

ENVELOPE EXTRACTION WITH THE HILBERT TRANSFORM

INTRODUCTION - BASIC IDEAS

Hilbert transform theory tells us many things that are useful in signal processing, communications, and sound synthesis [1]. One of these things is that the envelope of a signal $x(t)$ can be derived from $x(t)$ by first taking the Hilbert transform $\hat{x}(t)$ and then computing the envelope $e(t)$ as:

$$e(t) = \sqrt{x^2(t) + \hat{x}^2(t)}$$

This is pretty simple - just the Pythagorean theorem. Simple, that is, assuming we have the Hilbert transform of our signal. The Hilbert transform is obtained, not surprisingly, using a Hilbert transformer [2].

Fig. 1 illustrates a simple Hilbert transformer setup for envelope extraction. Here we have chosen a length 43 Hilbert transformer designed with Matlab's "firls" function as:

```
hh = firls(42,[0 0.03 0.1 0.9 0.97 1],[0 0 1 1 0 0],'Hilbert')
```

This impulse response is listed in Fig. 2, and the magnitude response of this Hilbert transformer is shown in Fig. 3, and is seen to be a reasonably good approximation to 1 from about 0.06 to 0.44. The exact details of this design are not essential to the main theme of this note.

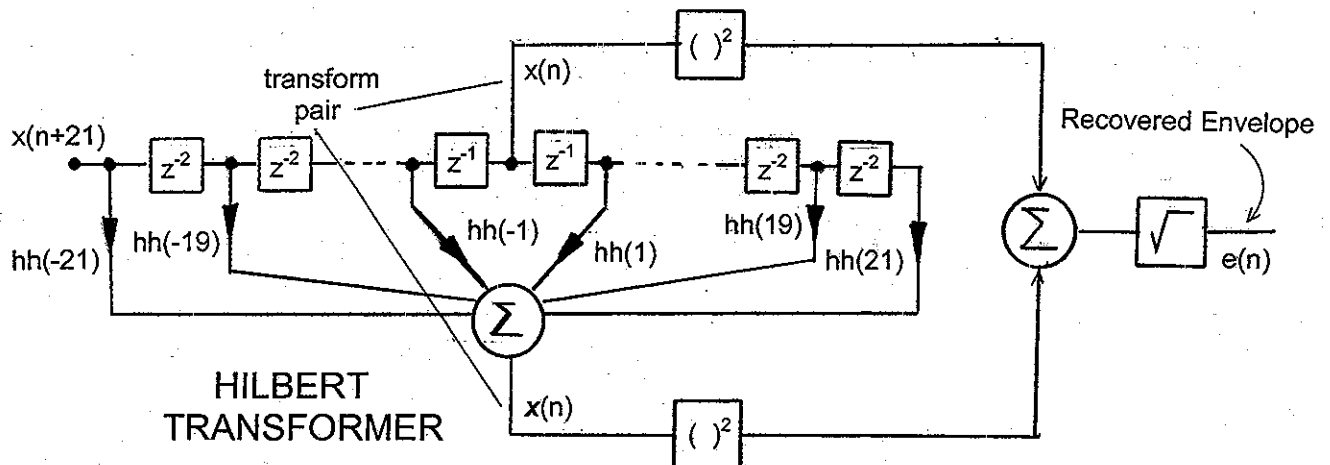


Fig. 1 Envelope Extraction Using Hilbert Transform

n	hh(n)	n	hh(n)
-21	0.0060	1	0.6227
-19	0.0123	3	0.1722
-17	0.0197	5	0.0656
-15	0.0263	7	0.0142
-13	0.0296	9	-0.0135
-11	0.0265	11	-0.0265
-9	0.0135	13	-0.0296
-7	-0.0142	15	-0.0263
-5	-0.0656	17	-0.0197
-3	-0.1722	19	-0.0123
-1	-0.6227	21	-0.0060

all other hh(n)=0

Fig. 2 Impulse Response of Hilbert Transformer

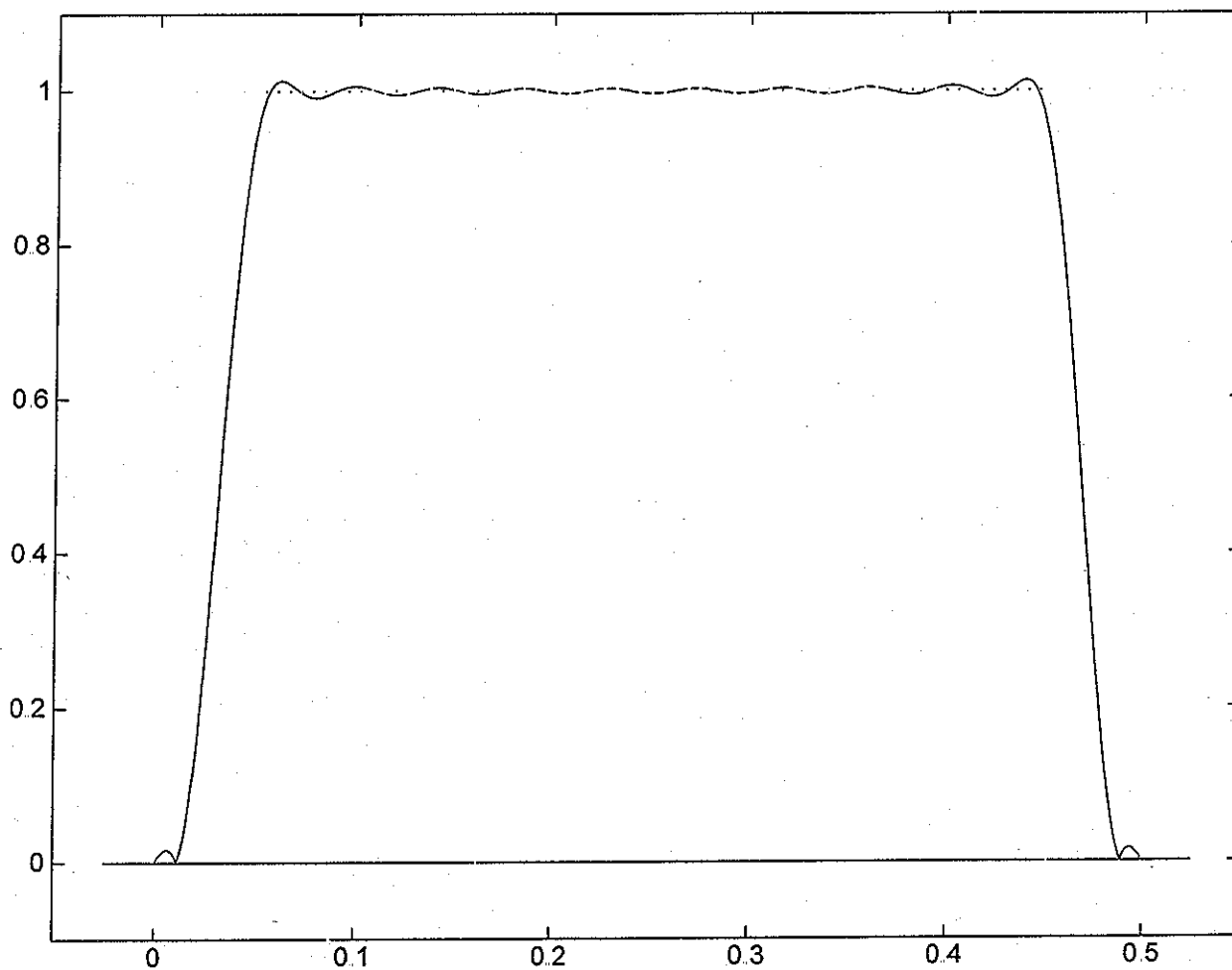


Fig. 3 Magnitude Response of Hilbert Transformer (the phase difference is exactly 90°)

It is of course useful to generate a signal to test that has a known envelope. Fig. 4 shows an envelope that we have generated that can be imposed on 10,000 samples of another signal. This envelope is of the ADSR form (Attack-Decay-Sustain-Release) that is popular in music synthesis. We note that the Attack and Decay have 1000 samples each, the Sustain is 6000 samples long, and the Release is 2000 samples long. Fig. 5 shows this envelope multiplying a sinusoidal waveform. Indeed it looks like little more than a dark splotch since the period of the sinusoid is very small relative to the envelope length.

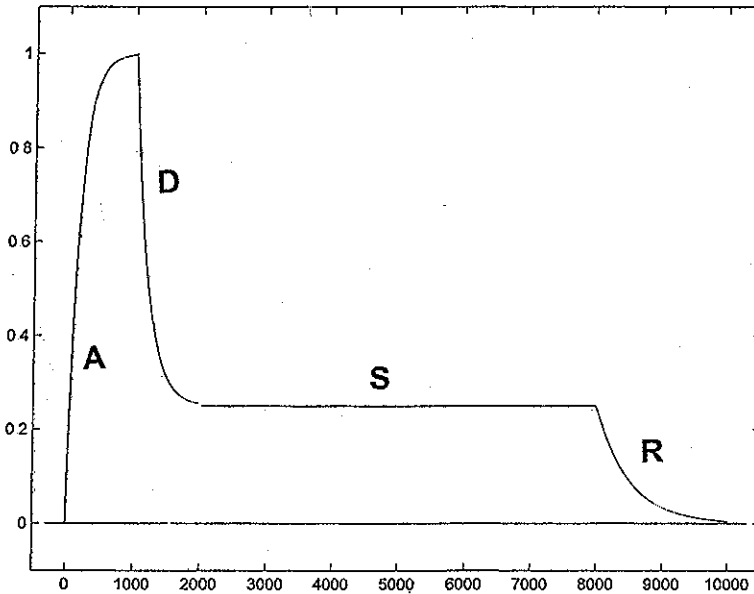


Fig. 4 A typical envelope of the Attack, Decay, Sustain, Release (ADSR) type as is often used in music synthesis can be our test envelope

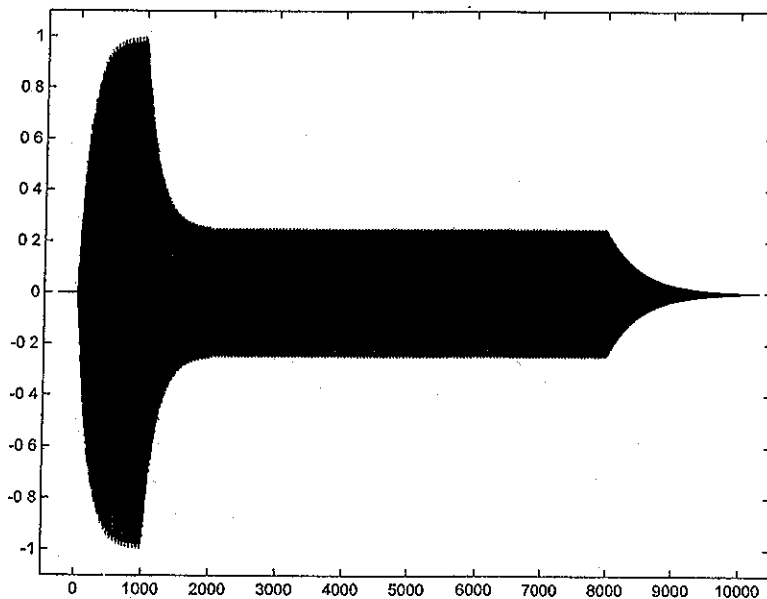


Fig. 5 The ADSR is imposed upon a sinusoidal waveform of a frequency that is very high relative to what could be resolved in this plot. That is, the sinusoidal waveform is multiplied point-by-point with the ADSR envelope

Fig. 6 shows a particularly successful recovery of the envelope using the Hilbert transformer as described. (In fact, it is so successful that we will be obliged to come up with less successful examples just to show that we didn't cheat.) The level of success is related to how well the Hilbert transformer approximates unity gain (Fig. 3). In producing Fig. 6, we have chosen a frequency (0.07) such that we achieve a particularly favorable case. This we can see by looking at a zoomed-in portion of the magnitude of the Hilbert transformer, as seen in Fig. 7, where we see that the actual magnitude at the frequency we are using, 0.07, is very close to 1.

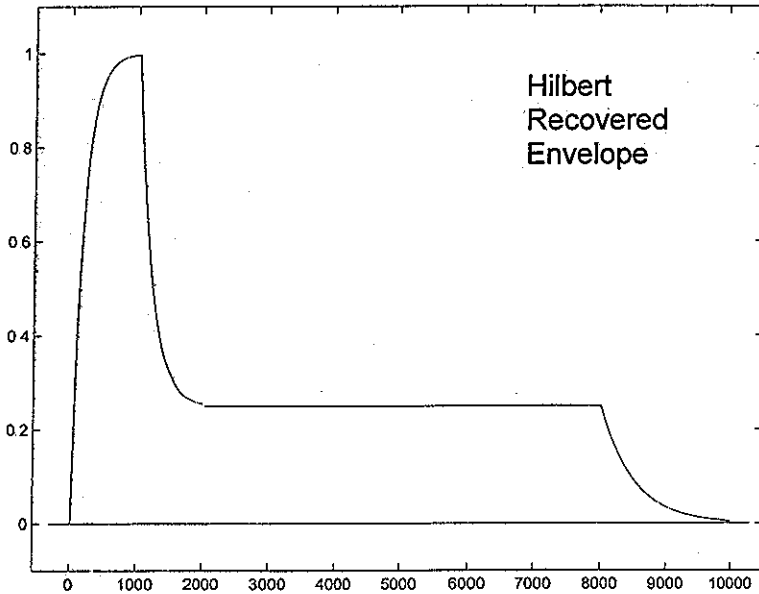


Fig. 6 Here we have a virtually perfect recovery of the ADSR envelope by using the Hilbert transform method. The frequency of the sinusoidal waveform is 0.07.

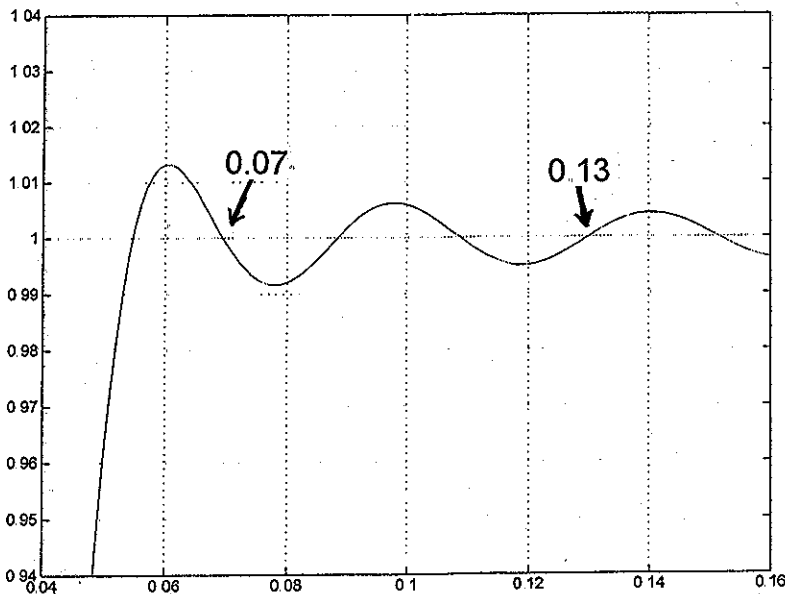


Fig. 7 A detail of the Hilbert transform magnitude from Fig. 3. Note that the magnitude at 0.07 and at 0.13 (frequencies we are using for some of our test) are very close to one, while at a frequency of 0.06 it is high by more than 1% and at 0.05 it is low by about 4% and is on the way down further.

We can also see from Fig. 7 that other frequencies have magnitudes differing from 1, and for example, the frequency 0.06 has a magnitude that is about 1% high. Fig. 8 shows the case of recovering the envelope if this frequency is used, and we see a mild ripple in the recovered envelope (seen best at the top of the attack). We note also that the response for a frequency of 0.05 is about 4% low, since here we are headed out of the low end of the usable bandwidth. Fig. 9 shows a case where the first 5000 samples are at a frequency of 0.06 while the last 5000 are at a frequency of 0.05. We note a blurring beyond sample 5000, which is due to the 4% amplitude error. This is not due to the change of frequency by itself. Fig. 10 shows a change of frequency from 0.07 to 0.13 at sample 5000, and we note only the expected transient glitch as the length 43 Hilbert transformer is reloading with the new frequency. Thus, we expect frequency changes *per se* to not be a major complication.

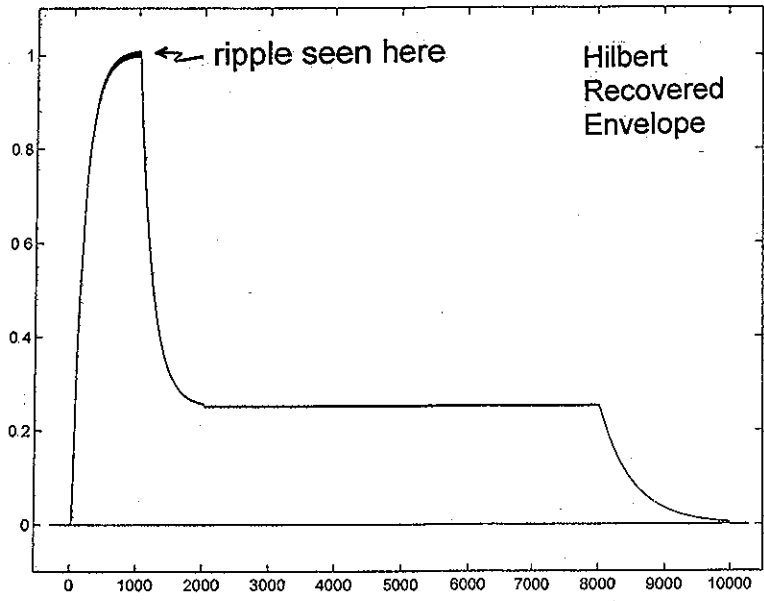


Fig. 8 Here the frequency is 0.06 instead of 0.07, and we see some ripple in the envelope

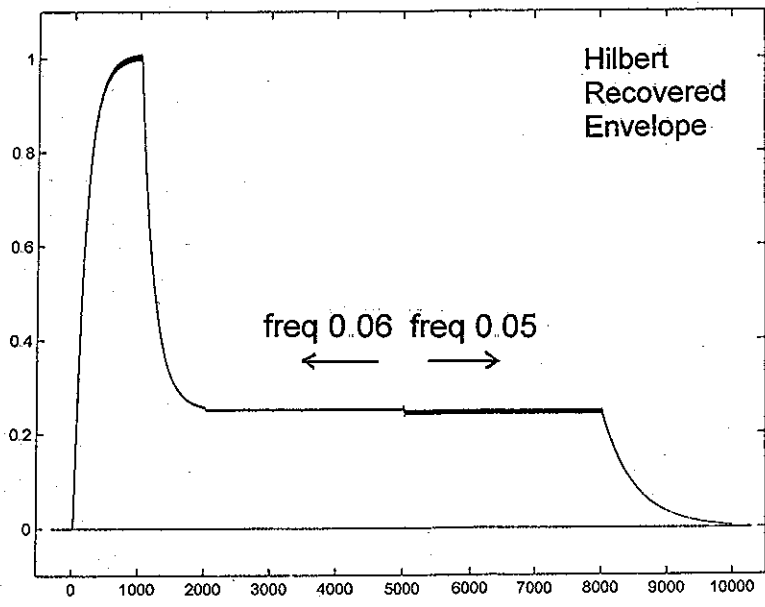


Fig. 9 The first 5000 samples are at a frequency of 0.06 (as in Fig. 8) while samples 5001 on are at a frequency of 0.05 (just below the edge of the Hilbert transformer's bandwidth)

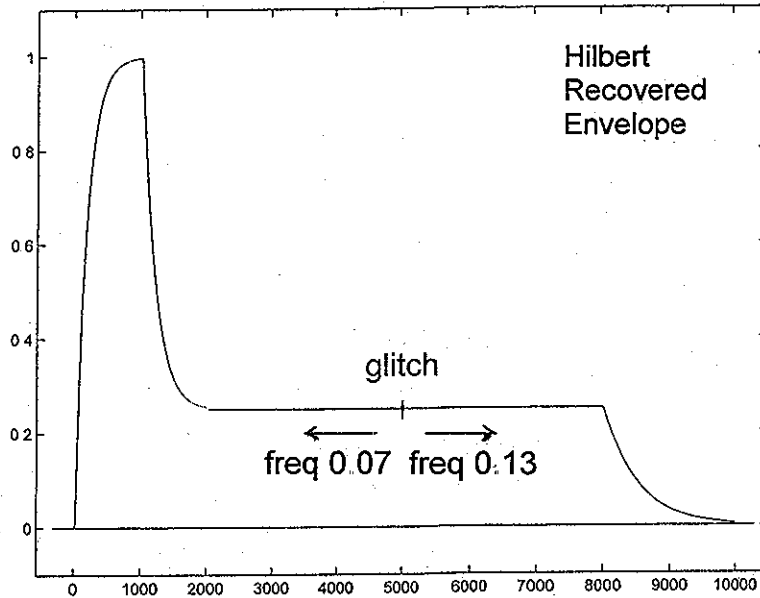


Fig. 10 Changing the frequency at the midpoint, from 0.07 to 0.13 causes a "glitch" as the Hilbert transformer is refilled with the new frequency, but otherwise, both frequencies are well suited to excellent envelope recovery (see Fig. 7)

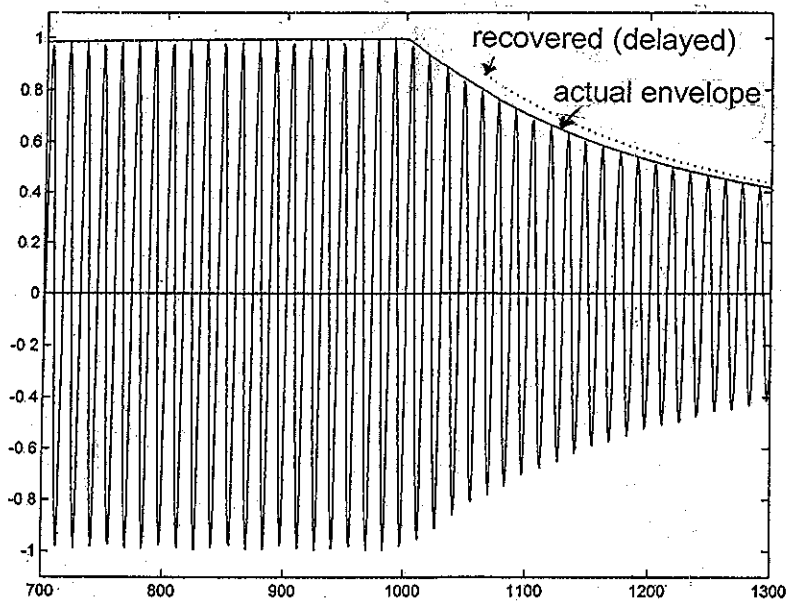


Fig. 11 The actual envelope recovered (dotted) is slightly delayed relative to the original envelope (solid) due to the delay of the Hilbert transform.

Another point which is hard to see on comparing Fig. 4 and Fig. 6 is that the recovered envelope is slightly delayed. In fact, this delay is exactly the delay of 21 from the Hilbert transform used. Fig. 11 perhaps makes this point more clearly by zooming in on a region from samples 700 to 1300, thus including the transition from attack to decay at sample 1000. [In addition to better showing the delay, this view also shows the expected relationship between the actual sinusoidal waveform and the envelope.]

BUT NOT SO FAST!

Here we do seem to have an excellent example (Fig. 6). Further, the faults we have shown in subsequent examples are not fundamental, and could likely be reduced sufficiently by choosing a better Hilbert Transformer. Unfortunately for many applications, there is a fundamental problem. This has been long noted [3] and involves the case of multiple frequencies in the input. Does this still work with multiple frequencies? In a word: No. At least it does not work if we insist that the envelope recovered be the one we impose on the waveform in the first place, or if our notion of what the envelope is does not fit what the Hilbert Transform method gives us.

Fig. 12 shows an example where two sinusoidal waveforms of different frequencies, (0.07 and 0.13) are input. These are the same frequencies as in Fig. 10, but here, instead of switching from one frequency to the other at the midpoint, both frequencies are present for all 10,000 samples. This is a very different result. Instead of a reasonably clean smooth recovery of the envelope, we see a mess. Although the envelope is still apparent, it is our eye that is recovering it as the top of the plot.

The notion that perhaps the Hilbert Transformer is finding an envelope that is different from the one imposed is illustrated in Fig. 13. In Fig. 13a we show a case of two frequencies where the two are here very close together, 0.09 and 0.091. It is still fairly apparent that the original imposed ADSR envelope is having some influence, but for the most part, the output is dominated by a lobed structure. The origins of the lobes are immediately apparent by considering the Hilbert transform input (the actual enveloped waveform) in Fig. 13b. In fact we well understand the lobes as being produced by the summation of two close sinusoidal waveforms in a familiar "beating" pattern. It is not at all unreasonable in this case to consider the actual envelope as being the product of the beat envelope (that is just due to the summation) which is then subject to the imposed ADSR. In this case, we believe we have recovered the true envelope, but clearly not the one we imposed. But the Hilbert transform method was "honest."

In order to relate Fig. 13a to Fig. 12, it will be useful to move the frequencies a bit further apart (to 0.09 and 0.094), and we have the recovered envelope shown in Fig. 13c, which is starting to look like Fig. 12. Of course, it is still beating in the details, and this is seen in Fig. 13d, which shows a smaller portion of the recovered envelope, along with the corresponding Hilbert transform pair that produces it. We note how the envelope still hugs the waveform.

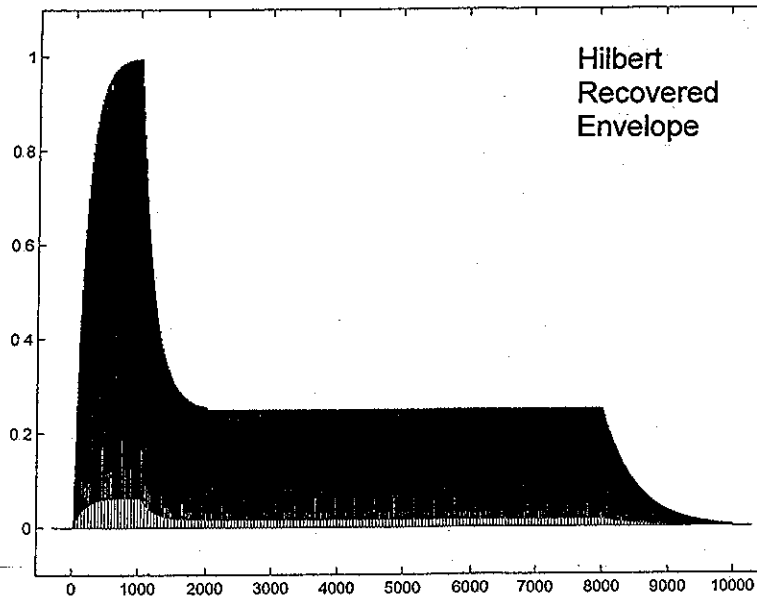


Fig. 12 Here we have two frequencies, 0.07 and 0.13, as in Fig. 10, but they are present simultaneously for all 10,000 samples. Clearly our hoped-for envelope recovery (recovery of the ADSR) has failed.

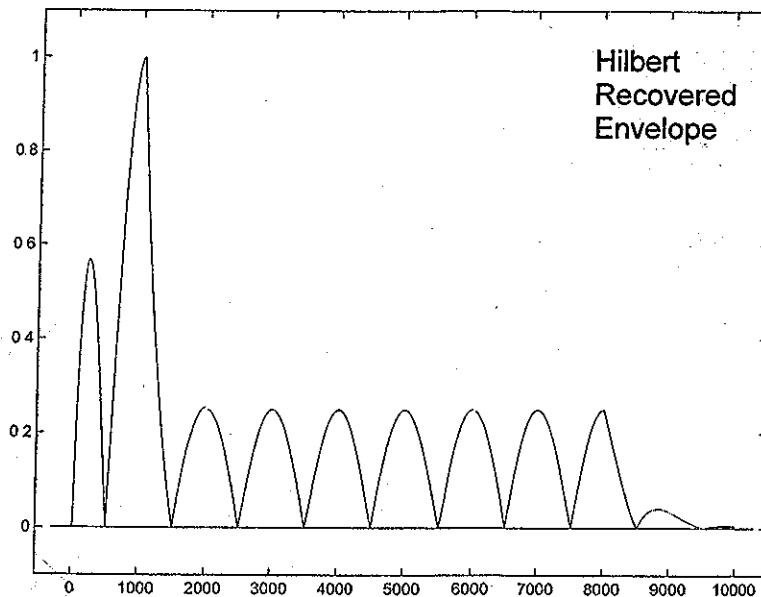


Fig. 13a For two close frequencies, 0.09 and 0.091, the recovered envelope is lobed according to the "beat envelope" of the summation, as well as the imposed ADSR. Probably there is a very real sense in which this is the "right answer."

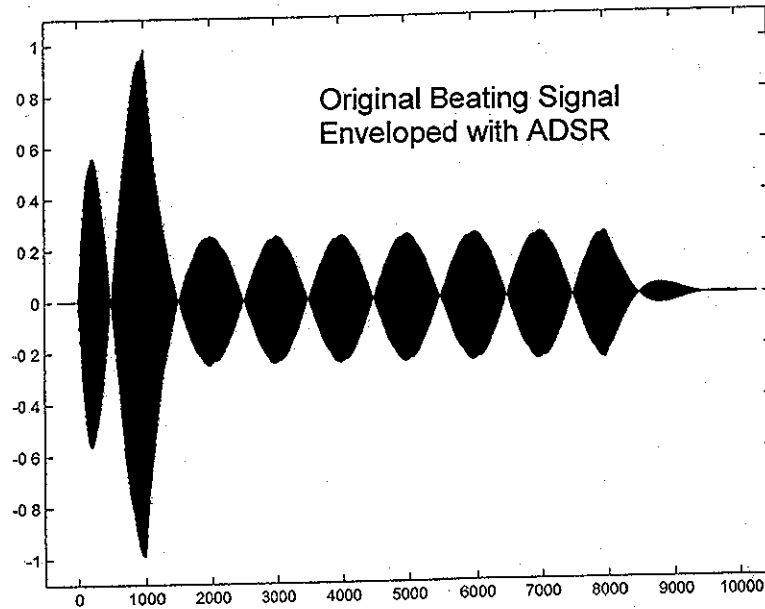


Fig. 13b The lobed structure found in the envelope of Fig. 13a is clearly the result of the actual amplitude beating that is going on.

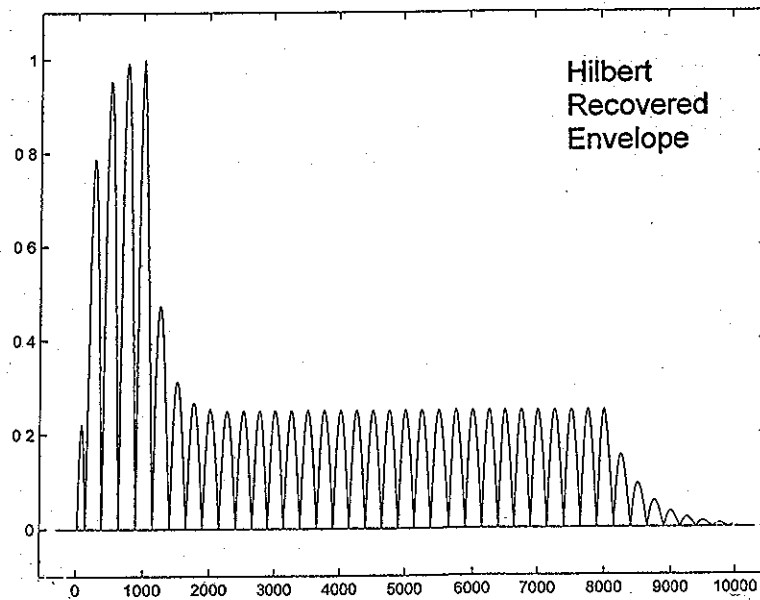


Fig. 13c With the frequencies a bit further apart, 0.09 and 0.094, the lobed structure in the recovered envelope gets more narrowly spaced, and starts to resemble Fig. 12.

In fact, the waveforms of Fig. 13d remind us of how the method worked with one frequency component. In Fig. 13e we return to just one component (0.092) and show the same portion of the output as the on in Fig. 13d for two components. Of course, this is just the familiar $\sin^2(\theta) + \cos^2(\theta) = 1$.

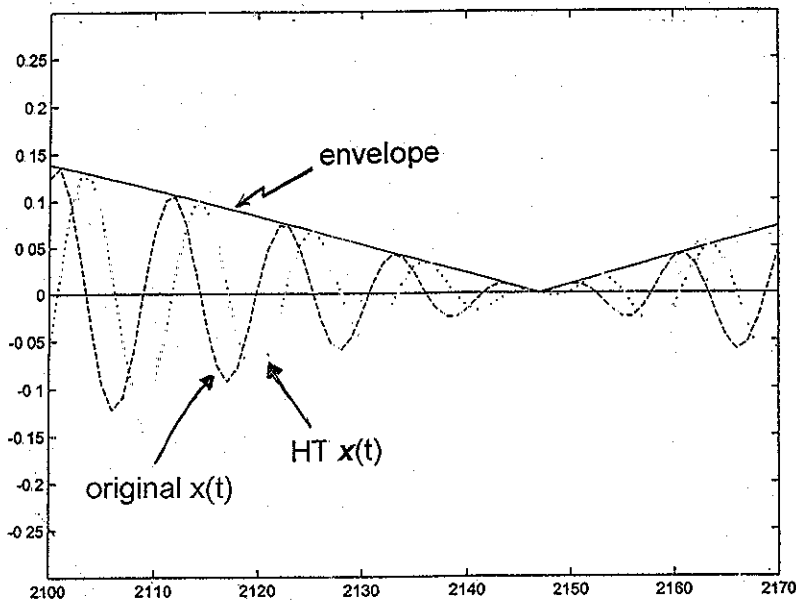


Fig. 13d A blow-up of a portion of Fig. 13c shows the recovered envelope as the V-shaped solid line. This envelope is recovered from the original signal (dashed line) and from its Hilbert transform (dotted line).

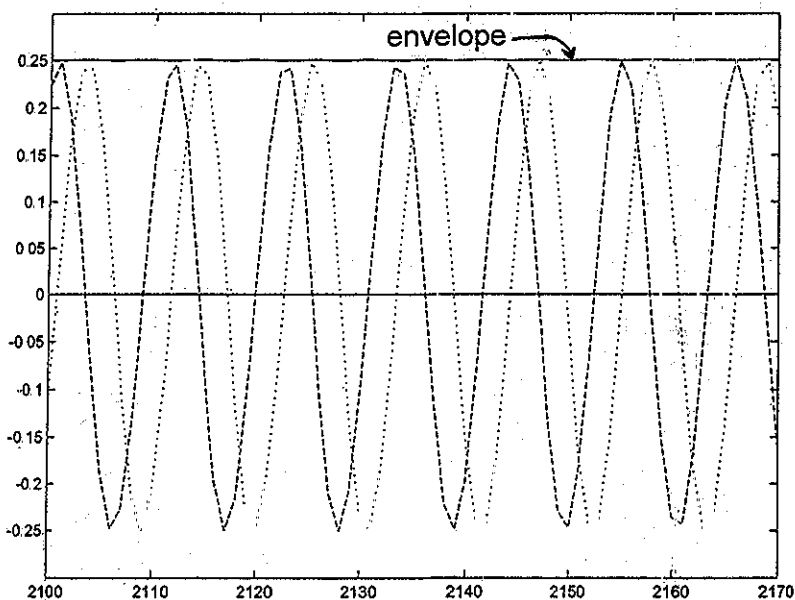


Fig. 13e Here we are reminded that in the single frequency case (0.092 here) we are basically just doing $\sin^2(\theta) + \cos^2(\theta) = 1$. Compare to Fig. 13d above.

ANOTHER PROBLEM – “PINGING WAVEFORMS”

Above we have seen three notions of an envelope: (1) the ADSR test envelope we generated, (2) the envelope the HT method returns, and (3) the envelope we really thought was the right one. In order to see how to bring these ideas together, or to see how they may need to remain apart, we need to bring in the human listener, and an associated scale of time, that is appropriate. This is a common problem when we try to apply mathematical calculations to the world of human perception. Time scale matters. Here an example using a “pinging waveform” will be useful.

In the case of audio event, the applicable time scales would be first the human time constant (about 50 milliseconds = 1/20 seconds) and second the normal duration of acoustic “events” of interest. These events would be things like individual musical notes, the bark of a dog, a word of speech, or a clap of thunder; events with durations of perhaps a quarter second or more - long (by an order of magnitude or more) with respect to the human time constant. In such cases, we are usually interested in describing how loud some sound is, as a function of time, as it progresses from start to finish. A conflict appears when the overall event is composed of smaller events (usually quasi-periodic) each of which seems to have amplitude variations.

Such a case is likely to appear when we have an excitation and a resonance (perhaps multiple resonances) involved. Ian Fritz was kind enough to suggest this as a challenging task for any envelope follower scheme. We will begin with an example where our main goal is to illustrate the concept with clear figures, although the time scales may not be realistic. Fig. 14a shows a waveform generated by pulses that are 250 samples apart which excite a digital resonator (high-Q digital filter) which is 2nd order with complex poles at a radius of 0.98 and angles $\pm 20^\circ$, which is then enveloped by the ADSR of Fig. 4. A zoomed-in portion is seen in Fig. 14b. Fig. 14c shows the recovered Hilbert envelope, with Fig. 14d showing a zoom-in of this. Essentially we see that we have recovered the envelopes of the individual “pings” corresponding to the excitation pulses. We did not get the ADSR back.

At this point we note that the spacing of the excitation pulses (and the recovered envelopes of Fig. 14c) is pitch related, and not event related. For clarity of plotting, we have shown relatively few (actually 40) pings inside the ADSR envelope. We are to understand that the pings are not heard individually, but rather as a contribution to musical timbre (tone color) of the sound. The “event” is the full 10000 samples corresponding to the ADSR length. The Hilbert transform method has given us the wrong envelope(s) ; at least, not the ADSR.

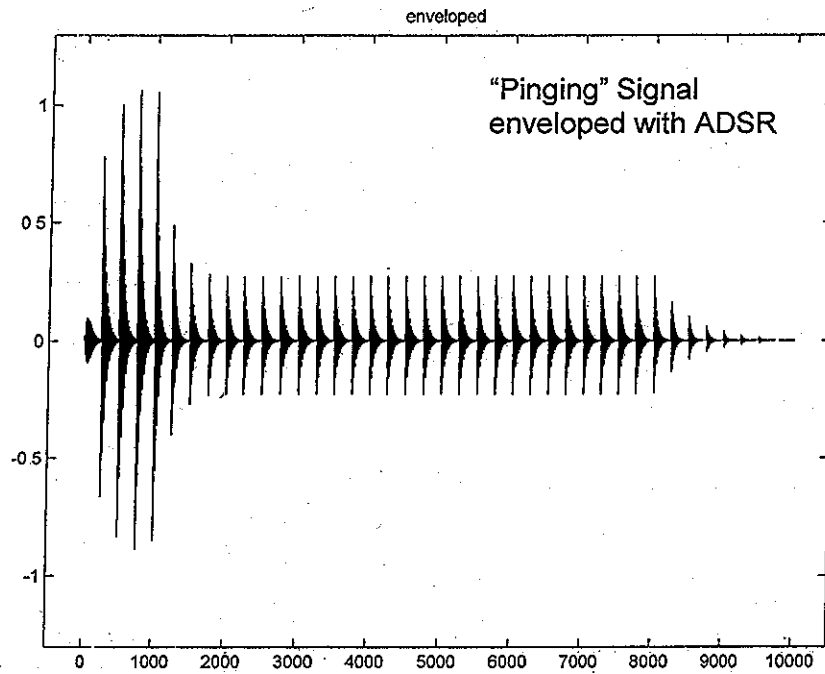


Fig. 14a Here a waveform is generated by pulses spaced 250 samples apart which excite a resonator causing ringing or "pings." The waveform has been enveloped with the ADSR of Fig. 4.

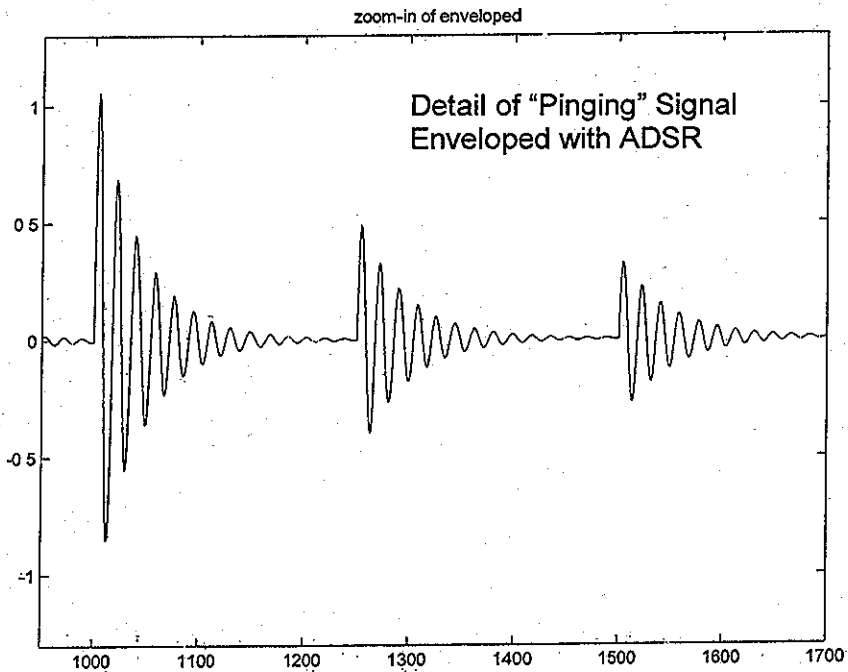


Fig. 14b Here we have a detail (zoom-in) from Fig. 14a and we can better see the details (exponentially decaying sinusoids) of the individual pings.

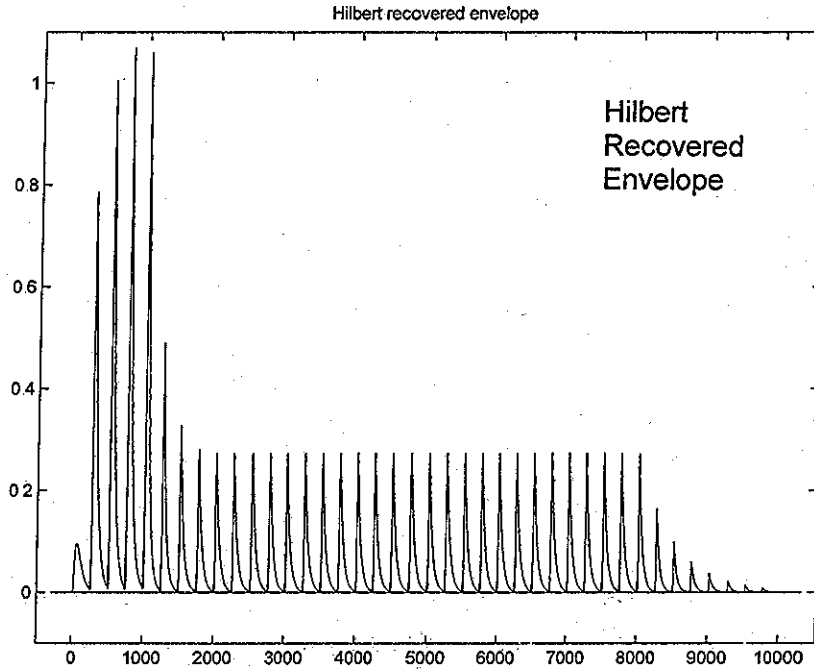


Fig. 14c The Hilbert transform method recovers the envelope of Fig. 14a. Note that we get the envelopes of the individual pings, not the ADSR we might have been hoping for.

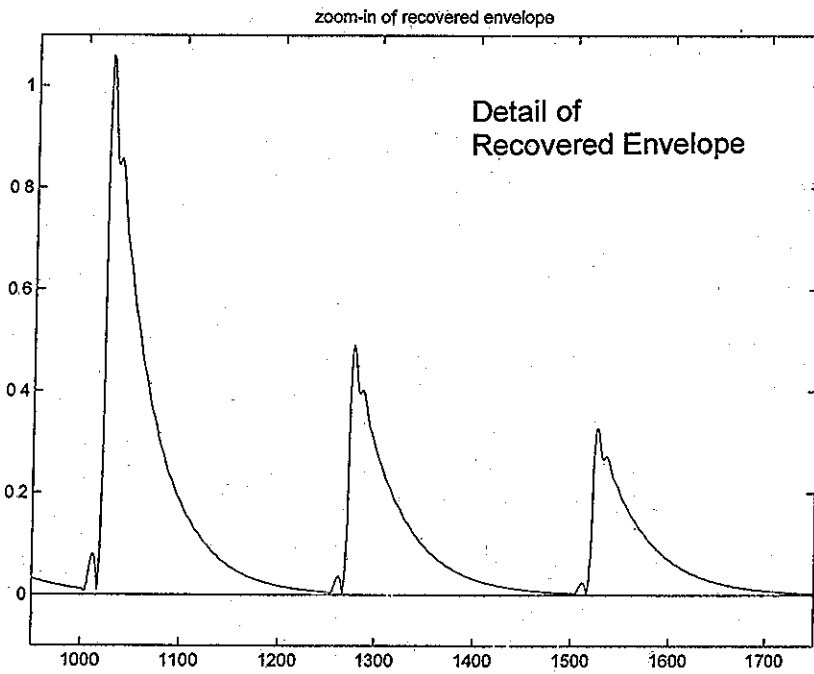


Fig. 14d A detail of Fig. 14c. Compare to Fig. 14b.

LOW-PASS AVERAGING OF THE ENTIRE EVENT

We were hoping for a robust envelope detector that would not force us to employ low-pass averaging with its associated delays and blurring of components that should be sharply defined in time. Also, the Hilbert transformer approach was computationally intense. Thus since the Hilbert transform approach did not solve the problem, and since we apparently still will need low-pass averaging, should we bother with the Hilbert transformer? After all, we could just rectify the original signal and low-pass, a very standard method of envelope recovery. Fig. 15a shows the original signal of Fig. 14a rectified (full-wave rectified or "absolute value"). Fig. 15b is a zoom-in of Fig. 15a. Note the striking similarity of Fig. 14c and Fig. 15a, and in their more detailed versions (Fig. 14d and Fig. 15b). Also note the differences. Fig. 14d shows the envelopes of the component pings, while Fig. 15b shows the rectified pings themselves.

Now, the question at hand relates to the addition of low-pass averaging. Intuitively we might expect better results by low-passing the component envelopes (Fig. 14c) rather than the rectified signal (Fig. 15a). However a small amount of thought will also suggest that in this case we may not see much advantage, as anything capable of low-passing the pings is certainly capable of low-passing the cycles within the pings themselves.

Accordingly, Fig. 16a and Fig. 16b show nearly identical results. Both are fairly ugly-looking in fact. But note that this ugliness is not "noise" but the actual pings that are not sufficiently filtered which cause spiking in the envelope. This is really a consequence of our wanting to plot the details of the pinging clearly. Later we will look at a more realistic case where the pings are denser.

What we seem to learn is that the use of the Hilbert transform to recover the component ping envelopes is likely not worth the effort, relative to the simple application of a full-wave rectifier. And as Ian suggested, this is a fairly tough envelope to coax out of the signal.

ANOTHER LOOK AT MULTIPLE FREQUENCIES – HARMONICS IN THIS CASE

We were initially delighted with the envelope recovery in the case of a single frequency (e.g., Fig. 6), and disappointed to see a mess for multiple frequencies (e.g., Fig. 12). Why does it now work for the sum? Because it is non-linear of course (the squaring and square roots)! Yet we still think we can "see" the envelope in Fig. 12, and we have our low-pass averaging to use as we have seen above.

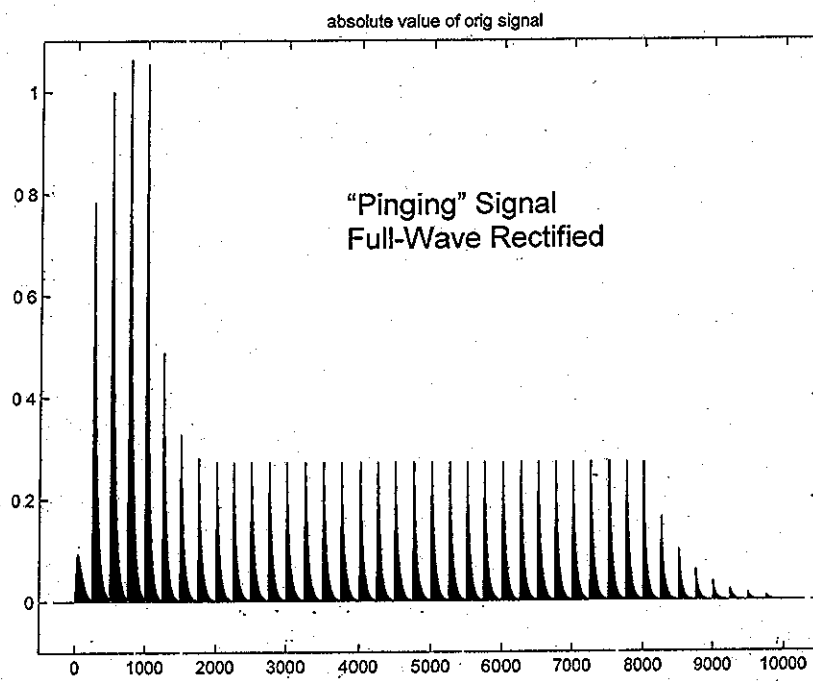


Fig. 15a Absolute value (full-wave rectification) of original signal.

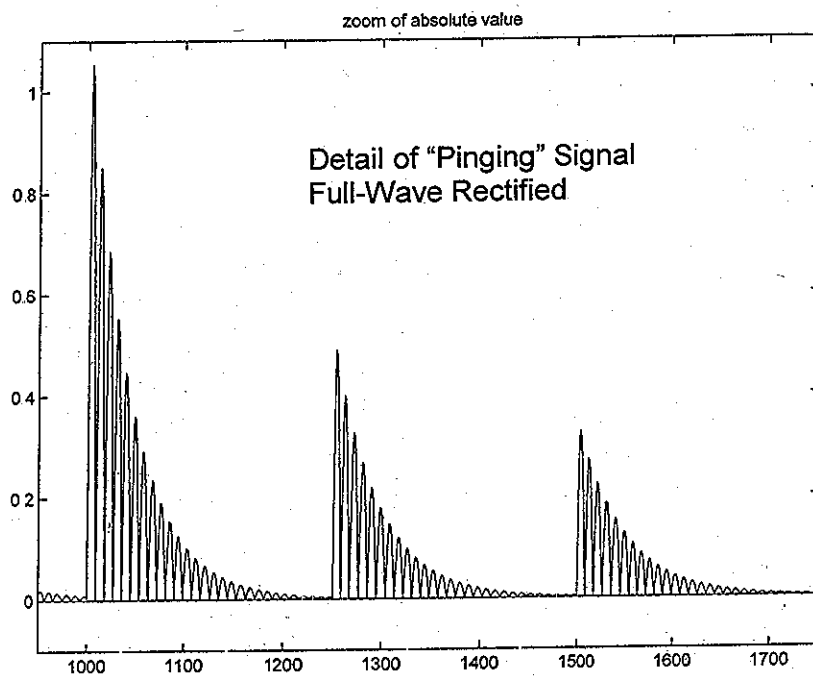


Fig. 15b A detail from Fig. 15a shows the rectified pings.

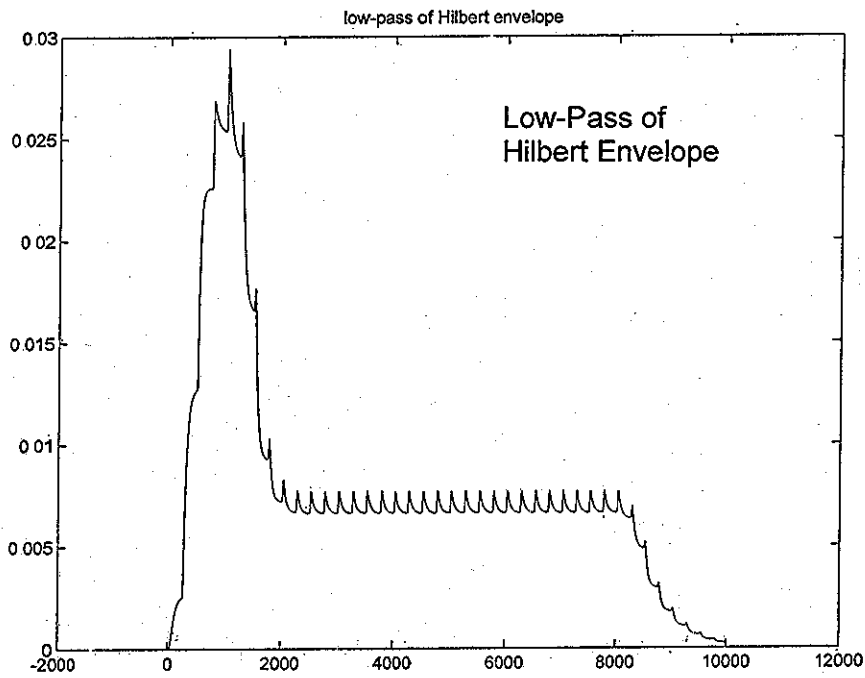
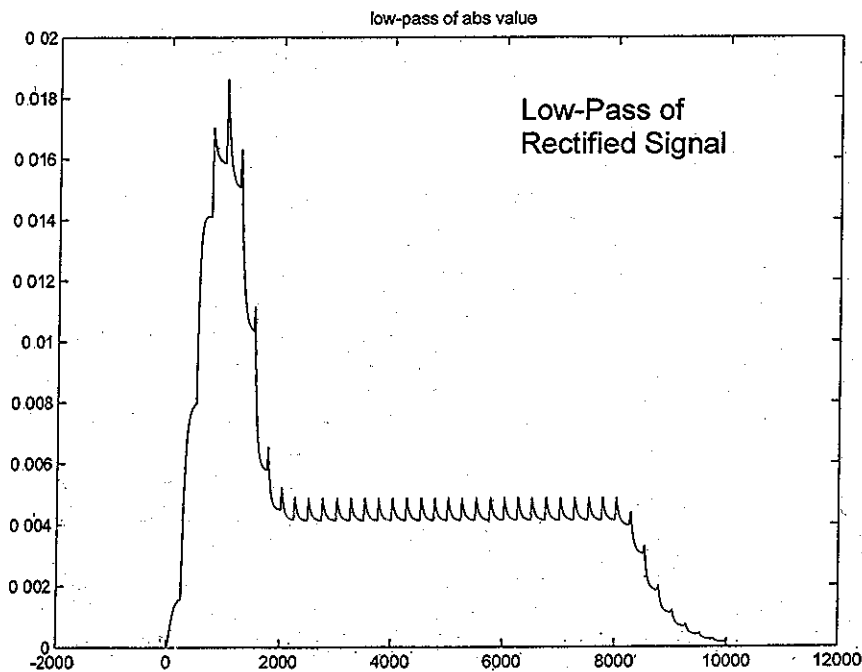


Fig. 16a Low-pass filtering of the Hilbert transform envelope(s) of Fig. 14c. The digital low-pass is a length 521 FIR filter with cutoff of 0.00015.



NOTE:
 Fig. 16b has less amplitude than Fig. 16a because of the lobed structure of the rectified signal (Fig. 15b) as compared to the Hilbert envelope (Fig. 14c)

Fig. 16b Low-pass filtering of the absolute value of the original signal (Fig. 15b). The digital low-pass is again a length 521 FIR filter with cutoff of 0.00015.

Fig. 17a shows an enveloped waveform consisting of four harmonics (0.06, 0.12, 0.18, and 0.24), and a zoom-in is shown on Fig. 17b. Fig. 17c shows the Hilbert envelope (zoomed in Fig. 17d). Compare this to Fig. 17e which is the full-wave rectified signal (zoomed in Fig. 17f). Low-passed versions of Fig. 17c and Fig. 17e are shown in Fig. 18a and fig. 18b respectively. Here the low-pass cutoff was set to about 0.0075, and the results are quite good, and we might well say, equally good. Accordingly we probably would reject the Hilbert transform approach in favor of the much simple full-wave rectifier.

Just below we want to take another look at the resonator case. For the moment we need to caution that we must not make final judgments based on a comparison of Fig. 18a and Fig. 18b (good) with Fig. 16a and Fig. 16b (kind of ugly). The reason is that the low-pass filters in the two cases are quite different. The cutoff for Fig. 16a and Fig. 16b is about 0.00015, as compared to 0.0075 (50 times higher) for Fig. 18a and Fig. 18b. In Fig. 16a and Fig. 16b, we had to set the cutoff very low so as to low-pass the pings. Yet we did not succeed totally in doing this (hence the spikes). If we had made the cutoff even lower, we might have gotten rid of most of the spikes, at the expense of blurring the whole envelope into one blob.

ANOTHER LOOK AT THE RESONATOR

We have just mentioned we intend to look at the resonator in a more fair-minded manner. This we do at the expense of producing figures that do not demonstrate the pinging of the resonator so well. Yet this will show several interesting results.

For Fig. 19a, we again have a waveform produced by inputting pulses into a resonator and imposing an ADSR envelope. Indeed this already looks strange because the negative portions are only about half the amplitude of the positive sides. Why is this?

What we have done here is first to increase the pulse density of the excitations by reducing their spacing from 250 to just 25. At the same time, we have reduced the radius of the resonator poles from 0.98 to 0.90. This means that the resonator is of a much lower Q and does not ring as much. The inset of Fig. 19a shows the impulse response of the resonator. (Compare to Fig. 14b.) Note that the positive excursion is much greater than the negative excursion due to the increased damping. Since Fig. 19a is made up of 400 of these rings, we understand the favoring of positive amplitudes relative to negative ones. Fig. 19b is a zoom-in of Fig. 19a, and we notice this imbalance, and the fact that the excitations are dense enough that ringings are not completed before new excitations are started. Of course, it is always true that the ringings are not complete (they are exponential), but here, we see it clearly (again compare to Fig. 14b).

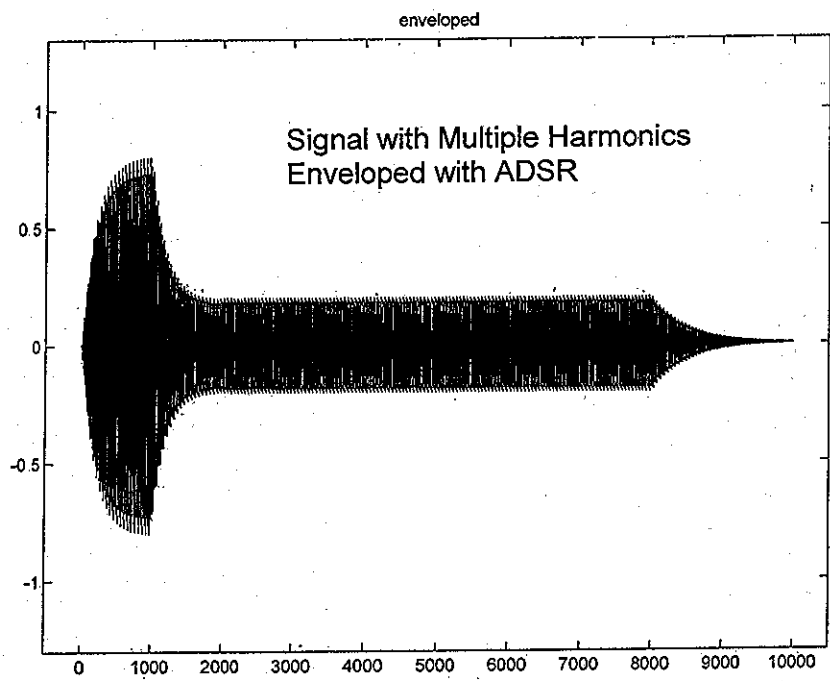


Fig. 17a A signal consisting of four harmonics with an imposed ADSR.

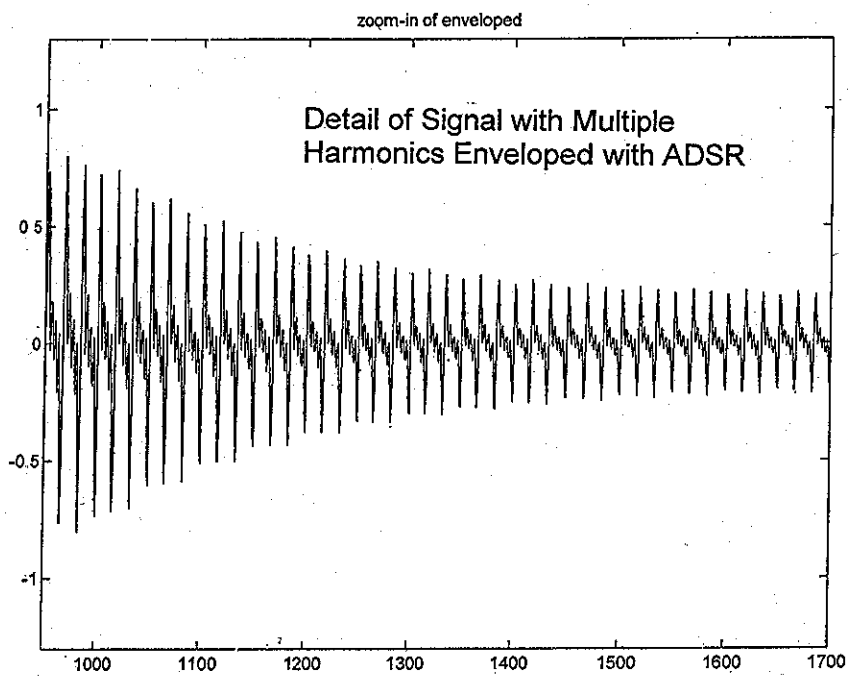


Fig. 17b A zoom-in of Fig. 17a

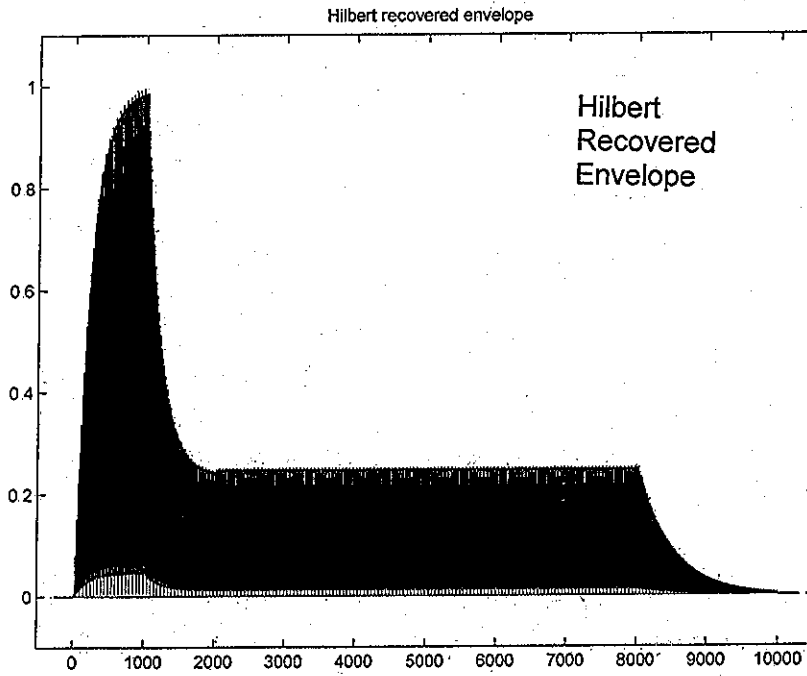


Fig. 17c The Hilbert recovered envelope of Fig. 17a

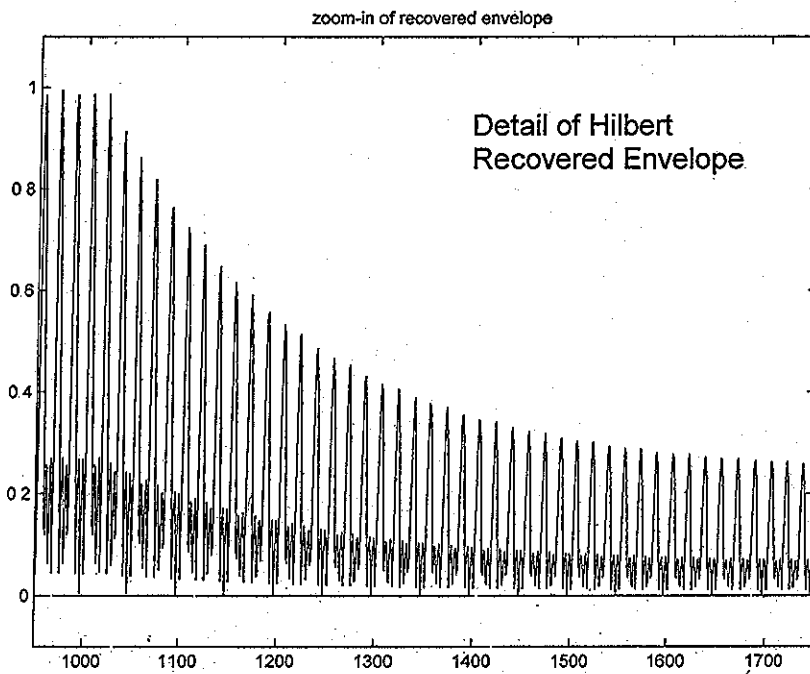


Fig. 17d A zoom-in of Fig. 17c

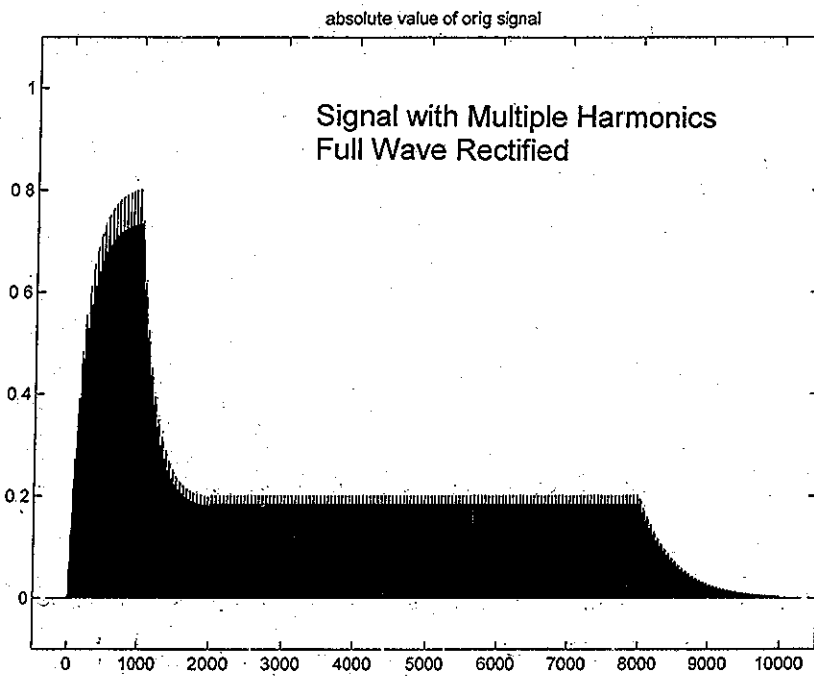


Fig. 17e The absolute value of Fig. 17a

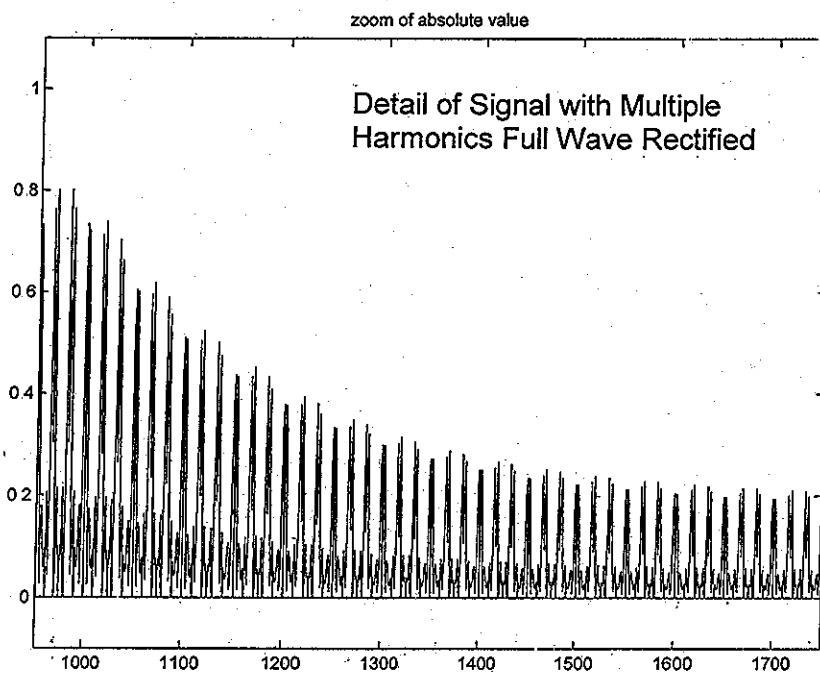


Fig. 17f A zoom-in of Fig. 17e

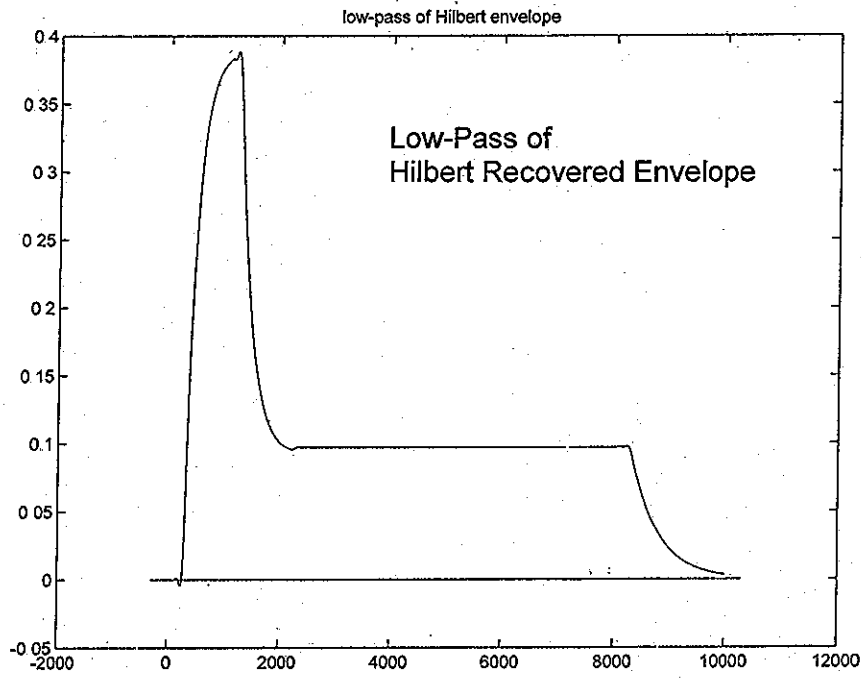


Fig. 18a Low-pass of Hilbert envelope

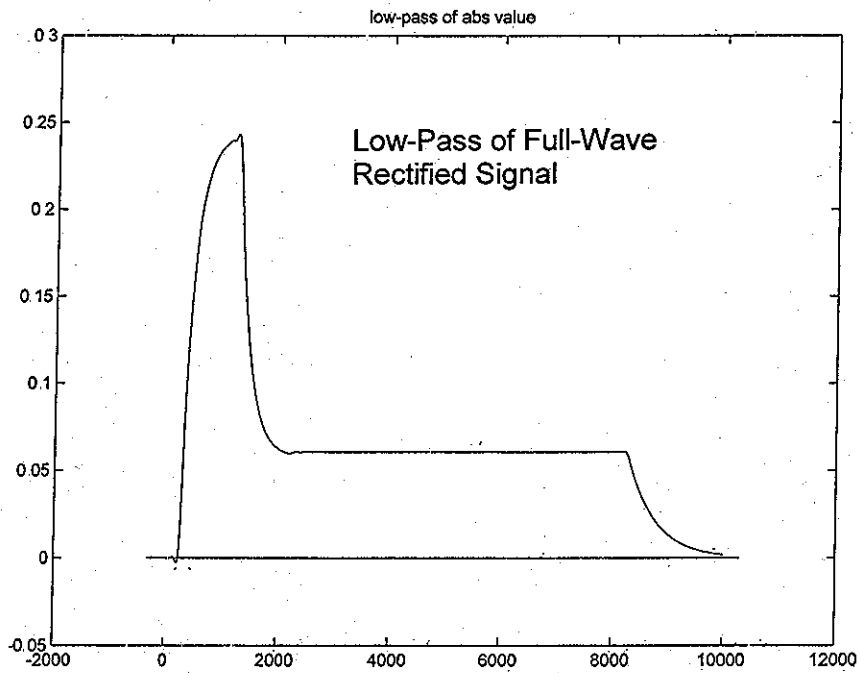


Fig. 18b Low-pass of absolute value

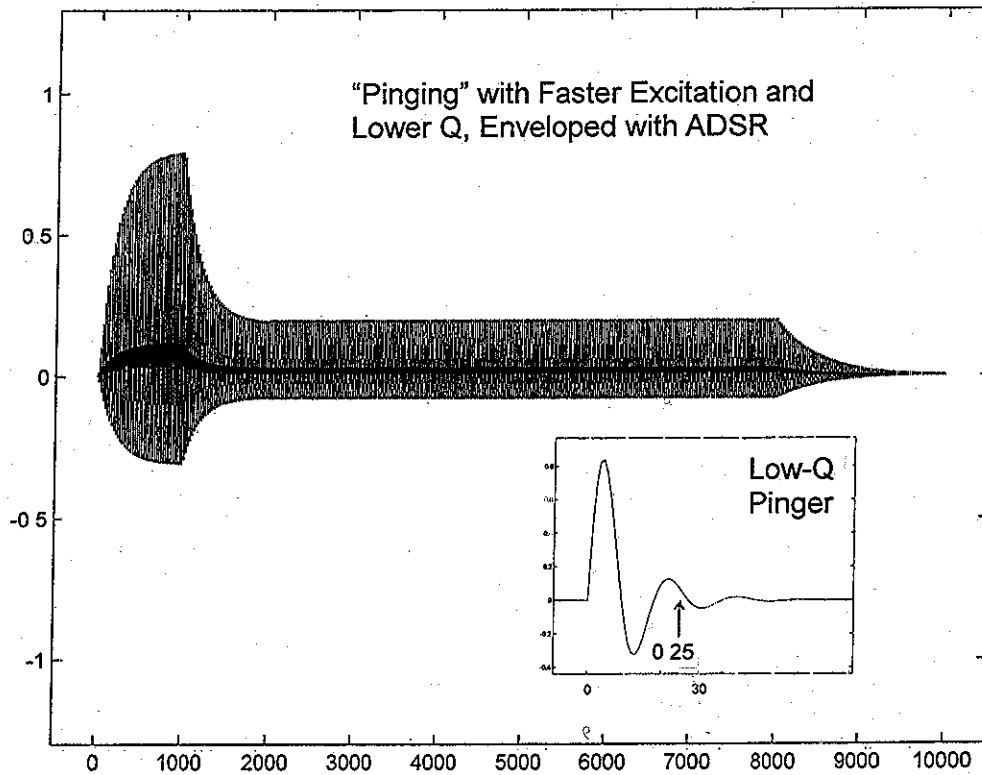


Fig. 19a Pulses into a resonator. Here the spacing of excitation pulses is 25 samples. The resonator's impulse response is shown in the inset.

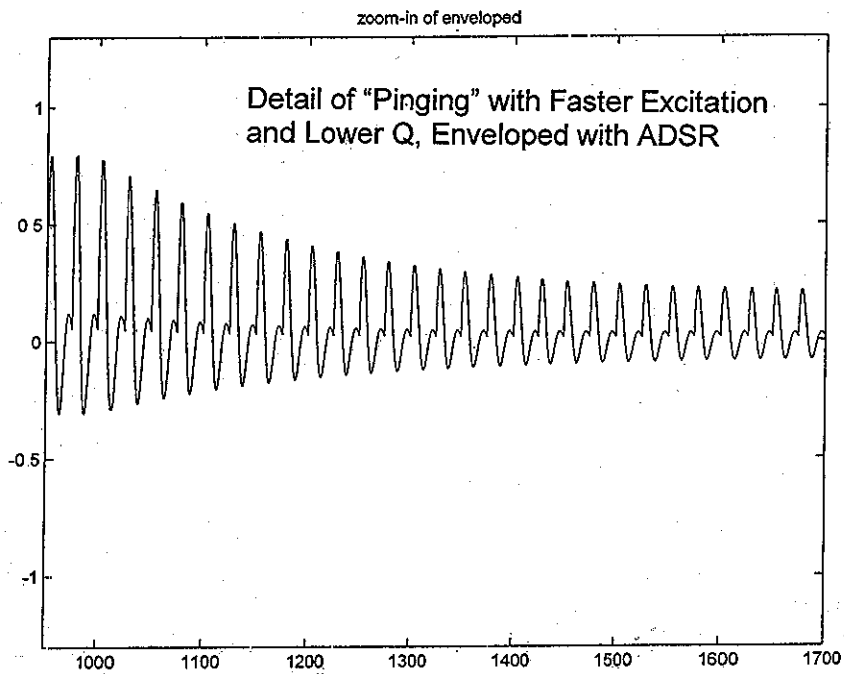


Fig. 19b Zoom-in of Fig. 19a. Note that new excitation pulses arrive before the resonator's impulse response (inset of Fig. 19a) becomes small.

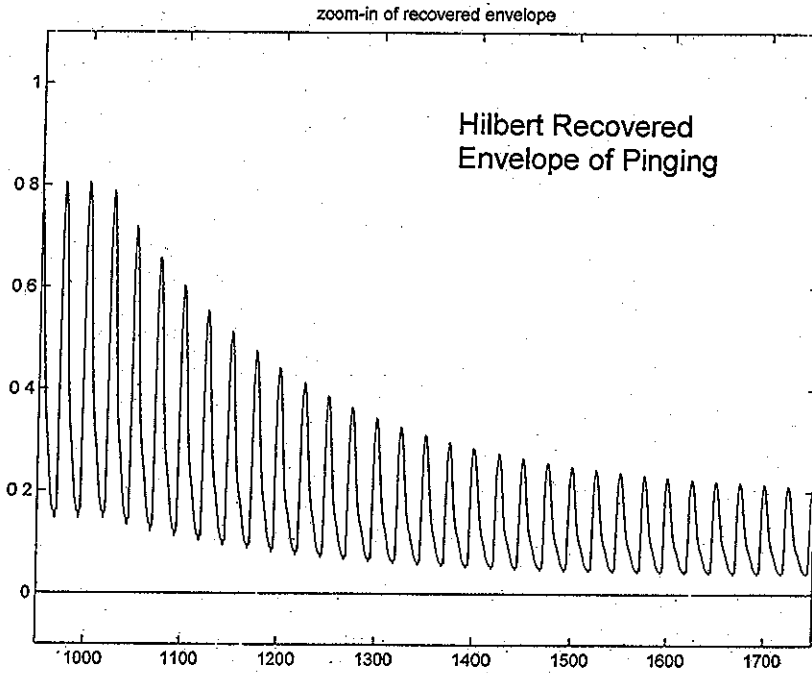


Fig. 19c Zoom-in of Hilbert recovered envelope for waveform of Fig. 19a. Note that the method seems to be struggling between recovering the envelopes of the component "pings" (about 31 here) and the general trend of the ADSR (a "dc" level from about 0.5 to 0.1 from left to right).

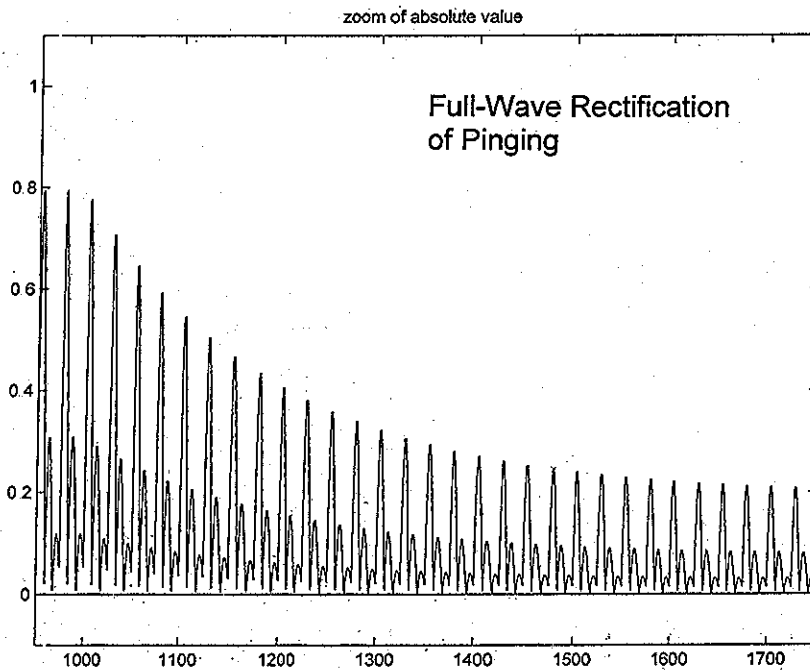


Fig. 19d Zoom-in of absolute value of the waveform of Fig. 19a. Note the unequal lobe sizes due to the differences of positive and negative ring lobes.

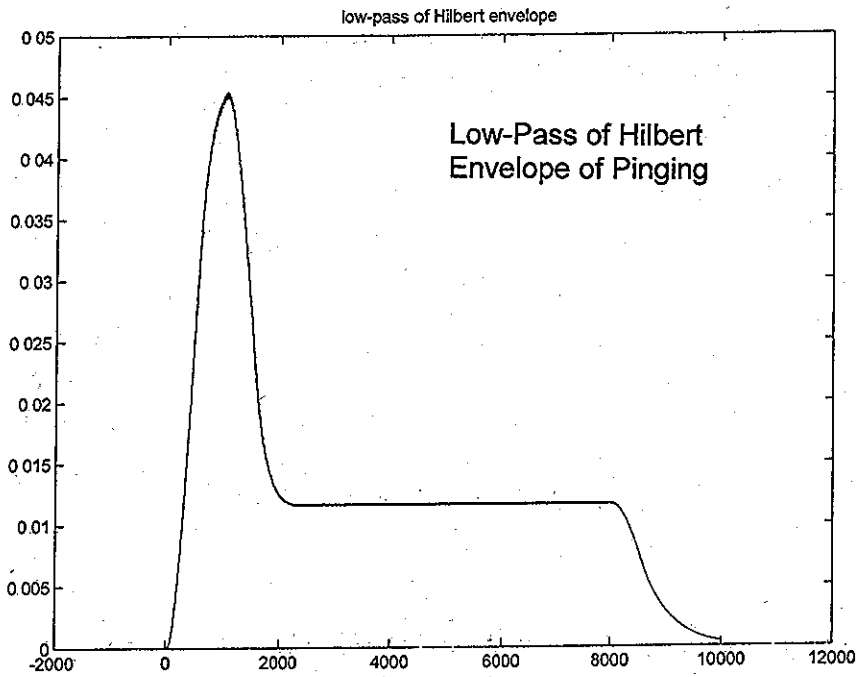


Fig. 19e Low-pass averaging of the recovered Hilbert envelope from Fig. 19a (cutoff 0.0075).

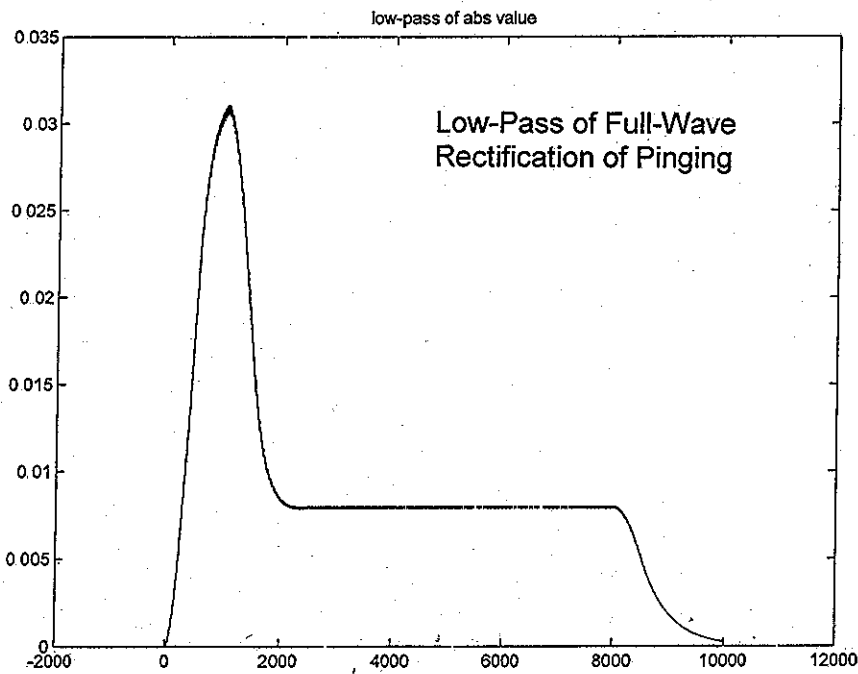


Fig. 19f Low-pass averaging of absolute value of Fig. 19a (cutoff 0.0075).

Fig. 19c shows a zoom-in of the recovered Hilbert envelope while Fig. 19d shows the corresponding zoom-in of the absolute value of the signal. This result is quite revealing in that there is a noticeable difference between these two. Above we noticed that the Hilbert transform method could recover an "event" envelope (the ADSR) for a single frequency, but tended to recover the envelope of beat cycles, or component pings in other cases. Here (Fig. 19c) we may have trapped it in an ambiguous case. We do see the envelopes of the component pings (31 from left to right) but also a falling dc level from about 0.5 to 0.1 going left to right. This we can attribute to having made the distance between excitation pulses small relative to what we would normally think of as the duration of the resonator ringing.

In contrast, Fig. 19d is just the absolute value. Normally, taking an absolute value reduces the residual ripple that remains after low-pass filtering. (After all, this is why it is easier to build a power supply with a full-wave rather than with a half-wave rectifier.) In Fig. 19d, we see the large positive peaks (0.8 on the left side), but the negative peaks (as rectified) are only about 0.3 (again, on the left). In some sense, this result resembles a half-wave rectification more than a full wave one (compare to Fig. 15b). Intuitively, we would expect Fig. 19c to be somewhat easier to low-pass filter than Fig. 19d.

Fig. 19e shows the low-pass averaging for the Hilbert envelope case and Fig. 19f shows the low-pass averaging for the absolute value case. The quality of the envelope recovery looks similar in both these plots. In fact, if we look closely we see that there is slightly more residual ripple in the absolute value case (this can probably be seen best at the very peak of the envelopes). An analysis of the details of the data (not just the plots) shows that for the Hilbert envelope case, the residual ripple is about 1.2% while for the absolute value case, it is about 1.9%.

CONCLUSIONS

We see that it is possible, by proper choice of low-pass averaging, to reasonably recover the envelope from the absolute value, much as we have always supposed, or actually done. Further, in the case of single frequencies, we can use the Hilbert transform approach without any low-pass averaging. In the case of multiple frequencies (which actually includes the resonator case) we may find interesting results using the Hilbert transform approach. Combining the Hilbert transform approach with low-pass averaging, we may find results that are better than low-pass averaging of the absolute value. But the improvement may be marginal (1.2% rather than 1.9% in our example) at considerable computational costs from a practical point of view.

REFERENCES:

- [1] B. Hutchins, "An Introduction to Hilbert Transform Theory and Applications," Electronotes, Vol. 14, No. 134, pp 7-17, Feb. 1982. See also the references listed there – in particular, the books by F. Haber and D.J. Sakrison. Hilbert transform is covered in many books on communications systems as well as books on signal processing.
- [2] A.V. Oppenheim, R.W. Schaffer, & J.R. Buck, Discrete-Time Signal Processing, Second-Edition, Prentice-Hall (1999), Chapter 11 "Discrete Hilbert Transforms." Many other DSP books include Hilbert Transform Design. See also our own filter design notes, Section 5b, in Electronotes No. 199.
- [3] N. Bodley, "Is this a filterless envelope follower," Electronotes, Vol. 4, No. 32, pg 3, Dec. 15, 1973. Anon, "Not an Envelope Follower, But:" Electronotes, Vol. 5, No. 34, pg 18, January 25, 1974.