Multiplexing and Demultiplexing of Control Signals by means of the STFT

Marlon Schumacher, MUMT605b, McGill University, Winter2009, Prof. P. Depalle

Abstract— This report describes methods for frequency-multiplexing and demultiplexing sinusoidal control-signals in order to use an audio-signal as a transmission medium. Time-domain and Frequency domain techniques are discussed which allow the generation and acquisiton of the real control-signals, and the estimation of their respective magnitude, phase and frequency by using properties of the Phase Vocoder.

I. INTRODUCTION

The project presented in this paper was motivated by a practical application: The use of digital signal processing techniques to efficiently and conveniently transmit a number of sinusoidal control signals such as time-varying voltages[1], typically obtained from sensor-circuits in the context of gestural control, as a composite signal using an audio-device as the only sensor interface. Rather than discussing the implementation of an existing concept, the development and considerations on a new application for generation, transmission and acquisition of control signals using amplitude modulation techniques are presented.
The use of audio-signals as a transmission medium for gesture data have been described to exhibit many advantages compared to the use of microcontrollers delivering event-based data via digital protocols [3, 5, 6]. Besides obtaining many channels of gesture data at audio-rate in high-resolution and the ease of using the commercial built-in low-latency audio devices almost every personal computer nowadays is equipped with, the probably most important aspect is the flexibility of interpreting gestures using digital signal processing techniques in software, performed natively on a generic computer's CPU [11]. In the context of gestural control, the isochronous treatment of gestures as signals should provide for the tightest possible coupling between sensor-data and sound synthesis parameters.

[1] Please note that this does not include sensing of impacts, or other abrupt discontinuities in the signal which might be critical due to creating high frequency components.

Frequency-division multiplexing (FDM) techniques like amplitude modulation (AM) have since long been used for transmission of multiple signals over a shared medium and are useful in this context due to their straightforward implementation and simple circuit-design compared to time-division multiplexing (TDM). Further, using properties of the phase-vocoder should provide an elegant and efficient way for multiplexing and demultiplexing of multiple AM-signals which are evenly distributed along the frequency spectrum of the transmission channel, with carrier frequencies being harmonics of the STFT's analysis frequency (f0).
The following sections will briefly review the considered amplitude modulation and demodulation techniques, then describe the concepts behind the project, present an implementation and conclude with a discussion and suggestions for future work.

II. AMPLITUDE MODULATION TECHNIQUES

There are various existing techniques for AM classified by the International Telecommunications Union (ITU), which can be roughly subdivided into two major categories: Double Sideband Modulation (DSB) and Single Sideband Modulation (SSB). For the purposes of this project both the modulator as well as the carrier signal are modelled as sinusoidal signals (sinusoidal amplitude modulation).

Double Sideband Amplitude Modulation

Early applications of Double Sideband Amplitude Modulation (DSB-AM) date back to as early as the 1870ies when used for acoustic telegraphy, a predecessor of the electric telephone, rendering it the oldest form of amplitude modulation[2]. Let us consider the carrier as a sinusoidal signal $c(t)$:

$$c(t) = A_c \cdot \sin(\omega_c t + \phi_c)$$

[2] US patent 166,095 -- *Electrical Telegraph for Transmitting Musical Tones* -- Elisha Gray, July 27, 1875

and let $m(t)$ represent the sinusoidal modulator:

$$m(t) = A_m \cdot \cos(\omega_m t + \phi_m)$$

Then the AM-signal is created by adding a constant D to the modulator m(t) and multiplying both with the carrier signal $c(t)$:

$$y(t) = [D + m(t)] \cdot c(t)$$

Setting $A_c = 1$ and $\phi_c = 0$ the AM-signal can be re-written using trigonometric identites as:

$$y(t) = D \cdot \sin(\omega_c t)$$
$$+ \frac{A_m}{2}\left[\sin\left((\omega_c + \omega_m)t + \phi_m\right) + \sin\left((\omega_c - \omega_m)t - \phi_m\right)\right]$$

This representation nicely shows the three components of the AM-signal (the carrier component and the two modulator components above and below the carrier). This process is also known as heterodyning. Since multiplication in the time-domain corresponds to convolution in the frequency domain, the multiplication of two real signals yields positive and negative sidebands, i.e. the AM-signal's spectrum consists of the modulator's positive and negative frequency-components centered around the carrier's positive and negative frequencies.

It becomes apparent that DSB-AM is both dynamically and spectrally inefficient: The spectral bandwidth of the modulator signal needed for transmission is twice its original (baseband) bandwidth since both positive and negative frequency components are shifted up to the positive carrier frequency. This means that in the best case 50% of the bandwidth of the transmission channel can be used for the modulator signals, the other 50% are wasted for redundant information. Moreover, most of the AM-signal's energy is concentrated in the carrier signal, which does not contain any valuable information (besides for certain demodulation schemes its frequency and phase). The remaining dynamic range is shared by the two sidebands of the modulator, from which only one is needed. Please note, however, that by setting D=0 the carrier can be eliminated, i.e. 'suppressed'.

Single Sideband Amplitude Modulation

Single Sideband Amplitude Modulation (SSB-AM) uses electrical power and bandwidth more efficiently by suppressing one of the sidebands and in most schemes avoiding the transmission of the carrier. There are several ways to suppress one of the sidebands [7]:

1) Bandpass filtering

The most straightforward way is bandpass-filtering: One of the sidebands (mostly the upper one) are removed via filtering techniques. Often the carrier is reduced or completely removed as well in order to increase the effective power available for transmission of the information. This is called Single Sideband Suppressed Carrier Amplitude Modulation (SSBSC-AM).

2) Hartley Modulator

This method uses phase discrimination to suppress the unwanted sideband and the carrier: By phase-shifting the sinusoidal carrier and modulator signals by 90 degrees (π/2), the imaginary parts of their complex representation as analytic signals are obtained; this is also called a hilbert-transform (the concept of analytic signals will be discussed in more detail later). By multiplying the carrier and modulator with their respective phase-quadrature components and subtracting or adding them to the DSB-AM signal (that is the modulated carrier), the upper or lower sideband and the carrier are suppressed. This is due to the property of anlytic signals having no negative frequency components, thus multiplying/ dividing two complex exponentials, which adds or subtracts their respective phases, yields either an upper or lower sideband. This could be regarded as frequency-shifting the modulator signal up to the carrier frequency. For two cosine waves representing the carrier and modulator, the Hartley Modulator can be expressed as

$$D + m(t)$$

and

$$y(t) = A_m \cos(\omega_m t)\cos(\omega_c t) - A_m \sin(\omega_m t)\sin(\omega_c t)$$
$$= A_m \cos(\omega_m - \omega_c)t$$

SSBSC-AM outperforms DSB-AM in terms of spectral efficiency and power usage by at least a factor of two. The higher complexity in terms of multiplexing and demultiplexing is justified due to the use of digital signal processing techniques in software; thus SSBSC-AM appears to be a well-suited method for multiplexing control signals in the context of this project. Please note that the two described AM techniques are antagonistic examples amongst various other techniques, please see the appendix for a table of AM schemes as designated by the International Telecommunications Union (ITU) in 1982.

III. Amplitude Demodulation Techniques

Let us consider the AM signal given as:

$$y(t) = \left(D + m(t)\right)\cos(\omega_c t)$$

where $D + m(t)$ is often referred to as "envelope" of $\cos(\omega_c t)$. If it is possible to extract this envelope from the AM signal, the original message can be recovered.

Envelope Detector

The simplest way of retrieving the envelope of the AM signal is using an envelope detector: The AM signal is first rectified and then lowpass-filtered:

$$y(t) = \left[D + m(t)\right] + y(t-1)$$

This method is computationally efficient but has several drawbacks: When using a composite AM signal a steep bandpass filter is needed to isolate the frequency-band of interest, otherwise several signals might be demodulated. It is also more susceptible to noise compared to other demodulation techniques and cannot be used for overmodulated signals (the carrier wave must not be modulated more than 100% with respect to its nominal level). Most important for this project, however, is the fact that it needs a carrier and therefore does not work for SSBSC-AM.

Product Detector

The product detector uses a similar technique for the demodulation of AM signals as is used for their generation: The AM signal is multiplied with a sinusoid of same frequency and phase as the carrier. This conceptually separates the modulator from its carrier (shifts the sideband down to DC and the carrier up to twice its frequency). The high-frequency components are then removed with a lowpass-filter to retrieve the original modulator signal:

$$y(t) = \left(D + m(t)\right)\cos(\omega_c t)\cos(\omega_c t) + y(t-1)$$
$$= \left(D + m(t)\right)\left(\frac{1}{2} + \frac{1}{2}\cos(2\omega_c t)\right) + y(t-1)$$

The product demodulator has several advantages over an envelope detector though with the penalty of increased complexity for demodulation: A stable oscillator is needed for demodulation whose frequency must exactly match the carrier frequency to avoid distortions (frequency-shifts). Also, the phase of the demodulating oscillator must match the carrier's phase (coherent demodulation), otherwise the S/N ratio decreases due to phase cancellations.

IV. Conception

The concept for transmitting a number of control signals through an audio-channel via frequency-multiplexing instead of time-multiplexing is justified by two considerations: Firstly, using time-multiplexing in the analog domain is susceptible to phase-distortions caused by the analog circuitry or the A/D conversion, thus we cannot be sure to be able to properly demultiplex the signal. Secondly, using frequency multiplexing techniques provides the flexibility to use entirely passive sensor-circuits, as for example in [3].

In the context of this application, however, the important specifity lays in that the modulator signals are sinusoidal control-signals. In contrast to event-based transmission schemes (such as MIDI, for example), where events are typically transmitted asynchronously in the form of quantized scalar values, when using signals we don't need to be limited to the amplitude of the real signal; rather it is desirable to measure and use local features of the control-signal, such as its magnitude, phase and frequency, yielding additional degrees of freedom in the sense that more independent variables can be transmitted through the control-signal. In order to estimate magnitude and phase, and to facilitate mathematical manipulations, the real control-signal x(t) can be modelled as a complex one z(t), such that

$$\mathrm{Re}\{z(t)\} = x(t)$$

This complex representation of a real-valued function is called the 'analytic signal'.

Analytic Signal

A signal which has no negative frequency components can be represented in continuous time as

$$z(t) = \frac{1}{2\pi}\int_0^{\infty} Z(\omega)e^{j\omega t}\,d\omega$$

Where $Z(\omega)$ is the complex coefficient (magnitude and phase) of the complex sinusoid $e^{j\omega t}$. Recalling Euler's identity any real sinusoid can be converted to a complex sinusoid by generating the imaginary part as the phase-quadrature component from the in-phase component via a phase-shift of 90 degrees (or π/2):

$$A \cdot e^{j\omega t + \varphi} = A \cdot \cos(\omega t + \varphi) + jA \cdot \sin(\omega t + \varphi)$$

For more complicated signals which can be expressed as a sum of sinusoids, a filter can be used which shifts each sinusoidal component by a quarter cycle (positive frequencies by –π/2, negative frequencies by π/2), which is called a *Hilbert Transform Filter*. This filter can be used to create the complex analytical signal $z(t)$ from a real signal $x(t)$ and its hilbert transform $\hat{x}(t)$, with the property of "filtering out" all negative frequencies by phase shifting. Given a complex signal it is possible to estimate the magnitude as the square root of the sum of the squares of the real and imaginary parts:

$$A_z(t) = \sqrt{x^2(t) + \hat{x}^2(t)}$$

and likewise the phase defined by the arctangent of the imaginary part over the real part:

$$\phi(t) = \mathrm{a}\tan\left(\frac{\hat{x}(t)}{x(t)}\right)$$

Accordingly, it is also possible to estimate the instantaneous frequency by differentiating the phase angle (measuring the angular difference between successive samples and divide by the time-interval between them):

$$\omega(t) \triangleq \phi'(t) = \frac{d}{dt}\phi t$$

Combining the product detector with a hilbert transform filter we can retrieve the real modulator-signal $y(t)$, its magnitude $A_m(t)$, the instantaneous phase $\phi_m(t)$ and the instantaneous frequency estimate $\omega_m(t)$ from an AM-signal $x(t)$ as shown in Fig.1 below:

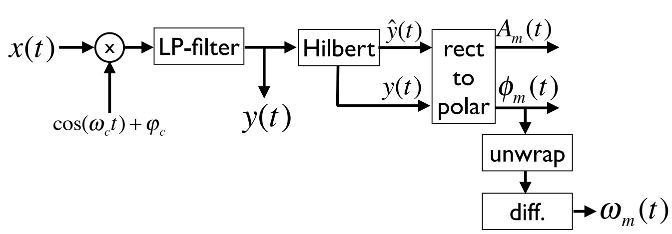


Fig. 1

Short-time Fourier Transform

The main idea behind the project is that the processing required for SSBSC-modulation and demodulation as shown in Fig.1 can be accomplished for many channels simultaneously using properties of the Phase-Vocoder. Let us recall the mathematic expression of the STFT X(n,k) of the signal x(n) as a function of frequency k and time n

$$X(n,k) = \sum_{m=-\infty}^{\infty} x(m) w_a(n-m) e^{-j\frac{2\pi}{N}km}$$

where w_a is the analysis windowing function. This equation can be regarded from two complementary perspectives, the fourier transform interpretation and the filterbank intepretation. The latter can be expressed as:

$$X(n,k) = \sum_{m=-\infty}^{\infty} \left(x(m) e^{-j\frac{2\pi}{N}km} \right) w_a(n-m)$$

In this representation the STFT is described as a heterodyne filterbank; Drawing from [1] in this perspective the internal operation of a single phase-vocoder channel consists of: 1) heterodyning the input with a sine and a cosine wave, 2) lowpass filtering the results, 3) converting the two filtered signals to polar coordinates, 4) unwrapping the phase values, 5) differentiating the phase-values, and 6) frequency shifting the instantaneous frequency estimate up to the channel's center frequency. Please note that the frequency estimation refers to the difference between the heterodyning sinusoid and the input signal, which is a welcome property in this context since before the last step (6) we obtain the instantaneous frequency estimate relative to DC, which is the instantaneous frequency of the modulator.

1) Demodulation

When setting the carrier frequency of a SSBSC-AM signal to the center-frequency of the bins in the STFT the carrier-signal will be heterodyned down to DC. What is then retrieved after convolution with the lowpass filter w_a (analysis window function) is the upper side band of the modulator, i.e. the control signal. Now we can get the real part of the control signal $y(t)$, its magnitude $A_m(t)$, the instantaneous phase $\phi_m(t)$ and the instantaneous frequency estimate $\omega_m(t)$. Fig. 2 shows a diagram of a single filter-channel of the phase vocoder (in continuous time) used for demodulation.

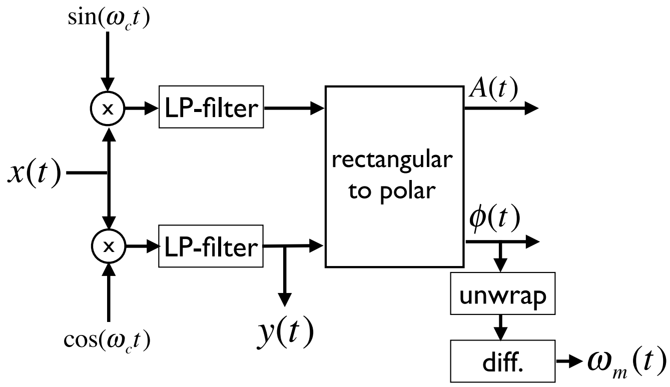


Fig. 2

It is interesting to compare Fig.2 to Fig.1. In this perspective, the difference lies in representing the modulator as a complex signal before or after the lowpass-filtering. Using the phase-vocoder for demodulation of SSBSC-AM signals should provide an efficient method for demodulating a large number of control signals simultaneously, though with a tradeoff in terms of bandwidth, number of channels and temporal resolution, which will be discussed later.

2) Modulation

The inverse STFT equation is given by

$$x(n) = \sum_{m=-\infty}^{\infty} w_s(n-m)\frac{1}{N}\sum_{k=0}^{N-1} X(m,k)e^{j\frac{2\pi}{N}km}$$

Where $X(m,k)$ is the complex sinusoid $A \cdot e^{j\frac{\phi m}{N}}$

By providing a magnitude and phase-value for every bin in successive STFT-frames, the sinusoidal control-signals can be modelled directly in the frequency domain. Applying the inverse FFT these sinusoidal control-signals are then heterodyned, that is frequency-shifted, into the frequency-bands of the respective bins and summed, resulting in an efficient way for generating multiple SSBSC-AM signals with carriers evenly spaced across the frequency spectrum.

3) Constraints

It is well-known that in order for the phase-vocoder to work properly a number of constraints need to be fulfilled; the components to be analyzed must be spaced widely enough apart such that only one component contributes to each frequency-band [8]. The bandwidth B_w of a frequency-band is determined by the windowing function's shape (window coefficient C_w) as well as the frequency resolution given by Samplerate Fs divided by the windowsize M. In order to avoid leakage into other bins the bandwidth of the control-signals must not exceed

$$C_w \cdot \frac{Fs}{M}$$

Further, in order to avoid ambiguities when estimating the instantaneous frequencies the hopsize H must fulfill the constraint

$$H \le \frac{M}{2C_w}$$

The ubiquitous tradeoff between between temporal and frequential resolution applies also here in the sense that larger windowsizes introduce more latency. Obviously, when modulating control-signals, changes in amplitude and frequency can only happen at framerate instead of samplerate; it might be interesting to investigate if this property avoids exceeding the bandwidth of the control-signals, e.g. due to modulating the respective magnitudes and phases to fast, resulting in amplitude or frequency modulation of the control-signals. Further, it is important to bear in mind the role of the analysis window for estimation of magnitude and instantaneous frequency: For a given sinusoidal signal such as

$$x(n) = A(n)\cos\left(\frac{2\pi}{N}knT + \theta(n)\right)$$

the magnitude estimate is attenuated by the gain of the lowpass filter (the window) as the input signal's frequency increases. Accordingly, by setting the carrier frequencies exactly between respective bins this effect can be inverted. Also, the phase, and therefore the instantaneous frequency estimate, is smeared in the sense that sudden changes in the input signal result in more gradual changes in the phase-vocoder [Dolson].

V. IMPLEMENTATION

The discussed concepts for SSBSC-AM and demodulation of sinusoidal control-signals have been implemented as a set of abstractions in Max/MSP. Tools for asynchronous demodulation and tracking of magnitude, phase and instantaneous frequency have been written which are available as part of McGill's digital orchestra toolbox[3]. Difficulties were encountered implementing the phase-vocoder for de/modulation using Max/MSP's pfft~ object, due to having no explicit control of the overlapping process and thus synchronicity issues arise when providing values to the

[3] http://www.idmil.org/software/dot

distinct STFTs with sample-accuracy. Although at the time of writing this report there is no phase-vocoder-based implementation yet, the concept has been proven by emulating the concepts for a number of channels in the time-domain, please see the appendix for a screenshot of an MSP-patch implementing demodulation of a sinusoidal SSBSC-AM signal as depicted in Fig.1.

VI. Conclusion & Future Work

The use of SSBSC-AM using properties of the phase-vocoder has proven to be an efficient and convenient way for transmitting sinusoidal control-signals through an audio-channel: In addition to the low-latency, high-resolution transmission using an audio-device and interpreting the signals in software, additional variables can be used, such as the control-signal's magnitude, phase and instantaneous frequency. In fact, the presented concepts have already found an application in a digital musical instrument for detection of plucking-gestures performed on string-like stretch-sensors: The sudden discontinuity in the signal when plucking (in contrast to fast pulling or stretching) causes a spike in the instantaneous frequency estimation which has proven to be more reliable compared to techniques based on tracking or differentiating the amplitude of the real signal (such as envelope followers or Schmitt-triggers, for example).

Whilst using the phase-vocoder is more efficient for multiplexing and demultiplexing of a large number of control signals with similar bandwidth, time-domain techniques might be more useful when a few number of control-signals with varying bandwidths are used. Another interesting property to investigate would be the use of specific window functions which might be employed for conditioning of control signals, in the sense of frequency-dependent smoothing or filtering. A Max/MSP-based implementation of the phase-vocoder is planned in which the overlapping is performed explicitely instead of relying on the pfft~ object, also a poly~-based implementation is envisaged.

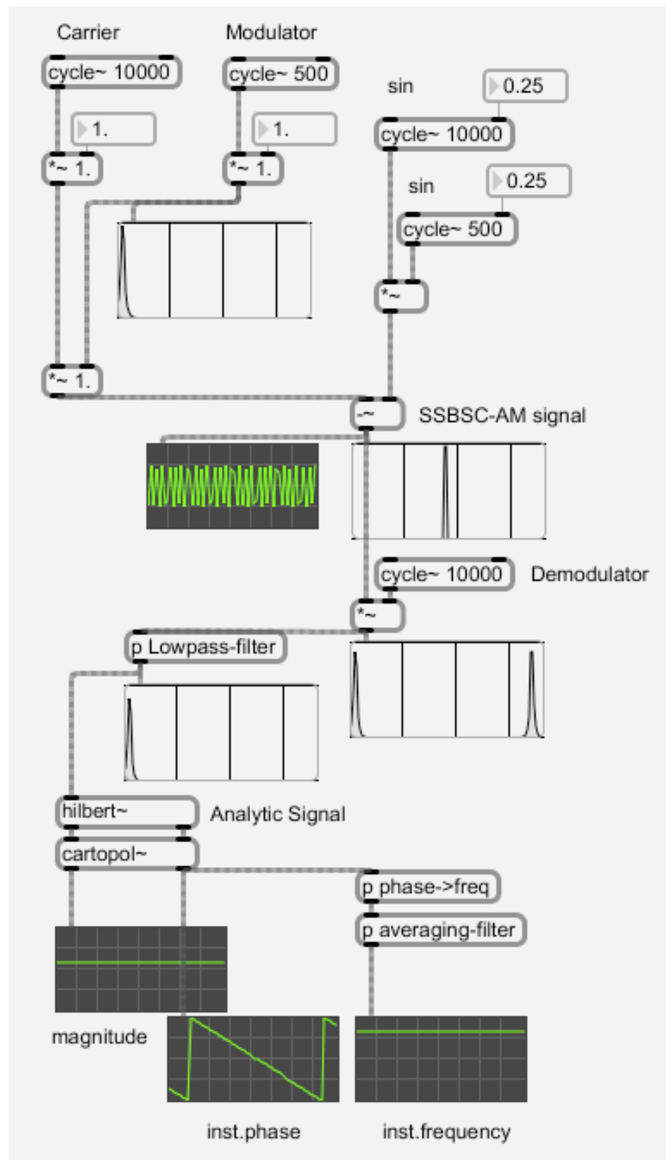
An interesting field for future work would be the use of subcarriers or orthogonal techniques as applied in Discrete Multitione Transmission (DMT) used for DSL-modems, or Orthogonal Frequency Division Multiplexing (OFDM) used for digital audio broadcasting, in which the modulated signals no more need to be evenly distributed across the frequency spectrum.

Appendix

1) AM-schemes designated by ITU 1982

designation	Description
A3E	double sideband DSB
R3E	double sideband reduced carrier DSBRC
H3E	single sideband full-carrier SSB
J3E	single sideband supressed carrier SSBSC
B8E	independent sideband ISB
C3F	vestigial sideband VSB
Lincompex	linked compressor and expander

2) SSBSC-AM Demodulation in MSP



REFERENCES

[1] Dolson, M. "The phase-vocoder: A Tutorial" Computer Music Journal vol. 10, no.4 pp. 14-27, 1986

[2] Gish, W. C., "Analysis and Synthesis of Musical Instrument Tones" AES Convention 61, New York 1978

[3] Jones, R. "Controlling a physical model with a 2D force matrix", Proc. of NIME, New York 2007

[4] Moorer, J.A., "Signal Processing Aspects of Computer Music: A Survey", Proc. of the IEEE, Vol 65, No. 8, 1977

[5] Neville, B. "Gesture analysis through a computer's audio interface: The Audio- Input Drum." MaSc. Thesis, University of Victoria, Victoria, BC, 2006

[6] Nevile, B., Driessen, P., and W. A. Schloss. "A New Control Paradigm: Software-Based Gesture Analysis for Music," *Proceedings of IEEE Pacific Rim Conference*, Victoria, Canada, August 2003.

[7] Newkirk, David and Karlquist, Rick (2004). Mixers, modulators and demodulators. In D. G. Reed (ed.), *The ARRL Handbook for Radio Communications* (81st ed.), pp. 15.1–15.36. Newington: ARRL. ISBN 0-87259-196-4.

[8] Puckette, M. S., "Phase-Locked Vocoder", Proc. of IEEE Conference on Applications of Signal Processing to Audio and Acoustics, 1995

[9] Smith, J.O. "Mathematics of the Discrete Fourier Transform (DFT)", second edition, online book, accessed April 13th, 2009

[10] http://en.wikipedia.org/wiki/Analytic_signal accessed April 13th, 2009

[11] Zicarelli, D. 1991. "Communicating with Meaningless Numbers." *Computer Music Journal* 15(4):74-7