nor the linear interpolation with delay brings any additional, beneficial effect in such filter banks. The analysis of these operations is similar to the one based on Fig. 4 above, and it results in the same drawbacks as the zero-order hold implementation of filter banks. Moreover, the additional lowpass filtering typical of these operations is simply not needed since the required filtering is accurately and thoroughly performed by the analysis and synthesis filters.

References

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

2. J. Lynn Smith, "The Broken Nyquist Barrier," *IEEE Signal Processing Magazine*, pp. 14-18, May 1996.

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

4. S. Lawrence Marple Jr., "Restoring the Nyquist Barrier," *IEEE Signal Processing Magazine*, pp. 24-28, January 1996.

5. Ronald E. Crochiere and Lawrence R. Rabiner, *Multirate Digital Signal Processing*, Prentice Hall, Englewood Cliffs, N.J, 1983.

6. Alan V. Oppenheim and Alan S. Willski, *Signals and Systems*, Prentice Hall, Englewood Cliffs, N.J, 1983.

—Grigore Braileanu
Professor, Electrical Engineering
Deptartment, Gonzaga University,
Spokane, WA 99258

Exploding the Nyquist Barrier Misconception

In a recent "Forum Feedback" [1], Smith continues to proclaim the novelty of an approach [2] that he purports to "break the Nyquist barrier," in spite of the revelation [3] that his approach is simply a special two-filter case of well known analysis and synthesis filter banks performed with sample-and-hold waveforms. Compare Smith's Fig. 4(b) in reference 2 with the conventional filter banks of Fig. 4 in reference 3 and observe that they are identical. Smith also makes further observations [1] to which I will respond with additional commentary. However, the primary purpose of this contribution to "Forum Feedback" is to expose a fundamental misconception regarding the universality of the sampling theorem as taught in most digital signal processing textbooks. It is this misconception that led Smith to prematurely claim victory over a perceived impenetrable Nyquist barrier.

Digital filter banks have been used since the early 1970s [4] for multi-rate applications in which a single large bandwidth signal is partitioned by filtering into a group of smaller bandwidth signals. Appropriate filter-output sample rates are selected for each analysis filter in the bank to ensure that the Nyquist criteria (no spectral aliasing) is satisfied (more on this issue of appropriate sample rates later). These filter-output sample rates will be reduced (decimated) relative to the sample rate required to satisfy the Nyquist criteria for the unfiltered broadband signal, reflecting the associated reduction in bandwidth by passage of the broadband signal through each filter. Filter banks therefore provide a tradeoff, exchanging the broadband signal with a high sample rate for a bank of filtered signals with lower sample rates. At no time is the Nyquist "barrier" broken (Nyquist criteria violated) in such filter banks. In his "novel" scheme, Smith uses a filter bank of two filters that split the bandwidth of the original signal in half between them. In order to satisfy the Nyquist criteria, each filter output therefore needs only to use half the sample rate as that required to sample the original unfiltered signal. To support his claim of breaking the Nyquist "barrier," Smith has incorrectly and misleadingly associated the sample rate at a filter *output* with the original signal used as *input* to each filter. The correct associations would pair *output* sample rate with the filter *output* signal and would pair *input* sample rate with the filter *input* signal.

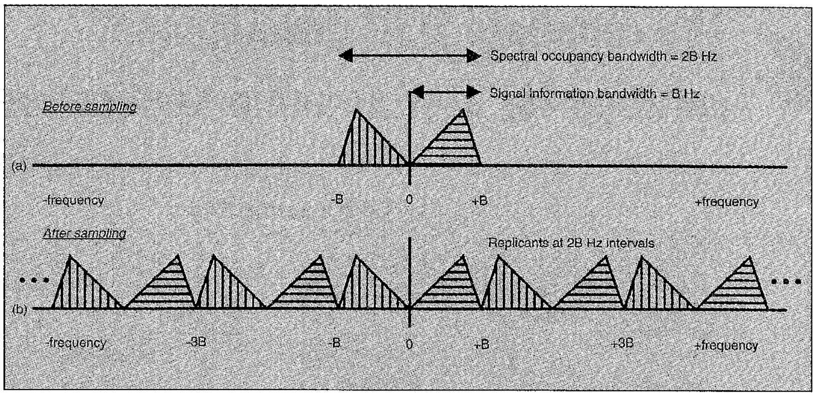
It is difficult to take Smith's continued claims of novelty seriously because he refuses to confront evidence to the contrary from the literature, such as the previously cited references [3] or the paper by Vaughan et al. [5]. It is further unfortunate that Smith simply dismisses [1, page 14, column 1] the Vaughan et al. paper without investing the time to study the paper in detail for the insight that it provides.

Contrary to Smith's opinion that my quoted statement, "Results of data analyzed by software simulation tools are meaningless" [3], was a distorted synopsis of his original article [2], it was actually intended as a commentary that such tools can't be trusted if used in a manner similar to that of Smith for deriving fallacious conclusions.

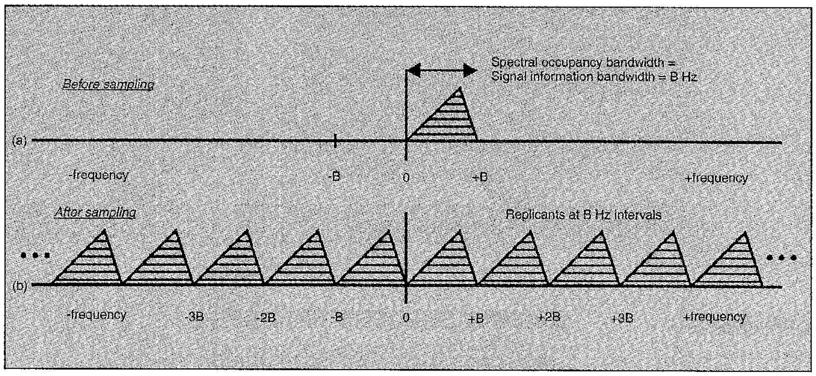
In an attempt to distinguish his methodology from that of Succi [6], Smith makes the assertion [1, page 14] that the sample-and-hold process he uses "does not actually have a spectrum because it is not equivalent to a function." This is nonsense! Derivations and illustrations were provided in reference 3, in particular Figs. 2(i), 2(j), 4(i), and 4(j), that replicated the sample-and-hold waveforms of Smith's original paper [1] and depicted the associated spectra. The sample-and-hold spectrum was shown [3] to be a distorted version of the sample-only spectrum. The distortion is an effect of the convolution between the undistorted sample-only spectrum and a sinc function related to the hold operation.

On page 16 of reference 1, Smith states that "if I am simulating a sample-and-hold process ... the sampling rate of the transform must be significantly higher than the process I am trying to simulate; otherwise alias of the tool distorts my study." While most readers would understand the association of a sampling rate with a *temporal-domain signal*, it is not clear what Smith means by the sampling rate of a *frequency-domain transform*. The reciprocal of the interval between FFT bins? The reciprocal of the interval between periodic replications of the base transform? Or something else? Why does the transform need to be sampled at all? A sampled signal actually has a continuous spectrum of frequency [9], although often an FFT is used to evaluate the continuous spectrum at a set of uniformly spaced discrete frequencies.

Smith argues that he believes aliasing did not distort his study, yet this is precisely what I contend did distort his study. This contention is based on observations regarding Fig. 2(d) in reference 3 [page 26, right column]. I must say I was astounded with Smith's claim



1. Frequency domain effects of sampling a real baseband (lowpass) bandlimited signal. (a) Spectrum of real-value continuous-time bandlimited signal of bandwidth B Hz. (b) Periodic spectrum created by sampling the continuous-time signal at a rate of 2B sps.



2. Frequency domain effects of sampling a complex baseband (analytic) bandlimited signal. (a) Spectrum of complex-value continuous-time analytic signal derived from the original real-value signal (Figure 1). (b) Periodic spectrum created by sampling the analytic signal at a rate of B sps.

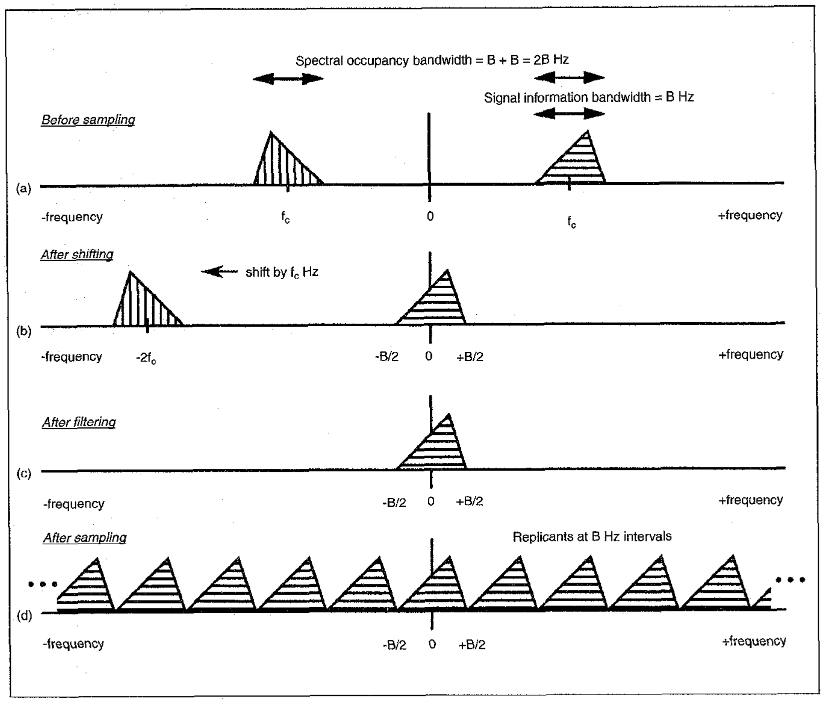
that he "wanted the alias effects and purposely created them" and that he "could recover a ... signal despite the alias effect." This is nonsense! Figure 2 and the associated discussion on page 26 of reference 3 clearly show that Smith's demonstration signal is heavily aliased, especially when viewed with logarithmic scaling. The sinc function distortion of the spectrum due to the hold operation further prevents exact recovery of the original pre-sample-and-hold signal spectrum. There is no way that Smith's synthesis approach [2] could exactly recover the original signal as it existed prior to the sample-and-hold operations. It is fortunate that telephone engineers do not design frequency division multiplex (FDM) telephone channel filter banks with this attitude of intentional aliasing in mind, otherwise we would be experiencing considerable interchannel crosstalk in telephone lines today!

Smith quibbles over possible differences in the definition of "Nyquist frequency" and how that could be used to spin a different interpretation of his results. I certainly agree with Smith that there is inconsistency with use of the adjective Nyquist to define terms in the signal processing literature, which at times leads to legitimate confusion if terms are not carefully defined in each technical publication. However, I do not believe this has any bearing on the interpretation of Smith's approach as anything but a special-case filter bank with two filters. In my previous correspondence [3], I tried to be pedagogical in distinguishing between two frequency entities. One frequency entity is associated only with the sampling rate of a digital system, which I termed the "foldover or folding frequency" [7, page 48] [8, page 31]. The second frequency entity is associated only with the spectral content of a signal, which I linked to the "Nyquist frequency." I am not alone in trying to intentionally distinguish these two frequency entities. Oppenheim and Schafer [9, page 87] called these the "Nyquist rate" and the "Nyquist frequency," respectively. Here the term Nyquist is used twice, illustrating the very inconsistency that has often caused confusion in the literature. Preventing this confusion was the motivation for my selection of the term "foldover frequency" rather than "Nyquist frequency." The Oppenheim and Schafer definitions do, however, introduce notional differences for the terms "frequency" and "rate" with which I strongly agree. The term "frequency" is reserved for describing attributes associated *only* with the *spectral content* of *signals*. The term "rate" is reserved for describing attributes associated *only* with the *sampling periodicity* of digital processing *systems*. For example, a digital processing system of a given sample "rate" can unambiguously process signals possessing a certain range of "frequencies" (more on this topic later). Thus, the ambiguity of using the term "frequency" to describe both signal and system attributes is avoided. To further emphasize the notional differences, it is recommended that a signal "frequency" should be specified in units of Hertz (Hz) whereas a system "rate" should be specified only in units of samples per second (sps).

In the final paragraph of reference 1, Smith expresses "a new point" he wishes "to raise" that "the hold operation implies an additional filtering effect which is a sinc function depending on sample hold period." Smith should have given credit to Marple [3, page 27, middle column] for acquainting him with this property instead of taking all the credit for the observation. This sinc function phenomenon related to hold operations has actually been known since at least the 1960s by designers of pulse-amplitude-modulation (PAM) communication systems [11, page 460].

Sampling Theorem Misconception

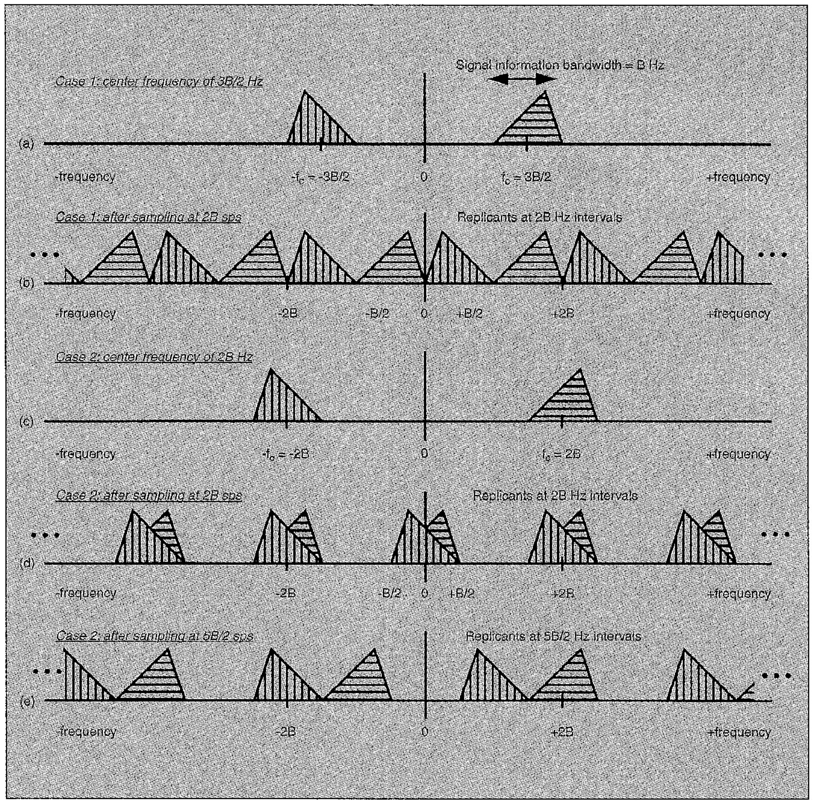
At the heart of the issue with Smith's claims of breaking the Nyquist barrier is a widely prevalent misconception in our community regarding the universal applicability of the sampling theorem as commonly presented in almost all signal processing textbooks. This classic sampling theorem is premised on the assumption that the continuous-time signal being sampled is real-valued and bandlimited. In order to reconstruct without distortion the original continuous-time signal from its samples, the key tenets of the classic textbook sampling theorem require: (1) a sampling rate of at least twice the signal bandwidth (equivalently, the highest signal frequency [Nyquist frequency] cannot exceed half the signal sampling rate) and (2) an infinity of signal samples (due to the infinite summation range of the interpolation formula for reconstructing the continuous-time signal from its samples).



3. Frequency domain effects of basebanding and sampling a real bandpass signal. (a) Spectrum of real-value continuous-time bandpass signal of bandwidth B Hz. (b) Spectrum of bandpass signal after complex demodulation to baseband (frequency shift of f_c Hz). (c) Spectrum of demodulated signal after low-pass filtering (d) Periodic spectrum created by sampling the demodulated and filtered signal at a rate of B sps.

Three points need to be made here. First, it is insufficient to simply sample a finite segment of a signal with a sampling rate of at least twice the highest frequency if the sampling theorem is to be satisfied. An infinity of samples, at least in theory, is also required in order to satisfy the tenets of the sampling theorem. Smith's demonstration signal [2] involves a finite number of samples, which in itself violates the sampling theorem tenets. Second, there are alternative sampling theorems that apply under different signal conditions than assumed for the classical textbook sampling theorem, two of which will be discussed subsequently. One of these alternatives is applicable to Smith's filter-bank approach. Because the sampling theorem is not unique, the factor of two relating sampling rate and highest signal frequency is therefore not a universal immutable law applicable to all signal conditions. Third, the sampling rate does *not* establish a barrier to the *highest permissible signal frequency*. The sampling theorem appropriate for given signal attributes may limit the highest signal frequency to be *less* than the sampling rate, *equal* to the sampling rate, or *more* than the sampling rate. The classical textbook sampling theorem limits the highest signal frequency to be *less* than the sampling rate. The two signal classes and their associated sampling theorems to be described subsequently will be demonstrated to permit the highest signal frequency to be *equal to* or even *higher* than the sampling rate. The true barrier to the minimum sampling rate, common to all sampling theorems, is actually the *signal bandwidth*, rather than the *highest signal frequency*. This author therefore believes that a more appropriate set of metrics for common use across all sampling theorems should be *Nyquist band* and *Nyquist bandwidth*, rather than simply a *Nyquist frequency*. The Nyquist band is simply a *range* of frequencies, for a specified sample rate, that a continuous-time signal may contain, which will not cause spectral aliasing if the signal is sampled. The Nyquist band requires two numbers for its specification, namely the lowest and highest frequencies in the band. The Nyquist bandwidth is simply the difference between these frequencies.

Consider first the principles behind the classical textbook sampling theorem. Figure 1(a) depicts the magnitude spectrum of a real-value continuous-time signal that is bandlimited to B Hz. Note that the signal's spectral extent is presumed to lie between 0 Hz and B Hz, so that the *signal information bandwidth* (a single number) and its highest frequency are one and the same, specifically B Hz. Due to the position of the signal spectrum, some textbooks will refer to a bandlimited signal as a lowpass signal or a baseband signal. Because the signal is assumed to be real-valued, it will have a symmetric spectrum, as shown in Fig. 1(a), that extends over a total *spectral occupancy bandwidth* of $2B$ Hz. The basis for any sampling theorem is the requirement that the spectrum of the sampled signal not be aliased, otherwise distortionless reconstruction of the continuous-time signal from its samples will not be possible. In the case of the classic textbook sampling theorem, sampling a signal at a minimum rate of $2B$ sps will yield the spectrum shown in Figure 1(b), which has periodic replicants of the original



4. Frequency domain effects using direct sampling of a real bandpass signal. Case 1: (a) Spectrum of real-value continuous-time band-pass signal of bandwidth B Hz and center frequency 3B/2 Hz. (b) Nonaliased periodic spectrum created by sampling the bandpass signal at a rate of 2B sps. Case 2: (c) Spectrum of bandpass signal of bandwidth B Hz and center frequency 2B Hz. (d) Aliased periodic spectrum created by sampling the bandpass signal at a rate of 2B sps. (e) Nonaliased periodic spectrum created by sampling the bandpass signal at a rate of 5B/2 sps.

signal spectrum at intervals of $2B$ Hz. Sampling at $2B$ sps is required to avoid aliasing (overlap of replicants) because the spectral occupancy bandwidth is $2B$ Hz. Thus, the classical textbook sampling theorem, which is more appropriately described as the *real-bandlimited-signal sampling theorem*, has a Nyquist bandwidth of B Hz over the Nyquist band of 0 to B Hz if the sample rate is $2B$ sps.

Consider now the analytic signal derived from the real-value bandlimited signal by either a temporal-domain operation of Hilbert transforming (approximated by a filter in practice) or a frequency-domain operation (Fourier transforming, zeroing the negative frequency transform terms, and inverse transforming). The analytic signal, which is complex-value, has identical signal information and spectral occupancy bandwidths of B Hz, as illustrated in Fig. 2(a). Therefore, the analytic signal can be sampled at a rate of B samples per second to produce a periodic spectrum that avoids aliasing, as shown by the replicants in Fig. 2(b). The nonaliased periodic spectrum of Fig. 2(b) forms the basis for the *complex-bandlimited-signal sampling theorem*. Although real-value bandlimited signals have a Nyquist bandwidth of *half* the sample rate, complex-value bandlimited signals have a Nyquist bandwidth *equal* to the sample rate. Both the real signal and its complex analytic cousin possess *identical* signal-information bandwidths (equivalently, both have *identical* highest frequencies), yet the complex signal requires *half* the sample rate as that of the real signal.

Bandpass signals are a third signal category that have their own sampling theorem, namely the *real-band-pass-signal sampling theorem* [10, section 14.1]. Actually, there are two possible real-bandpass-signal sampling theorems, depending on whether downconversion is performed prior to sampling or as part of the sampling process. The latter case is the one most often associated with the bandpass sampling theorem. For purposes of illustration, consider a real-value bandpass signal with signal information bandwidth of B Hz and center frequency of f_c Hz, yielding the magnitude spectrum shown in Figure 3(a). If the signal is downconverted to baseband by complex demodulation with the complex sinusoidal $\exp(-j2\pi f_c t)$, then the spectrum shown in Fig. 3(b) will be the result. A lowpass filter of $B/2$-Hz bandwidth applied to the demodulated signal will then yield the complex-value baseband signal with a spectrum centered at 0 Hz, as depicted in Fig. 3(c). Note that the baseband signal is representative of the original bandpass signal because it contains the identical signal information bandwidth. The demodulated and filtered baseband signal may then be sampled at a rate of B sps to produce a nonaliased periodic spectrum with replicants at intervals of B Hz, as shown in Fig. 3(d). Thus, the Nyquist bandwidth associated with a bandpass signal that is basebanded and sampled at B sps is B Hz.

As an alternative to complex demodulation, a bandpass signal may be directly sampled and the downconverted baseband signal culled from the periodic spectrum as the spectral replicant closest to 0 Hz in the sampled bandpass signal spectrum. The connection between sampling and demodulation is presented in more detail under the topic of impulse train modulation [9, page 81; 10]. Consider first the case of a real bandpass signal of signal information bandwidth B Hz and center frequency $3B/2$ Hz. This signal has a magnitude spectrum as shown in Fig. 4(a) with a total disjointed spectral occupancy bandwidth of $B + B = 2B$ Hz. If this bandpass signal is sampled at a rate of $2B$ sps, the resultant nonaliased periodic spectrum with replicant periodicity of $2B$ Hz is shown in Figure 4(b). For this case of a bandpass signal, a B Hz Nyquist bandwidth is possible if sampled at $2B$ sps. Consider next the case of a bandpass signal with the identical signal information bandwidth of B Hz, but having a center frequency of $2B$ Hz as shown in Fig. 4(c). If a sampling rate of $2B$ sps is used, then aliasing occurs in the sampled signal spectrum as indicated in Fig. 4(d), so it is not an appropriate sampling rate to preserve the signal information bandwidth. If a sample rate of $5B/2$ sps is selected instead, then a nonaliased spectrum will result, as shown in Fig. 4(e). Thus, $5B/2$ sps is an appropriate sampling rate for this specific case of a bandpass signal. As stated previously for bandpass signals [3], the required sampling rate to avoid aliasing will always lie somewhere between one and two times the spectral occupancy bandwidth (between $2B$ and $4B$ sps) in order to preserve the signal information bandwidth without aliasing.

—Larry Marple
ORINCON Corporation
San Diego, CA

References

[1] J.L. Smith, Forum Feedback, *IEEE Signal Processing Magazine*, May 1996.

[2] J.L. Smith, "Breaking the Nyquist Barrier," *IEEE Signal Processing Magazine*, pp. 41- 43, July 1995.

[3] S.L. Marple Jr., "Restoring the Nyquist Barrier," *IEEE Signal Processing Magazine*, pp. 24-28, January 1996.

[4] R.W. Schafer and L.R. Rabiner, "Design of Digital Filter Banks for Speech Analysis," *Bell System Technical Journal*, vol. 50, pp. 3097—3115, December 1971.

[5] R.B. Vaughan, N.L. Scott, and D.R. White, "The Theory of Bandpass Sampling", *IEEE Transactions on Signal Processing*, vol. 39, pp. 1973-1984, September 1991.

[6] G. Succi, *IEEE Signal Processing Magazine*, pp.16-19, November 1995.

[7] S.D. Stearns and D.R. Hush, *Digital Signal Analysis*, second edition, Prentice Hall, Englewood Cliffs, NJ, 1990.

[8] R.K. Otnes and L. Enochson, *Digital Signal Analysis*, Wiley, 1972.

[9] A.V. Oppenheim and R.W. Schafer, *Discrete-Time Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1989.

[10] E.O. Brigham, *The Fast Fourier Transform and its Applications*, Prentice Hall, Englewood Cliffs, NJ, 1988.

[11] J.B. Thomas, *An Introduction to Statistical Communication Theory*, Wiley, 1969.

The Complex (and Circular) Argument Continues

While I would have been laughing at anyone who thought that complex notation was useless, unfortunately Bernard Picinbono (July 1996, "Forum Feedback" [1]) meant to be writing about me. I would like to clear up the misrep-