AN APPROACH TO LINEARIZATION

by

Eugene Wong
Department of Electrical Engineering and Computer Sciences
and Electronics Research Laboratory
University of California, Berkeley, California 94720

In a large variety of signal processing situations, nonlinearities play a deleterious role, and their presence severely limits the performance of the signal processing system. In the case of a zero-memory (or instantaneous) nonlinearity, (i.e., the output y (t) at time t is a function of the input x (t) at the same instant of time), one might think that linearization can be achieved simply by cascading the nonlinearity with its inverse. To some extent this approach is useful. However, there are a number of serious objections to this approach of linearization. First of all, the nonlinearity may not have an inverse, e.g., $F(x) = x^2$. Secondly, the exact form of the nonlinearity may not be known, and , if known, may drift in time. Thirdly, nonlinearities are seldom truly zero-memory. More often, they are better represented by an instantaneous nonlinearity followed by a low pass filter with bandwidth W. In such cases, cascading it with an inverse may do more harm than good, depending on the signal bandwidth relative to W.

Another approach to linearization is to modulate the input signal using a mode of modulation, such as FM, PWM, PCM. These modulated signals are relatively immune to nonlinearities because the amplitude of the modulated signal carries no information. Aside from the added complexity of

The research reported herein was initiated by the author as a consultant to the Ampex Corporation. Additional work was supported in part by the U.S. Army Research Office, Durham, Contract DAHCO4-67-C-0046.

modulating and demodulating, the principal defect of this approach is bandwidth expansion. Again, if we represent the nonlinearity by a zero-memory nonlinearity followed by a low-pass filter with bandwidth W, then the bandwidth of the input signal prior to modulation must be substantially less than W. How much less depends on which modulation scheme is used and on the desired signal-to-noise ratio. In other words, linearization by modulation is at the expense of useful bandwidth and/or signal-to-noise ratio.

Consider a zero memory nonlinearity represented by

$$y(t) = f(x(t)) \tag{1}$$

where $f(x)$ is an arbitrary real-valued function. Suppose $x(t)$ is sinusoidal and $f(-x) \neq -f(x)$, then even harmonics will be present in $y(t)$. One easy and well known way of getting rid of the even harmonics is to connect two such nonlinearities in parallel, so that the over-all output is now

$$y(t) = g(x(t)) = f(x(t)) - f(-x(t)) \tag{2}$$

Since $g(x)$ satisfies $g(x) = -g(-x)$, no even harmonic will be present in $y(t)$. Relative magnitudes of odd harmonics to the fundamental remain unchanged. The approach proposed in this paper is a generalization of this simple idea.

Suppose that K identical instantaneous nonlinearities are connected in parallel in an arrangement so that the k^{th} nonlinearity is preceded by and ideal phase shifter with phase shift $(k-1)\pi/K$ and followed by phase shift of $-(k-1)\pi/K$. Let the input be a sinusoid

$$x(t) = x_o \cos(\omega t + \phi) \tag{3}$$

and let f denote the nonlinearity. The function $f[x_o \cos(\omega t + \phi)]$

is periodic with period $^{2\pi}/\omega$ and under very general conditions admits a Fourier series representation of the form

$$f[x_o \cos(\omega t + \phi)] = a_o + \sum_{n=0}^{\infty} [a_n \cos n\omega t + b_n \sin n\omega t] \tag{4}$$

The output y(t) of the parallel combination of K nonlinearities can now be written as

$$\begin{aligned} y(t) &= \sum_{K=0}^{K-1} \Bigg\{ \sum_{n=0}^{\infty} a_n \cos [n(\omega t + \frac{k}{K} 2\pi) - \frac{k2\pi}{K} \\ &\qquad + \sum_{n=1}^{\infty} b_n \sin [n(\omega t + \frac{k}{K} 2\pi) - \frac{k}{K} 2\pi] \Bigg\} \\ &= \sum_{n=0}^{\infty} a_n \operatorname{Re} \left\{ e^{in\omega t} \sum_{k=0}^{K-1} e^{i \frac{k(n-1)}{K} 2\pi} \right\} \\ &\qquad + \sum_{n=1}^{\infty} b_n \operatorname{Im} \left\{ e^{in\omega t} \sum_{k=0}^{K-1} e^{i \frac{k(n-1)}{K} 2\pi} \right\} \end{aligned} \tag{5}$$

The sum $\sum_{k=0}^{K-1} e^{ik\left(\frac{n-1}{k}\right)2\pi}$ is given by

$$\begin{aligned} \sum_{k=0}^{K-1} e^{ik\left(\frac{n-1}{k}\right)2\pi} &= K \quad , \text{ if } \frac{n-1}{k} = \text{integer} \\ &= 0 \quad \text{otherwise} \end{aligned} \tag{6}$$

Therefore , (5) now becomes

$$\begin{aligned} y(t) = K \sum_{m=0}^{\infty} \Big[a_{mK+1} & \cos(mK+1)\ \omega t \\ + b_{mK+1} & \sin(mK+1)\ \omega t \Big] \end{aligned} \tag{7}$$

In other words only harmonics of order $mK+1$ are present in $y(t)$, and the relative magnitudes of these components are unchanged. For K=2, this result reduces to the one mentioned earlier, and only odd harmonics will be present. The number K need not be large to achieve a drastic reduction in harmonics. For example, if K=4 the remaining harmonics are 5th, 9th, 13th, etc.

There are a number of difficulties in implementing the scheme proposed in this paper. The two most serious ones are: (a) the requirement of having K identical nonlinearities, and (b) the requirement of constant phase shifters. The first difficulty can probably be circumvented by using time-division multiplex. The second problem is quite serious, but might be made easier by hetrodyning.