



ELECTRONOTES 229

Newsletter of the Musical Engineering Group

1016 Hanshaw Road, Ithaca, New York 14850

Volume 24, Number 229

March 2017

PITCH VS. FREQUENCY

-by Bernie Hutchins

INTRODUCTION

Often it is said that frequency and pitch [1] are the same thing. Engineers call it frequency and musicians call it pitch. Indeed, we use the same units, Hertz, (Hz); formerly both were known by the much more sensible description “cycles per second or CPS”. AC power has a frequency of 60 Hz. A CD has a sampling rate of 44.1 kHz. An orchestra tunes to a pitch of A=440 Hz when the oboe sounds during tune-up (before the conductor appears). Why, by the way, the oboist? Who put him/her in charge and why? The answer is not exceedingly good, but it is because the oboe sound is harmonically rich, penetrating, and the pitch is very strong. Why not a flute, which is a purer tone (more sinusoidal); one frequency instead of a fundamental and multiple harmonics? The oboe is better for the intended purpose: it's EASIER to match to. There is a lesson there.

One often-cited distinction is the claim that frequency is an objective attribute while pitch is a subjective attribute. (True, but subject to over-interpretation.) The reason is that we can often make fairly good objective definitions that allows us to take in data on a tone of some complexity, and calculate a number for pitch that is, more or less, exactly what we wanted. In the case of the oboe, for example, this is what we can do, although the oboe (and most other musical tones), are a combination of several or many frequencies. One frequency has one pitch. Many combinations of frequencies can have the same pitch. So pitch is a more general abstraction. A full determination of either frequency or pitch is of course (like most quantities) subject to some measurement errors. The measurement of frequency is usually more straightforward (a physical measure). The analytic determination of pitch will often require a more nuanced consideration, to allow for the subjective nature of the perception. We want the same answer the ear/brain would give, and we don't know

exactly how the pitch perception works in all cases. We easily construct frequency meters (counters). When we use a frequency counter (digital readout), it is generally the case that we expect the input frequency to have been set and to remain constant, such as turning a frequency dial of a function generator and then taking our hand off the knob before starting to count. Then the counter counts some feature (like well-defined sign changes) for a full second and displays the result. Things like pitch meters or pitch-to-voltage converters are still difficult, even to this day. Among the complications mentioned above, the pitch often may not be even relatively constant at all for more than a tenth of a second or so (such as a rapid flourish of notes on a piano). And, the pitch, as a right answer, might well be a curve, (a function of time) and not a single number.

Rather than try too hard to relate pitch to frequency, we generally find it useful to relate to a repetition rate of a waveform (the reciprocal of the periodicity time). This is often the same as the lowest frequency. For example, a waveform consisting of four frequencies, 200 Hz, 400 Hz, 600 Hz, and 800 Hz has a pitch of 200 Hz, the lowest frequency in this case, which is also the repetition rate (see examples developed later). We immediately dissuade ourselves of this simple interpretation by considering the case where we remove the 200 Hz component (keeping 400 Hz, 600 Hz, and 800 Hz) and still hear a pitch of 200 Hz quite clearly. There is no spectral energy at all at the 200 Hz pitch in this case. This is the so-called “missing fundamental”. On the other hand, we could have a tone produced by some oscillator that drives one of more “resonators” at higher frequencies. In this case, the pitch generally corresponds to the repetition rate (driving oscillator) even though the spectrum overall may have substantial energy at higher frequencies that are NOT harmonics of the drive rate (examples to come).

TEST SIGNALS FOR PATTERN REPETITION - ADDITIVE

In electrical engineering in general, we are often interested in using the notion of frequency. As an example, a radio station is assigned a particular carrier frequency. In the engineering of audio signals, intended for the ear as a receiver, often it is pitch that gets our attention.

Specifically in the case of music synthesis, the most central topic of our publications here, we are often interested in pitch as the carrier of a musical melody. That is, we intend that our equipment and methods impose a series of pitches as a synthesized musical signal. In this case, we need to know how pitch is implanted in a signal. We are thus trying to understand the pitch of a wide variety of waveforms that repeat a pattern. The production of these “test tones” for pitch studies is quite analogous to producing the tones to offer as music. The possible approaches, additive synthesis, subtractive synthesis, and modulation synthesis are very similar, if not identical [2].

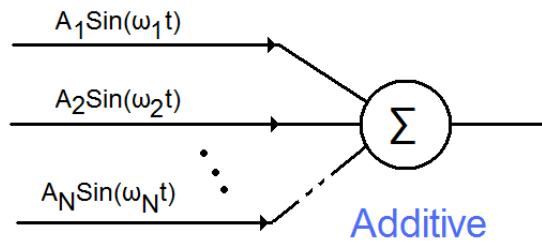


Fig. 1

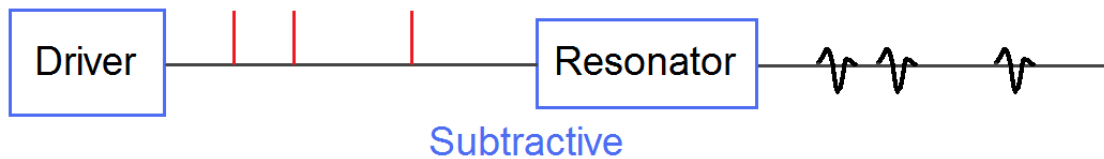
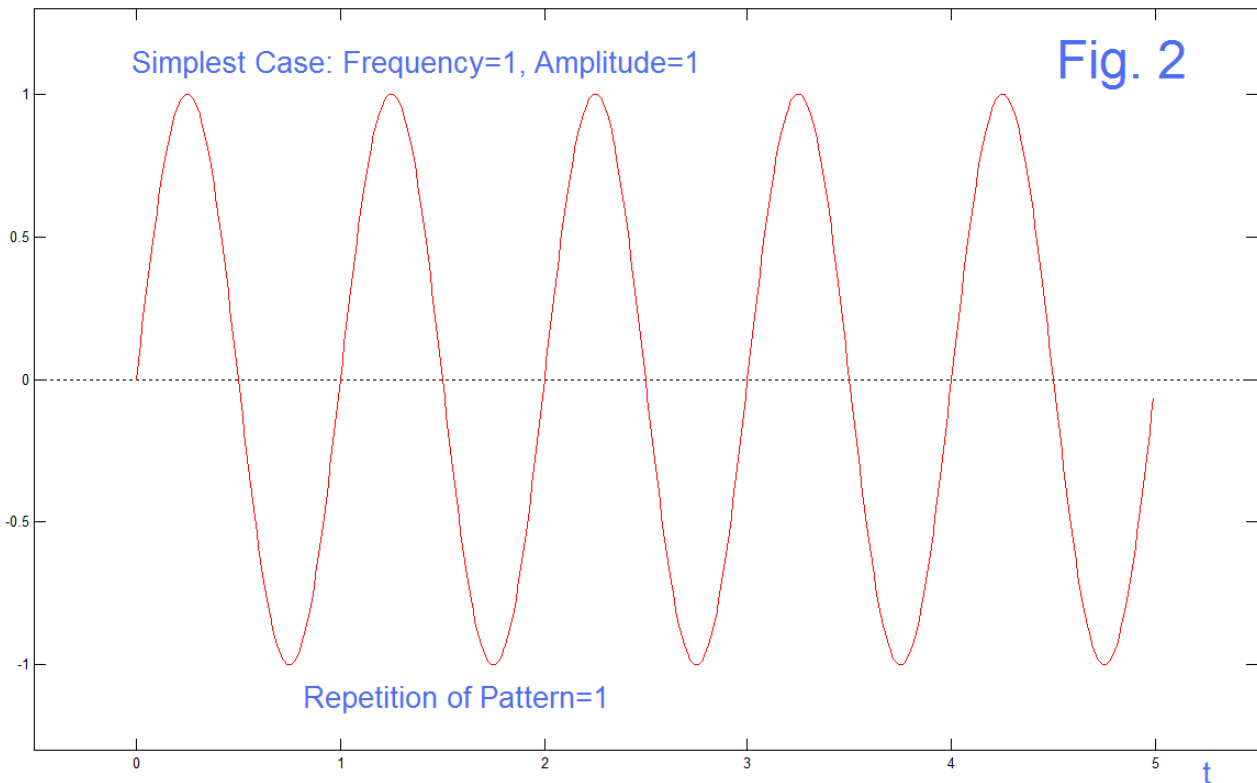


Fig 1 (top) shows the simple (Additive) approach where a series of sinusoidal signals are added (any arbitrary phases could be used). Here note that the signals being added are exactly periodic. The sum may also be periodic, if the frequencies are all integer multiples of some fundamental (which need not itself be one of the frequencies). The pattern repeats with the period of this fundamental frequency (Fig. 3). If the frequencies are not harmonics, the sum will not be periodic (Fig. 5), at least not exactly. To the extent that pitch is determined by repetition rate of a pattern, there is no pattern in this non-harmonic case. It is probably clear to the reader that the ear/brain in seeking periodicity may accept (or tolerate) some degree of approximation. It may in fact find an acceptable notion of a pitch, even if the sound is very rough or uncharacteristic of a harmonic sound.

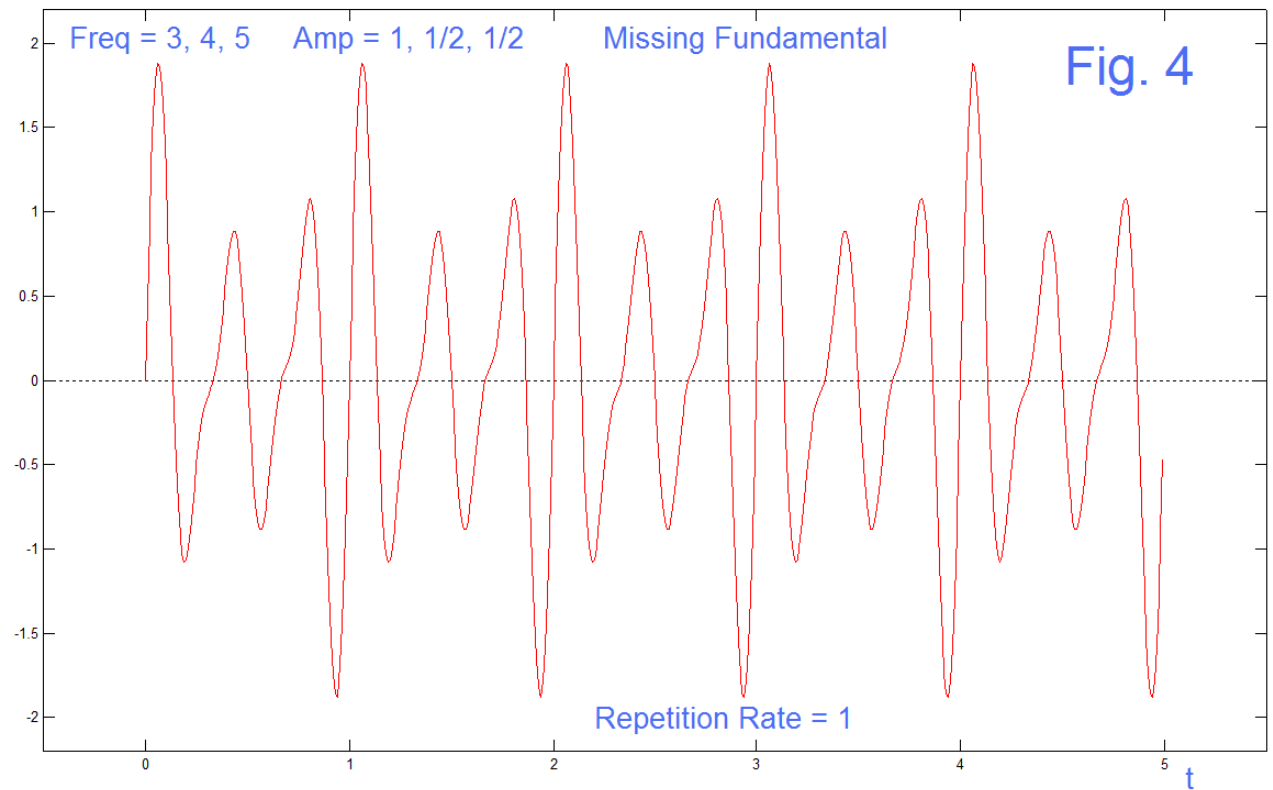
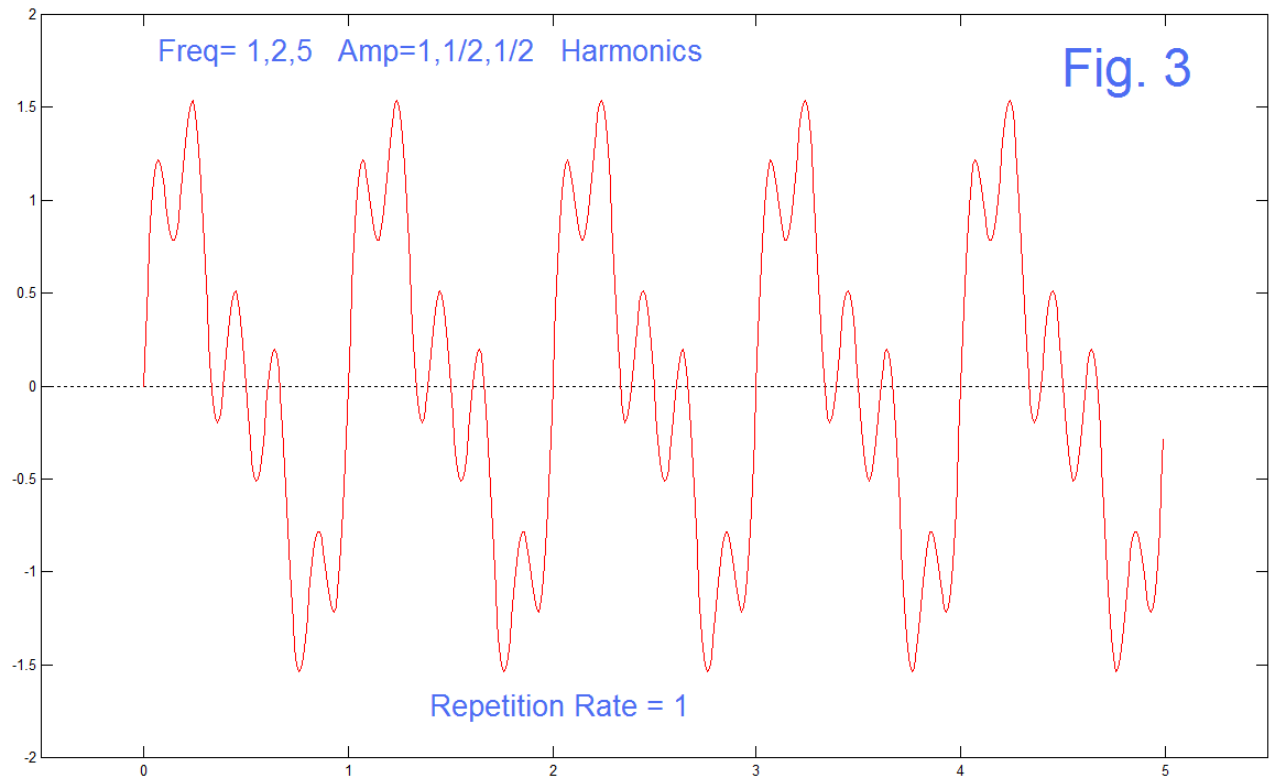
The second scheme in Fig. 1 (Subtractive) is shown here in terms of a “Driver” and a “Resonator”. The resonator is a device that responds to an input event from the driver giving a characteristic waveform. We look at this as something quite general, although the “ringing” or “ping” of a band-pass filter is traditional. The driver likewise is intended to be something very general, but is often an oscillator. [In the usual music synthesizer, the driver would be a voltage-controlled oscillator (VCO) while the resonator would be a voltage-controlled filter (VCF). Here we tend to suppose that the resonator (the filter) is fixed and NOT a tracking VCF; the latter would be the case with most music synthesizers.

Is the output of the driver/resonator perfectly periodic? If we assume (or assure) that the resonator’s response decays rapidly enough, then the resonator events are identical and have a spacing that is the same as the driver events. That is, the output is periodic if the driver is. That’s rather obvious and simple enough. The interesting thing is perhaps that the characteristic output of the resonator itself need NOT have components that are related to the repetition rate of the driver. This resonator response will affect the timbre (tone color), and perhaps the strength of a pitch impression, but not generally the pitch.



The simplest case (using Fig 1, top, Additive) of a signal with a clear frequency and a perfectly equivalent clear pitch is a sinewave (Fig. 2). Here the signal is generated as $\text{Sin}(2\pi t)$, a frequency of 1 Hz. Note as well that the repetition rate is 1. This is no surprise. We show here 5 full cycles mainly to better make the point about repetitions. We should perhaps also note that when it comes to listening to the sine, a signal as short as just 5 cycles is not adequate (the ear/brain requires more – we used something like 150 cycles in our tests). Also, let's admit that we are looking at a pitch of 1 Hz as a normalized value (and this will continue below). We can't actually hear a pitch as low as 1 Hz. The point is basically the famous “uncertainty relationship” (best known in physics perhaps) that says that if you want to know a pitch fairly well (a frequency) you will need to have a long enough signal (many cycles). That is $\Delta f \Delta t > (\text{some constant})$.

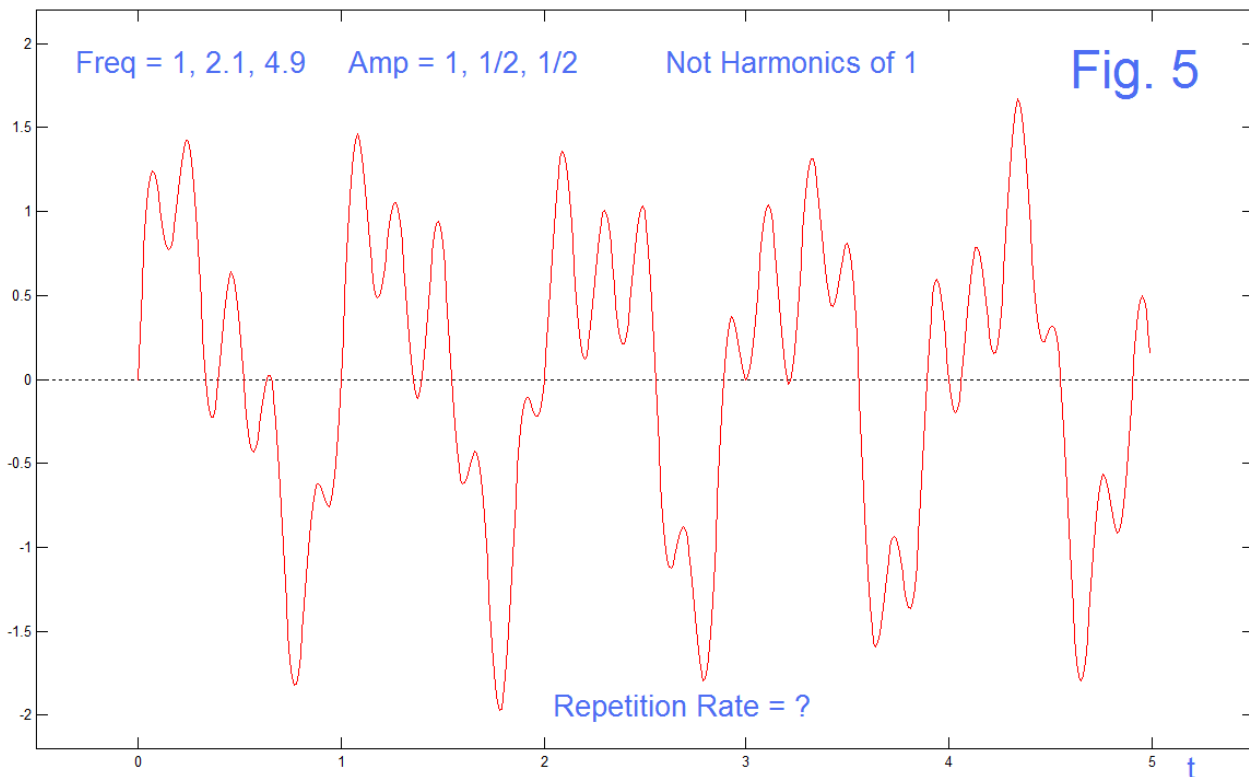
Perhaps surprisingly (as we hinted above) a sine wave does not have the strongest sense of pitch, In fact, it is often a poor choice for pitch matching experiments, especially as we may be trying to match a test tone that has strong harmonics. (Something like a sawtooth or narrow pulse is often a better choice.) Fig. 3 shows a tone formed from a fundamental of frequency 1 (as in Fig. 2) with added harmonics of $f=2$ (second harmonic) and $f=5$ (fifth harmonic), the harmonics having amplitudes of $1/2$, just as an example. Note that the waveform is more visually complex. In fact, we can easily “see” the fifth harmonic as varying five times as fast as the fundamental. None the less, the waveform has five full cycles, and the repetition rate of the pattern is still 1. The pitch, despite the higher frequencies, is still 1.



Note that what we have done here is Fourier synthesis: a sum of sinewaves. We are accustomed to the reverse process of Fourier analysis where a periodic waveform is represented by a "Fourier Series" of harmonics. In a general case of Fourier Series, we

have what is usually an infinite set of components. Some may be missing (like all the even harmonics of a square wave). But we might suppose that with synthesis that if we wanted a pitch corresponding to the fundamental, we had best start with that fundamental. Nobody says you can't sum sinewaves omitting the fundamental – just that we wonder if we could possibly have a pitch of 1 in such a case? Yes we can – it's the famous case of the "missing fundamental". Fig. 4 shows the case where we use harmonics 3, 4, and 5. Thus we not only omit the fundamental, but the second harmonic! The experimental result (in fact, one exploited by pipe organ builders hundreds of years ago!) is that you do hear the pitch of 1 despite no Fourier energy at all at that frequency. So here is a good example of where frequency and pitch depart. Note that the one thing that clues us in here is that the repetition rate of the pattern is again, 1.

At this point, we want to consider non-harmonic sinusoidal components. We learn some things of great interest in looking at this. Fig. 5 shows the case where we have three frequencies: 1, 2.1, and 4.9. Note that these are roughly 1, 2, and 5, but are not of course, exact integer multiples of 1. A quick comparison of the first cycle of Fig. 5 with the first cycle of Fig. 2 shows they are somewhat similar. The subsequent cycles are not the same as the first, although there are approximately 5 of them as in Figs. 2, 3, and 4. We certainly can't talk about any repeating pattern based on the evidence here. So we probably would want to run a listening test at this point. Very roughly, it has the pitch of 1, although rough is the word that well-describes the sound itself. We probably anticipated a crude approximation to the first examples. No repetition rate. But WAIT!



The waveform of Fig. 5 does have a repeating pattern – we just have not shown enough of it here. The frequencies 1, 2.1, and 4.9 are the 10th, 21st, and 49th harmonics of a fundamental of $f=0.1$! So shouldn't the pitch be 0.1? After all, we have agreed that a fundamental and other harmonics can be omitted. Well, the problem is that this not a rule, but rather an observation of how the hearing mechanism has apparently evolved to accommodate reasonable variations which are accepted as being the same thing, apparently to the advantage of the hearer.

If you want to argue that 0.1 Hz is the pitch, then we can write two decimal places (or three, or four) instead of just one. Eventually, the claimed pitch will be too low to consider. Before that even, the spacing between harmonics is excessive for hearing the tone as having a single pitch – you will “hear out” the individual frequencies.

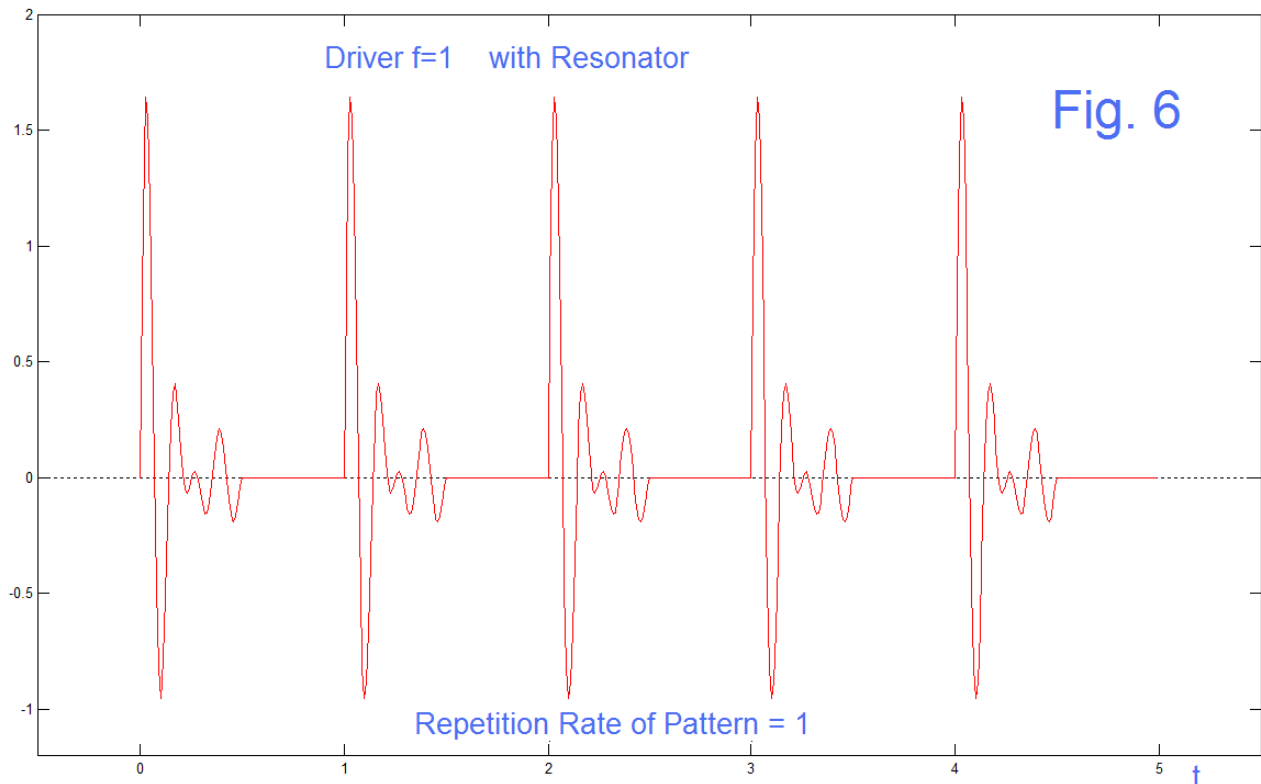
So we don't want to push this idea too far. On the other hand, few musical instruments produce perfect harmonics in the additive manner suggested. (See the resonator concept of a driven case just below here.) There are imperfections. The ear/brain, in many cases, does successfully allow for reasonable approximations, as though the harmonic case is a template to be fitted. For example, a lot of percussion instruments (like bells) have imperfect harmonics. Such sounds also generally decay relatively rapidly (Δt is small) so that in a precise assessment of pitch, the notion of achieving a very narrow sense of pitch (a strong pitch) would be forbidden by the uncertainty principle (Δf must be large).

PERIODIC EXCITATION OF RESONATOR - SUBTRACTIVE

We move now to the case of a periodically driven system (the Subtractive case of Fig. 1). We are still interested in the repetition rate of a pattern as determining the pitch. In the additive examples, the details of this repetition were determined by the components added. Here the construction is more direct. You decide what the pattern is going to be, and attach the repetitions to some driving signal. Fig. 6 shows an example. Rather than a description of what the pattern is, we will give the code for the reader to study:

```
tt=0:.01:.49;
ss=sin(2*pi*6.08765*tt)+sin(2*pi*8.054321*tt);
ex=exp(-5*tt);
ss=ss.*ex;
dr=[1 zeros(1,99)];
dr=[dr dr dr dr dr];
ss=conv(ss,dr);
ss=ss(1:500);
```

Even if you don't use Matlab, you can likely see that we are forming a length-50 sequence of two sine waves, applying an exponential decay (and truncating it at length-50), and then repeating it at impulses spaced at length-100. Hence a pattern is constructed and repeated. Two things too note: We have formed our pattern with two sine waves that are not related



harmonically to the rate of the driver. That is, the frequencies 6.08765 and 8.054321, not harmonics of 1. We could have chosen harmonics, but the more general case is more interesting. Also note that, because we truncated the decaying exponential (and sine waves) at length-50, the resonator patterns fit entirely within the spaces between the impulses (impulses not shown). We could have chosen resonator patterns longer than 100. If we did we might not have repeats that are entirely typical until the middle of the tone. But in the case of Fig. 6 it is exactly repeating, and we can suppose that other cases will at least be good approximations. What have we got? A periodic waveform that can be understood in terms of a Fourier Series!

Accordingly, the waveform of Fig. 6 is composed of a fundamental of frequency $f=1$, and harmonics of $f=2, f=3, \dots$ on to infinity. (The series does not truncate because of the flat region of all zeros.) Is there anything special about the frequencies 6.08765 and 8.054321? These are fairly close to 6 and to 8, and we might expect more substantial spectral support in the vicinities of 6 and 8. That is, the harmonics 6 and 8 would be expected to be larger than others.

A clue of what is going on can be found by observing from the generating code in line 7 that a convolution of time waveforms is involved. This is a filtering – the multiplication of the spectrum of one of them by the spectrum of the other. One of them can be considered a frequency response. Which one? Either one. Taking the filter input to be the impulse train, its spectrum is flat (also an impulse train). This pulse train is convolved (in time) with ONE cycle of the resonator. Thus the frequency response of the resonator (resonators often ARE filters) determines the shaping of the Fourier Series. This is pretty much the

same thing as finding the Fourier Series coefficients by taking the Fourier Transform (continuous time) of a single cycle and sampling it.

Reversing the analysis we still convolve the resonator shape (“impulse response”) with an impulse train. But in this case, the “filter” is not the resonator but rather has an impulse response that IS an impulse train (thus a summer followed by a time delay of 1, with a feedback of +1). The impulse response of the resonator is thus input and recirculated forever (again we get Fig. 6). The frequency domain impulse train IS the frequency sampler for the Fourier Series. We don’t usually think about this alternative view, but they have to give equivalent results of course.

But these shows us clearly why the repetition rate of the pattern is 1, and thus the pitch is 1. It is just another way to construct a set of harmonics (Fourier Series coefficients). The details of the resonator’s transform determine the weights.

We want to look a bit more at the way a resonator leads to a pitch at the repetition rate of the driver, and how we can eventually get a pitch that corresponds to the resonator itself. First however, in Fig. 7 through Fig. 10 we look at a case where we can better see how the frequency response of the resonator determines a spectrum.

In Fig. 7, we begin with a pulse train (height 10) separated by 40 samples, and a sampling rate of 6000 Hz. The repetition rate is thus $6000\text{Hz}/40 = 150\text{ Hz}$ and indeed the signal has a strong pitch of 150 Hz. In the case of a listening test, some 12,000 samples are constructed, not just the 200 samples of Fig. 7a. The spectrum shown in Fig. 7b is obtained as the FFT of all 12,000 samples, adjusted so that the lower side of the FFT (all we need) runs from 0 to 3000 Hz. Indeed the FFT is flat. This we knew would happen because a FFT of a pulse train in time is a pulse train in frequency (with proper assumptions). The fact that the spectrum is flat does not contradict the strong pitch, since the spectrum is composed of just harmonics of 150 Hz.

Next we choose a resonator that is simpler than that of Fig. 6. We want it shorter than length-40, and we choose it to hint at something like a 6th harmonic. Specifically we use the length-12 sequence [1 1 1 -1 -1 -1 1 1 1 -1 -1 -1] which is samples of two cycles of a square wave. This is nothing more than a length-12 impulse response of a FIR filter. Every time an impulse arrives, it outputs the length-12 square sequence (instead of the decaying sinusoidal shapes of Fig. 6). In the case of 12,000 samples, it is no surprise that we hear the same 150 Hz pitch, as in Fig. 7. The repetition rate is 150 Hz and the harmonics are multiples of 150 Hz. (The pitch is just as strong, or stronger.) We can arrive at the signal of Fig. 8 either by convolution (FIR filtering is convolution) or we can just construct the signal with program code – to the same result (Fig. 8a).

Likewise we take the FFT of Fig. 8a and this is shown as Fig. 8b – a spectrum more interesting than that of Fig. 7b. We do only see harmonics, but they now have different amplitudes. We did not expect additional harmonics since this is a linear filtering. It is a filter (resonator) driven by a periodic waveform and the pitch is that of the driving waveform.

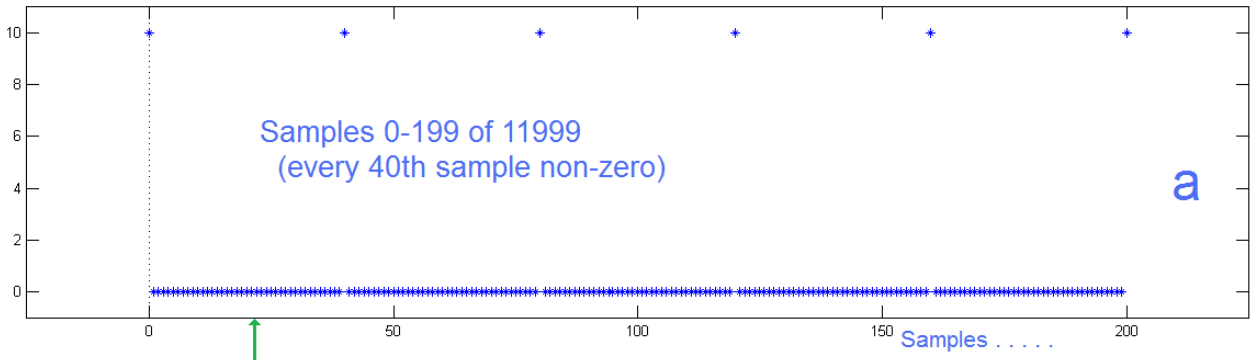


Fig. 7

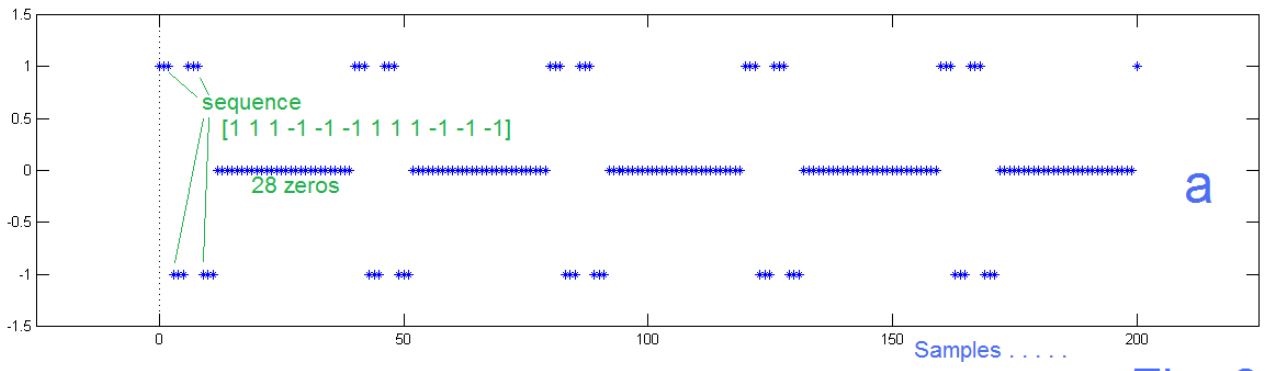
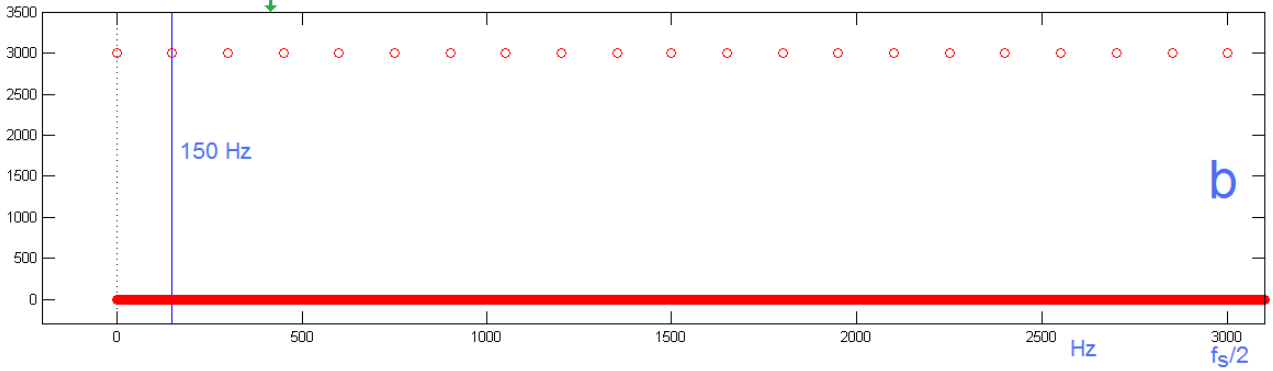
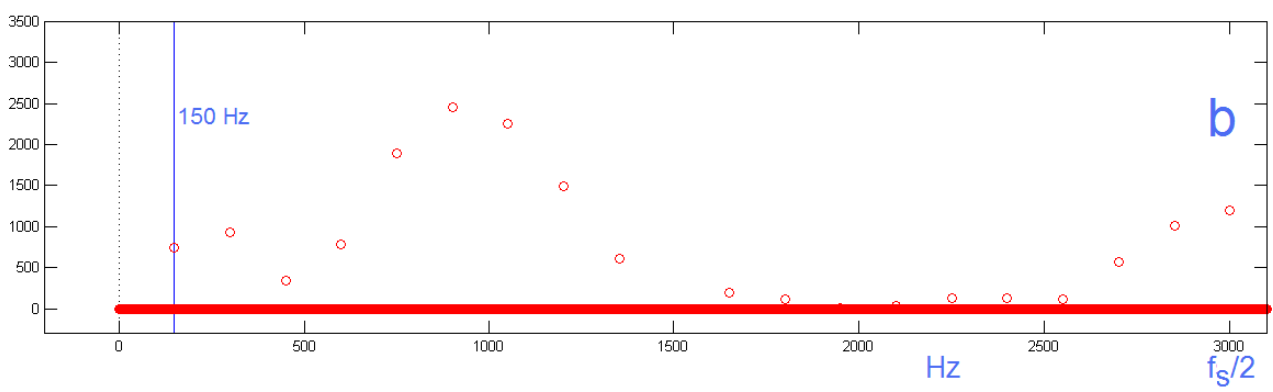
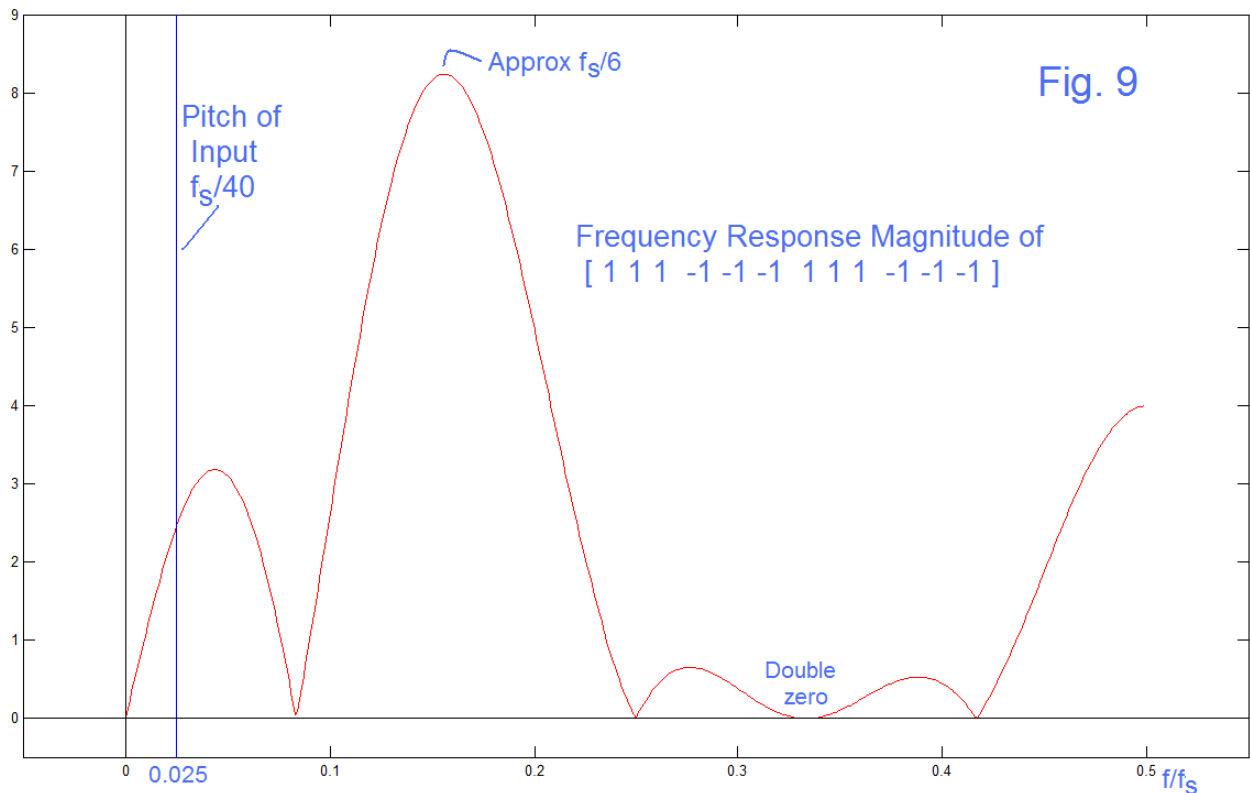


Fig. 8

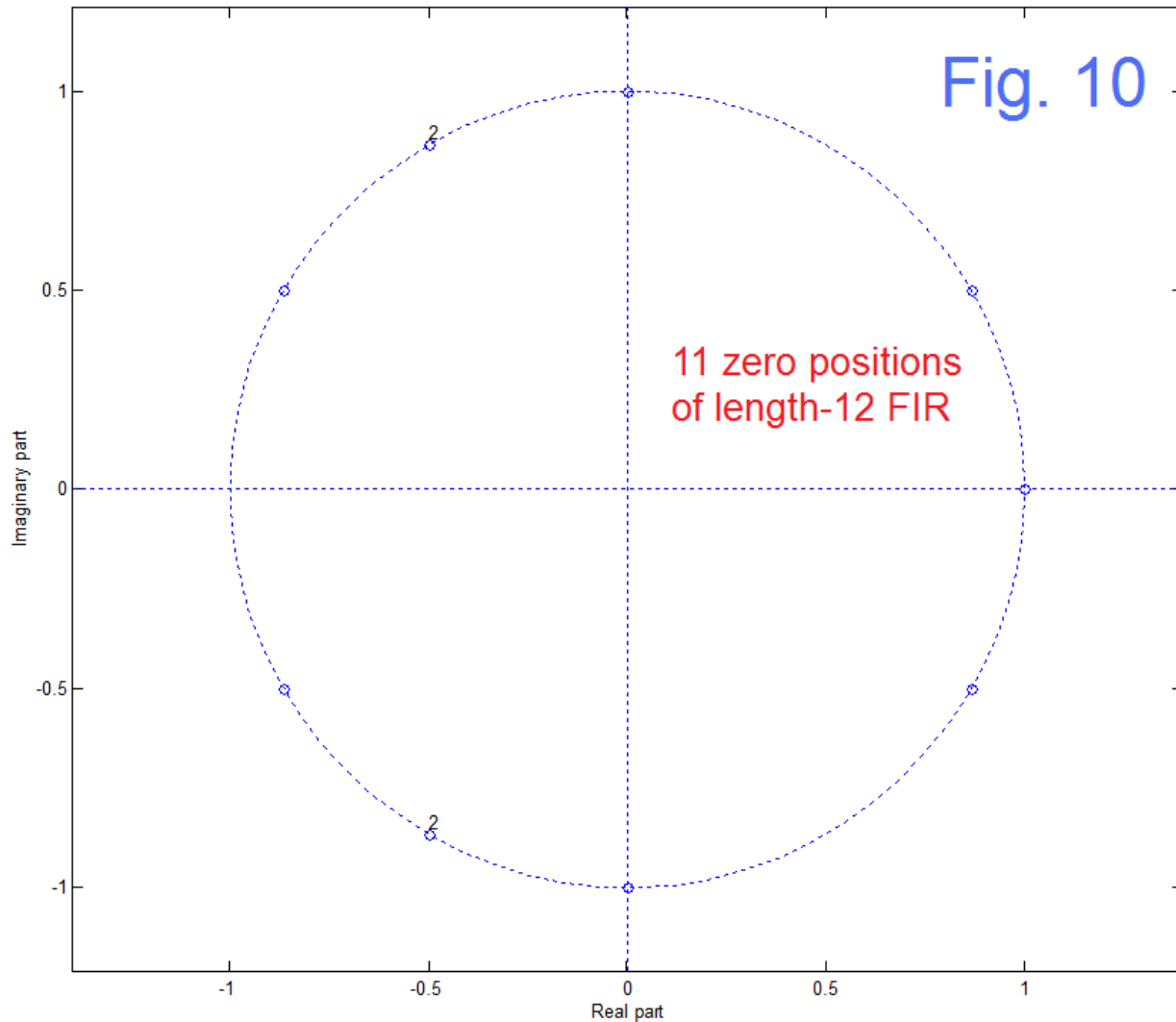


Is this right? Well, if we have calculated correctly, it should be. Still a simple test suggests itself – this is the filtering of a flat spectrum (Fig. 7b) and the result (Fig. 8b) should therefore “image” the filter’s frequency responses, and we have chosen a simple filter (although we do want to calculate the frequency response (as opposed to just hinting that it favors a region around $f_s/6$); like with Matlab’s *freqz*). Fig. 9 shows the magnitude of the frequency response and we note that indeed, Fig. 8b appears to be the filtering of Fig. 7b as seen in Fig. 9. Most signal processing engineers prefer to look not just at the frequency response but also at a plot of the zeros (and poles if any) of a filter, and the zeros of the filter are shown in Fig. 10. This helps us understand the peak around $f_s/6$, and the curious double-zero (flat at zero response) dip at $f_s/3$.



There is nothing truly surprising about the pitch with a resonator turning out to be the frequency with which the resonator is driven. Further, the analysis shows that the spectral content of the output is that of the input multiplied by the frequency response (filtering) of the resonator and this is consistent with our notions of subtractive synthesis and the pitch of harmonic series. Here we considered that the resonator differs from traditional subtractive synthesis (music) in that the filter does not track the excitation rate.

Thus the main point we want to emphasize here relates to the fact that driven sounds (bowed strings, and winds, including the human voice) have a pitch determined by the rate driven (not by the resonance), and a spectrum determined by both the excitation and the resonance. Note that the resonances may control the setup rate of the excitation (see APPENDIX), by a coupling mechanism.



AN ARTIFICIALLY AMBIGUOUS PITCH

What would happen if the resonator produced a segment between excitations that was just what we would have had for some higher rate excitations? This is almost certainly not going to happen naturally. Our interest is in what we can learn about pitch perception by writing code that does just this. In particular, to what extent could this “bridging” be made imperfect without evoking the driving excitation pitch? Not much - as it turns out.

The series of experiments shown in Fig. 11 (a-f) show an excitation consisting of a train of impulses that repeat once every 100 samples. Played at a 10,000 Hz sampling rate, this has a pitch of $10,000/100 = 100$ Hz. To each of the length-100 cycles we add a segment of a sine wave, chosen as a 300 Hz sinusoidal waveform (cosine actually). The program used here was named *pitch8.m* in Matlab (being the 8th program I wrote for studying pitch here). The program computes and displays a single cycle of the impulse driver, with a percentage of the sinusoidal resonator (in length). The code appears below.

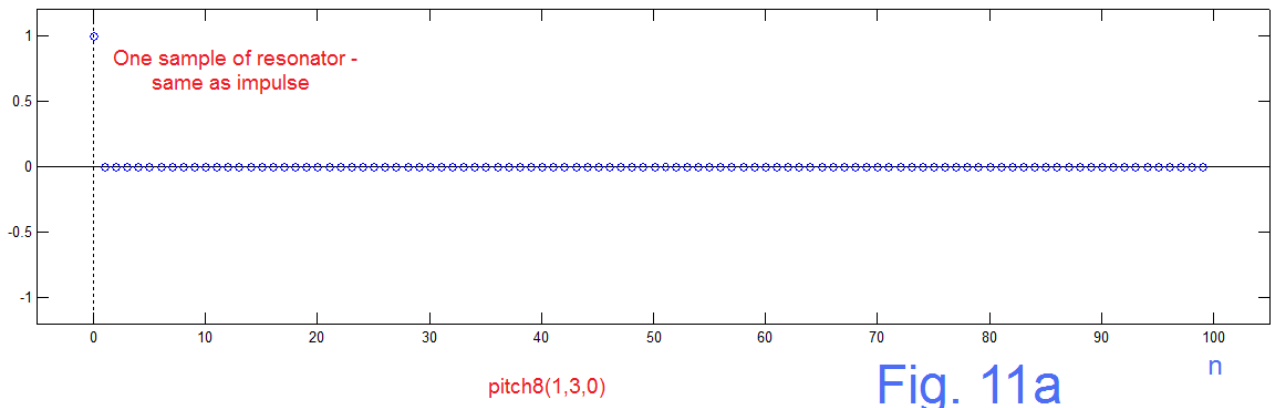


Fig. 11a

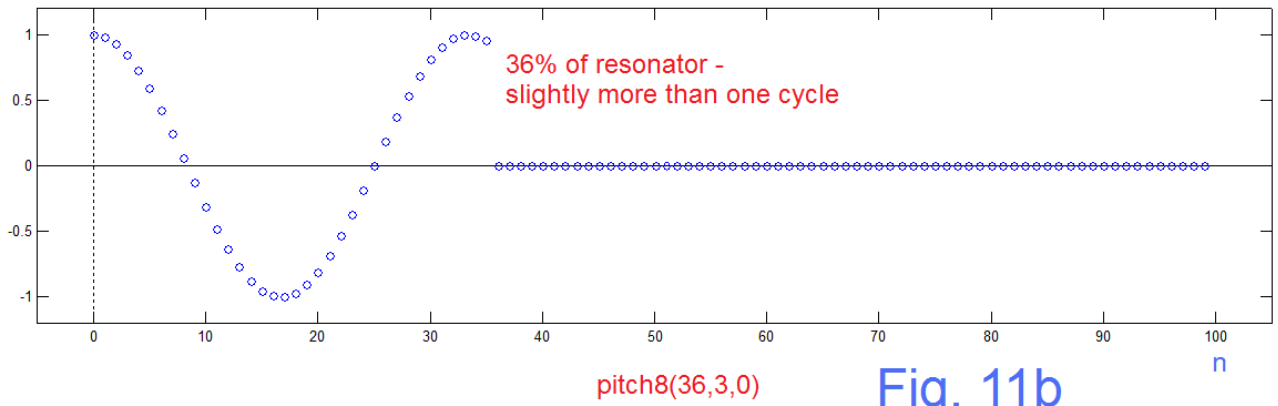
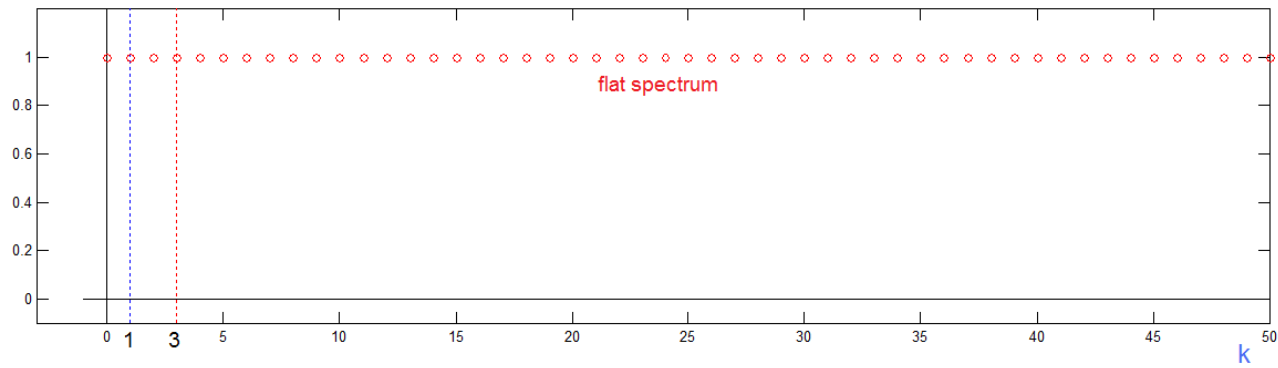
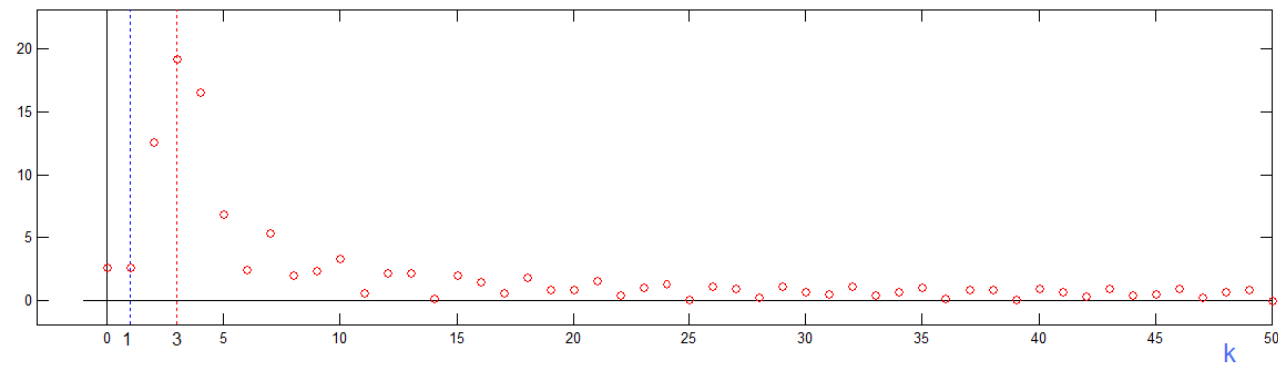


Fig. 11b



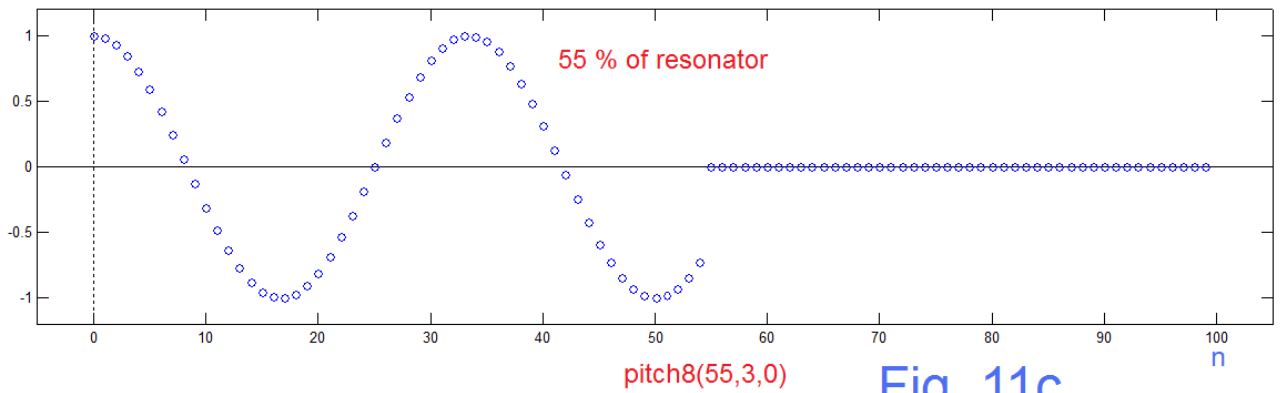


Fig. 11c

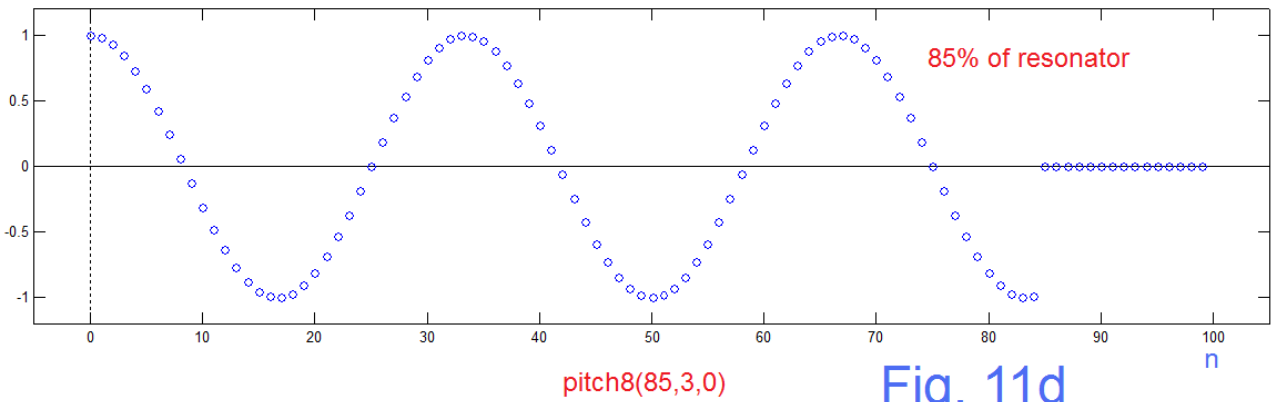
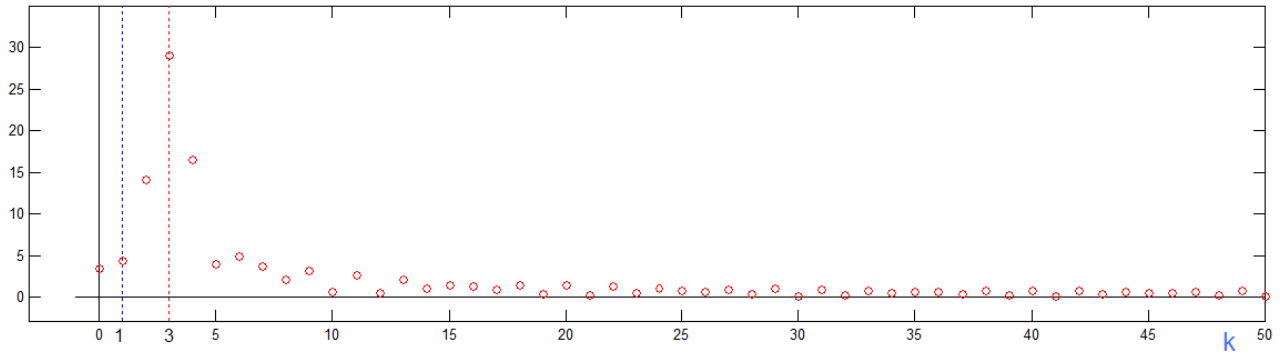
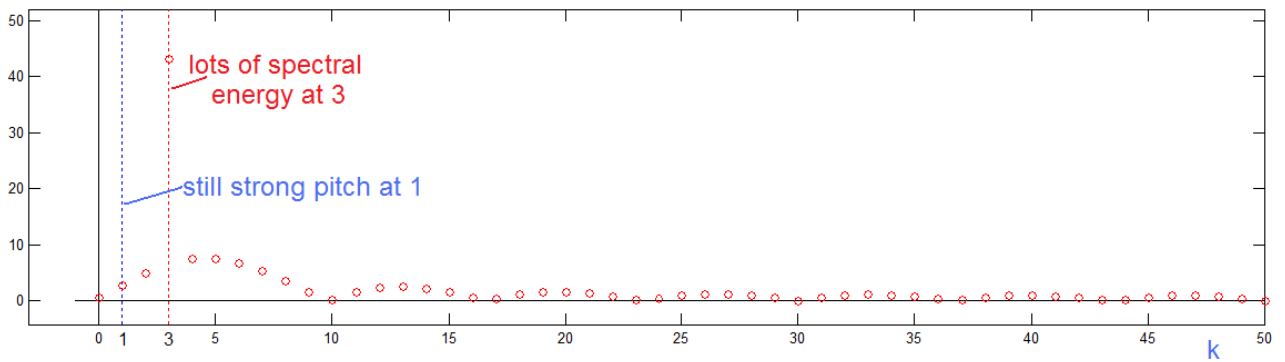


Fig. 11d



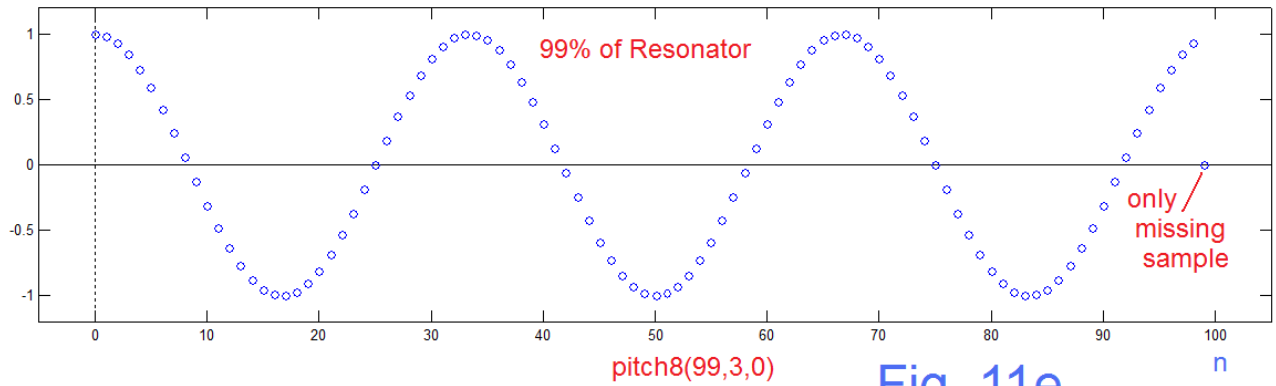


Fig. 11e

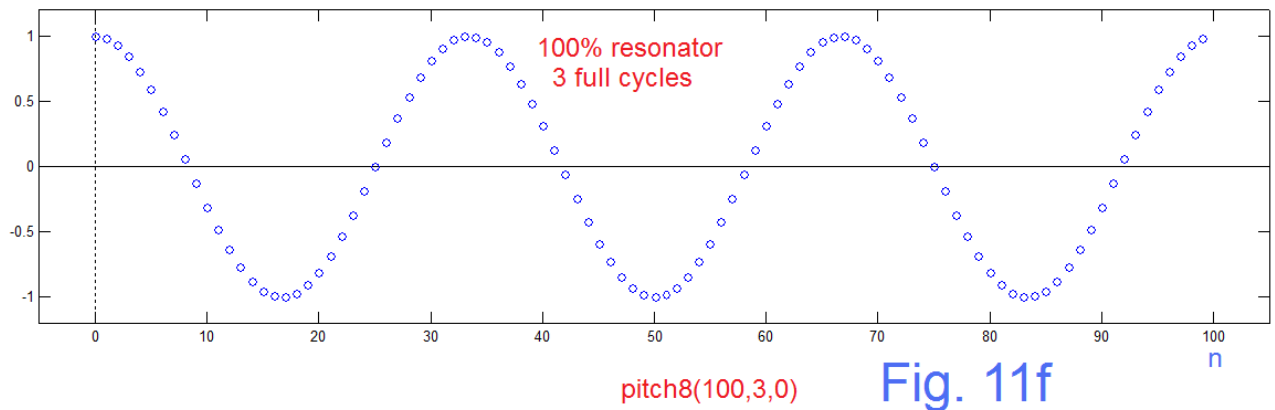
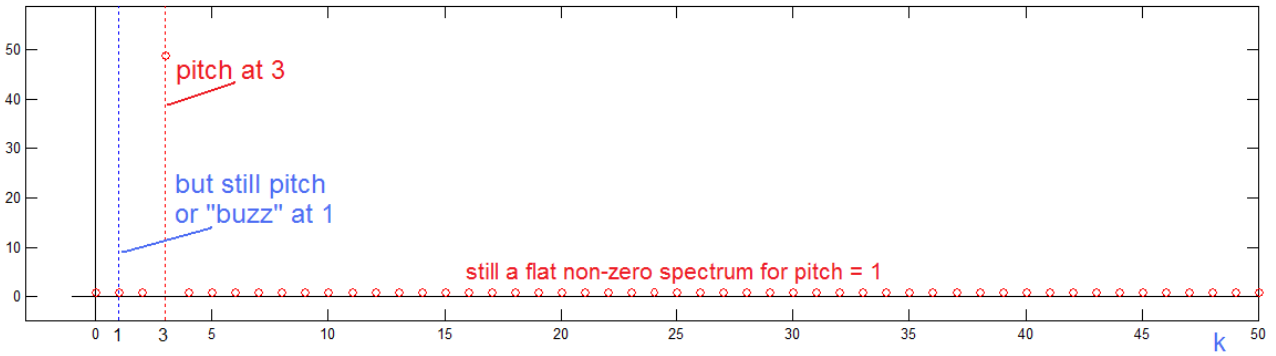
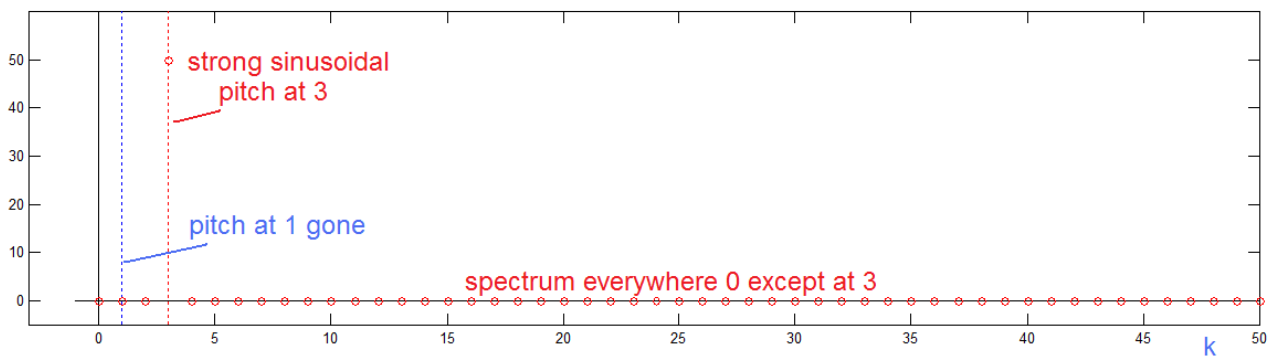


Fig. 11f



The extremes are Fig. 11a (1% - amounting to just the impulse train) and Fig. 11f (100%). In between we have four other resonator lengths. We emphasize that the use of the cosine is contrived for the purpose. The time sequences are in the top panels of the figures, and the corresponding magnitude FFTs are in the bottom panels (only 51 points are needed). As in the examples above, the plots are far too short for listening tests. Thus they are repeated 144 times for 14400 samples or 1.44 seconds of sound each. Fig. 11a has a very strong pitch of 100 Hz and Fig. 11f has a very strong pitch of 300 Hz.

Note well that while the 300 Hz pitch is the frequency of the resonating sine wave in Fig. 11f, and is seen as the only FFT component there, the 100 Hz pitch of Fig. 11a is represented by a flat spectrum (100 Hz plus a whole bunch of harmonics). This is because the 100 Hz case is a train of impulses, not a sinusoidal waveshape. We thus look for support for a 300 Hz tone as coming from a single FFT spike at $k=3$, while the spectrum for 100 Hz is a flat “comb” or components spaced at integers.

While we have spoken of pitch as being subjective, we need to keep in mind that this subjectivity looks at times a lot like an uncertainty or ambiguity. This results in a sense of pitch that varies in strength. It also means that the sense of pitch can be influenced by context (the history of presentations leading up to a particular test signal.)

For example, if we begin with Fig. 11a (1%, or the pulse train itself – using the cosine of 0), we hear a strong pitch of 100 Hz. Moving on to Fig. 11b, we hear a pitch of 100 Hz that is at least as strong as Fig. 11a. We note that Fig. 11b uses just over a full cycle of a 300 Hz as its “resonator” portion, and this is reflected in a 300 Hz peaking on the FFT as shown. Yet, the pitch remains 100 Hz in this presentation, and we are not especially aware of the 300 Hz component. That is until we play the 300 Hz tone (Fig. 11f) first, and then Fig. 11b. It is not true that the pitch changes to 300 Hz in this order or presentation. It is still a strong 100 Hz pitch. What does change is that we are more aware of the 300 Hz component. We start to “hear it out” as they say. In other words, a hint of more than one component is evident. It is no surprise that this “hint” gets stronger as the % of the resonator increases (Fig. 11c – Fig. 11e). Further, if we make the resonator portion even higher harmonics (like 400 Hz, 500 Hz, or 600 Hz say), the notion that there are two components with two different pitches (albeit harmonically related) is all the more evident.

Here we are learning about what the ear hears in terms of pitch, and this contrasts with what we might first “suppose” should happen. Fig. 11d (85% resonator) visually suggests that the 300 Hz is becoming dominant. We see more than two full cycles of the 300 Hz component, and the FFT has a spike at 300 Hz more than 5 times the surrounding components (it has finally appeared as an apparent winner). Yet the pitch is a strong 100 Hz and the 300 Hz component is not especially noticeable until prompted by first hearing Fig. 11f.

Here is the Matlab code for Fig. 11 a-f.


```

function pitch8(n,f,ea)
%
t=0:.01:.99;
si=cos(2*pi*f*t);
g=[ones(1,n) zeros(1,(100-n))];
ex=exp(-3*(100/n)*t);
em=ones(1,100);
if ea >0; em=ex; end
em;
s=si.*g;
s=s.*em;
figure(1)
subplot(211)
plot([-10 110],[0 0],'k')
hold on
plot([0 0],[-2 2],'k:')
plot([0:99],s,'ob')
hold off
axis([-5 105 -1.2 1.2])
S=abs(fft(s));
subplot(212)
plot([-1 110],[0 0],'k')
hold on
plot([0 0],[-100 500],'k')
plot([1 1],[-100 500],'b:')
plot([f f],[-100 500],'r:')
plot([0:99],S,'or')
hold off
axis([-3 50 -0.1*max(S) 1.2*max(S)])
s=[s s s s s s s s s s s s];
s=[s s s s s s s s s s s s];
sound(s,10000)

```

PITCHES OF A LESS “FRIENDLY” NATURE

Above, with the exception of Fig. 5, which has approximate harmonics and sounds rough, the tones offered are “friendly” in that in the company of recordings of acoustic instruments (perhaps trumpets, violins, oboes) they might be a bit on the bland side, too regular to sound like acoustically generated sounds (when extended for a full second or more), but not much of a puzzle. There exist however sounds for which a clear pitch is more obscure and which do not sound at all like known acoustic instruments. Such tones are generally of an artificial nature (produced by analog circuits or by digital synthesis). They serve first as curiosities and then, potentially, as test tones to probe theories of pitch perception. Here we will look (1) at tones that are obtained by a frequency shifting (of all harmonics by the same number of Hertz), (2) of the pitches of filtered noise, and (3) of pitches at band edges.

FREQUENCY SHIFTING

The title of this discussion being Pitch vs. Frequency, we would suppose that the notion of a pitch shift should differ from that of a frequency shift. The major issue here is that a circuit (or calculation) used for shifting one or multiple frequencies is straightforward (“single sideband modulation”); and we have a good idea what a frequency is in terms of a repetition of cycles and devices such as Fourier analyzers. Thus we know what the frequencies are before and after shifting. Pitch on the other hand is more subjective, and can be ambiguous, area often subject to context, and less determined by “rules” or formulas. True enough, a pure tone of say 440 Hz has both a frequency and a pitch of 440 Hz. Here we jump to an example.

Suppose we have the simultaneous presentation of sinewave components of 300 Hz, 600 Hz, and 900 Hz. This has a nice clear pitch of 300 Hz. If we were to shift all three components up by a factor of 305/300: 300 Hz to 305 Hz, 600 Hz to 610 Hz, and 900 Hz to 915 Hz we would have an equally good example of a signal of pitch equal to 305 Hz. Such a situation would be obtained with a musical instrument playing a different scale tone, or even naturally by a Doppler shift (about 12.5 miles/hour speed). A pitch shifting.

But a frequency shifter simply shifts all frequencies by the same amount. This would mean that a 5 Hz frequency shift would produce a complex tone with components of 305 Hz, 605 Hz, and 905 Hz. From the point of view of a pitch shift, the second and third harmonics (of 305 Hz) would be flat. As mentioned above, an automatic viewing through a “missing fundamental” viewer would call this a fundamental of 5 Hz (missing) and all harmonics of 5 Hz missing except the 61st, 121st, and 181st, which is absurd in that the ear does not ever handle a pitch as low as 5 Hz, nor would it allow for so many missing harmonics. Instead, the ear (and brain) would likely consider it an imperfect rendition of a harmonic tone – but of what fundamental?

Well perhaps the pitch is 300 Hz – since the difference between the three components is still 300 Hz, just as it was in the original case. Or perhaps the pitch is 305 Hz, the lowest component. But you probably suppose that the pitch perception mechanism is looking for some notion of a best fit. That is, a pitch of $605/2 = 302.5$ Hz. That is, the pitch assuming the middle frequency (605 Hz) is the correct 2nd harmonic. This would make the first component (305 Hz) a slightly sharp fundamental (above 302.5 Hz) and the third component (905 Hz) a slightly flat third harmonic (below 907.5 Hz). Although pitch matching is difficult, this last case seems to be what is found experimentally.

This three component experiment was popular because it was quite easily obtained using amplitude modulation (AM). The AM carrier became the middle component with equally spaced upper and lower sidebands (spaced at the modulation frequency) tracking this center. As such, the spectrum could be displaced (thus frequency shifted).

The procedure suggested at this point is to go to an audio test experiment listed here as **Pitch2**. This program has a number of adjustable parameters. By adjusting amplitude coefficients **g**, we can turn on/off up to 7 components of the test signal. The frequencies here were chosen 300 Hz apart, with an offset that varies from 0 Hz to 25 Hz, 2.5 Hz between trials. Thus the first trial has three components of 300 Hz, 600 Hz, and 900 Hz which is familiar. The second trial has frequencies of 302.5 Hz, 602.5 Hz, and 902.5 Hz. On the one hand, this seems a small difference. On the other hand, it sounds quite a bit different (not so much in pitch), but in what we can call “roughness”.

```
% pitch2.m
fo=[300 600 900 1200 1500 1800 2100]

g=[1 1 1 0 0 0 0]

t=0:.0002:2.9998;
for os=0:2.5:25
f=fo+os
s=zeros(1,15000);
for k=1:7
s=s+g(k)*sin(2*pi*f(k)*t);
end
figure(1)
plot(t,s)
sound(s,5000)
pause
end
```

It may be the case that the frequency shifted signals are perceived as being unnatural and in fact annoying. Setting aside this prejudice, what you will likely hear as being clear is a stepwise general upward pitch with each of the 11 trials. The lower frequency (from 300 Hz to 325 Hz in the 11 trials) is likely “heard out” in the corresponding examples. Persons who have been involved with our music synthesis efforts for years will likely soon recognize the sounds as those of modulated examples. (Except possibly for some bird songs, modulated sounds are quite unfamiliar in general.) The sound synthesizer user recognizes them as raw material leading to rich timbres and “clangorous” (like percussive) sounds.

Figures 12a through 12d show three seconds of sound each of four examples. The actual samples are so close together (some 15,000 of them) that we only see blue “blotches” which identifies the envelopes. This tells us what we need to consider. The cases are for the offsets of 0 Hz (a), for 2.5 Hz (b), for 7.5 Hz (c) and for 20 Hz (d). Note first of all that the case of no offset (Fig. 12a) has a uniform amplitude and is just the familiar harmonic case, as a baseline here.

In contrast, the case where the components are offset by 2.5 Hz (Fig. 2b) has a pronounced amplitude variation and that this variation is close to the offset frequency. This is very similar to a “beat” (for the same reason) and the rate is low enough to be generally annoying. We easily follow the amplitude changes going up and down. However, when the offset becomes 7.5 Hz (Fig. 12c) the depth of the amplitude variations is much as it was with the 2.5 Hz offset, but the variation is about three times as fast (as we might have

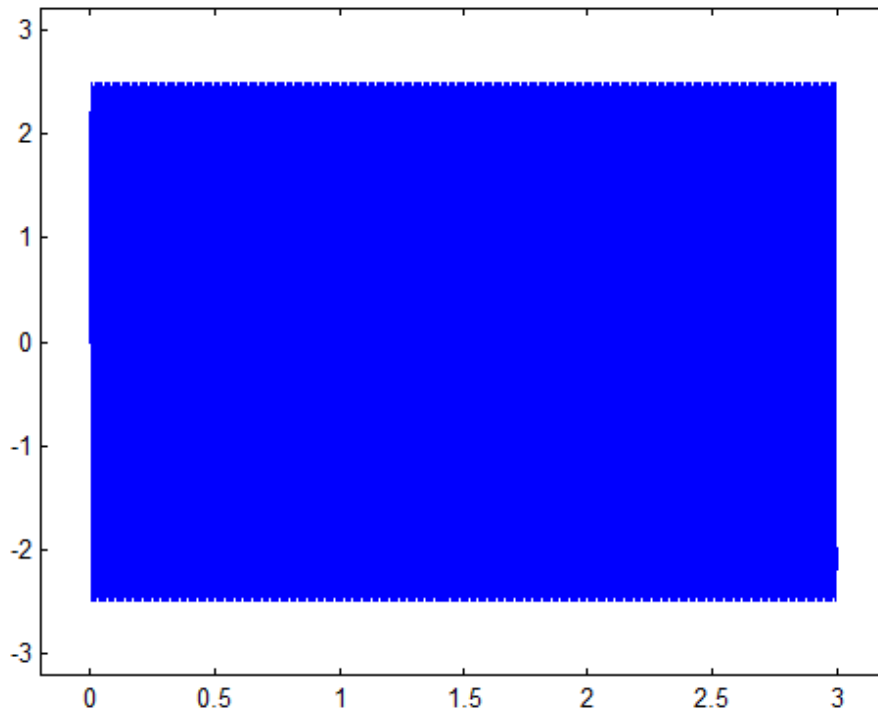


Fig. 12a

Components: 300Hz, 400Hz, 500Hz

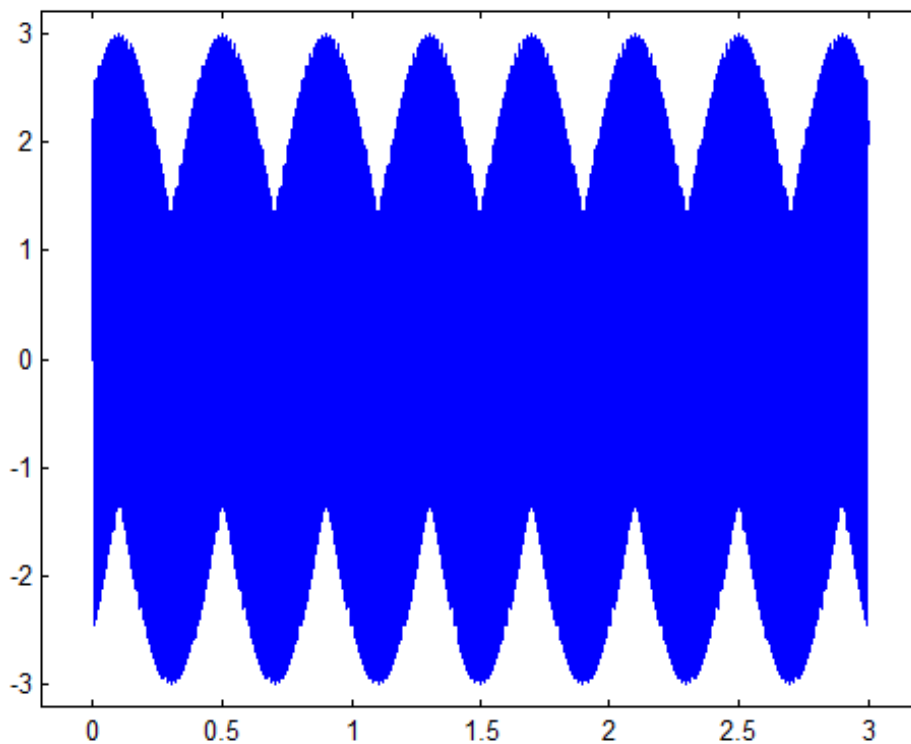


Fig. 12b

Components 302.5Hz, 602.5Hz, 902.5Hz

expected) and this puts it at about 7.5 Hz where we are in the range of ordinarily musical vibrato frequencies. This is the transition range where the variations are too fast to be followed individually but too slow for the general modulation impression. In consequence,

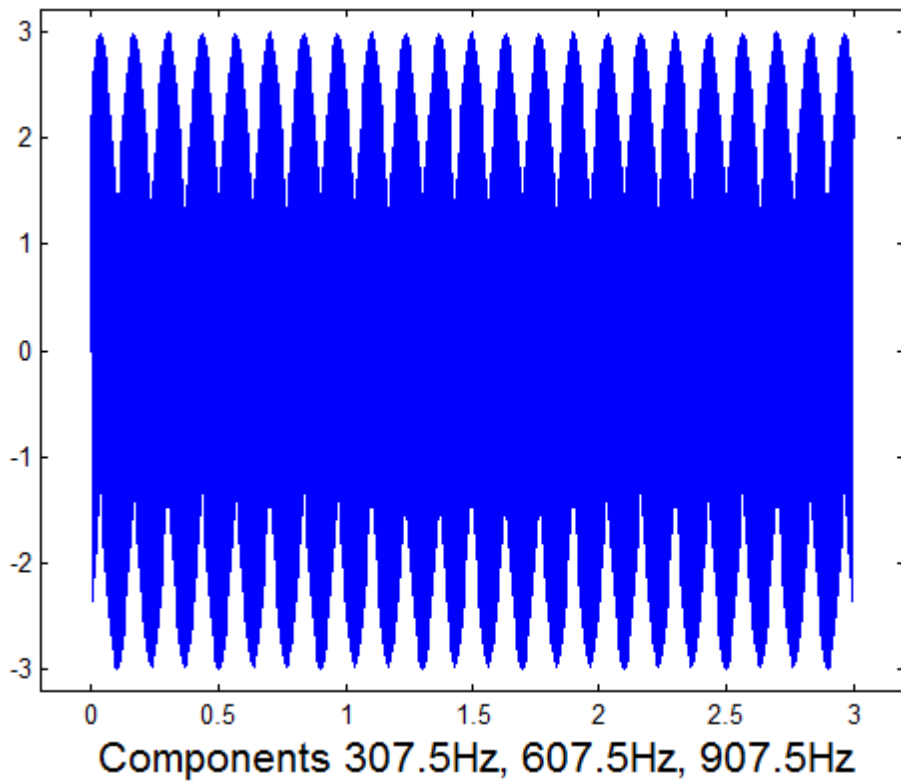


Fig. 12c

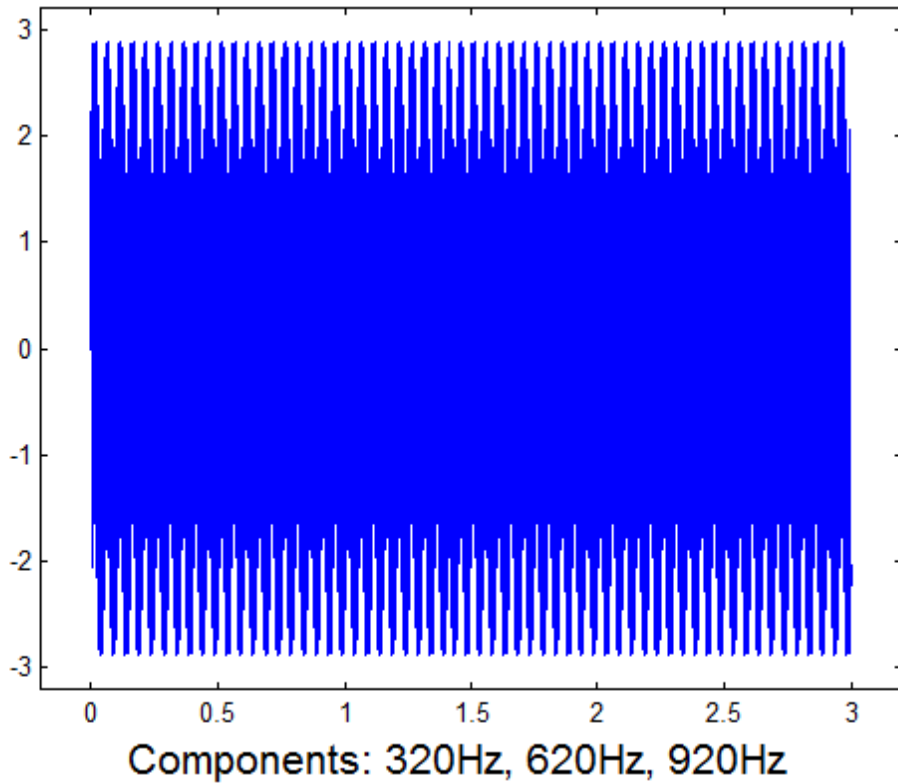


Fig. 12d

the sound is accepted as conventionally musical. By the time the offset reaches 20 Hz, Fig. 12d, we see that the amplitude variations get even faster to the point where it looks relatively uniform. The sound of Fig. 12d is decidedly in the range of modulation effects,

and is well above what one expects from vibrato. While mathematically the same as the acceptable vibrato case, normal musical vibrato is determined by what a human player can do physically. It is done by periodic motion of a hand, a finger, or of throat muscles. This is limited to perhaps 8 Hz (try shaking your hand faster). It is an acceptable and often quite lovely enhancement of the expression elements of a musical tone.

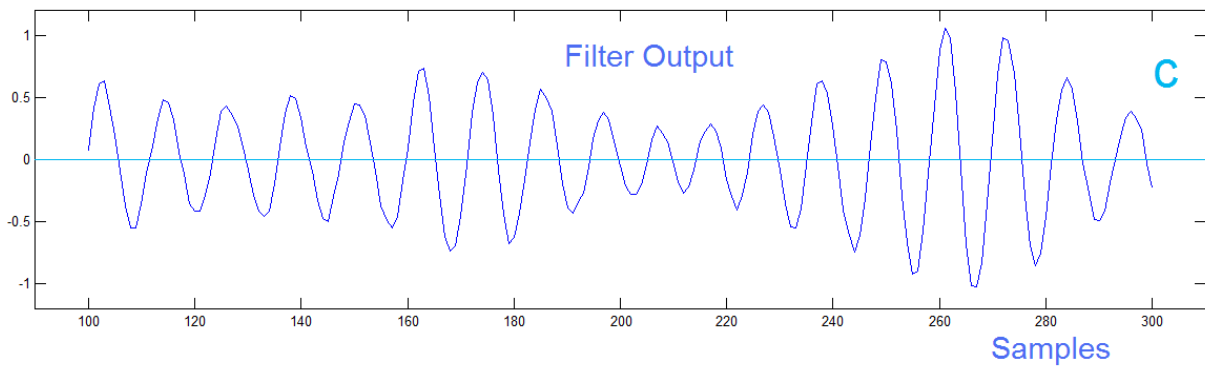
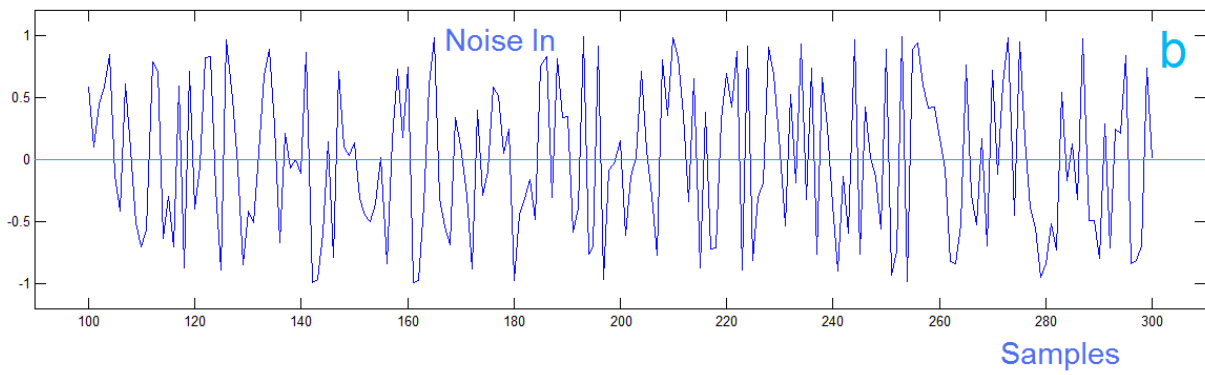
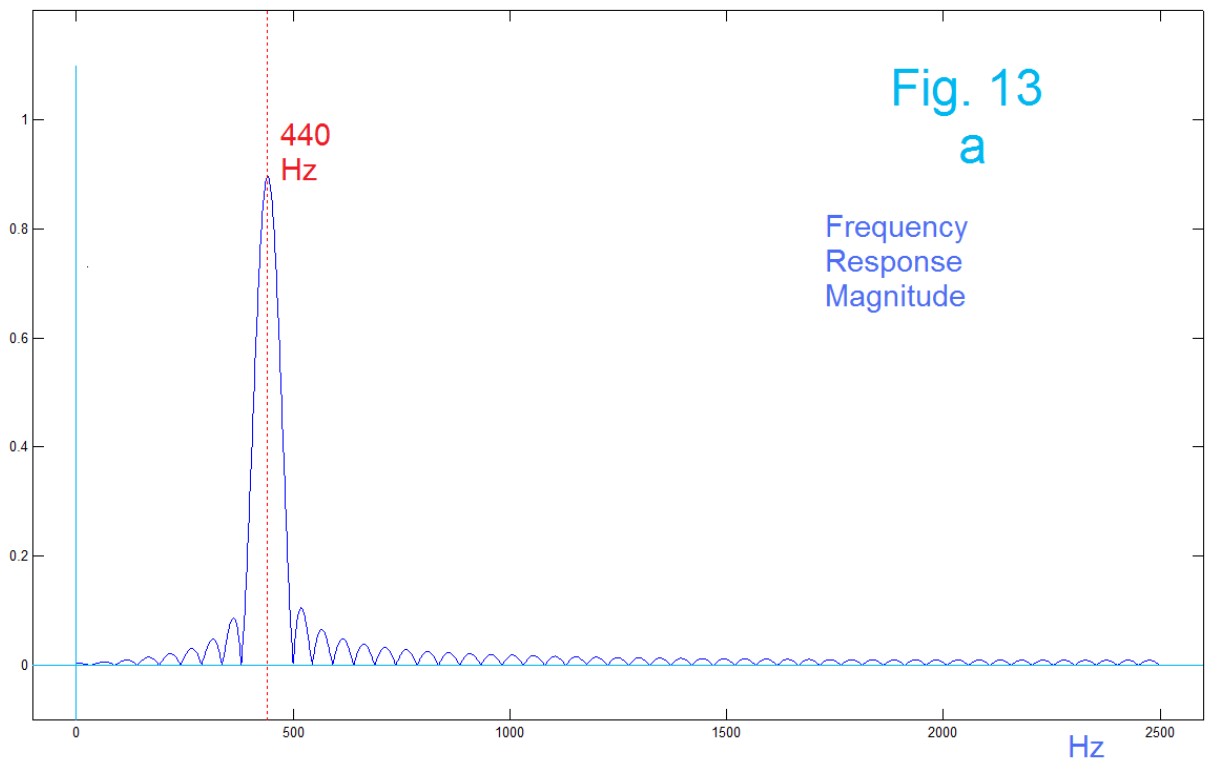
The question of the pitch of modulated signals is quite broad and available for experimentation. Here we used frequency shifting in an AM-like comparison, but other types of modulation such as FM are common. One thing to keep in mind is that modulated tones (and indeed, we suppose, unmodulated ones) in a musical context are not heard so much as experimental tones for pitch perception study as they are heard shaped in amplitude and in spectral aspects (like by filtering) as acceptable musical objects. [Careful study is suggested. For example, while I found that the pitch of the series of offsets seemed to allow a pitch at 300 Hz plus the offset to be “heard out” or matched, by the offset of 20 Hz, expecting a pitch match to 320 Hz, there was a better pitch match to the 22.5 Hz offset. So something subtle is going on, which I did not have time to investigate.]

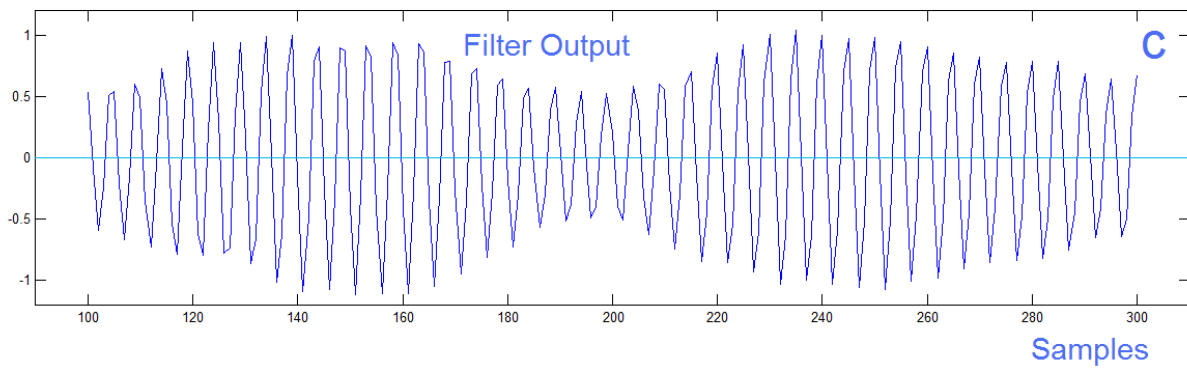
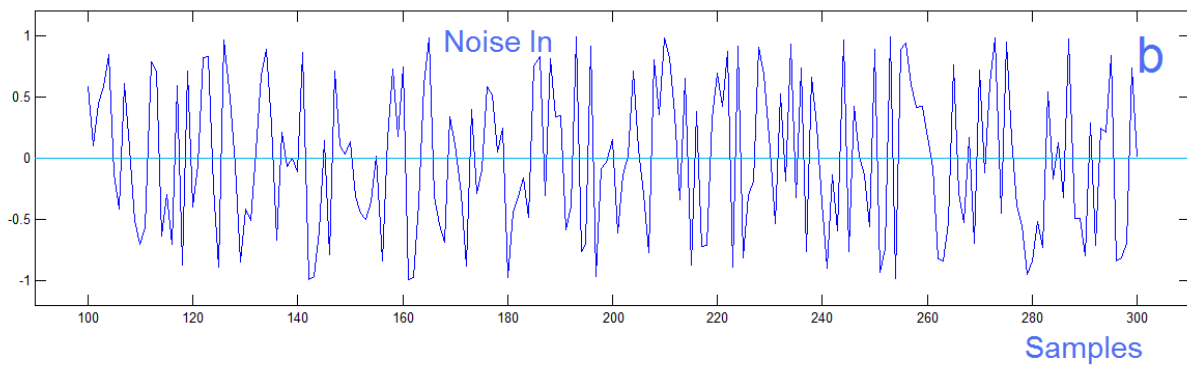
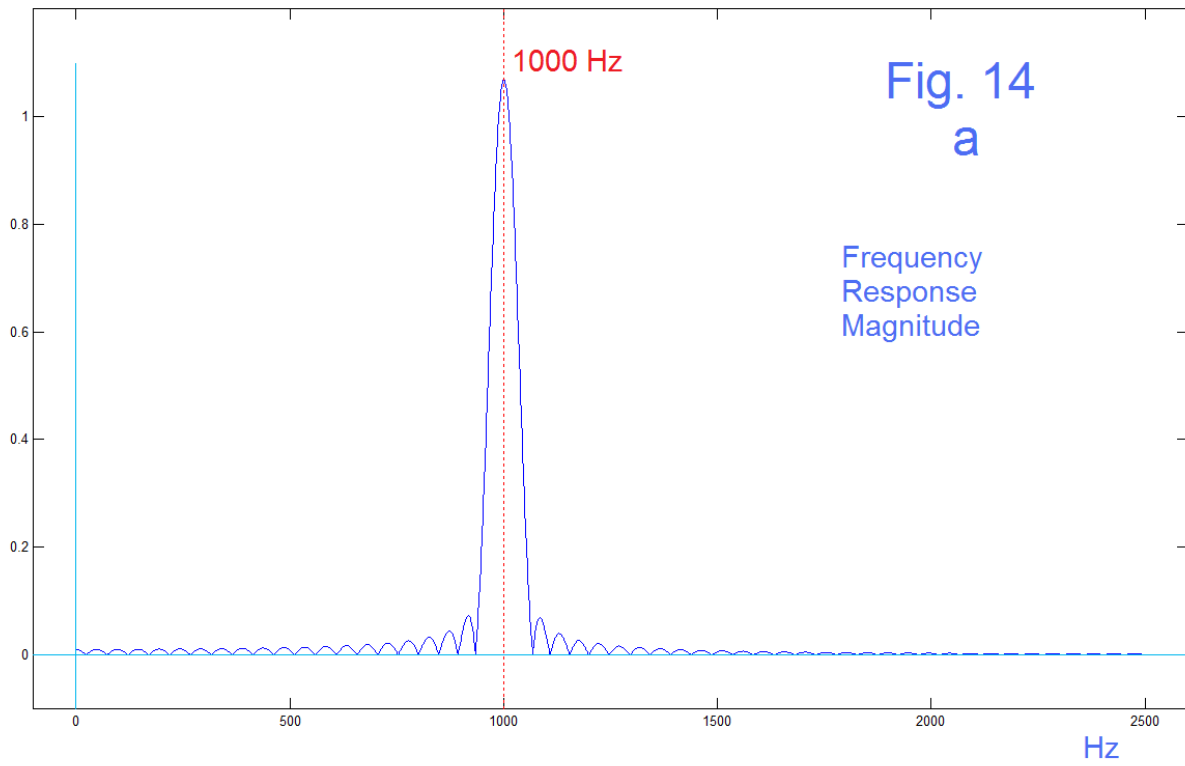
BAND-PASS NOISE, AND EDGE PITCH

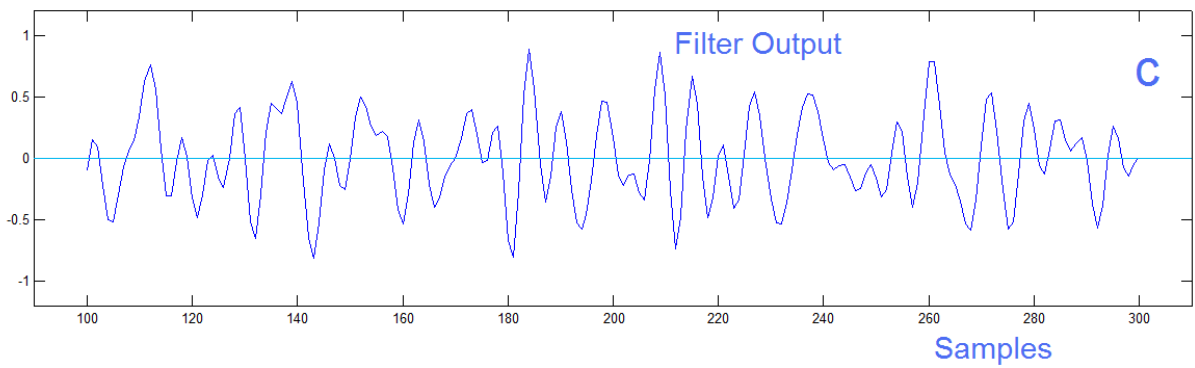
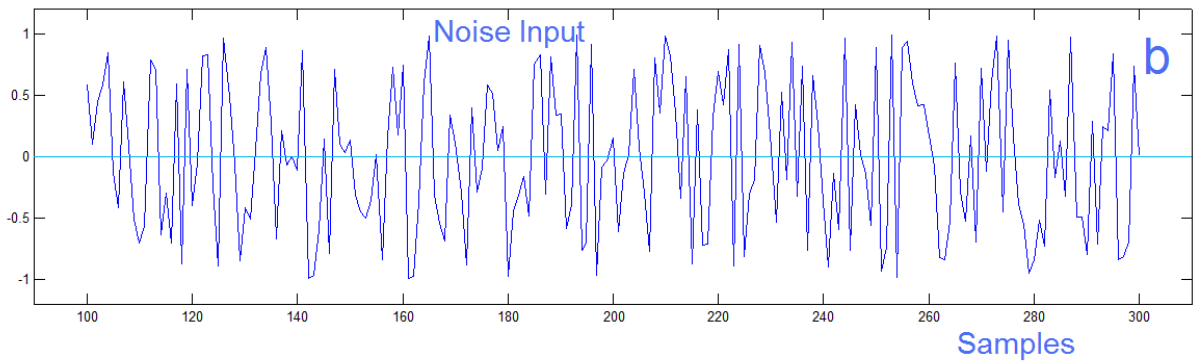
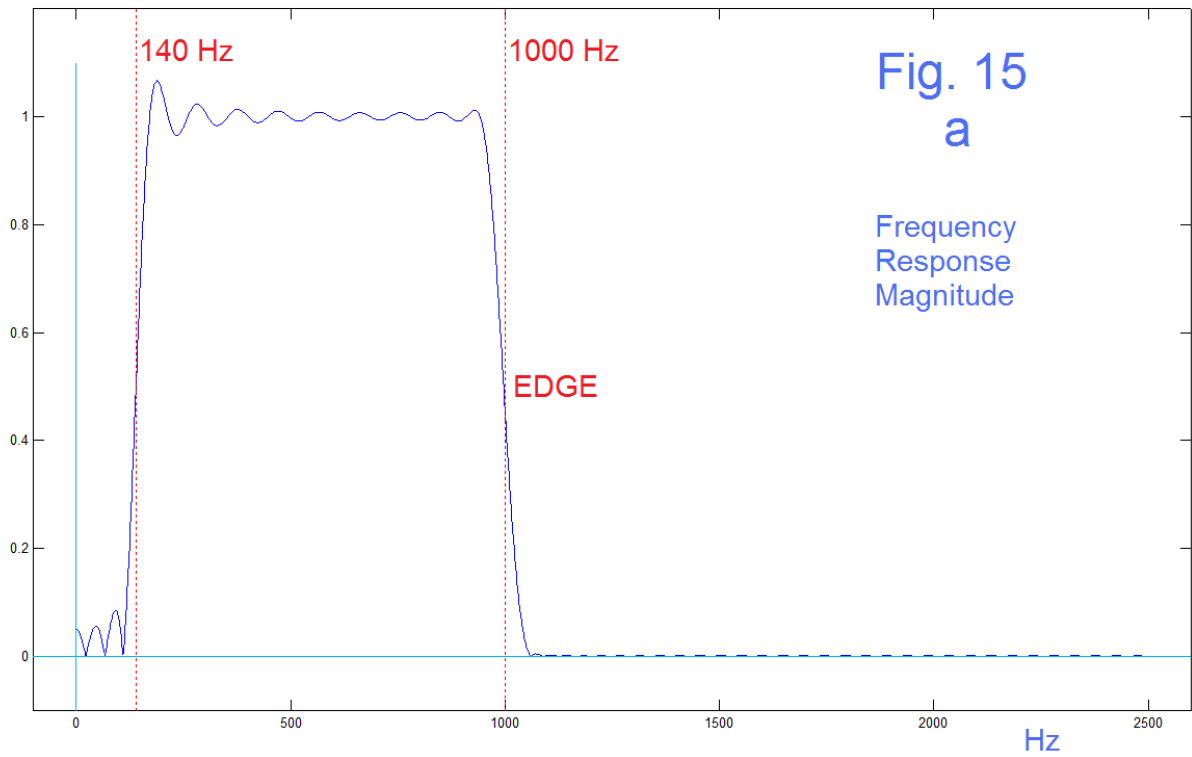
White noise is generally reputed to have no actual pitch. It’s the random “hissing” sound similar to that of air escaping from a tire. It is well known that filtering a white noise (making its spectrum non-flat) results in a “colored noise” that may vary from a vague whistling (as of wind blowing around a corner) to an actual pitch that easily carries a melody. Such use of filtered noise dates back to the very early days of analog music synthesizers.

A wide variety of filters can be employed to color the noise. Even delaying the noise and adding it to itself has a weak pitch corresponding to the reciprocal of the delay times. This relates to the notion that a repetition rate determines pitch – yet it only repeats once. This is better understood in terms of a periodic frequency response (comb filter) [3]. The early attempts to use filtered noise related mostly to recursive filters (filters with poles). For the purposes of the study that follows, we will be using moderate-length FIR (Finite Impulse Response) digital filters designed by the Matlab *firls* function, borrowing the main ideas from a previous note [4]. The magnitude responses are indicated in the top panels (a) of Fig. 13, Fig. 14, and Fig. 15.

The test signals here are about 3 seconds long at a sampling rate of 5000 Hz. These 3 second signals are used so that listening is comfortably done. On the other hand, we can’t expect to resolve the 14,000 samples in a plot, so only a representative portion (200 samples, or just 0.04 seconds) is plotted [middle (b) and lower panels (c)]. These plots allow us to display the randomness of the input noise (same in all three test cases) and the degree of repetition on the filtered noise.







We will look at the three figures (13-15) in some detail along with audio examples of four cases.

Input Noise

<http://electronotes.netfirms.com/orignoise.WMA>

440

<http://electronotes.netfirms.com/mid440.WMA>

1000

<http://electronotes.netfirms.com/high.WMA>

140-1000

<http://electronotes.netfirms.com/toy.WMA>

The first audio case (orignoise.WMA) is just the white noise input, and corresponds to the (b) panel of Figures 13-15. All are the same, and we remark again that we print in the plots only 200 of the approximately 14,000 samples. The audible result is a “hiss” with no notable sense of pitch. The spectrum of the noise is not shown, although it should be similar to many published results. NO SINGLE noise example will be truly flat. This SOUND is the input to our test filters and is the baseline for the other results.

Fig. 13 and Fig. 14 show sharp bandpass filters (FIR length 101) along with the resulting filter outputs. The first thing to note is that the outputs are smoother and show considerable evidence of a sinusoidal component (although of varying amplitude). Fig. 13c shows a reasonable 440 Hz centered noise while Fig. 14c shows a reasonable 1000 Hz centered noise. Listening to these (mid440.WMA and high.WMA) we rather easily match each to the corresponding pitches. Further, to the extent that we believe we can “count cycle” in the plots, we confirm the same pitches. In Fig. 13c we count 17.5 cycles in the 200 samples shown. The 200 samples at a 5000 Hz sampling rate give 0.04 sec. Thus $17.5/0.04$ gives a frequency of 437.5 Hz, a very good verification of what should have been about 440 Hz. We don’t expect perfect agreement, although Fig. 14c seems to show rather exactly 40 full cycles which would be $40/0.04 = 1000$ Hz. These two examples are unremarkable and correspond to conventional music synthesis techniques.

What is new here is Fig. 15 where we use a broad bandpass rather than a sharp one. The output (Fig. 15c) has some indication of sinusoidal segments, but clearly the period varies quite a bit. If we count it anyway, I get about 23 cycles (I would gladly accept counts from say 20 to 26). This would be a pitch of 575 Hz. The average of the bandedges would be $(140 + 1000)/2 = 570$ which seems better than we deserve! (Perhaps the geometric mean of $\sqrt{1000 \times 140} = 374$ should have been considered, and would have been less agreeable.) In the listening test (toy.WMA), no strong pitch jumps right out. Certainly, there is nothing that suggests 570 Hz. It sound, if anything, like a whistle of a wind gust. Not noise, but not very musical.

When pitch matching, pretty much by definition we “prime” our perception mechanism with a test signal. (A few individual with exceptional musical ears, and perfect pitch, might be able to just call out pitches.) Here we could use a tone generator, but the filtered noises already presented offer an interesting comparison. It is my finding that if I listen to the 1000 Hz noise, and then the broad-topped case, I believe there is a clear but not strong pitch at 1000 Hz in the broad-top. (One does not hear the 440 Hz pitch, in contrast.)

What this seems to be is the classic “edge pitch” characterized by hearing a pitch at a position where there is a sharp transition (1000 Hz in this case) [4, 5]. This is hard to explain. Perhaps, just perhaps, we might envision a broad top as a series of sharp bandpass responses close together. In the middle, each peak has side peaks tending to hide it. On the high side, the highest peak is exposed. Or perhaps it is a general proclivity of the perceptual systems to seek out edges.

REFERENCES

[1] References on Pitch

- [1a] B. Hutchins, "The Ear – Part 1: Basic Ideas of Pitch Perception," *Electronotes*, Vol. 10, No. 92, August 1978
- [1b] B. Hutchins, "The Ear – Part 2: An Observational Basis for Pitch Perception Theories," *Electronotes*, Vol. 10, No. 93, September 1978
- [1c] B. Hutchins, "The Ear – Part 3: Models and Phenomenon Vs. Place and Fine-Structure Theories," *Electronotes*, Vol. 10, No. 94, October 1978
- [1d] B. Hutchins, "The Ear – Part 4: Recent Developments Regarding Pitch Perception," *Electronotes*, Vol. 11, No. 100, April 1979
- [1e] Roederer, J.G., *The Physics and Psychophysics of Music*, 3ed, Springer (1995)
- [1f] Hartmann, W.M., *Signals, Sound, and Sensation*, AIP Press (1997)
(Our pal Bill Hartman's great book – seems free online too.)
- [1g] F. Wightman & D. Green, "The Perception of Pitch", *Amer. Scientist*, Vol 62, April 1974, <http://leachlegacy.ece.gatech.edu/ece4445/downloads/pitch.pdf>
- [1h] Heller, E.J. *Why You Hear What You Hear*, Princeton U. Press (2013)
- [2] B. Hutchins, "Reviewing the Current State of Music Synthesis", *Electronotes* Volume 23, Number 220 January 2014 <http://electronotes.netfirms.com/EN220.pdf>
- [3] See [1f] Chapter 15
- [4] B. Hutchins, "Edge Pitch, Tinnitus, And The Hum - A Quick Look (and Listen)", *Electronotes* Webnote, ENWN-45, 12/12/2016
<http://electronotes.netfirms.com/ENWN45.pdf>
- [5] Houtsma, A.J.M., Chapter 8 Pitch Perception, pg 283 (1995)
<http://web.mit.edu/hst.723/www/ThemePapers/Pitch/Houtsma95.pdf>
- [6] B. Hutchins, "A White Noise Curiosity," *Electronotes*, Volume 22, Number 208 January 2012. <http://electronotes.netfirms.com/EN208.pdf> ; B. Hutchins, "More Concerning Non-Flat Random FFT," *Electronotes* Application Note No. 416, Nov 7, 2014
<http://electronotes.netfirms.com/AN416.pdf>

APPENDIX – FEEDBACK CONTROLLING EXCITATION

We have taken the view of a periodic excitation signal being filtered as it passes through a resonator. Thus we had the view that the source of excitation and the mechanism of filtering were separate and definable objects. The interaction of the two was thought of as restricted to the output of the former being the input to the latter. Likely this is rarely true. In many cases there is a “coupling” of the mechanisms. In only a few cases; a guitar perhaps, where a fret-defined string vibrates transmitting sound to a “box” (guitar body) do we see this. It is probably generally true of electronic music synthesis where we sometimes struggle to implement (simulate) couplings for an acceptably realistic imitation of an acoustical instrument.

A trumpet (or similar wind instrument) is a good example of a case where the excitation couples with the resonance. A beginning trumpet player struggles to make his/her lips “buzz” into the mouthpiece at the right rate such that the horn itself cooperates and produces a sustained tone. In due order, the player comes to terms with the instrument and manages to live with the pitch choices the instrument allows. At length, a skilled player learns to impose small variations to be intoned, as with an ensemble, or with a sequence of notes. That is, the player was “forced” by a resonant mode to play a G instead of the C below or the C above (by tightening or loosening the lips). If you want to play the F below G, you have to press down the first valve to make the overall pipe longer - you can't do it with the lips alone, although you can get from G down to F[#] by liping. This means that you have to have the main tuning slide roughly correct, and you have to push down the right valves, but to a small degree (perhaps up to a half tone), intonation is still up to the lips of the player. So much as we suggested that during pitch matching, a subject could “ride” an analog knob, with aural feedback, a veteran player achieves precisely correct intonation by aural feedback and a slight physical adjustment.

What we have not really addressed is why resonance supports only certain pitches. It is a simple issue of supporting a standing wave. Any standing wave will dissipate due to radiation of usable sound and air friction, and needs to be re-supplied with appropriately timed small pulses of pressure. It is the periodic pressure of the standing wave against the lips, already more-or-less adjusted to open and admit a pulse (by the skill of the player), that regulates the exact timing. Thus the resonance interacts with the excitation.

String instruments have their own regulating feedback. It is well known how plucked strings work. Depending on the point along the length where they are plucked, they produce a fundamental and a series of overtones. We say “overtones” because they do not produce harmonics exactly, but instead, the overtones tend to go slightly sharp due to an end effect of string stiffness. The plucked string (“pizzicato”) is essentially like a guitar, and the decaying vibration is duly filtered/radiated by the wooden body of the instrument.

Famously string instruments are also capable of sustained (rather than plucked, decaying) tones through the replacement of radiated and frictionally lost energy by bowing. This is often thought of as a “stick-and-slip” mechanism between the rosined bow and the string. As the bow moves, it sticks to the string and displaces it by some relatively small increment, then abruptly jumps off and the string snaps back by the same increment. It might seem like this would produce a sawtooth-like displacement, with the pitch determined

by the speed of the bow. It does not behave in this way – the pitch depends on the resonant frequency of the string. So it must be that the natural vibration causes the string to snap off the bow at about the point where it was getting ready to slip. This is very much like the instance of the pulses of pressure from the standing wave in the trumpet causing the lips to open. It is the responsibility of the talented player to assure that bow pressure and speed support the sustained tone.

A curious case of great interest is the piano. The piano is a percussion instrument. It is not that different from a traditional percussion instrument such as a marimba. The marimba has tuned metal or wood bars struck by a mallet, while the piano has strings (some singles, some pairs, and some triples) struck by felt “hammers” activated by the keys. When you press a key, the hammer flies up and strikes the string(s). It bounces off and that’s it. The energy imparted to the strings decays away.

But – you protest – the piano “sings”. The tones, while decaying, actually hang around a good while, at least as compared to most standard percussive instruments. First, it is true that the piano has a “sustain” pedal, but this just lifts automatic dampers. We are very much accustomed to sounds instigated by a hammer-like impulse followed by an exponential decay with a single decay time constant. What is different about the piano is that there are two time constants – a faster one and a slower one.

Electronically it is quite easy to achieve two different time constants. A time constant is often seen as a product of a resistor times a capacitor (physical units in seconds). Since capacitors are pretty much fixed, electronically we can change an RC product by a variable resistance, even a voltage-controlled resistance. Let us hope that all readers here have opened up a piano out of curiosity. Thus you have noted the very long heavy (wrapped) single strings for the low notes, the pairs of strings in the middle range, and the short triples on the high side. Perhaps intuitively, we recognize that the lower strings naturally vibrate for longer periods of time. Also perhaps we note that the high strings seem to sing longer than we might expect. It turns out that the multiple strings are not so much for adding loudness. Instead, the paired (or tripled) strings start out being hit by the hammer and produce vibrations that are almost certainly of slightly different frequency, but initially in phase. The energy decays rapidly, and the two (or three) strings wonder out of phase. In wondering out of phase, they slightly influence the string supports to move such that the two frequencies become coupled. They lock out of phase, and thus come to radiate energy at a lower rate. Still the ear has an extremely large dynamic range and the piano sound is heard to keep singing, long after we might have otherwise expected it to do so.

We perhaps should mention that the piano has notes with pitches down to $440\text{Hz}/16 = 27.5\text{ Hz}$ (the lowest A), but you can’t really hear this pitch as a pure tone! It seems quite apparently heard due to the harmonics (missing fundamental). This is the same as the organ builder’s trick of hundreds of years ago. In addition, a very popular instrument (the violin) in fact has a weak fundamental (it’s too small for its range), and is likewise supported by harmonics.

* * * * *