Sinewave Additive Synthesis Revisited

3128 (Z-1)

Robert C. Maher
University of Nebraska-Lincoln
Lincoln, NE 68588-0511, USA

4E/Z-1

Presented at the 91st Convention 1991 October 4–8 New York

AES

AN AUDIO ENGINEERING SOCIETY PREPRINT

Sinewave Additive Synthesis Revisited

Robert C. Maher

Digital Audio Signal Processing Laboratory
Department of Electrical Engineering
University of Nebraska-Lincoln
209N WSEC
Lincoln, NE 68588-0511

E-mail: eerdrcm@engvms.unl.edu
Voice: 402-472-3771, Fax: 402-472-4732

0: Abstract

Sinusoidal additive synthesis is a well-known technique in the audio and electro-acoustic music fields. However, recent developments in digital signal processing methods allow the use of additive synthesis for precise editing, splicing, noise reduction, and other digital audio applications. Several new signal processing methods based on an extended sinewave model are described which allow for flexible manipulation of audio material.

1: Introduction

Many audio signal processing tasks require modification of the time, frequency, phase, and amplitude characteristics of an input signal. Examples include equalization (modification of frequency content), additive mixing (amplitude modification), delay or rate compression (time scale modification), etc. Ideally an audio signal processing system should incorporate a flexible time-frequency representation to allow *independent* manipulation of all aspects of the input signal. This paper describes the use of an additive synthesis method to provide the desired flexible formulation.

Additive synthesis is a widely used technique for the generation and decomposition of complex waveforms [1-9]. Several extensions and generalizations of the technique have appeared over the last few years which allow for precise, time varying control over signal parameters [10-15]. These new formulations can be implemented as part of an audio analysis/modification/synthesis framework with unexpectedly good results in many cases of practical interest.

This paper begins with a description of the standard additive synthesis formulation and time varying extensions. Several examples of typical processing situations are presented next, followed by a brief conclusion.

2: Additive Synthesis (and Analysis)

Additive synthesis is a method for generating complex waveform *f(t)* by adding together a set of simple waveforms, typically sinusoids, with independent amplitudes, frequencies, and phases, viz.,

$$f(t) = \sum_{k=1}^{K} A_k \cdot \sin(\omega_k t + \phi_k) \tag{1}$$

where K is the total number of components and the constants A_k, ω_k, and ϕ_k are the amplitude, angular frequency, and initial phase angle of the kth sinusoidal component. Note that equation (1) forms a standard Fourier series if the component frequencies are harmonically related (i.e., $\omega_k = k \cdot \omega_1$). Any *periodic* waveform can be represented in Fourier series form, although the total number of components K may be infinite.

However, most signals of practical interest are not strictly periodic because the amplitude and the frequency content generally varies with time. Thus, equation (1) can be generalized to include time varying components:

$$f(t) = \sum_{k=1}^{K} A_k(t) \cdot \sin(\theta_k(t)), \tag{2a}$$

where

$$\theta_k(t) = \phi_k(t) + \int_0^t \omega_k(\tau)\, d\tau. \tag{2b}$$

The term $\phi_k(t)$ is included as an optional "phase adjustment" parameter. *Note that the correct relationship between the argument* $\theta_k(t)$ *of the sine function and the time varying frequency* $\omega_k(t)$ *involves a time integral!*

The instantaneous frequency of component k at time t is defined to be the time derivative of the phase function $\theta_k(t)$, which is given by

$$\frac{d}{dt}\{\theta_k(t)\} = \frac{d}{dt}\{\phi_k(t)\} + \omega_k(t) \quad . \tag{3}$$

A schematic representation of the additive synthesis equations (2a) and (2b) is given in **Figure 1.**

It should be noted that while the equations above are presented in continuous-time form, discrete-time versions can obtained also. Thus, the additive synthesis procedure can be implemented in a straight-forward manner on a digital computer or signal processor. Note also that the amplitude and frequency functions appear separately as functions of time in the synthesis equations so that (at least in principle) independent modifications can be made to the amplitude, frequency, and time scales of each component.

For additive synthesis to be useful, of course, it is necessary somehow to obtain the functions $A_k(t)$, $\omega_k(t)$, and $\phi_k(t)$ in the first place. For the cases of interest here the functions are obtained using a sinusoidal analysis procedure. Several approaches for sinusoidal analysis have been used over the years, including parallel filter banks, spectrographs, swept bandpass filters, and digital methods based on the discrete Fourier transform (DFT) and short-time Fourier transform (STFT) [3-9]. The early techniques were generally *pitch synchronous*, meaning that the signal to be analyzed was assumed to be harmonic with known fundamental frequency. For example, the length of the discrete Fourier transform block was chosen to be the same number of samples as the waveform period. This restricted the analysis techniques to monophonic, harmonic signals of constant frequency (see **Figure 2**). Nonetheless, the analysis/synthesis results provided an essentially perfect reconstruction of original signal, both mathematically and perceptually [5, 6].

3: The 'MQ' Formulation

A more general analysis/synthesis framework, published first by McAulay and Quatieri [10], has also been found to be useful for speech, music, bioacoustic sounds, etc. [11-15]. The McAulay and Quatieri (or *MQ*) representation can be considered a generalization of simple Fourier analysis to include time-variant spectra and possibly non-harmonic partials, as in equation (2) above. In the MQ analysis procedure the digitized input signal is divided into overlapping sections called *frames*. Each frame is multiplied ("windowed") by a lowpass window function to reduce spectral leakage, followed by calculation of a high resolution DFT using a zero-padded Fast Fourier Transform (FFT) algorithm. The magnitude of the DFT is computed and all "peaks" in the magnitude spectrum are identified using interpolation ([11]) and attributed to *underlying sinusoidal components* at those frequencies. The amplitude, frequency, and phase corresponding to the spectral peaks are then calculated and recorded. The analysis and peak-picking process is repeated for each of the input frames and the spectral peak information (amplitude, frequency, and phase) is matched from frame to frame in order to follow changes in the input signal. The matching process results in connected chains (or tracks) of peaks from frame to frame. The peak tracks are "born" and "die" as the spectral content of the signal varies with time. The peak matching process has the drawback that the computation load of the MQ analysis procedure is high if the number of peaks varies substantially from frame to frame. However, a nice feature of frame to frame matching is that it ensures at least first order amplitude, frequency, and phase continuity. The MQ analysis system is depicted in **Figure 3**, and an example of the analysis output is shown in **Figure 4**.

The input signal can be regenerated by an additive synthesis technique using the amplitude and frequency information obtained for each frame and a smoothly interpolated phase function, as shown in **Figure 5**. Because the FFT returns phase values reduced to the principal value range ($-\pi$ to π), the *actual* absolute phase is not known. Instead, the MQ procedure tests which phase extension (by multiples of 2π) results in the phase function with the smallest curvature under the constraint that the slope of the phase function must equal the measured frequency (phase derivative) at each block, with cubic spline interpolation between blocks.

While the MQ process does not necessarily form a mathematically perfect analysis/synthesis system, the resynthesis results have been found to be excellent for many musical input signals [14, 15]. Signals such as broadband noise that are not well described as a sum-of-sinusoids are synthesized with somewhat diminished fidelity, but again the results can be surprisingly good for complex sonic textures.

The implementation of the MQ analysis/synthesis process used to generate the examples in this paper utilized 32-bit floating point numbers for all numerical parameter storage and arithmetic. This choice was made to suit the research needs of this project at the expense of inefficient data storage and additional computation overhead. Appropriate data range choices should allow a smaller and faster integer implementation of most aspects of the process. Indeed, the original algorithm of McAulay and Quatieri (for speech processing) was realized in a realtime, 16-bit integer implementation [10].

4: Examples of 'MQ' Additive Synthesis for Audio Processing Tasks

In this section several illustrative examples of additive synthesis applied to audio processing tasks are presented. These include time scale, amplitude scale, and frequency scale modifications, smooth-phase splicing and signal interpolation, broadband noise removal, and creative *musique concrète* applications and sound effects generation.

Process 1: Time Compression/Expansion

Time compression and expansion refers to a change in the temporal evolution of a signal without a corresponding change in the frequency spectrum. This can be accomplished within the MQ framework by changing the implicit time spacing between the analysis frames during synthesis. Moreover, the compression or expansion can be made to vary with time as well, creating gradual *accelerando* and *ritardando* effects. The amplitude and frequency data is simply compressed or stretched in time, but the phase information must be treated carefully: the phase of a sinusoidal component is the *time integral* of the frequency (see equation 2b), so the original phase information from the analysis must be replaced with new phase data obtained by integrating the compressed or stretched frequency information [16]. An example of time varying compression/expansion is shown in **Figure 6.**

Empirical results indicate that the maximum useful time compression is about 0.25, meaning that the original signal is compressed in time to one quarter of its original length. Time expansion by up to a factor of 3 or 4 typically provides satisfactory results. For expansion beyond a factor 3 the sinewave basis of the MQ process becomes increasingly apparent: very short "on-off" sinusoidal components that were nothing more than clicks in the original analysis are stretched into noticeably pitched tones by an extreme time expansion.

Process 2: Frequency Shifting

Frequency shifting is accomplished in a similar manner to time scale compression/expansion. The frequency and phase information is adjusted to provide a change in the frequency spectrum while retaining the amplitude and temporal evolution information. Simply multiplying all the frequency information by a constant amount and compensating the component phases shifts both the frequency of each component *and* the overall spectral envelope. This results in a change in the resonant or formant characteristics of the shifted signal, but small frequency shifts are often found to be acceptable. More sophisticated methods for detecting and compensating for the spectral envelope via homomorphic estimation or other techniques are possible if the need arises [16]. A simple example of frequency shifting is shown in **Figure 7**.

Process 3: Splicing and Interpolation

The frame-to-frame matching of the MQ process has several benefits for additive synthesis with modifications. Consider the example of **Figure 8a** in which a gap exists in the signal simulating a lost or corrupted segment of data. The MQ analysis data for the example is show in **Figure 8b**. The gap information can be estimated by extending the MQ amplitude, frequency, and phase functions across the gap in order to synthesize the missing segment, as shown in **Figures 8c and 8d**. A similar technique can be used to avoid transitional artifacts during a cross fade between two or more signals.

Another application of the MQ process is in waveform *looping* for sample-lookup synthesizers. In this case the signal to be looped is first sent through the MQ analyzer, then resynthesized with the required duration and phase adjustments to match the signal at the required loop boundaries (**Figure 9**). Thus, problems associated with waveform loop point discontinuities can be significantly reduced or avoided entirely.

Process 4: Noise Removal

Since the MQ analysis procedure identifies only spectral peaks in each analysis frame, it is possible to set a minimum threshold below which no peaks are selected. If the threshold is selected to be above the level of any unwanted noise components but below the level of the desired signal information the MQ process behaves as a multiband noise gate [17]. Additional automatic noise removal qualities can be included by eliminating any sinusoidal components or component dropouts with duration less than some preselected amount, thereby reducing the presence of impulsive clicks and pops. Interactive editing of the MQ analysis data is also possible if the tedium of manual examination of the analysis output is deemed necessary. An example of noise removal on a short segment of input signal is shown in **Figure 10**.

Process 5: Musique Concrète and Special Effects

The techniques described above can be combined with additive mixing, filtering, equalization, etc. to produce a wide variety of special sound manipulation effects for technical or creative purposes [12-15, 18]. In particular, musical compositions involving manipulated recordings of real-world sounds (musique concrète) can benefit greatly from the processing methods available in the MQ repertoire [19].

5: Conclusions

This paper has described the use of an extended formulation of the well-known sinusoidal additive synthesis technique in several modern digital audio applications. The primary advantage of this approach is the flexibility of the sinusoidal representation: careful transformations and manipulations of the analysis data produce predictable and controllable results. Widespread use of this procedure is possible as fast programmable digital processors and architectures become increasingly available for digital audio purposes. Work by the author is continuing on fast implementations of this process.

6: Acknowledgements

The author is pleased to acknowledge Ariel Corporation for providing assistance to the University of Nebraska Digital Audio Signal Processing Laboratory. Additional assistance from both the University of Nebraska Research Council and the Department of Electrical Engineering (Prof. Rod Soukup, Chairman) is also greatly appreciated.

7: References

[1] Ohm, G.S., "Uber die Definition des Tones, nebst daren Geknupfter Theorie der Sirene und Ahnlicher Tonbildender Vorrichtungen," *Ann. Phys. Chem.*, vol. 59, pp. 513-565, 1843.

[2] von Helmholtz, H.L.F., *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, (tr. J. Ellis, 1954); Dover, 1885.

[3] Freedman, M.D., "A Method for Analyzing Musical Tones," *J. Audio Eng. Soc.*, vol. 16, no. 4, pp. 419-425, 1968.

[4] Beauchamp, J.W., "A Computer System for Time-Variant Harmonic Analysis and Synthesis of Musical Tones," in *Music by Computers*, H. von Foerster and J.W. Beauchamp, eds., pp. 19-61; New York: Wiley, 1969.

[5] Risset, J.-C., and Mathews, M.V., "Analysis of Musical-Instrument Tones," *Physics Today*, vol. 22, no. 2, pp. 23-30, 1969.

[6] Moorer, J.A., "On the Segmentation and Analysis of Continuous Musical Sound by Digital Computer," Ph.D. Dissertation, Stanford University, Stanford, CA (also Dept. of Music Rep. STAN-M-3), 1975.

[7] Alles, H.G., "Music Synthesis Using Real Time Digital Techniques," *Proc. of the IEEE*, vol. 68, no. 4, pp. 436-449, 1980.

[8] Dodge, C., and Jerse, T.A., *Computer Music: Synthesis, Composition, and Performance*; New York: Schirmer Books, 1985.

[9] Moore, F.R., *Elements of Computer Music*; Englewood Cliffs, NJ: Prentice-Hall, 1990.

[10] McAulay, R.J. and Quatieri, T.F., "Speech Analysis/Synthesis Based on a Sinusoidal Representation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, no. 4, pp. 744-754, 1986.

[11] Smith, J.O., and Serra, X., "PARSHL: An Analysis/Synthesis Program for Non-harmonic Sounds Based on a Sinusoidal Representation," *Proc. 1987 Int. Computer Music Conf.*, pp. 290-297; San Francisco: Computer Music Assn., 1987.

[12] Maher, R.C., and Beauchamp, J.W., "An Investigation of Vocal Vibrato for Synthesis," *Applied Acoustics*, vol. 30, no. 2-3, pp. 219-245, 1990.

[13] Maher, R.C., "Evaluation of a Method for Separating Digitized Duet Signals," *J. Audio Eng. Soc.*, vol. 38, no. 12, pp. 956-979, 1990.

[14] Serra, X., "A System for Sound Analysis/Transformation/Synthesis Based on a Deterministic plus Stochastic Decomposition," Ph.D. Dissertation, Stanford University, Stanford, CA (also Dept. of Music Rep. STAN-M-58), 1989.

[15] Serra, X., and Smith, J.O., "Spectral Modeling Synthesis: A Sound Analysis/Synthesis System Based on a Deterministic plus Stochastic Decomposition," *Computer Music Journal*, vol. 14, no. 4, pp. 12-24, Winter, 1990.

[16] Quatieri, T.F., and McAulay, R.J., "Speech Transformations Based on a Sinusoidal Representation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, no. 6, pp. 1449-1464, 1986.

[17] Moorer, J.A., and Berger, M., "Linear-Phase Bandsplitting: Theory and Applications," *J. Audio Eng. Soc.*, vol. 34, no. 3, pp. 143-152, 1986.

[18] Maher, R.C., "Computer Processing of Audio Signals by Exclusion Filters," *J. Acoust. Soc. Am. Suppl. 1*, vol. 88, p. 188, 1990.

[19] Maher, R.C., "A Meditation on a Theme by Dr. Martin Luther King, Jr.," (LP recording), part of *New Electronic Works from the University of Illinois*, (University of Illinois, Urbana, 1989).

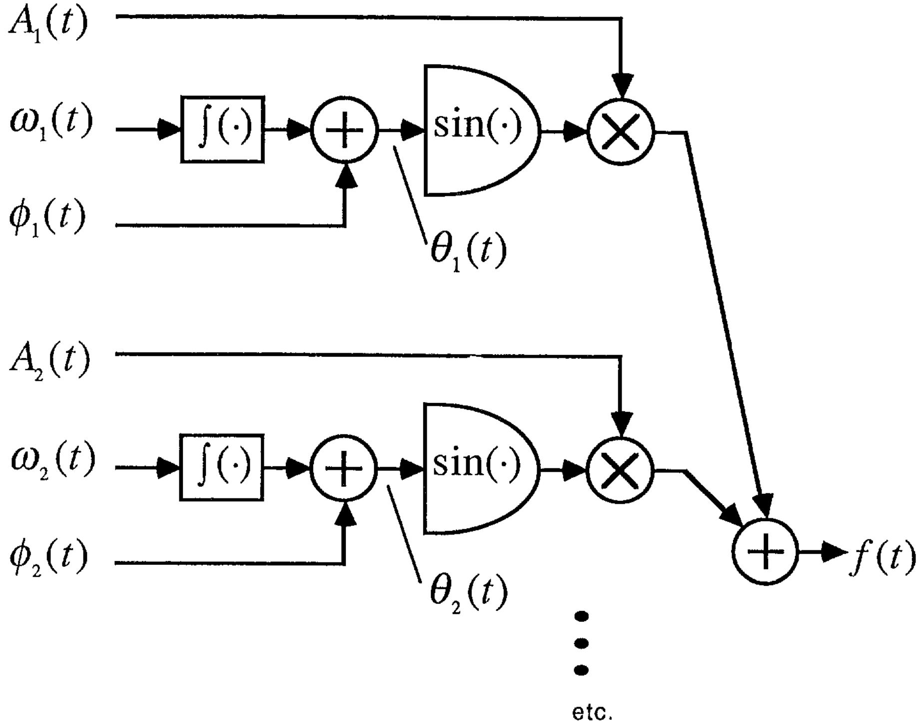

Figure 1: Schematic representation of additive synthesis equations 2a and 2b. Each sinusoidal component has independently varying amplitude $A_k(t)$, angular frequency $\omega_k(t)$, and phase offset $\phi_k(t)$. For harmonic synthesis $\omega_k(t) = k \cdot \omega_1(t)$.

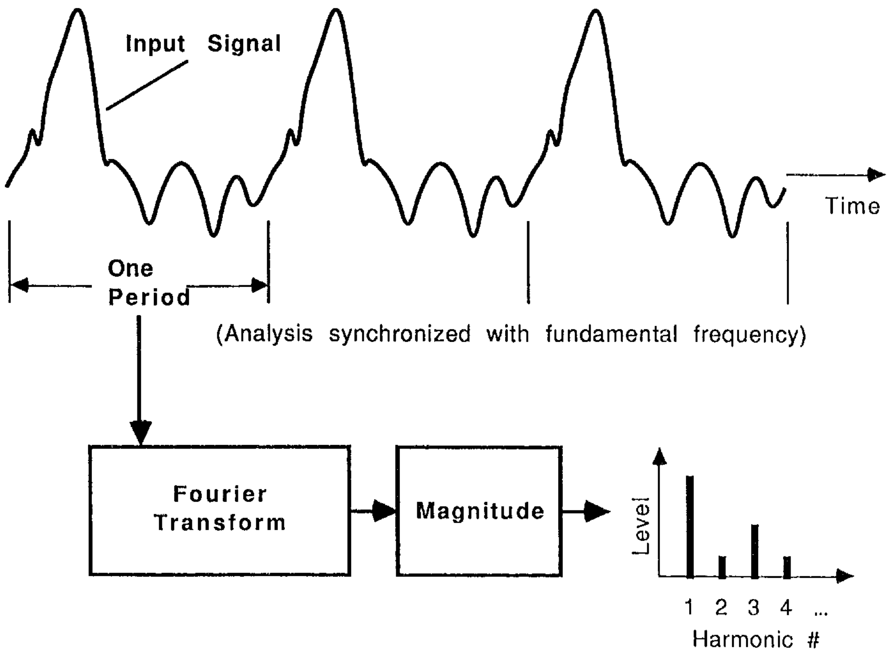

Figure 2: An example of a pitch-synchronous analysis framework. The input signal is assumed to be harmonic with known fundamental frequency so that spectral levels from a Fourier Transform correspond to harmonics of the signal. Modification to include signal weighting (windowing) prior to the transform is also possible.

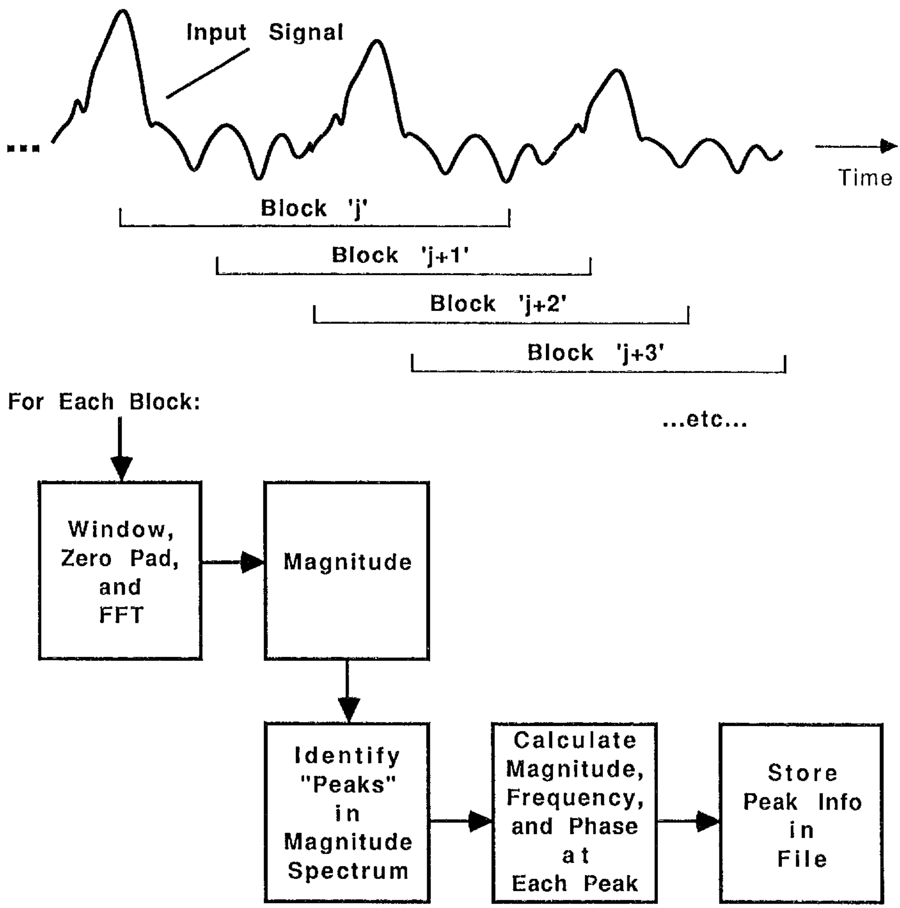

Figure 3: Functional schematic of the McAulay-Quatieri (MQ) signal analysis procedure. The input signal is divided into overlapping blocks with length greater than the longest waveform period (lowest frequency) to be processed. A zero-padded FFT is used to increase the density of spectral samples.

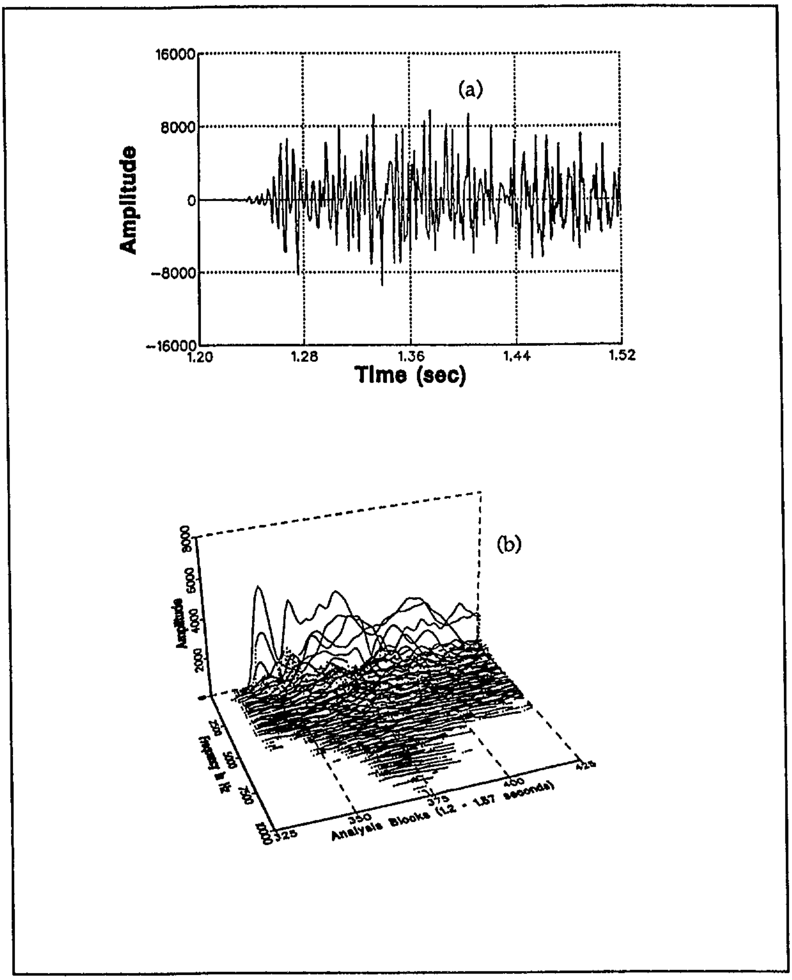

Figure 4: Example of MQ analysis:
(a) Input signal.
(b) MQ analysis output.

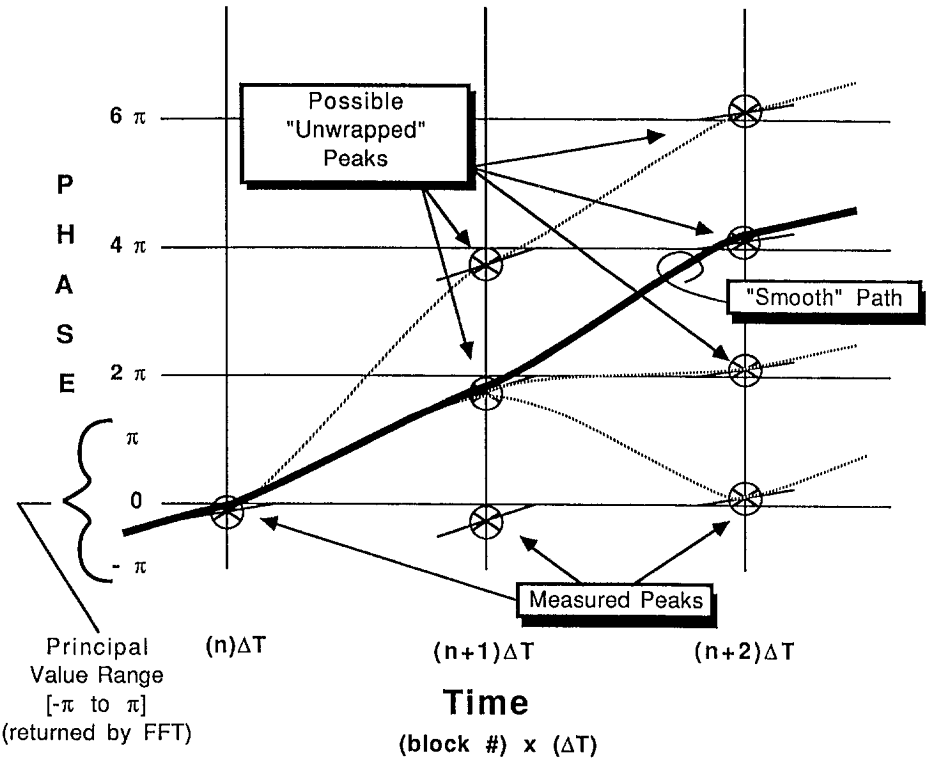

Figure 5: Synthesis from the MQ analysis data using smooth phase interpolation. Each spectral peak is measured in terms of magnitude, frequency, and phase. Phase measurements from the Fourier transform are obtained in the principal value range (-π to π) and must be *unwrapped* to the correct extension. The frequency is the *slope* of the curve on the phase vs. block plot above. A "smoothest trajectory" rule is used to determine a reasonable estimate of the actual, non-principal value of the phase.

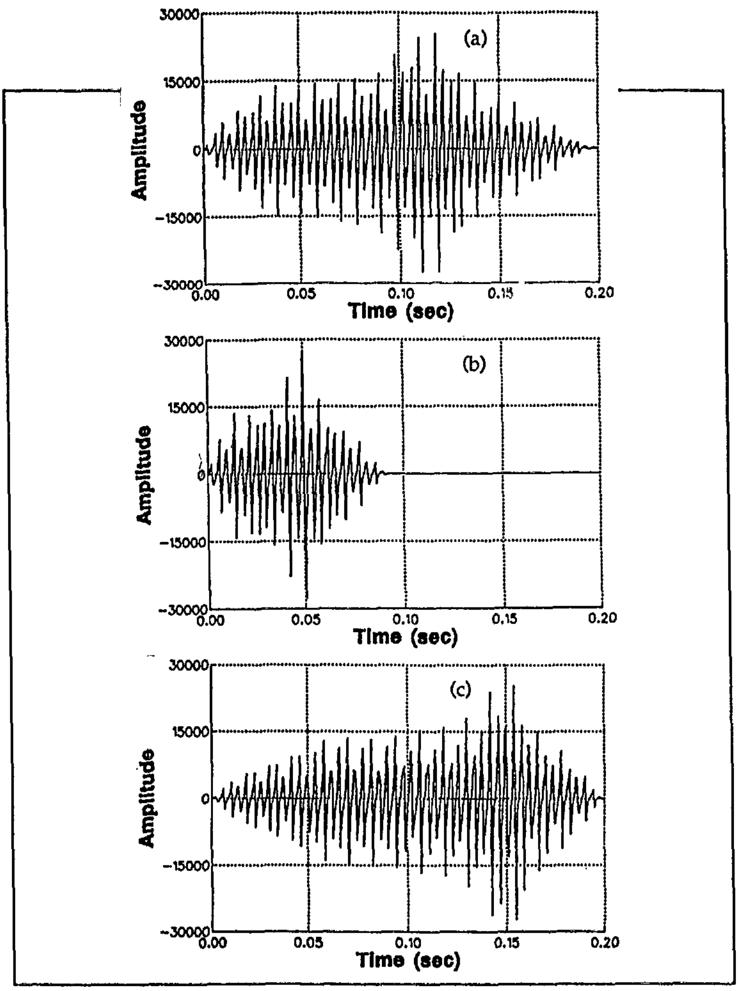

Figure 6: Time scale compression and expansion.
a) Original signal.
b) Signal compressed in time by a factor of 2.
c) Signal with time varying expansion/compression.

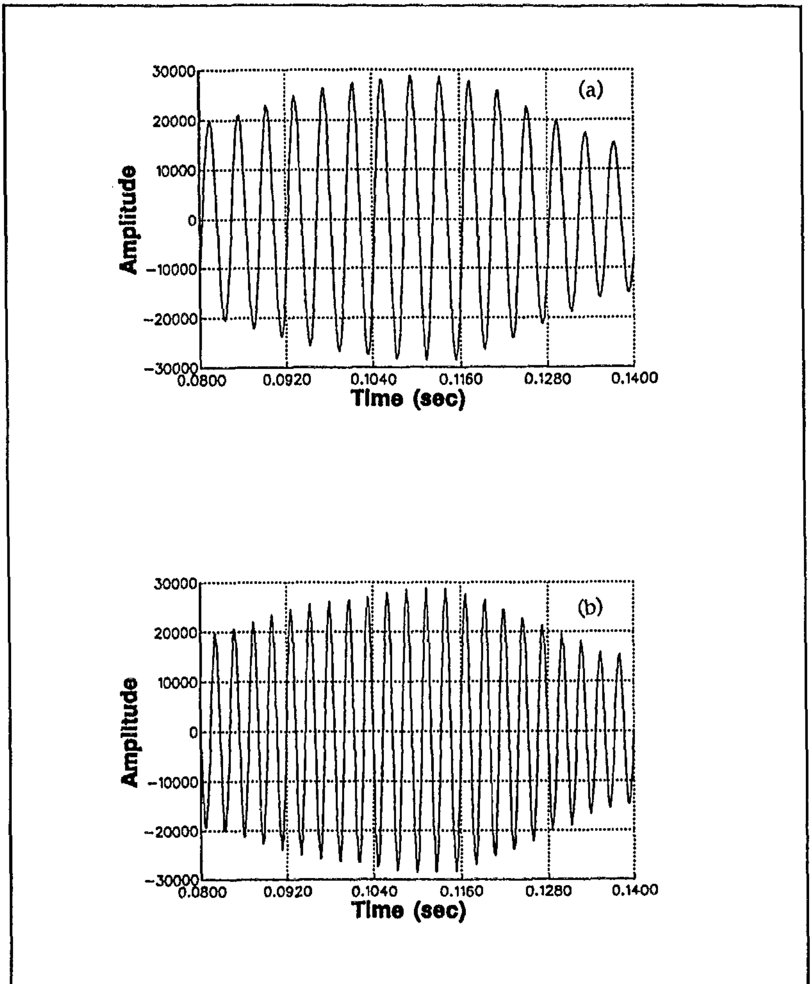

Figure 7: Simple example of frequency shifting.
a) Original signal segment.
b) Segment with frequency multiplied by 1.5
(note the unchanged time envelope).

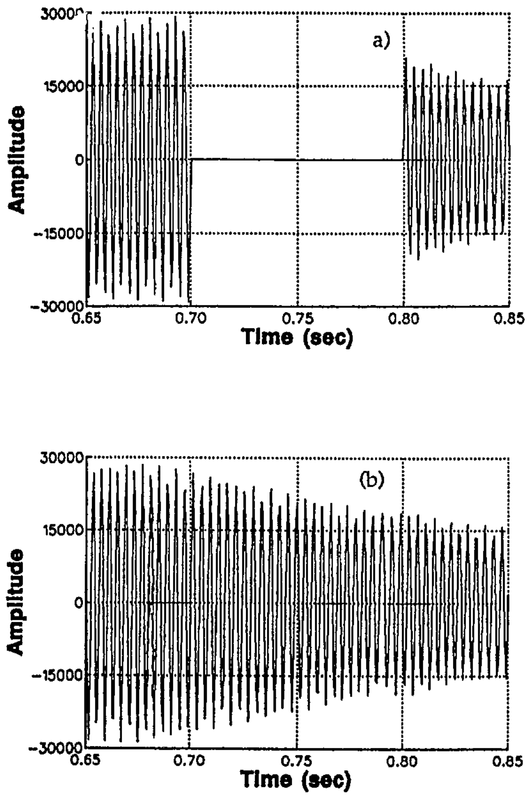

Figure 8: Resynthesis to fill a signal "gap".
a) Original signal with gap to simulate missing or corrupted data.
b) Resynthesis to fill gap by extending analysis information across the missing region.

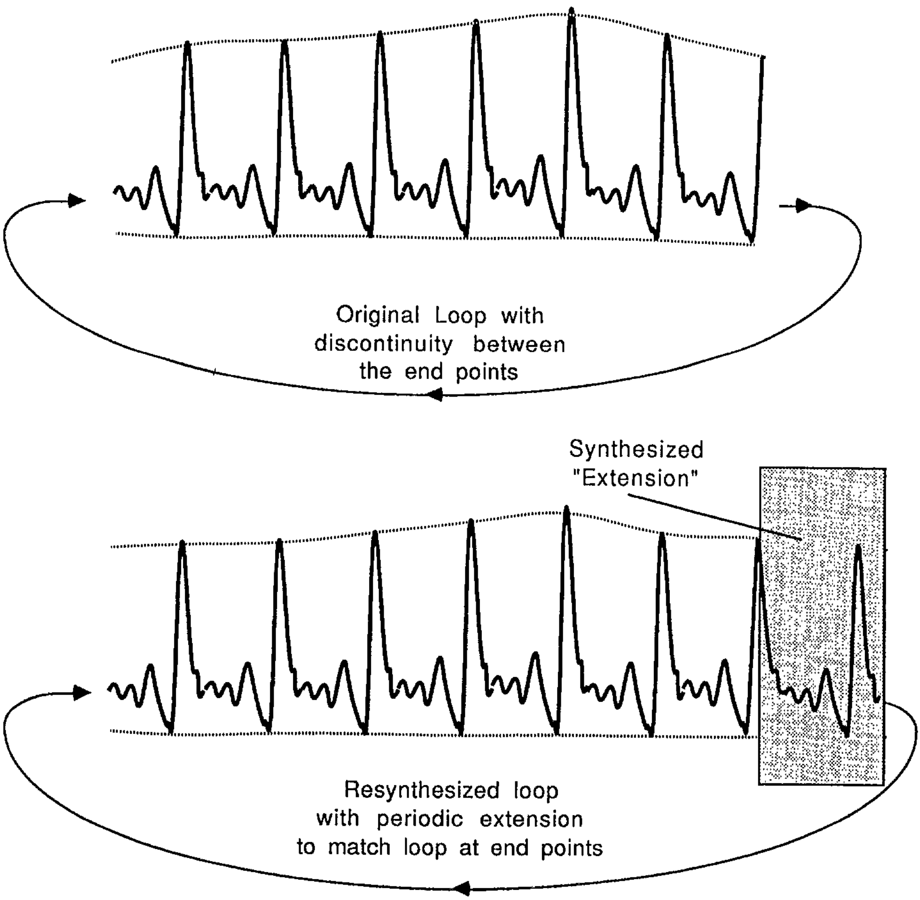

Figure 9: Example of MQ additive synthesis to generate a signal extension to match the ends of a waveform loop.

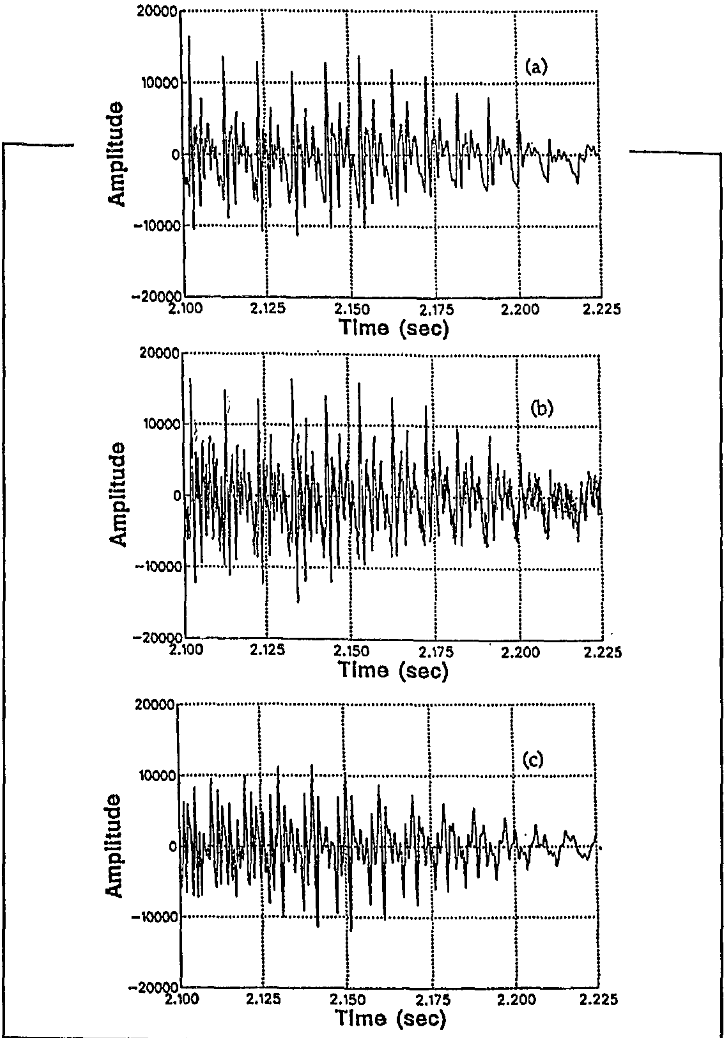

Figure 10: Example of broadband noise removal from a noisy recording of speech.
a) Original signal without noise.
b) Signal with additive noise (broadband).
c) Resynthesis with spectral threshold to suppress noise.