To appear in *IEEE Potentials*

Signals and Systems II

Part IV: DSP approaches to IQ modulation and demodulation

Jeffrey O. Coleman
Naval Research Laboratory, Radar Division

This six-part series is a mini-course, focused on system concepts, that is aimed at the gap between Signals and Systems and the usual first DSP course.

Part III discussed analytic signals and linear data modulation. This fourth article in the series is about analog and DSP-based IQ demodulation. Figures are numbered in one sequence across the entire series, and gaps appear in their numbering in some individual articles. Figures are posted for instructional use on the author's website: http:// alum.mit.edu/ www/ jeffc

PART IV

Analog IQ modulation and demodulation

Figure 23 from Part III, reproduced here, presents two demodulator structures that might naturally be termed the shift-then-filter and filter-then-shift approaches, on the left and right respectively. As pictured, the latter requires an analog analytic filter, but in fact few engineers alive today have ever seen an all-analog filter-then-shift in-phase and quadrature (IQ) demodulator. Those who have are likely to speak of it as using a *phasing method* and a *Hilbert transformer*, a 90° phase shifter, reflecting the clumsier approach to analysis of an earlier era.

In contrast, analog shift-then-filter IQ demodulators and analog IQ modulators certainly remain in widespread use, though in ever declining abundance. But why declining?

Mixers are analog devices that approximately multiply signals by sinusoids. Their imperfections contribute spurious frequency-shift impulses at harmonics of the desired shift frequency, as illustrated in the Fig. 25 demodulator. *Double-balanced mixers* greatly suppress the even-order carrier harmonics, and if the carrier frequency is chosen carefully, filter stopbands may be able to suppress the spurious signal terms created by the other spurious impulses. Mixers also introduce signal nonlinearities, a topic beyond the scope of this series.

However, even if mixers were replaced with perfect analog multipliers, a spurious impulse at the negative of the shift frequency would effectively arise from inevitable imperfections in the gain ratio of the I and Q signal paths and in the phase difference between the two carriers. In 1938 an unfortunate pilot became famous as Wrong-Way Corrigan by flying from New York to Ireland after misreading his compass on a flight supposedly headed to California. This impulse is the Wrong-Way Corrigan of frequency shifting—it shifts the right amount but in the wrong direction and therefore always contaminates the desired output signal with a little of that signal's conjugate. This spurious conjugate term is so difficult to adequately suppress in demanding applications that analog IQ modulators and demodulators are becoming endangered species.

For all these reasons, analog IQ converters are rapidly being replaced in practice with the DSP-based alternatives of the sort discussed next.

A wideband IQ downconverter

The Fig. 26 filter-then-shift sampling demodulator is typical of modern systems. It is a wideband demodulator, as the passband signal's bandwidth is a substantial fraction of its center frequency. Here the analytic-filtered signal is spectrally convolved with an infinite set of impulses. Looked at as frequency shifts, these create the desired output, a sampled version of the signal that was presumably originally input to some modulator. But here the spectral convolution itself is technically not sampling, as the uniformly spaced impulses do not include the origin. This family of frequency shifts is factored on the right for practicality into sampling at a "noble rate," decimation by two, and a frequency shift.

A frequency shift of a discrete-time signal is straightforwardly realized, amounting to a simple multiplication of impulse areas or the samples that represent them by the values taken by the complex exponential at the impulse times. Certain frequency relationships are especially simple, however. The Fig. 26 shift is by an odd multiple of half its input sample rate, so its input samples are simply multiplied by ± 1 in the realization.

A narrowband IQ downconverter

The bandlimited input to the narrowband IQ downconverter in the top half of Fig. 27 allows all-digital analytic filtering, but the embedding of the digital filter's narrowband input signal in wideband noise means that most of that digital filter's frequency-response period must be stopband. Because that stopband is so wide relative to the passband, here three digital filters, each operating at a different sample rate, are used together to obtain an analytic-filter frequency response as the product of their individual responses. The upper filter is a shaping filter and establishes the desired passband and transition bands using a small frequency-response period. Periodically repeating passbands are then masked out by two masking filters.

In the lower system, factoring of sampling has enabled noble reordering for practicality. The first masking filter applies a complex impulse response to a real input, and the other two filters apply real impulse responses to complex inputs.

The filter-response shapes are sketched differently in the upper and lower systems in Fig. 27. This is not to suggest that one be changed into the other but instead simply economizes on space while presenting two design options. The functions of the filters in both cases are the same, so pretend that the upper and lower responses are identical, as given either in the upper or in the lower system.

These two filtering design options, upper and lower, attempt to minimize filter computation in different ways. The frequency responses sketched for the upper filter set respect that the realization cost of each filter output sample is roughly proportional to the fraction of the frequency-response period over which the response is strictly specified. The high- and middle-speed masking filters have no passbands specified and fractional stopband widths of only 10% and 20% respectively, so while their larger frequency-response periods imply more frequent output samples, each sample's computation is extraordinarily simple. The shaping filter's low sample rate provides its computational economy, and it can be designed so that the product of all three filters has the desired nearly flat passband.

The lower filter set aims for computational efficiency in realization by making each of the three filters *halfband*. The "percentage specified" rule of thumb above and a halfband filter's computational efficiency are properly both topics for a DSP course, but halfband filters are so important in applications that examining the basics is justified anyway.

Halfband filters

The frequency response sketched in the second line on the left in Fig. 15 is that of a halfband filter because it displays halfband symmetry: two frequency-shifted copies of the complex frequency response (not just its magnitude) sum to a constant if the shifts differ by half the sample rate.

Halfband filters are suitable only for certain applications because of their implicit restrictions on passband and stopband boundaries. Typically a halfband filter has a well-defined stopband, and halfband symmetry implies that it is less than half the sample rate in width. Halfband symmetry then further implies a passband of the same width as the stopband, with the two transition bands separating them having midpoints separated by half the sampling rate.

Two closely related terms need not concern us further here but are good to be aware of. First, shifting and summing N copies of the periodic frequency response of an *Nth-band filter* yields a constant when the shifts are by $0 \ldots N-1$ times $1/N$ of its period. Occasionally Nth-band filters are termed *Nyquist* filters. Second, a periodic function in the time domain displays *half-wave symmetry* if zero results when it is summed with copy of itself shifted by half a period.

Halfband symmetry affects the efficiency of a filter's realization in two ways, one helpful and one unhelpful. It is certainly unhelpful in making the frequency response as close to unity in the passband as it is to zero in the stopband. Because of this, in most applications designing a halfband filter with enough stopband rejection results in unnecessary passband precision, and like all forms of filter overspecification, that increases complexity.

But halfband symmetry is most helpful in zeroing almost half the samples of the impulse response! The associated convolution terms need not be computed in the realization, so implementation complexity is reduced by nearly half.

The helpful trait is typically the dominant one, so when a halfband filter is application compatible, it is very often the most efficient approach. Generally the only way to be sure, however, is to carry out a design in some detail for both halfband and non-halfband approaches and compare.

Why does a halfband filter have all those zero impulse-response samples? The somewhat roundabout argument begins with the observation that a halfband filter is a special case of an Nth-band filter, which in turn is a discrete-time (d.t.) version of a Nyquist filter. It should be no surprise then that the transformations used in Fig. 14 from Part II, reproduced here, to derive Nyquist filtering now reappear in d.t. form in Fig. 15. In the latter the system on the left can be transformed using the most noble identity once and commutativity of multiplication twice to obtain the equivalent system on the right. If the filter possesses halfband symmetry, the ***parenthetically grouped*** frequency response on the right will be a simple gain, and the output signal, not shown, will look like the input signal give or take a scale factor.

The rest of the argument is in the time domain. The key is that the most noble identity applied in Fig. 15 only because we began on the left with an unusual input signal: its spectral period implies that only its even-numbered samples can be nonzero! The system output has the same property, because the last step is a decimation by two that discards odd-numbered samples.

Now consider two special cases. First, if the filter impulse response comprised a single unit impulse at an even-numbered sampling instant, the filter convolution would simply shift the even-sample-only input by an even number of samples to yield an even-sample-only filter output that would survive the subsequent decimation unchanged. Second, if the filter impulse response instead comprised a single unit impulse at an odd-numbered sampling instant, the filter convolution would shift the even-sample-only input by an odd number of samples to yield an odd-sample-only filter output that would then be lost in the decimation.

Actually neither special case applies, and we must assume the filter impulse response contains both even- and odd-numbered samples, resulting in both even- and odd-numbered filter output samples. But outputs due to odd-numbered filter impulse-response samples are all lost to decimation. The outputs due to even-numbered filter impulse-response samples all survive decimation to form the system output.

Now we are at the heart of the matter, because on the right in Fig. 15 we have argued that if the filter is halfband, the system output is the same as the system input. This means that only one of the filter's even-numbered impulse-response samples can be nonzero! If a second one were nonzero, a second copy of the input would propagate through the filter to the output with some other net delay, and the system would be more than a simple gain!

The more-or-less mathematics-free argument above is actually much more involved than a straightforward mathematical development, which the reader is encouraged to work out: try stating the halfband condition in the frequency domain using a spectral convolution with two impulses and then converting that statement to the time domain.

Part V will continue with discussions of frequency conversion and its use in larger systems.

Read more about it

It was Dan Scholnik's MSEE thesis work (many years before his Ph.D.) at Michigan Tech that started my thinking down the path that led to this series, and the example designs published in the later papers below inspired this article's example IQ demodulators.

D. P. Scholnik, "Quadrature demodulation and modulation using bandpass sampling and reconstruction," Master's thesis, Michigan Technological University, Houghton, MI, USA, Aug. 1997.

Dan P. Scholnik and J. O. Coleman, "Integrated I-Q demodulation, matched filtering, and symbol-rate sampling using minimum-rate IF sampling," in *Proc. 1997 Symp. on Wireless Personal Communication*, Blacksburg, VA, June 1997.

D. P. Scholnik and J. O. Coleman, "Simple, exact models of sample-interleaving demodulators/modulators for quadrature bandpass sampling/reconstruction," in *Proc. 1997 Conf. on Information Sciences and Systems (CISS '97)*, Baltimore, MD, Mar. 1997.

About the author

Jeffrey O. Coleman (S'75–M'79–SM'99) joined the Radar Division of the Naval Research Laboratory (NRL) in Washington DC in 1978 then left it in 1985 for graduate studies, for a stint with The Boeing Company, and for a faculty position at Michigan Technological University from which he returned to NRL in 1997. His 1975/1979/1991 SBEE/MSEE/PhD degrees are from the Massachusetts Institute of Technology, Johns Hopkins University, and the University of Washington respectively, and his research is on theory and design methods in DSP. More: http://alum.mit.edu/www/jeffc

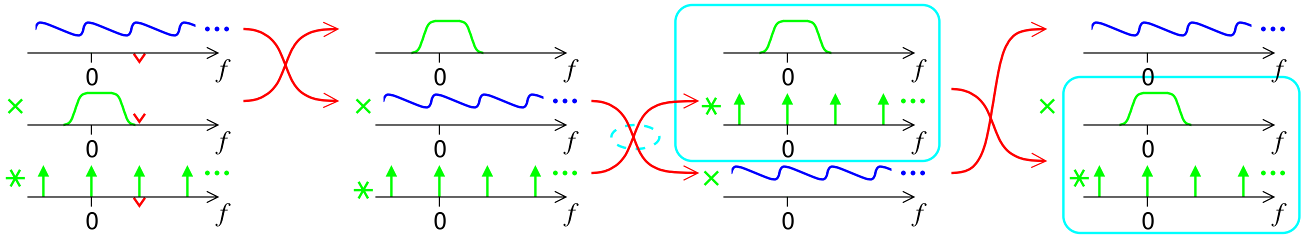

Fig. 14 (from Part II) The *most noble identity* (center) and product reordering (other red arrows) permit the top-to-bottom order of operations to be altered according to the *parenthetical groupings* and lead, on the right, to the Nyquist criterion.

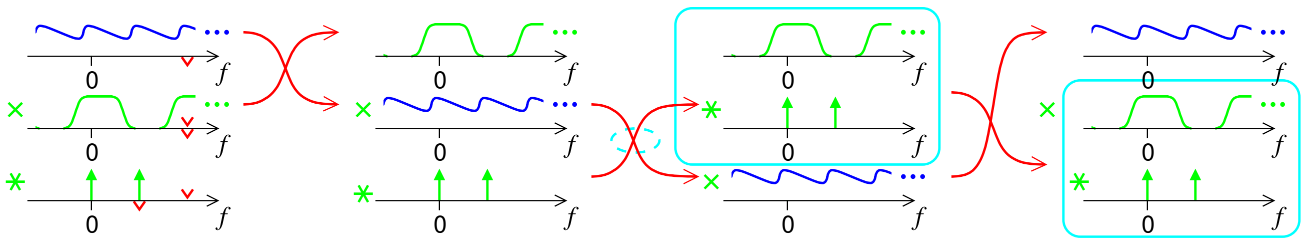

Fig. 15 Only one even-numbered sample of a halfband filter's impulse response is nonzero, the sample at the origin. This derivation parallels that of Fig. 14 and in fact Nth-band and Nyquist filters are closely related.

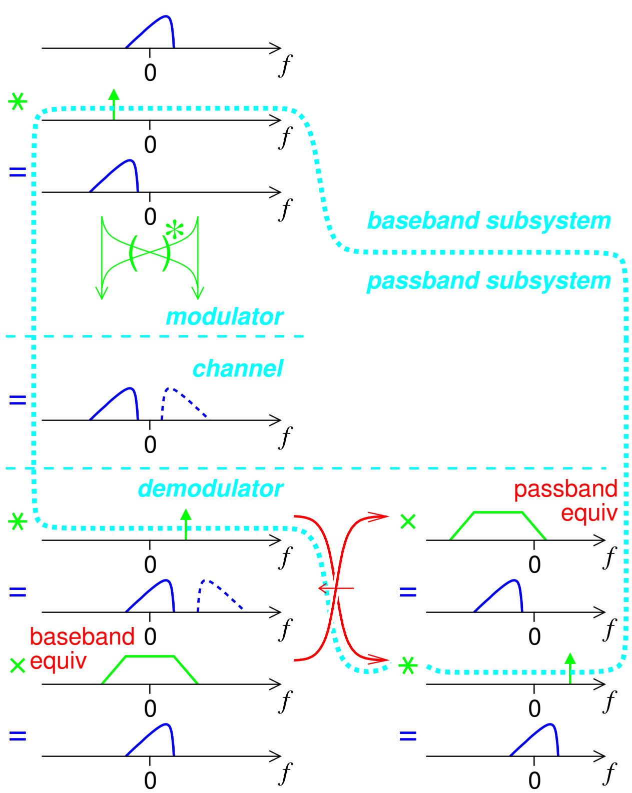

Fig. 23 (from Part III) A classic linear or IQ modulator and the usual shift-then-filter IQ demodulator on the left and an alternate, filter-then-shift IQ demodulator on the right.

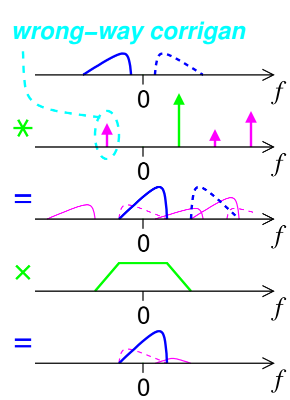

Fig. 25 In analog demodulators, spurious **harmonics** in the frequency shifting lead to spurious output terms.

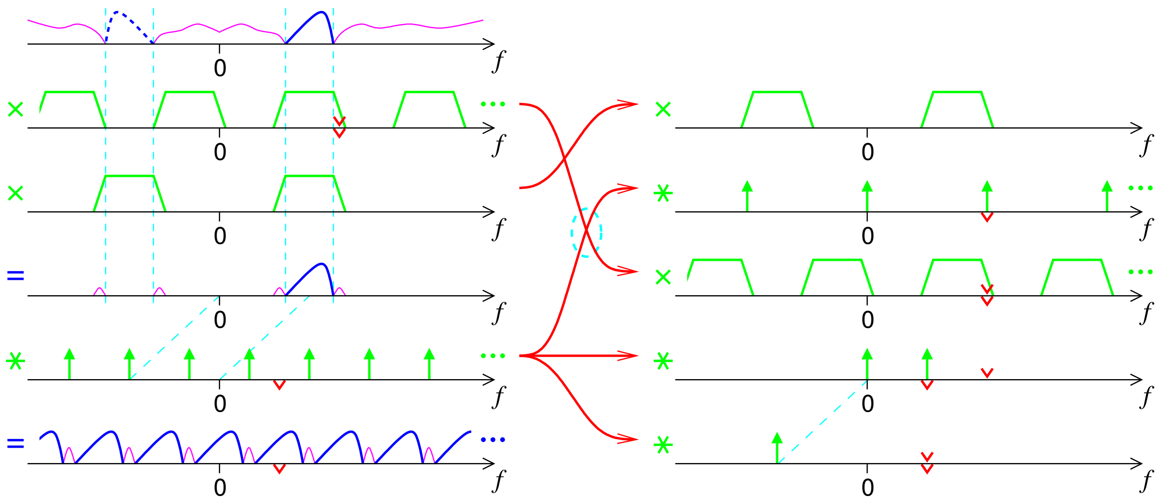

Fig. 26 A wideband filter-and-shift sampling demodulator.

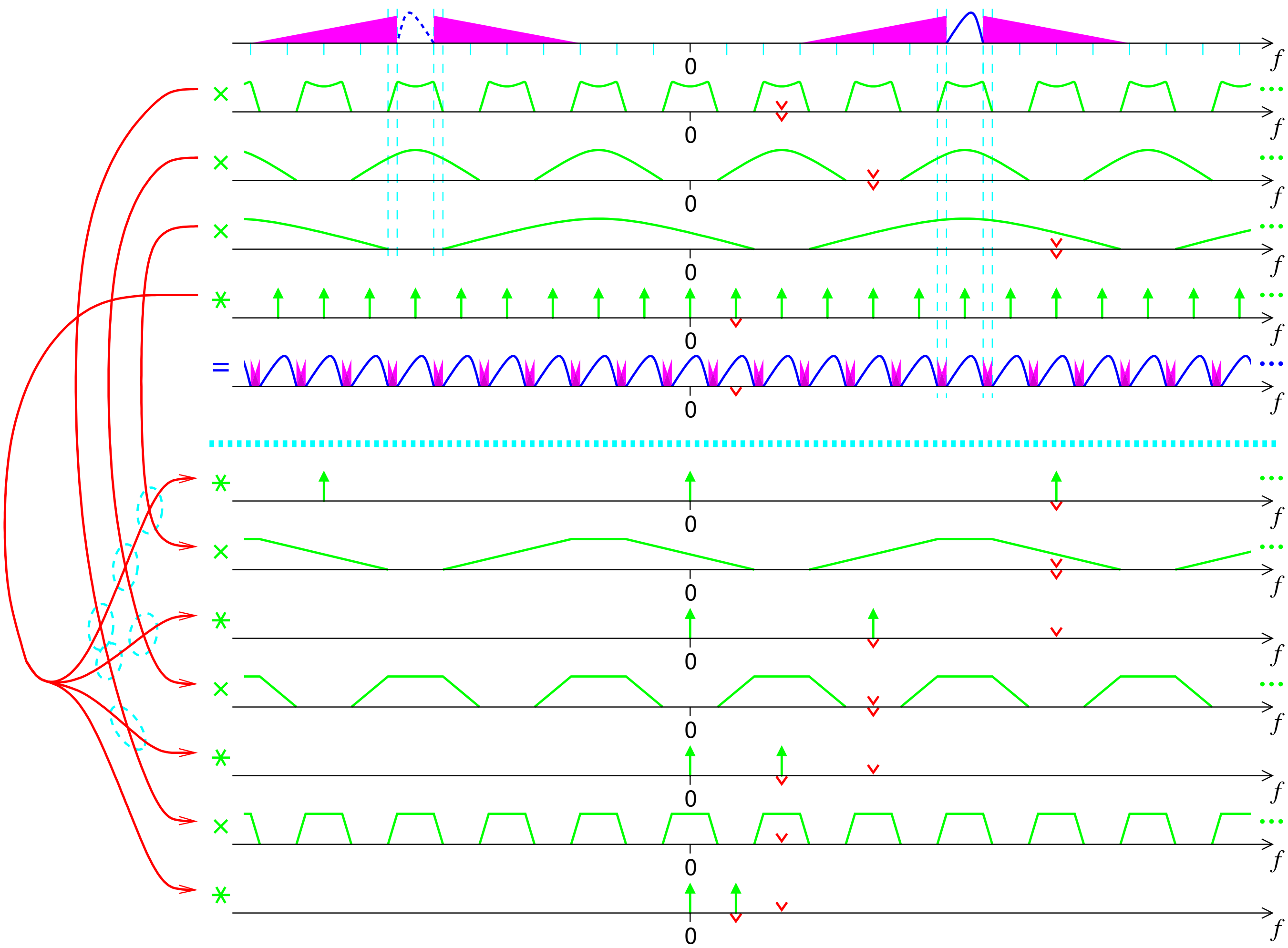

Fig. 27 This narrowband IQ demodulator rejects bandlimited but wideband input noise. Factorization of sampling and using the most noble identity six times transform the upper system into the lower. The differing filter responses in the upper and lower diagrams here illustrate a design choice rather than a transformation. Choose one or the other to use everywhere based on computational efficiency.