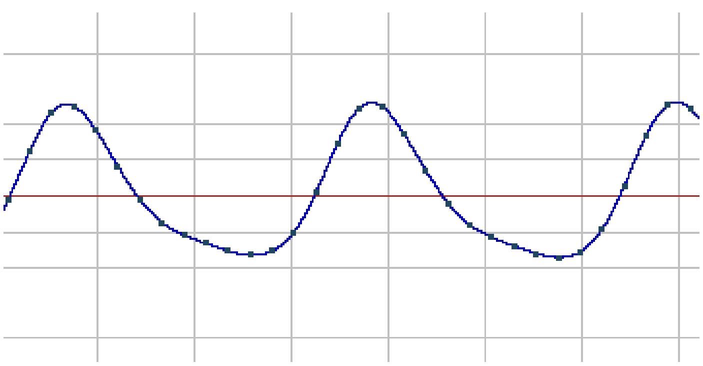


Figure 10: The relatively symmetrical signal of the first part of the call in wave-view.

King Parrot



Another bird that truly does produce harmonics is the King Parrot and we see below in Figure 11 that it is done the same way (by distorting the bottom half of the waveform but even more so than the Pied Butcherbird). One of the Butcherbirds in the yard of yours truly, has recently taken to imitating the King Parrot.

Figure 11: The King Parrot call in WAV view showing how the harmonics are generated

The White-throated Treecreeper

You saw an example of the use of the independent syrinxes in the middle of the Butcherbird call discussed here. But wait till you see this little guy. The White-throated Treecreeper produces these interesting “harmonics” as seen in Figure 12. Now you can see immediately that these appear to be a nice example of a fundamental (3526 Hz) with the second and third “harmonics”. Here is another rare instance where the FFT view, in its fuzzy way would agree. What the FFT can’t do however is explain how this little guy does it and the fact that these are not “harmonics” at all.

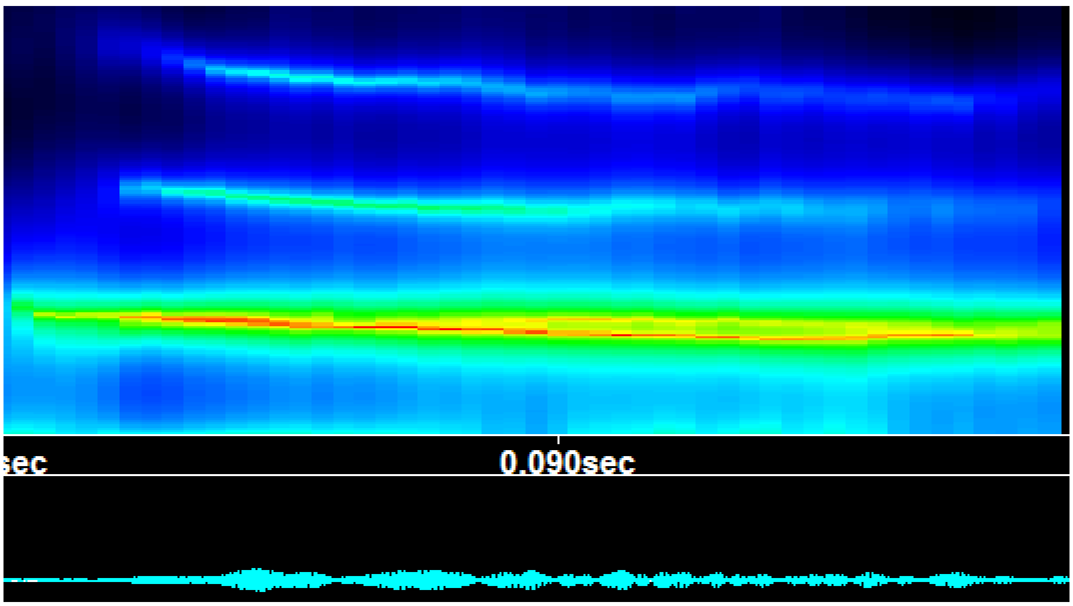


Figure 12: The White-throated Treecreeper generating some nice “harmonics”.

When we zoom in on the fundamental there is something unusual about it. There is not one signal there but two! See in Figure 13 that this little bird is generating in stereo two signals which vary from being almost the same frequency to being up to 600 Hz apart (you can see a hint of this in Figure 12 as well). The “harmonics” are generated by the beat frequencies (the sum and difference of the two) and so the “harmonics” are not harmonics at all but are inter-modulation products. Notice also the complex modulation envelope (in blue at the bottom) is similar to the envelope of the third part of the Butcherbird call in Figure 4. This is because both of the envelopes are the result of interference patterns between two signals of different frequencies (inter-modulating signals) which successively add in and out of phase. Notice again that one syrinx is dominant. Also note that the envelope of the WAV file is similar to the wobulated envelope of the Butcherbird. Either technique can produce such an envelope.

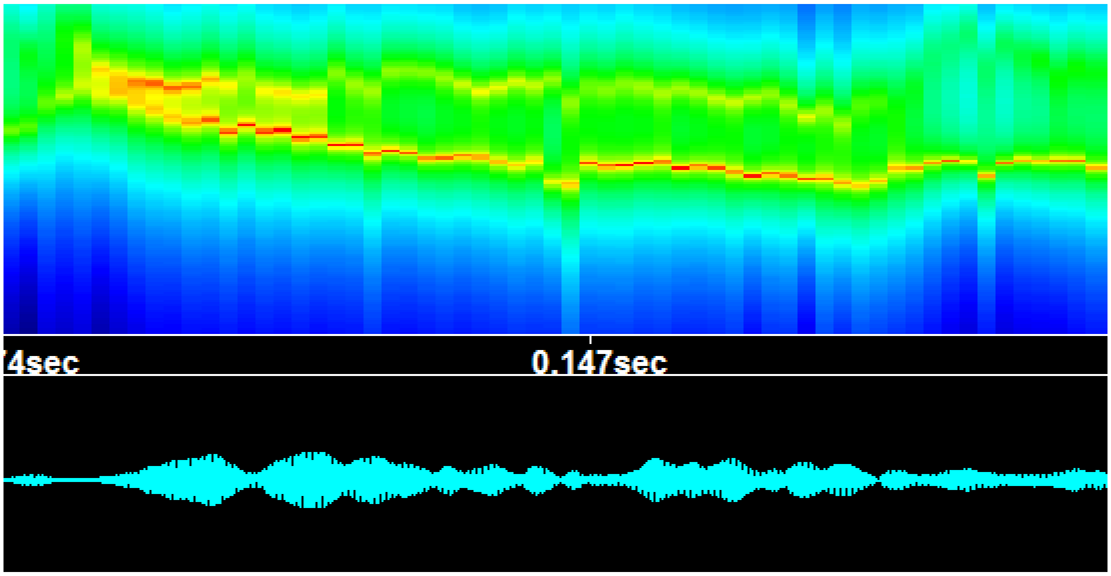


Figure 13: A zoomed-in view of the fundamental (3526 Hz) of the White-throated Treecreeper.

Finally we look at some calls that do not seem to generate "harmonics". In Figure 14 below we see the LPC spectrum of a King Parrot feeding call. There are two rather strong signals at 4876 Hz and another about 1 kHz lower. What is clearly missing here is any sign of inter-modulation products. Also notice that the envelope is rather fuzzy. These two things taken together suggest that this is not the call of one parrot, but of two. If the parrot was calling in stereo it could hardly do so without clear inter-modulation products which are not seen. Listening to the recording also confirms a strong likelihood of at least two callers. The spectrogram shows however that two are dominant in signal level.

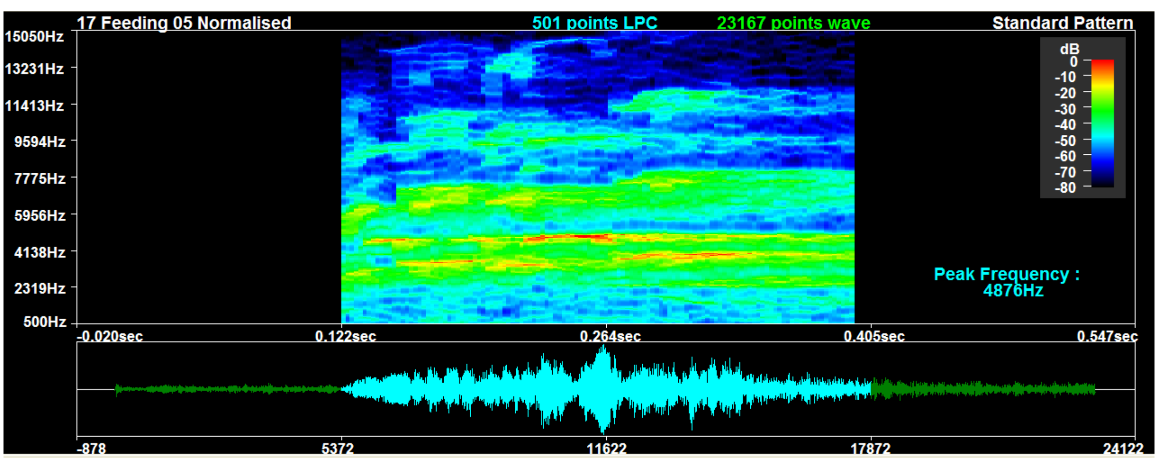


Figure 14: A King Parrot call?? More likely two King Parrot calls.

For comparison see the fuzzy FFT version of this spectrogram in Figure 15 below.

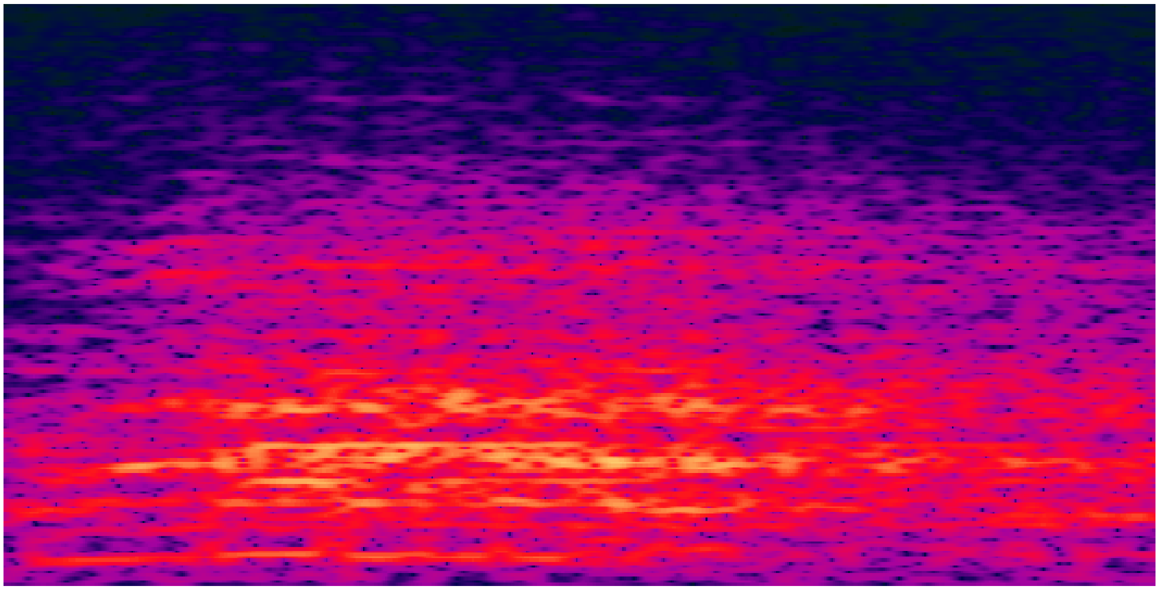


Figure 15: The FFT view of the call in Figure 14

Next we see in Figure 16 another example of this call. This time the envelope is very fuzzy and there does not appear to be any clear inter-modulation products. This is a flock of King Parrots each trying to be heard by calling at different frequencies.

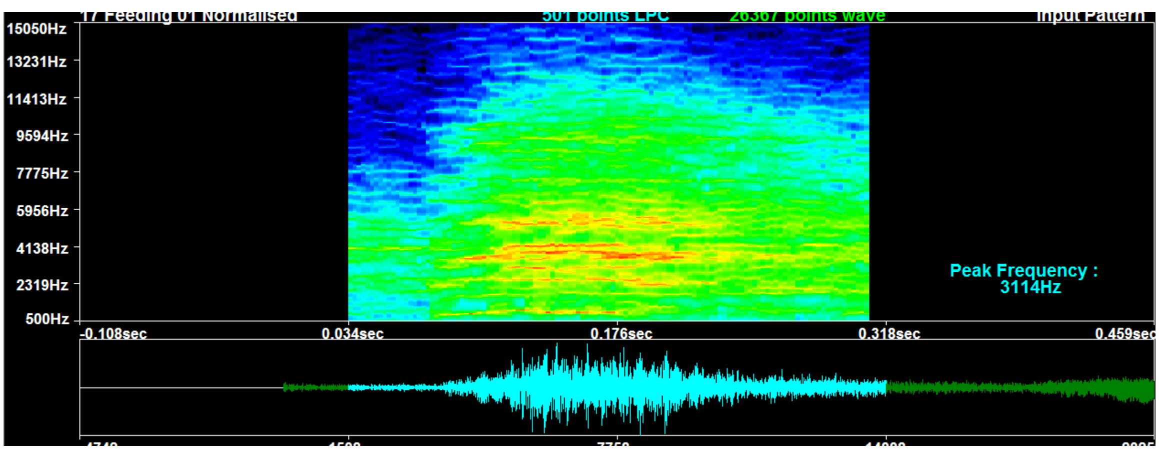


Figure 16: A lot of King Parrots calling at once, each “tuned” to a different frequency.

Conclusion

Here we have seen that spectrums that are usually interpreted as “harmonics” by bio-acousticians are not generally correctly described, even where we have deliberately chosen signals (calls) that look suspiciously like they have harmonics in both the FFT and LPC (if you don’t look too hard). However in most cases the “harmonics” are shown to be anything but harmonics. The Fourier Transform is also known as the Harmonic Transform because that it what it is.

In other unpublished studies, we have shown that the examples discussed above are far from being rare in birds, frogs and bats. It seems probable that birds and frogs generate these calls types to improve the chances of being heard. Pure tones, as they propagate, produce spatial interference patterns which cause them to reinforce and cancel at different locations. Where cancellation occurs, the calls will not be heard. The relative deafness of birds (compared to humans and other mammals) would suggest that they need to work hard to be heard.

Bats are another question. Bats also wobulate extensively and use other fancy modulation techniques. Why would bats generate such complex signals? Complex signals require complex decoders and when those signals are being used for location, it makes positioning more difficult (or does it, there are some interesting advantages of complex signals in the bat's regular environment). This is still being investigated and I hope to have some answers soon.

It has been shown that new science (learning things that were not previously known) can easily be done if the right techniques are used and properly applied. Avian calls are more complex that has to date been appreciated. All calls, and interpretations of those calls, need to be revisited with advanced spectrum analysis techniques, and to be re-evaluated. The examples used in this paper are not exceptions but are examples of common calls.

This paper is far from exhaustive and the reader is invited to download the SoundID software (www.soundid.net) and have a closer look at their own favourite calls to see what they are really all about. Much more remains to be learned about the true nature of bird and animal calls and even music and mechanical noise. Now that the proper tools are available there are no more excuses.