

DESIGN AND APPLICATION OF A HILBERT TRANSFORMER IN A DIGITAL RECEIVER

Matt Carrick (Northrop Grumman, Chantilly, VA, USA; matthew.carrick@ngc.com);
 Doug Jaeger (Northrop Grumman, Chantilly, VA, USA; douglas.jaeger@ngc.com);
 fred harris (San Diego State University, San Diego, CA, USA; fred.harris@sdsu.edu)

ABSTRACT

A common method of down converting a signal from an intermediate frequency (IF) to baseband is using a quadrature down-converter. One problem with the quadrature down-converter is it requires two low pass filters; one for the real branch and one for the imaginary branch. A more efficient way is to transform the real signal to a complex signal and then complex heterodyne the resultant signal to baseband. The transformation of a real signal to a complex signal can be done using a Hilbert transform. Building a Hilbert transform directly from its sampled data sequence produces suboptimal results due to time series truncation; another method is building a Hilbert transformer by synthesizing the filter coefficients from half band filter coefficients. Designing the Hilbert transform filter using a half band filter allows for a much more structured design process as well as greatly improved results.

1. INTRODUCTION

The digital portion of a receiver is typically designed to receive a signal at IF and down-convert it to baseband for additional processing. A common method of down-converting a signal is through the use of a quadrature down-converter, which both performs spectral translation as well as the conversion of a real signal to a complex signal. The down-conversion is operation multiplies one branch by a cosine carrier and another branch by a negative sine carrier which shifts the positive frequency segment of the signal's spectrum to baseband. The signal is then low pass filtered to remove the sum frequency terms. This processing chain is referred to as a quadrature demodulator [1] or quadrature down-converter. The quadrature down-converter method has the drawback that two low pass filters must be used after the down-conversion; one on the real and one on the imaginary branch.

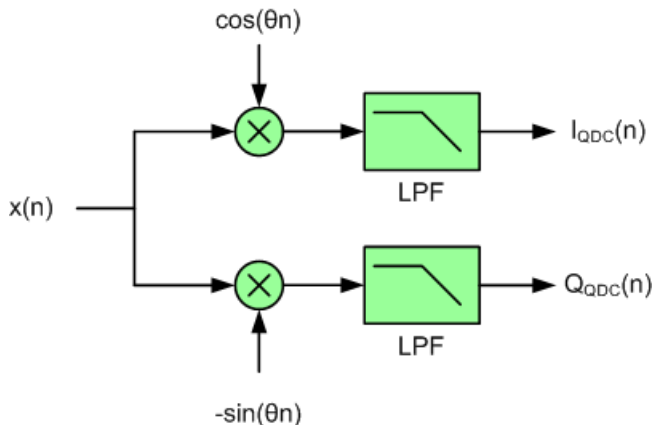


Figure 1: A quadrature down-converter

Another way of viewing the problem is that the quadrature down-converter not only extracts the desired segment of the spectrum it rejects the undesired spectral image, the spectral replica present in the Hermetian symmetric spectra of a real signal. Removing this image would result in a single sided spectrum which being non-Hermetian symmetric is the transform of a complex signal. The resulting single-sided spectrum can be complex heterodyned to baseband. This sequence of operations is the reverse order of operations for the quadrature down-converter. Instead of low-pass filtering the I and Q channels, the signal is first converted from a real signal to a complex signal using the Hilbert transform. The negative frequency image is removed by the Hilbert transform which allows the signal to be complex heterodyned to baseband [1].

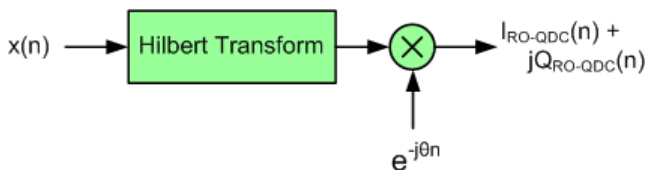


Figure 2: A reverse order implementation of the quadrature down-converter

It should be noted that the applications of the Hilbert transform within a digital receiver are not limited to this case of down-conversion; this example is simply providing the stimulus for designing the Hilbert transform filter.

2. HILBERT TRANSFORM BACKGROUND

A Hilbert transform is a convolutional operator which is commonly referred to as a wide band 90° phase shifter. The filter embedded in the Hilbert transform performs a $+90^\circ$ and -90° phase shift for negative and positive frequencies in a signal, respectively [2]. A continuous Hilbert transform is shown in equation 1, while the discrete representation of the Hilbert transform is shown in equation 2 [3]. The frequency response of the continuous filter embedded in the Hilbert transform is shown in equation 3. Here we see that spectral components on the positive frequency axis are phase shifted -90° while the spectral components on the negative frequency axis are phase shifted $+90^\circ$. Note that the phase shift preserves amplitude. This spectral response is illustrated in figure 3. As expected, the odd symmetric time signal $h(t)$ has an odd symmetric spectrum.

$$\hat{x}(t) = x(t) * h(t), \text{ where } h(t) = \frac{2}{\pi t} \quad (1)$$

$$\hat{x}(n) = x(n) * h(n),$$

$$\text{where } h(n) = \begin{cases} \frac{2}{\pi n} & n \text{ is odd} \\ 0 & n \text{ is even} \end{cases} \quad (2)$$

$$H(f) = -j \operatorname{sgn}(f) \quad (3)$$

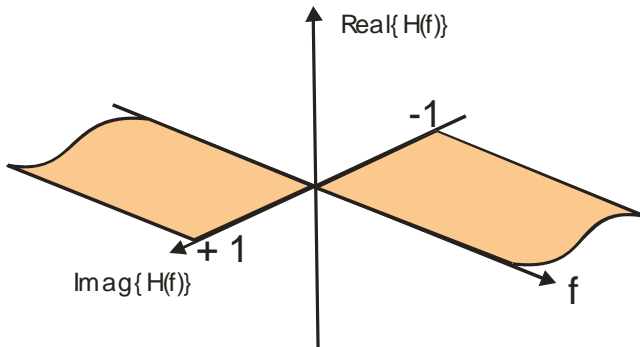


Figure 3: Frequency Response of the Hilbert transform Filter

The Hilbert transform is used to form the analytic signal, the complex time signal corresponding to the spectral content of the positive frequency axis. The analytic sampled data signal is described in equation 4.

$$a(n) = x(n) + j \hat{x}(n) \quad (4)$$

The block diagram demonstrating how to obtain the analytic signal with the Hilbert transform filter is shown in figure 4. Due to causality, when implemented, the output of the real branch will have to be delayed such that it is time aligned with the delayed output of the imaginary branch. One method of implementing this time delay is shown in figure 5.

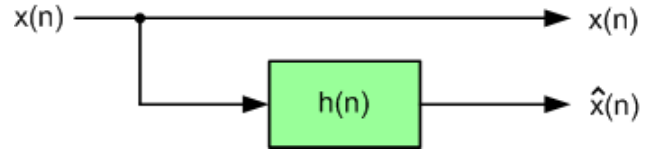


Figure 4: Block diagram for applying a Hilbert transform [3]

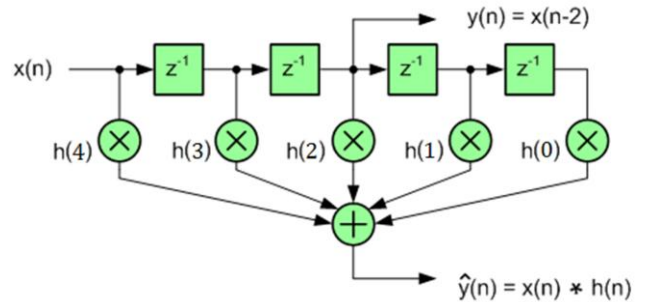


Figure 5: Time aligning the output of the real and imaginary outputs

The analytic signal formed by use of the Hilbert transform operation in figure 4 destructively cancels all of the negative frequency components in the signal. When we apply a complex heterodyne to the output of the analytic signal generator we translate to baseband the spectrum centered on the positive frequency axis and thus performs the same function as the quadrature down-converter. Comparing the two structures, figures 1 and 4, we see that the reverse order quadrature down-converter only requires a single filter versus the quadrature down-converter which requires two. Therefore, the reverse order quadrature down-converter should be used instead of the quadrature down-converter as long as the single Hilbert transform filter uses less than the number of taps in the two filters in the quadrature down-converter.

3. DESIGNING A HILBERT TRANSFORM FILTER USING A DISCRETE SEQUENCE

The discrete time series for the Hilbert transform was presented in equation 2. To determine a performance baseline for the filter let us use the filter coefficients directly from the discrete sequence. At this point we have not yet determined the number of taps required in this filter to meet

a given design specification, so for convenience of the discussion we choose a filter of length 31 taps. This is an odd number of weights chosen to assure a zero weight at the symmetry point of the filter.

Figures 6a and 6b show the impulse response and the frequency response of the Hilbert transform filter formed by truncating the impulse response to 31 weights. Notice the significant ripple in the amplitude of the spectrum about the nominal values of ± 1 . This ripple is Gibbs phenomena [3], the failure of the series to converge in the neighborhood of the discontinuity at normalized frequencies 0 and 0.5. This deviation from unity gain will prevent the destructive cancellation of the spectral components in the two paths that are supposed to exhibit equal amplitude with opposing phases. This imperfect cancellation due to the ripple leaves a residual image which would limit the utility of the process. The amplitude of the ripple is not reduced by increasing the filter length but rather is shifted towards the discontinuities. In the jargon of digital signal processing, the pass band ripple is attributed to the truncation of the infinitely long sequence described by equation 2. The amplitude of the ripple can be controlled by the severity of the truncation process. Applying a window other than the abrupt default rectangle window will result in a more gradual transition to the zero amplitude coefficient values which will significantly reduce the levels of pass band and stop band ripple levels.

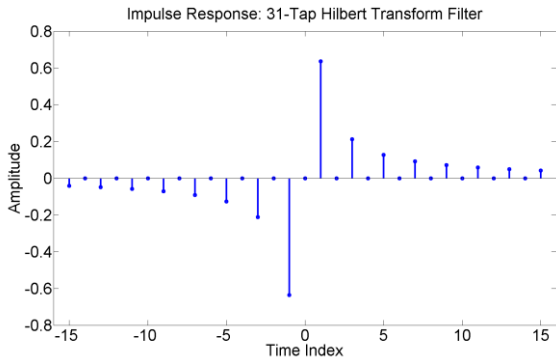


Figure 6a: Hilbert transform filter coefficients

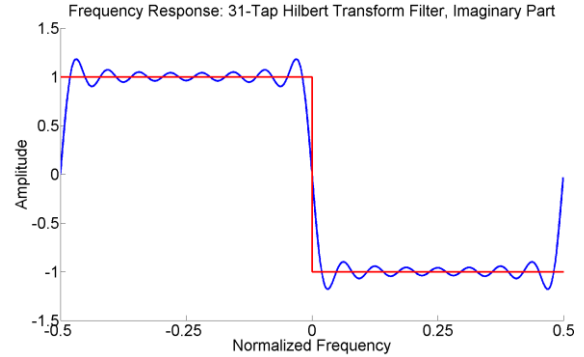


Figure 6b: Spectrum: Hilbert Transform Filter, ripple on order of 182 parts per thousand

The windowed filter coefficients in figure 7a are seen to more gently transition to zero than the rectangle windowed filter and correspondingly the spectrum in figure 7b shows a much smaller level of ripple relative to the nominal ± 1 response. The pass band ripple is now on the order of 6 parts per 1000 as opposed to the original 182 parts 1000 as shown in figure 7c.

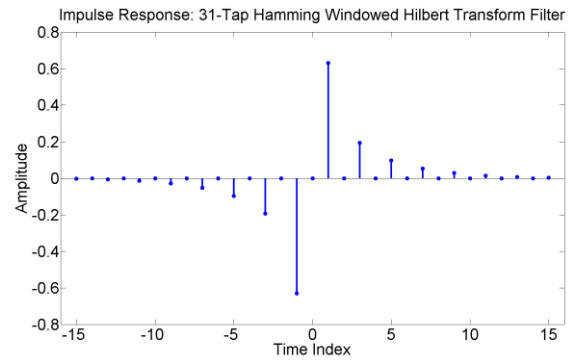


Figure 7a: Windowed filter coefficients

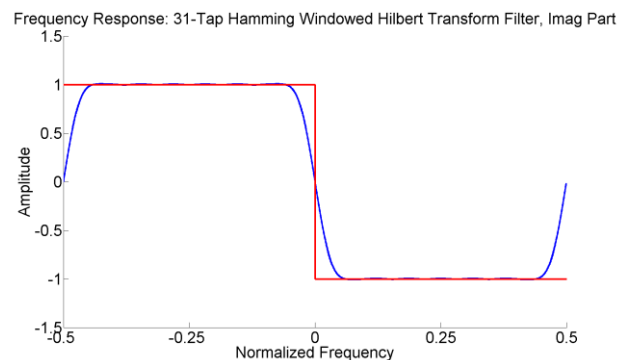


Figure 7b: Spectrum: windowed filter coefficients

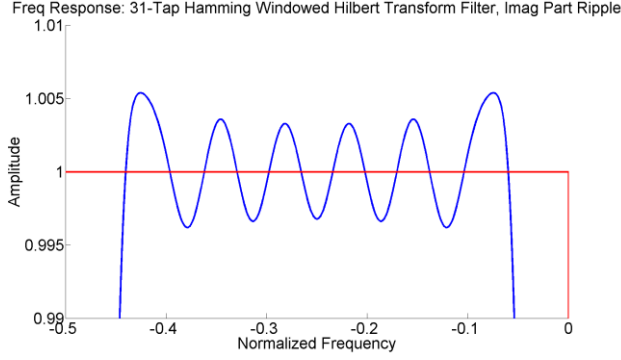


Figure 7c: The Pass Band Ripple: Windowed Coefficients on order of 6 Parts in 1000

Knowing that we can control the ripple level with smooth windows and that we can set that level to any desired value, we now address the matter of filter length and the particular windowing function. We ask, ‘what order filter should be chosen’ and ‘what is the optimum window to apply’? We choose to answer these questions by introducing a slight change in perspective of the filter design process that will guide us to obvious answers to these two questions.

4. SYNTHESIZING A HILBERT TRANSFORM FILTER WITH HALF BAND FILTERS

We now examine the structure of the Hilbert transform filter through its frequency domain characteristics. The Fourier transform of the Hilbert transform filter sequence was presented in equation 3 for the continuous filter and is presented in equation 5 for the sampled data filter [3].

$$H(\theta) = -j\text{sgn}(\theta), \quad \frac{\pi}{2} < \theta < \frac{\pi}{2} \quad (5)$$

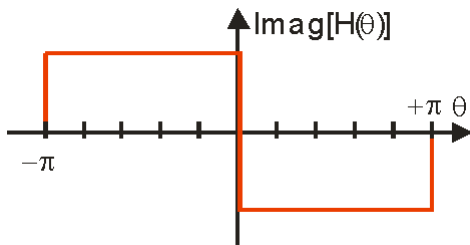


Figure 8: Fourier transform of the sampled Data Hilbert Filter

We approach the design of filter matching $H(\theta)$, the response in figure 8, by separating the positive and negative portions of the response. The positive portion of the response is $-G(-\theta)$ while the negative portion of the response

is $G(\theta)$ as shown in figure 9. A filter will be built for each response and then later combined to form the total response.

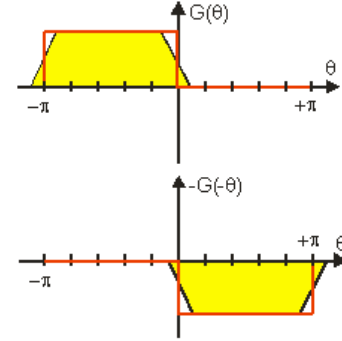


Figure 8: Fourier transform of the sampled Data Hilbert Filter

Looking at the piecewise spectrum in figure 9, each response can be described as a down-converted and an up-converted half band filter [4] scaled by +1 and -1 respectively. The equations describing the filters in figure 9 are given in equations 6 and 7. The positive frequency half band is up-converted to the quarter sample rate by the complex exponential $e^{jn\pi/2}$, while the negative frequency half band is correspondingly shifted by $e^{-jn\pi/2}$.

$$G(\theta) = H_{HB}\left(\theta + \frac{\pi}{2}\right) \quad g_{+\theta}(n) = h_{HB}(n)e^{-jn\pi/2} \quad (6)$$

$$G(-\theta) = H_{HB}\left(\theta - \frac{\pi}{2}\right) \quad g_{-\theta}(n) = h_{HB}(n)e^{+jn\pi/2} \quad (7)$$

Where $h_{HB}(n)$ is the baseband half band filter

To form the full response, $H(\theta)$, partial responses $-G(-\theta)$ and $G(\theta)$ must be summed and multiplied by j as shown in equation 8. The time domain equivalent is given in equation 9, which relates the half band filter coefficients to the Hilbert transform filter coefficients.

$$H_{HT}(\theta) = -j\left[H_{HB}\left(\theta - \frac{\pi}{2}\right) - H_{HB}\left(\theta + \frac{\pi}{2}\right)\right] \quad (8)$$

$$h_{HT}(n) = -j\left[h_{HB}(n)e^{+jn\pi/2} - h_{HB}(n)e^{-jn\pi/2}\right] \quad (8.1)$$

$$h_{HT}(n) = 2h_{HB}(n)\frac{e^{+jn\pi/2} - e^{-jn\pi/2}}{2j} \quad (8.2)$$

$$h_{HT}(n) = 2\sin\left(\frac{n\pi}{2}\right)h_{HB}(n) \quad (9)$$

The question of how to choose the filter order and the best windowing function has been converted to that of designing a FIR half-band filter, a task with which we are quite familiar. After designing the half band filter $h_{HB}(n)$

the Hilbert transform filter coefficients can be synthesized by multiplying the coefficients by a sine wave.

5. HALF BAND FILTER DESIGN

The analytic signal formed by the combination of the output of the Hilbert transform filter and its delayed input the delayed can be thought of as the result of a filtering operation designed to reject the negative frequency image in the double sided spectrum of the signal as seen in figure 10a. The baseband version of the same problem is shown in Figure 10b.

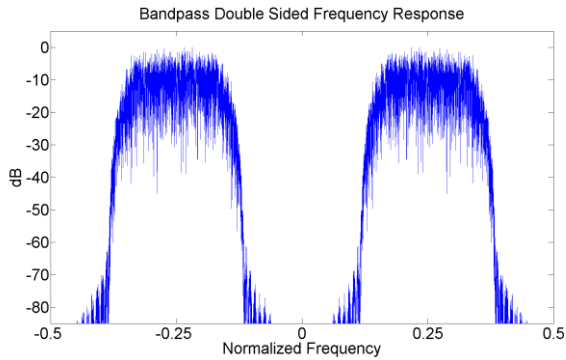


Figure 10a: Example of double sided spectrum

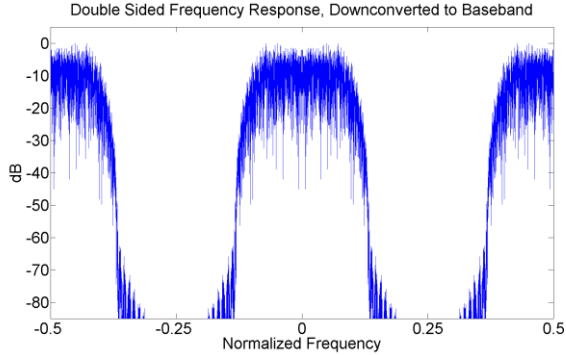


Figure 10b: Down converted double sided spectrum

By viewing the problem as that of rejecting the negative frequency image (figure 10a) or that of rejecting an image at the half sampling rate (figure 10b), the design of the Hilbert transform filter is exactly the same as a design of a standard FIR filter. That design requires definition of pass band boundaries, of stop band boundaries, hence transition band boundaries, and pass band and stop band ripple levels from which the prototype filter length can be estimated. Note that the spectrum of a half band filter is even symmetric about the origin and odd symmetric about the quarter sample rate ($\pi/2$) so that the stop band and pass band width are equal as are the pass band and stop band ripple levels.

To reject the image at half the sampling rate in figure 10b, we must design a low pass half-band filter. The filter

pass band width must equal the pass band width of the baseband signal which in this example is 0.1π radians/sample. Since the image at baseband also resides at the half sampling rate, the width of the stop band is the same as width of the pass band, which is also 0.1π radians/sample. Subtracting the bandwidth of the reflected image (0.1π) from the half sampling bandwidth (0.5π) provides us with the stop band boundary, which is 0.4π radians/sample. The transition bands is the difference of the stop band and pass band boundary edges which is 0.3π radians. An example is shown in figure 11.

This discussion calculating the pass band and stop band reflects the defining property of the half band filter: the transition band is always centered on the quarter sample rate and the width of the pass band is equal to the width of the stop band. A generalized example of this constraint is illustrated in figure 11.

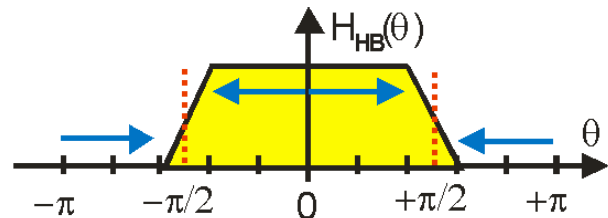


Figure 11: General frequency response of a true half band filter

The two remaining parameters to define for the half band filter are the stop band attenuation and the filter order. Specifying the desired level of stop band attenuation as the parameter, A , in dB, the filter order can then be approximated using equation 10 [4].

$$N \cong \frac{f_s A(\text{dB})}{\Delta f 22} \quad (10)$$

Where f_s is the sample rate, $A(\text{dB})$ is attenuation in dB, and Δf is the transition bandwidth. For our example, $f_s=2\pi$, $\Delta f=0.3\pi$, and $A(\text{dB})=80$, therefore $N \approx 25$

Now that all of the parameters have been defined, the half band filter can be designed. The coefficients can then be calculated with the standard MATLAB call to the Remez or firpm algorithm (with frequency vector normalized to $f_s/2$) as shown below.

$$\begin{aligned} f &= [0 \ 0.1 \ 0.4 \ 0.5]/0.5; \\ a &= [1 \ 1 \ 0 \ 0]; \\ h_{\text{HB}} &= \text{firpm}(N-1, f, a); \end{aligned}$$

Where $N-1$ is the filter order, 1-less than the number of filter coefficients, as estimated by equation 10

To ensure a zero a symmetry point of the filter, the order is decreased from 25 to 24 while still meeting the original specifications. The filter coefficients are then computed with the MATLAB code provided above. The response of the filter is overlaid with the down-converted spectrum in Figure 12. The passband of the signal is within the 3dB cutoff of the filter, and the image residing at the half sample rate is rejected by the sidelobes which are 80 dB down by the time the passband of the image is reached. It should also be noted that this filter is also the same filter that would be used in the quadrature down-converter; therefore the quadrature down-converter uses a total of twice the number of taps as compared to the Hilbert transform filter.

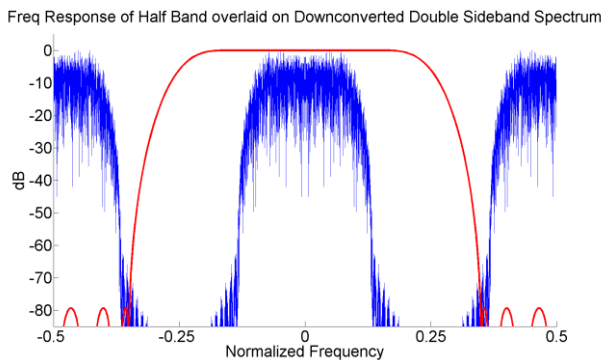


Figure 12: The frequency response of the half band filter coefficients is overlaid against the down-converted baseband signal

The half band filter coefficients are then synthesized into Hilbert transform filter coefficients using equation 9. The Hilbert transform filter coefficients are shown in figure 13 and the resulting frequency response in figure 14.

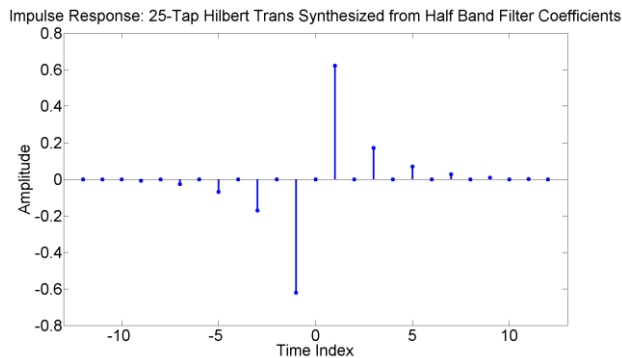


Figure 13a: Hilbert transform filter coefficients, synthesized from half band filter coefficients

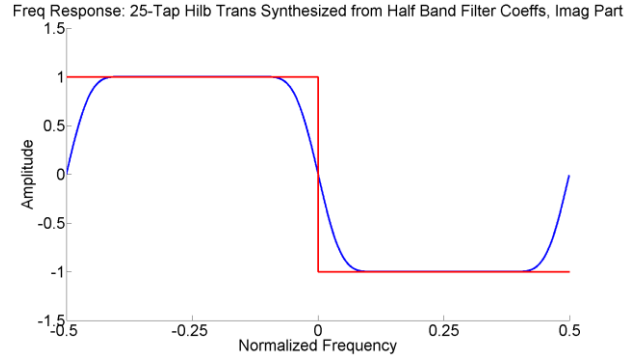


Figure 13b: The frequency response of the Hilbert transform filter coefficients

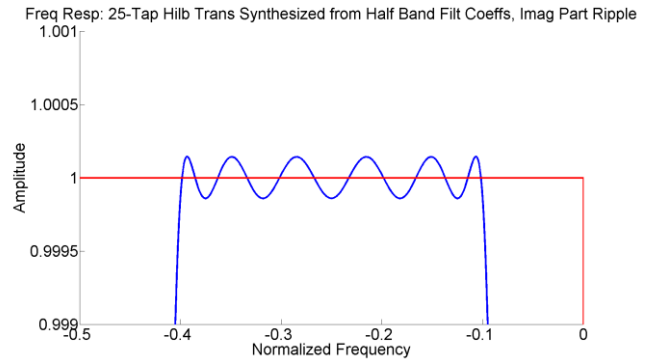


Figure 13c: The passband ripple is now less than 2 parts per 10,000

The filter coefficients pass the initial sanity check as the impulse response in figure 13a shows a very similar response to previous responses as in figures 6a and 7a. The frequency response in figure 13b is also very similar to previous responses in figures 6b and 7b. Figure 13c shows a substantial improvement in passband ripple, as it is now 2 parts per 10,000.

The half-band based filter design method makes designing the Hilbert transform filter much simpler as it is reduced to the design of a FIR filter. By reducing it to a FIR filter design problem, the design parameters can be fine tuned more easily which allows better results.

The attraction of this conversion is that the half band filter has zero-valued weights for all of the even-indexed offsets from its point of symmetry. This means that the Hilbert transform filter will similarly exhibit zero valued coefficients at the even indexed offsets from its point of symmetry. Interestingly, since the time series and spectrum are odd symmetric, the index at the point of symmetry must also be zero so that all the even indexed coefficients about its point of symmetry are equal to zero.

The design of a half band filter is covered in detail in [4] and will not be further discussed here. Once the half band filter coefficients have been calculated they are then turned into Hilbert transform filter coefficients using

equation 9. We note that due to finite arithmetic, the even indexed coefficients will not be precisely zero. We should of course set these not quite zero valued coefficients to zero. P.P. Vaidyanathan has shown a clever variation of the half-band filter design using the Remez algorithm that only computes the non zero valued coefficients and we then insert the missing zero valued samples [5].

6. HILBERT TRANSFORM COMPUTATIONAL COMPLEXITY

Comparing figures 1 and 2 we recognize that the quadrature down-converter requires two filters while the Hilbert Transform based reverse order quadrature down converter requires only a single filter. It certainly makes sense to consider use of the reduced complexity option. A second consideration is that since the Hilbert transform based image reject filter reduces the bandwidth by 2-to-1 we are permitted to also reduce the sample rate by the same 2-to-1 ratio. Reducing the sample rate of the analytic signal aliases its input quarter sample rate to its output half sample rate. The spectrum can be down-converted to baseband by the complex exponential $e^{j\pi n} = (-1)^n$, which simply alternates the signs of successive output samples. Since the half band filters used by the Hilbert transform filter and the I-Q quadrature down-converter are essentially the same filters, the number of multiplies needed to implement the quadrature down-converter is twice that of the reverse order quadrature down converter due to the required second mixer in the quadrature down-converter.

We have already observed the even indexed coefficients of the Hilbert transform filter are zero, which reduces the number of multiplies to half of the filter order, or in this case $24/2 = 12$. We can further reduce the number of multiplies by folding the Hilbert transform filter about its symmetry point and add the two samples that use the same coefficient either side of the symmetry point, bringing the number of multiplies to $N/4$, or $24/4 = 6$.

7. CONCLUSION

The Hilbert transform filter allows a real signal to be transformed into its complex representation. Using this method to convert the signal to its analytic complex representation and then performing a complex heterodyne to baseband requires approximately one-quarter of the resources as a quadrature down-converter. A Hilbert transformer can be quickly designed by generating the coefficients using the discrete sequence equation, although this produces poor results due to sequence truncation. These results can be improved by windowing the coefficients, but choosing the optimal windowing function requires additional information. Another method that will produce the same results synthesizes a Hilbert transformer filter from

a prototype half band filter. This design method is well structured, and easy to analyze.

8. REFERENCES

- [1] J. Barry, E. Lee, and D. Messerschmidt, *Digital Communication, Third Edition*, Kluwer Academic Publishers, Norwell, MA, 2004. Pages 20, 51.
- [2] J. Proakis, *Digital Communications, Fourth Edition*, McGraw Hill, New York, NY, 2001. Page 150.
- [3] A. Oppenheim, *Discrete-Time Signal Processing, Second Edition*, Prentice Hall, Upper Saddle River, NJ, 1999. Pages 363, 468, 789-792.
- [4] f. harris, *Multirate Signal Processing for Communication Systems*, Prentice Hall, Upper Saddle River, NJ, 2004. Pages 206-211.
- [5] P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice Hall, Upper Saddle River, NJ, 1993. Pages 153-154, 185.