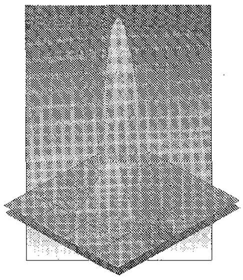


SP Forum

Larry Marple's title, "Restoring the Nyquist Barrier," was used without attribution to introduce the article by George Succi in the November 1995 *SP Forum*. The editor regrets that a note was omitted crediting Dr. Marple with this title, and notifying our readers that his article was forthcoming (see below).

Mark Fowler of Loral-Federal Systems and Bruce McKeever of Fujitsu both wrote to correct a printing error in the Succi article. In Fig. 2c, the triangular spectrum should be shifted so that its nulls occur at zero and multiples of F_s. Mr. McKeever continues, "By the way, using baseband lowpass filtering as Succi did for the Fig. 4 illustration, in the example of Fig. 2, would produce an inverted spectrum signal. It carries all the information, but as audio would sound strange. But it's still usable, for example, for single sideband multiplexing in analog communications."

Readers are encouraged to review the Editor's message for changes in the SP Forum department.

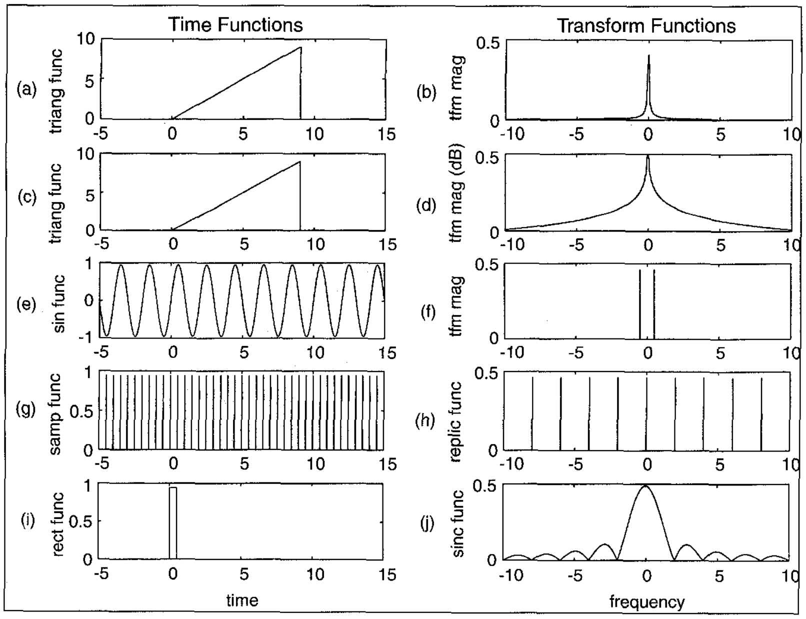
—Jack Deller

Restoring the Nyquist Barrier

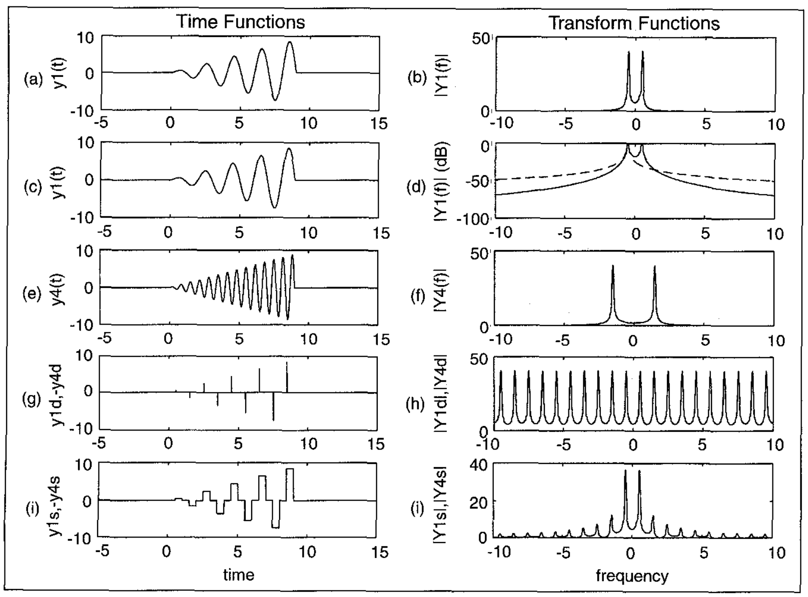
"Results of data analyzed by software simulation tools are meaningless." This was my first impression after reading the SP Lite article "Breaking the Nyquist Barrier" by Lynn Smith in the July 1995 issue [1]. This article contains a number of fundamental conceptual errors upon which I shall comment. The author has also rediscovered filter banks, despite extensive published art on this topic. However, beyond these conceptual and rediscovery issues, I was most struck by the dependence of the author on the use of a software simulation tool to justify the author's erroneous conclusions without an apparent full understanding of the graphical results that the tool produced. The Smith article reinforces a concern that I have been expressing to my colleagues in academia regarding the extensive use of DSP software simulation tools in virtual signal environments as a means for teaching signal processing. A selection of DSP software tools were highlighted in the article by Ebel and Younan [7] that appeared in the November 1995 *IEEE Signal Processing Magazine,* an issue dedicated, coincidentally, to signal processing education.

Specifically, there appears to be a growing dependence on these tools with canned experiments that fails to adequately prepare many students for solving real world signal processing problems. This is most manifest during technical interviews that I often conduct with new graduates who are candidates for employment. Without access to software tools during the interview, I have observed with increasing incidence that these graduates, when presented with situations involving typical signal processing applications of importance to my employer, are unable to confidently propose signal processing operations using only knowledge of basic signal processing principles. The most evident difficulty has been their inability to relate properties of continuous time-domain and spatial-domain signals with discrete-domain digital representations of and operations on those signals. Mathematical normalization of parameters (for example, the assumption of an unity sampling rate, or expressing frequency in radian units) often utilized in academic treatments of signal processing operations also handicaps students in forming an intuitive sense of time and frequency scale when confronted with actual signals and their transforms. Signal processing analysis and simulation software tools should be used knowledgably for purposes of productivity enhancement, and should not be used blindly without the capability to determine when the answer provided by the tool "looks right." This viewpoint is reminiscent of the debate concerning the introduction of hand calculators in public schools, in which it was argued whether hand calculators should be used by students as a substitute *before* learning the mathematical operations performed by the calculators or should be used only as productivity aids *after* they had substantial experience with the mathematical operations.

I would now like to demonstrate, by use of first principles, "restoration" of the Nyquist barrier of the demonstration signal used in the Smith article [1] by showing that it was never broken in the first place. I will do this armed only with four basic waveforms (depicted in Fig. 1), their transforms, and two variants of the convolution theorm. Specifically, if $x(t) \leftrightarrow X(f)$ designates the Fourier transform relationship between the temporal waveform $x(t)$ and its Fourier transform $X(f)$, while $y(t) \leftrightarrow Y(f)$ designates the Fourier transform relationship between



1. Four basic waveforms used to create the demonstration signals of Figs. 2 and 3. (a)& (c) Triangle temporal function (time-limited). (b) Magnitude of triangle function transform using linear scale (not bandlimited). (d) Magnitude of triangle function transform using logarithmic scale. (e) Sinusoidal temporal function. (f) Sinusoidal function transform (pair of impulse functions). (g) Sampling function (impulse function sequence). (h) Transform of sampling function (also an impulse sequence referenced as the replicating function). (i) Rectangular pulse temporal function. (j) Magnitude of rectangular pulse transform (a frequency domain sinc function).



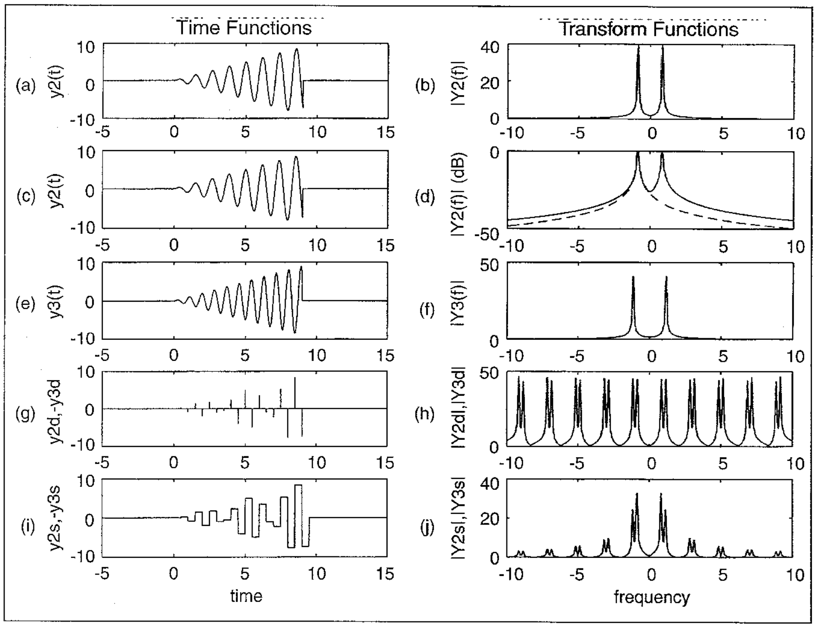
2. Construction steps to create Smith's y_1 and y_4 demonstration signals [1]. (a) (c) $y_1(t)$ obtained as product of Figs. 1a and 1c waveforms. (b) Magnitude of $y_1(t)$ transform, $|Y_1(f)|$, using linear scale. (d) Magnitude of $Y_1(f)$ using logarithmic scale (note: not bandlimited). (e) $y_4(t)$ waveform using 3X the sinusoidal frequency of $y_1(t)$. (f) Linear magnitude of $y_4(t)$ transform, $|Y_4(f)|$. (g) Sampled $y_1(t)$ or $-y_4(t)$ waveform obtained as product of Figs. 2a/2e and 1g; note that sampled waveforms $y_1d(t)$ and $-y_4d(t)$ are indistinguishable. (h) Magnitude of transform of $y_1d(t)$ or $-y_4d(t)$; note periodic structure of transform. (i) Sampled and held $y_1(t)$ or $-y_4(t)$ waveform obtained as convolution of Figs. 2g and 1i. (j) Magnitude of transform of $y_1s(t)$ or $-y_4s(t)$; note attenuation of spectral replicants beyond baseband spectrum.

the temporal waveform y(t) and its Fourier transform Y(f), then the following product (·) and convolution (*) relationships hold (see sections 4.5 and 4.6 in Brigham [2] for proofs):

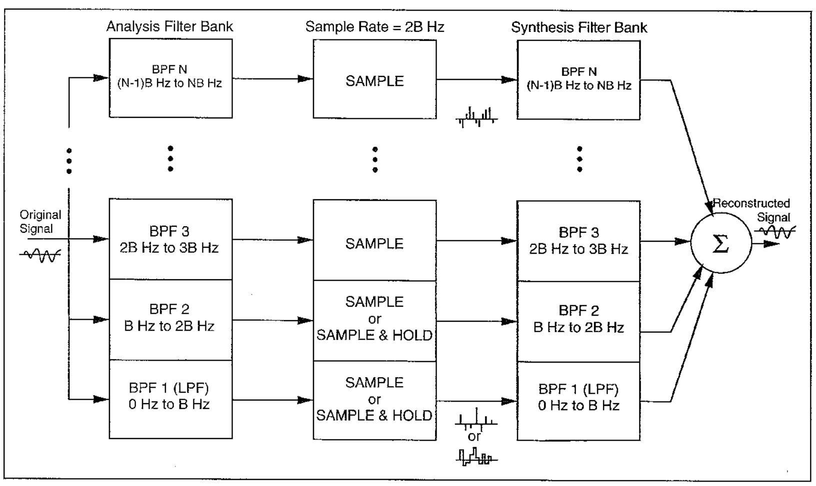
$$x(t) \cdot y(t) \leftrightarrow X(f) * Y(f)$$
$$x(t) * y(t) \leftrightarrow X(f) \cdot Y(f)$$

Note that lower-case notation is used for temporal functions and upper-case notation is used for transform functions. Although not necessarily regarded as a "devoutly-held" theorem in the signal processing community, these variants of the convolution theorem will nevertheless serve a divinely inspired purpose here.

The two fundamental flaws in the reasoning of the Smith article were (1) confusing the foldover frequency with the Nyquist frequency and not recognizing that these are different entities, and (2) failing to recognize the aliasing effects from sampling a demonstration signal that was not bandlimited. The spectral foldover frequency is determined only by the *selected sample rate* (specifically, it is half the sample rate) and it may be selected independently of the characteristics of the signal being sampled. The Nyquist frequency, on the other hand, is a function only of the signal (specifically, it is determined by the *bandwidth* of the signal) and it is independent of the selected sampling rate. The Nyquist frequency is a *lower bound* for the foldover frequency in the sense that failure to *select* a foldover frequency at or above the Nyquist frequency will result in spectral aliasing and loss of the capability to reconstruct a continuous-time signal from its samples without error (basis for the sampling theorem). The Nyquist frequency for a real-valued, *low-pass* signal is simply equal to the highest significant frequency component in the signal (the bandwidth being the range between 0 Hz and this highest frequency). The Nyquist frequency for a real-valued, *band-pass* signal sampled in quadrature lies between one and two times the bandwidth between the lowest and highest significant frequencies in the pass band [8]. The Nyquist frequency for a continuous-time signal which is



3: Construction steps to create Smith's y_2 and y_3 demonstration signals [1]. (a)(c) $y_2(t)$ obtained as product of Figs. 1a and a sinusoidal waveform of higher frequency than the 1c waveform. (b) Magnitude of $y_2(t)$ transform, $|Y_2(f)|$, using linear scale. (d) Magnitude of $Y_2(f)$ using logarithmic scale. (e) $y_3(t)$ waveform using 2X the sinusoidal frequency of $y_2(t)$. (f) Linear magnitude of $y_3(t)$ transform, $|Y_3(f)|$. (g) Sampled $y_2(t)$ or $-y_3(t)$ waveform obtained as product of Figs. 3a/3e and 1g; note that sampled waveforms $y_2d(t)$ and $-y_3d(t)$ are indistinguishable. (h) Magnitude of transform of $y_2d(t)$ or $-y_3d(t)$; note periodic structure of transform compared to Fig. 2h. (i) Sampled and held $y_2(t)$ or $-y_3(t)$ waveform obtained as convolution of Figs. 3g and 1i. (j) Magnitude of transform of $y_2s(t)$ or $-y_3s(t)$; note attenuation of spectral replicants beyond baseband spectrum and first aliased replicant.



4: Filter bank scheme for analysis and synthesis of signal of NB Hz bandwidth using N band-pass filters of B Hz each.

not bandlimited is infinity; that is, there is no finite sample rate that would permit errorless reconstruction of the continuous-time signal from its samples.

Consider now the four basic waveforms and their associated Fourier transforms depicted in Fig. 1, in which the time axis may be considered to be in units of seconds and the frequency axis may be considered to be in units of Hz. The basic waveforms are a triangle function waveform tri(t) of 9-second duration in Fig. 1a and repeated in Fig. 1c, an infinite-duration 1/2-Hz sinusoidal function waveform sin(t) in Fig. 1e, a sampling function samp (t) consisting of an infinite sequence of impulse functions spaced at 1/2 -second intervals (i.e., a 2 Hz sample rate) in Fig. 1g, and a rectangular pulse rect(t) of 1/2-second width in Fig. 1i. The magnitudes of the Fourier transforms (i.e., the spectra) of these waveforms are illustrated in the right column of Fig. 1. Shown are the triangle function transform TRI(f) plotted with linear scaling in Fig. 1b and with logarithmic scaling in Fig. 1d, the sinusoidal function transform SIN(f) in Fig. 1f (a pair of impulse functions at ± 1/2 Hz), the sampling function transform SAMP(f) (also an infinite sequence of impulse functions, spaced at 2 Hz intervals, called the replicating function) in Fig. 1h, and the rectangular pulse transform in Fig. 1j, which is a sinc function with zero crossings at 2 Hz intervals. The exact mathematical structure of these temporal functions and their transforms may be found in section 2.3 of Brigham [2] or chapter 2 of Marple [3].

One motivation for portraying the transform of the triangle function with both linear scaling and logarithmic scaling in Figs. 1b and 1d is to highlight the fact that the triangle waveform is not a bandlimited signal, which is more obvious with logarithmic scaling than with linear scaling. A second motivation is to illustrate that most real-world continuous-time signals, typically being time-limited to some observation interval, are not bandlimited. However, for purposes of digital processing, we must consider even non-bandlimited signals as "essentially" bandlimited in order to select a sampling interval for acquiring such signals via analog-to-digital conversion. Using the rule-of-thumb that the quantization noise floor of an analog-to-digital converter is roughly (6dB) x (# converter bits) below the peak signal level that can be quantized by the converter, the essential bandwidth of a signal may be considered as that portion of the signal spectrum that has magnitude above the quantization noise floor. Using the spectrum of the triangle function shown in Fig. 1d as an example, the essential bandwidth when using a 6-bit converter (producing a noise floor 36 dB below the peak) is approximately 1 Hz. Had an 8-bit converter been used, producing a noise floor 48 dB below the peak, the esssential bandwidth would now become approximately 9 Hz, a factor of nine increase by using a converter with only two additional bits of resolution. The 2 Hz sample rate, with a 1 Hz foldover frequency, that was selected by Smith would therefore be inadequate to capture the essential bandwidth of a triangle function if using sample-and-hold operations involving converters of more than 6 bits.

Let us now use the waveforms of Fig. 1 and the convolution theorem to compose the demonstration signals used in the Smith article [1]. The gated 1/2-Hz sinusoidal $y_1(t)$ of the Smith article is a composite signal obtained by multiplication (modulation) of the triangle and sinusoidal functions

$$y_1(t)=\mathrm{tri}(t)\cdot\sin(t)\leftrightarrow Y_1(f)=\mathrm{TRI}(f)*\mathrm{SIN}(f)$$

and the transform $Y_1(f)$ is therefore obtained as the convolution of the triangle function transform with a pair of impulse functions at $\pm 1/2$ Hz that represent the frequency location of the "carrier." Convolution of a transform with an impulse function simply replicates the spectrum of that transform positioned about the frequency location of the impulse rather than about 0 Hz. The transform $Y_1(f)$ is therefore obtained from the transforms depicted in Figs. 1a and 1f to produce the transform magnitude depicted in Figs. 2b (linear scaling) and 2d (logarithmic scaling).

Observe from Fig. 2d that $y_1(t)$ is not bandlimited, which means the Nyquist frequency is infinite, in contrast to the 1 Hz value indicated by Smith as the Nyquist frequency. Also, observe that the creation of the gated sinusoidal signal has distorted the original shape of the triangle function spectrum, depicted as the dashed response in Fig. 2d for purposes of comparison with the spectrum of the gated sinusoid. This is due to the interactive overlap of the two shifted replicant spectra. Thus, it will not be possible to recover the original spectrum of the triangle function after modulation. If the sinusoidal frequency is increased to 3/2-Hz to form the signal $y_4(t)$ of the Smith article, then the temporal waveform and frequency spectrum depicted in Figs. 2e and 2f are the result.

We can create a sampled version of y_1 or y_4 by simple multiplication with the 2-Hz sampling function samp(t) of Fig. 1g:

$$y_1d(t)=y_1(t)\cdot\mathrm{samp}(t)\leftrightarrow Y_1d(f)=Y_1(f)*\mathrm{SAMP}(f)$$

The sampled temporal waveform is the sequence of samples shown in Fig. 2g, with a spectrum as depicted in Fig. 2h formed by replicating the transform of Fig. 2b at each location of an impulse in the replicating function depicted in Fig. 1h. Note that the replication process in the frequency domain creates aliasing due to overlap of the spectra from adjacent replicants. Because the foldover frequency for a 2-Hz sampling rate is 1 Hz, sampling of the 3/2-Hz gated sinusoid $y_4(t)$ replicates the spectrum $Y_4(f)$ of Fig. 2f, creating an indistinguishable spectrum to that of the aliased spectrum of $Y_1(f)$ (both shown in Fig. 2h). The aliasing due to sampling $y_4(t)$ at 1/2-second intervals (replicating $Y_4(f)$ at 2-Hz intervals) has in effect demodulated (frequency shifted; aliased; folded over) the spectra centered at 3/2 Hz to a baseband spectra centered at 1/2 Hz.

To model the effect of the hold operation of the sample and hold, we need only convolve the sample sequence y_1d or y_4d with the rectangle function rect(t) of Fig. 1i:

$$y_1s(t)=y_1d(t)*\mathrm{rect}(t)\leftrightarrow Y_1s(f)=Y_1d(f)\cdot\mathrm{RECT}(f)$$

yielding the waveform $y_1s(t)$ or $-y_4s(t)$ as depicted in Fig. 2i. The correponding spectrum in Fig. 2j is obtained by weighting the periodic transform of Fig. 2h with the sinc function of Fig. 1j, which attenuates all replicants except those closest to 0 Hz. The hold operation may therefore be considered as a crude form of low-pass filtering that attenuates replicant spectral components above baseband. It should be noted that Succi [6], who also responded to the Smith article, did not account for this hold operation, so his figures labeled "spectra of digital sample and hold" are more accurately labeled "spectra of digital sample."

Waveforms $y_2(t)$ and $y_3(t)$ in the Smith article, which involve sinusoidal frequencies between the 1/2 and 3/2 Hz frequencies used for Fig. 2, produced the results depicted in Fig. 3. By examining Figs. 2b, 3b, 3f, and 2f, we see the waveform spectra follow predictable spectral peak shifts in which the two peaks in the spectra move away from 0 Hz proportionally to the sinusoidal frequency. However, the sampled and held versions of these signals do not provide such straightforward spectral shifts, but instead yield spectra with varying degrees of overlap due to aliasing about the 1 Hz foldover frequency, as indicated in corresponding Figs. 2j and 3j.

When we plot the transforms using only a linear scale, the peaks in the transforms of y_1, y_2, y_3, and y_4 only appear to be essentially bandlimited, and the downshifting of these peaks to a baseband region between ± 1 Hz due to the aliasing effect of sampling was exploited by Smith to develop the two-filter LSL+HSH scheme [1]. The sampling theorem for band-pass signals, with its effective downconversion of the band-pass signal spectrum to a baseband near or about 0 Hz due to positioning of a spectral replicant as a result of the sampling operation, is well known to the signal processing community; see for example section 14.1 of Brigham [2] or section 2.4 of Crochiere and Rabiner [4]. Sampling of band-pass signals forms the basis for filter bank schemes, well described in chapter 7 of Crochiere and Rabiner [4] or the text by Vetterli and Kovacevic [5]. For example, the analysis-synthesis scheme shown in Figure 4 uses a filter bank of N band-pass filters, each of uniform bandwidth B Hz, to sample and reconstruct a continuous-time real-value signal of bandwidth NB Hz using a sample rate of only 2B sps (most authors use Hz rather than samples per second, but sample rates are better distinguished from signal frequencies by use of sps). See the short note by Succi [6] for an explanation of band-pass sampling. Note that the band-pass filter for the 0 Hz to B Hz band is actually a low-pass filter. Because each band-pass signal represents a filtered portion of the overall signal spectrum limited to just B Hz bandwidth, it may be sampled at a rate of 2B sps (or B sps if sampled in quadrature) and no aliasing will occur. The reconstruction (synthesis) filter bank takes as input the sample sequence of weighted impulses as depicted in Figs. 2g or 3g. However, if one chooses to implement the filter bank using the sample-and-hold waveforms of Figs. 2i or 3i, rather than the sample-only waveforms of weighted impulse sequences shown in Figs. 2g or 3g, then one is limited to essentially a filter bank of N=2 filters due to the attenuation of the replicant spectra by the frequency-domain sinc function, as illustrated in Figs. 2j and 3j. In fact, the quality of the reconstructued waveform is degraded by use of the sample-and-hold waveform rather than the sampleimpulse waveform due to the attenuative effect of the frequency domain sinc function in the upper band of the two filter scheme.

—S. Lawrence Marple, Jr.
Acuson Corporation
Advanced Development
Mountain View, CA

References

1. J. Lynn Smith, "Breaking the Nyquist Barrier," IEEE Signal Processing Magazine, pp. 41-43, July 1995.

2. E. Oran Brigham, *The Fast Fourier Transform and its Applications*, Prentice Hall, Englewood Cliffs, NJ, 1988.

3. S. Lawrence Marple Jr., *Digital Spectral Analysis*, Prentice Hall, Englewood Cliffs, NJ, 1987.

4. Ronald E. Crochiere and Lawrence R. Rabiner, *Multirate Digital Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1983.

5. Martin Vetterli and Jelena Kovacevic, *Wavelets and Subband Coding*, Prentice-Hall, Englewood Cliffs, NJ, 1994.

6. G. Succi, "Restoring the Nyquist Barrier," *IEEE Signal Processing Magazine*, pp. 16-18, November 1995.

7. W. J. Ebel and N. Younan, "Counting on Computers in DSP Education," *IEEE Signal Processing Magazine*, pp. 38-43, November 1995.

8.. J. L. Brown, Jr. "On Quadrature Sampling of Bandpass Signals," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 15, pp. 366-371, May 1979.