A Tutorial on Digital Sound Synthesis Techniques

Article *in* Computer Music Journal · January 1992

DOI: 10.2307/3679529

1 author:

Giovanni De Poli
University of Padova

Giovanni De Poli
Centro di Sonologia Computazionale
Istituto di Elettrotecnica ed Elettronica
Università di Padova, Italy

A Tutorial on Digital Sound Synthesis Techniques

Introduction

Progress in electronics and computer technology has led to an ever-increasing utilization of digital techniques for musical sound production. Some of these are the digital equivalents of techniques employed in analog synthesizers and in other fields of electrical engineering. Other techniques have been specifically developed for digital music devices and are peculiar to these.

This paper introduces the fundamentals of the main digital synthesis techniques. Mathematical developments have been restricted in the exposition and can be found in the papers listed in the references. To simplify the discussion, whenever possible, the techniques are presented with reference to continuous signals.

Sound synthesis is a procedure used to produce a sound without the help of acoustic instruments. In digital synthesis, a sound is represented by a sequence of numbers (samples). Hence, a digital synthesis technique consists of a computing procedure or mathematical formula, which computes each sample value.

Normally, the synthesis formula depends on some values, that is, *parameters*. Frequency and amplitude are examples of such parameters. Parameters can be constant or slowly time variant during the sound. Time-variant parameters are also called control functions.

Synthesis techniques can be classified as (1) *generation techniques* (Fig. 1a), which directly produce the signal from given data, and (2) *transformation techniques* (Fig. 1b), which can be divided into two stages, the generation of one or more simple signals and their modification. Often, more or less elaborate combinations of these techniques are employed.

Computer Music Journal, Volume 7, No. 4,
Winter, 1983, 0148-9267/83/040008-19 $04.00/0,

Fixed-Waveform Synthesis

In many musical sounds, pitch is a characteristic to which we are quite sensitive. In examining the temporal shape of pitched sounds, we see a periodic repetition of the waveform without great variations. The simplest synthesis method attempts to reproduce this characteristic, generating a periodic signal through continuous repetition of the waveform. This method is called *fixed-waveform synthesis*.

The technique is carried out by a module called an *oscillator* (Fig. 2), which repeats the waveform with a specified amplitude and frequency. In certain cases, the waveform is characteristic of the oscillator and cannot be changed. But often it can be chosen in a predetermined set of options or given explicitly when required.

Usually, in digital synthesis the waveform value at a particular instant is not computed anew for each sample. Rather, a table, containing the period values computed in equally spaced points, is built beforehand. Obviously, the more numerous the points in the table, the better the approximation will be. To produce a sample, the oscillator requires the waveform value at that precise instant. It cyclically searches the table to get the point nearest to the required one. Sometimes a finer precision is achieved by interpolation between two adjacent points.

The distance in the table between two samples read at subsequent instants is called the *sampling_increment*. The sampling_increment is proportional to the frequency f of the generated signal according to the following formula (Mathews 1969):

$$sampling\_increment = \frac{N}{SR} f,$$

where N is the table length and SR the sampling rate.

In the oscillator, the frequency is usually speci-

Fig. 1. Classification of synthesis techniques. Generation techniques (a) and transformation techniques (b).

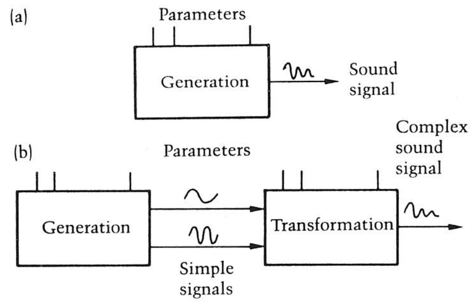
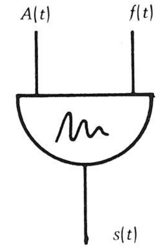


Fig. 2. Fixed-waveform synthesis oscillator.



fied as a sampling-increment and the algorithm that realizes it is as follows:

$$signal\ [t] := amplitude * table\ [phase], \quad \text{(Relation 1)}$$

and

$$phase := mod(n, phase + sampling\_increment), \quad \text{(Relation 2)}$$

where

Table contains one period of the waveform;
Phase is the theoretical position in the table of the sample to be extracted at the instant; and
Amplitude is the signal amplitude.

Relation 2 computes the phase value in the subsequent instant, approximating the frequency integration by a summation. The modulus operation keeps the phase inside the table length n.

It is noteworthy that the signal generated in this way is an approximation of the desired one (Mailliard 1976). The approximation depends on the table length, the interpolation method, and the signal frequency. For a sufficiently long table, it is fully satisfactory.

The results of fixed-waveform synthesis are of poor musical quality, as the sound does not present any variation along its duration. This technique can be changed by allowing the amplitude to vary in time. In real sounds, the amplitude is rarely constant: it starts from zero, reaches a maximum after a certain time (attack), remains nearly constant (steady state) and, after a certain evolution, it returns to zero (decay). This sequence of amplitude behavior is called the *envelope*. Thus, when the amplitude varies according to a control function, we have fixed-waveform synthesis with an amplitude envelope.

The envelope can be generated in many ways. In software-based synthesis, the most frequent method uses an oscillator module, seen previously, using a very low frequency equal to the inverse of the duration. In this case, it performs a single cycle and its waveform corresponds to the amplitude envelope.

By carefully analyzing natural periodic sounds, it has been shown that even the most stable ones contain small frequency fluctuations. These improve the sound quality and avoid unpleasant beatings when more sounds are present at the same time.

The fixed-waveform technique can also be modified so that the oscillator frequency can slowly vary around a value. This enables the production of a tremolo and, with wider variations, of a glissando or melodies.

The combination of these two variations constitutes fixed-waveform synthesis with time-varying amplitude and frequency. The waveform is fixed, while the amplitude and frequency vary. The partials are exact multiples of the fundamental, and they all behave the same.

Fixed-waveform synthesis is realized rather simply. Hence, it is often employed when good sound quality is not required. The constant waveform gives the sound a mechanical, dull, and unnatural character, which soon annoys the audience. Thus, in musical applications, fixed-waveform synthesis is not very effective when used alone. It is employed for its simplicity when timbral variety is not required, for example, for real-time synthesis on very limited hardware.

For economy, other methods of generating waveforms that do not use tables or multiplications have been devised. The simplest generates a square or (more generally) a rectangular wave, alternating sequences of positive and negative samples of the same value. The frequencies that can be obtained are submultiples of the sampling rate.

A sawtooth signal can also be generated by an accumulator to which a constant value is continuously added. The output increases linearly until it overflows and starts from the beginning. The signal frequency is proportional to the constant value. This method is used to produce linearly variable control signals. Every time the additive constant changes, the slope changes. Hence, functions composed of straight segments, such as envelopes, can be obtained.

This technique has been generalized recently by Mitsuhashi (1982*a*). A polynomial of degree N can be generated by putting N accumulators in cascade. The accumulators are initialized by the value of the forward differences, in decreasing order, of the polynomial to be generated (Cerruti and Rodeghiero 1983). The waveforms obtained exhibit great variety and, in certain conditions, they are periodic.

Granular Synthesis

The technique of fixed-waveform synthesis produces rather static sounds in time. Yet a fundamental characteristic of musical sound is its timbral evolution in time. A sound can be thought of as a sequence of elementary sounds of constant duration, analogous to a film, in which a moving image is produced by a sequence of images.

In computer music, the elementary sounds are called *grains*, and the technique of exploiting this facility is *granular synthesis* (Roads 1978). The grains can be produced by a simple oscillator or by other methods. The duration of each grain is very short, on the order of 5–20 msec.

There are two ways to implement granular synthesis. The first is to organize the grains into *frames*, like the frames of a film. At each frame, the parameters of all the grains are updated. This is the approach sketched by Xenakis (1971). The second way involves scattering the grains within a *mask*, which bounds a particular frequency/amplitude/time region. The density of the grains may vary within the mask. This is the method implemented by Roads (1978).

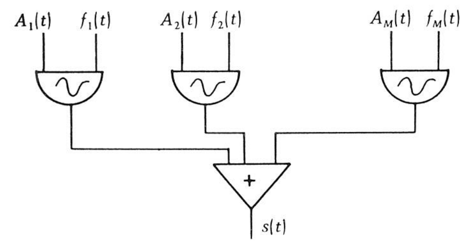
A problem with granular synthesis is the large amount of parameter data to be specified. In some other types of synthesis (additive and subtractive, to be discussed shortly), these data can be obtained by analyzing natural sounds. However, no analysis system for granular synthesis has been developed. Another possibility is to obtain the parameter data from an interactive composition system, which allows the composer to work with high-level musical concepts while automatically generating the thousands of grain parameters needed.

Additive Synthesis

In *additive synthesis*, complex sounds are produced by the superimposition of elementary sounds. In certain conditions, the constituent sounds fuse together and the result is perceived as a unique sound. This procedure is used in some traditional instruments, too. In an organ, the pipes generally produce relatively simple sounds; to obtain a richer spectrum in some registers, notes are created by using more pipes sounding at different pitches at the same time. The piano uses a different procedure. Many notes are obtained by the simultaneous percussion of two or three strings, each oscillating at a slightly different frequency. This improves the sound intensity and enriches it with beatings.

In order to choose the elementary sounds of additive synthesis, we first note that the Fourier analysis model enables us to analyze sounds in a way similar to the human ear and so to extract parameters that are perceptually significant. When we analyze a real, almost-periodic sound, we immediately notice that each partial amplitude is not proportionally constant, but that it varies in time according to different laws. In the attack portion of a note, some partials, which in the steady state are negligible, are often significant.

Fig. 3. Additive synthesis.



Any almost-periodic sound can be approximated as a sum of sinusoids. Each sinusoid's frequency is nearly multiple that of the fundamental, and each sinusoid evolves in time. For higher precision, the frequency of each component can be considered as slowly varying. Thus, additive synthesis consists of the addition of some sinusoidal oscillators, whose amplitude, and at times frequency, is time varying (Fig. 3).

The additive-synthesis technique also provides good reproduction of nonperiodic sounds, presenting in the spectrum the energy concentrated in some spectral lines. For example, Risset (1969) imitated a bell sound by summing sinusoidal components of harmonically unrelated frequencies, some of which were beating. In Risset's example, the exponential envelope was longer for the lower partials.

Additive synthesis provides great generality. But a problem arises because of the large amount of data to be specified for each note. Two control functions for each component have to be specified, and normally they are different for each sound, depending on its duration, intensity, and frequency. The possibility of data reduction has been investigated. At Stanford University, a first result has been obtained by representing the control functions of the amplitude and the frequency of each component by line segments, without affecting "naturalness" of the sound (Grey and Moorer 1977).

The next step has been to investigate the relations between these functions (Risset and Mathews 1969; Beauchamp 1975) or their relation to others of more general character (Charbonneau 1981). Additive synthesis is most practically used either in synthesis based on analysis (analysis/synthesis), often transforming the extracted parameters, or when a sound of a precise and well-determined characteristic is required, as in psychoacoustic experiments. In any case, in order to familiarize musicians with sound characteristics and frequency representations, the technique is also useful from a pedagogical point of view.

Additive synthesis can be generalized by using waveform components of other shapes besides sinusoids. To allow the reproduction of any sound, these waveforms have to satisfy specific mathematical properties. Walsh functions are an example of this kind of function; they are used for their simple hardware realization (Rozenberg 1979).

VOSIM

In the synthesis techniques already discussed, oscillators that periodically reproduce a given waveform are employed. Other synthesis techniques, instead of continuously repeating a given waveform, calculate it anew each period, with minor variations. The control of this calculation process allows continuous spectral variations. A common method of this type is the voice simulation (VOSIM) technique. A VOSIM oscillator has been devised in a project at the Institute of Sonology in Utrecht (Kaegi 1973, 1974; Kaegi and Tempelaars 1978).

The VOSIM waveform (Fig. 4) consists of a sequence of N pulses of shape $\sin^2$, of the same duration T, and of decreasing amplitude. The sequence is followed by a pause M. Each pulse's amplitude is smaller than the preceding one, by a constant factor b.

The VOSIM spectrum (Fig. 5a) is described as the product of two terms (Tempelaars 1976; De Poli and De Poli 1979). The first term S_1 (Fig. 5b) depends only on the pulse shape and limits the signal bandwidth to $2F$ (being $F = 1/T$). The second term S_2 (Fig. 5c) depends on the relationship between the individual pulse amplitudes. S_2 is periodic in the frequency domain with a period F, and it is sym-

Fig. 4. VOSIM oscillator: T is the duration of single pulse, M the rest between two sequences of pulses.

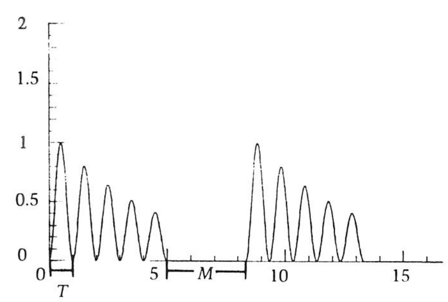
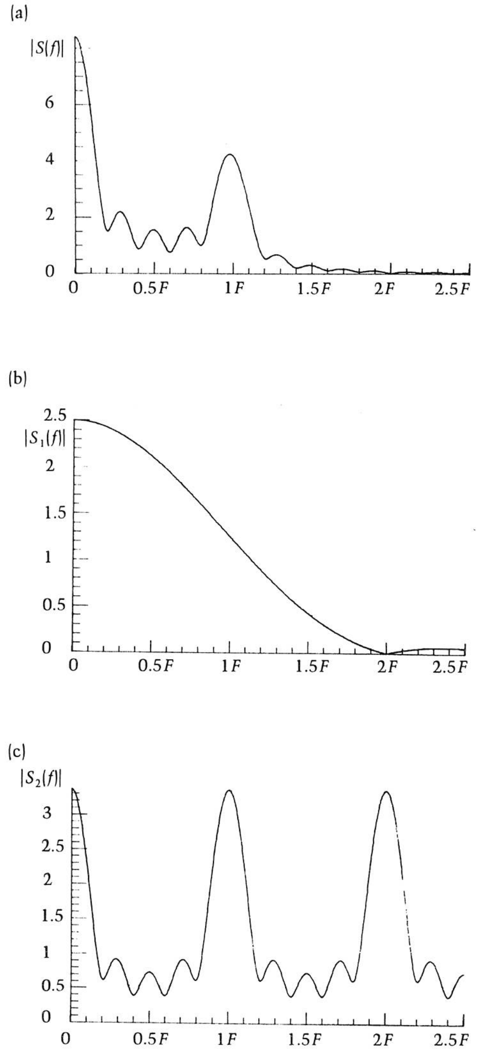


Fig. 5. Spectral envelope of a VOSIM oscillator (N = 5, b = 0, 8) (a). The envelope is the product of the terms S_1 (b) and S_2 (c).



metric with respect to $F/2$. When $b \approx 1$, its amplitude will be greater around the extremes of the period 0 and F. When $b \approx -1$, its amplitude will be greater in the central position around $F/2$. Thus, a characteristic formant in F or $F/2$ will result. The number of pulses N produces N oscillations in the S_2 term between 0 and F, with strong signals for b near $\pm F$.

This constitutes the spectral envelope of the repeated waveform. Taking α as the ratio between the signal period and a single pulse duration, the number of the harmonic corresponding to the formant is α if b is positive, and $\alpha/2$ if b is negative. Thus, by varying α, the formant shifts, and the relative amplitude of all the harmonics vary continuously but not homogeneously, following the spectral envelope. The signal and the formant frequencies can be separately controlled.

More kinds of sounds can be obtained by modulating (sinusoidally or randomly) the value of the time interval M between two consecutive pulse sequences. This means that α varies independently from T. In this case, the formant frequency remains constant while the harmonic amplitudes vary. Then the ear can easily perceive the spectral envelope and fuse the components together. This property makes the VOSIM oscillator effective in musical applications.

If α variation is strong, practically aperiodic sounds or colored noises are obtained. Adding several VOSIM oscillators allows one to control the position of the formants. This results in an additive synthesis of already complex sounds rather than of sinusoidal components. Instead of the frequency of partials, the position of the formants is controlled. This is a more relevant parameter, from an acoustic standpoint.

The *formant-wave-function synthesis* of Rodet (1980) is analogous to VOSIM, but it allows overlapping of single waveforms. This provides better control and generally richer sounds. Mitsuhashi (1982*a*) and Bass and Goeddel (1981) generalized the VOSIM model by including the case of pulses of any amplitude and using different elementary waveforms.

Synthesis by Random Signals

Up to now, we have considered signals whose behavior at any instant is supposed to be perfectly knowable. These signals are called *deterministic signals*. Besides these signals, *random signals*, of unknown or only partly known behavior, may be considered. For random signals, only some general characteristics, called statistical properties, are known or are of interest. The statistical properties are characteristic of an entire signal class rather than of a single signal. A set of random signals is represented by a *random process*. Particular numerical procedures simulate random processes, producing sequences of random (or more precisely, pseudorandom) numbers. The linear congruential method is commonly used to produce uniformly distributed numbers. From a starting value X_0, a sequence of random integers $X_0, X_1, \ldots, X_K \ldots$ is generated according to the relation

$$X_{K+1} = (a \cdot X_K + c)_{\mathrm{mod}\, m},$$

where m is the modulus and the maximum sequence period, and a and c are two specific integer constants.

The modulus operation can be avoided by choosing m as the maximum number representable in the computer, that is, $m = 2^b$, where b is the word length (bit number in a binary computer). So the numbers are automatically truncated. The choice of X_0, a, and c greatly affects the statistical characteristics of the generated sequence, and its acceptability has to be accurately verified by statistical tests. A general discussion of various distributions and the methods used to generate them can be found in Lorrain's paper (1980).

Random sequences can be used both as *signals* (i.e., to produce white or colored noise used as input to a filter) and as *control functions* to provide a variety in the synthesis parameters most perceptible by the listener.

In the analysis of natural sounds, some characteristics vary in an unpredictable way; their mean statistical properties are perceptibly more significant than their exact behavior. Hence, the addition of a random component to the deterministic functions controlling the synthesis parameters is often desirable.

In general, a combination of random processes is used because the temporal organization of the musical parameters often has a hierarchical aspect. It cannot be well described by a single random process, but rather by a combination of random processes evolving at different rates.

Linear Transformations

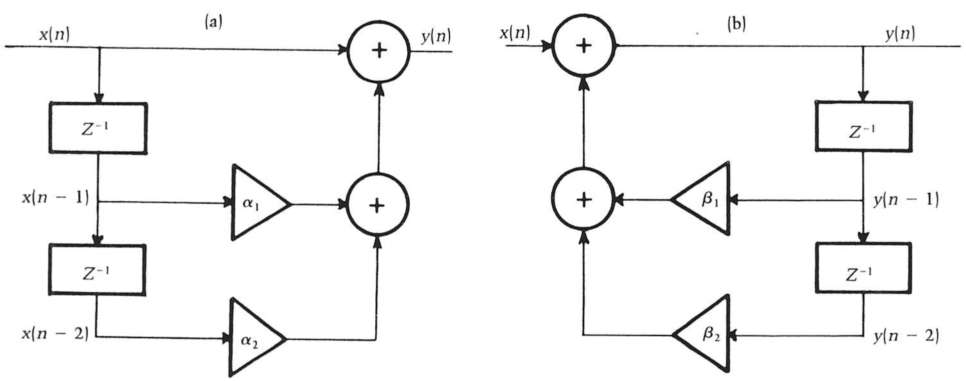
Let us now examine techniques for signal modification. A *transformation* is a set of rules and procedures transforming a signal called input to another signal called output. A transformation is linear if the superimposition principle is valid, that is, if the effect of the transformation caused by a two-signal addition is equal to the addition of the individual signal transformations applied separately. In particular, in a linear transformation a signal can be multiplied by a constant but not by another signal.

Digital filters are linear transformations that can be described by the following difference equation:

$$\sum_{K=0}^{N} a_K\, y(n-K) = \sum_{i=0}^{M} b_i\, x(n-i),$$

where a_K and b_i are the filter coefficients and $x(j)$ and $y(j)$ are the jth sample of the input and output signal. The value of the output sample is thus a linear combination of current instantaneous input with the preceding instant's input and output. When

Fig. 6. Finite-impulse-response (FIR) filter with two zeros described by the equation $y(n) = x(n) + \alpha_1 x(n-1) + \alpha_2 x(n-2)$ (a). Infinite-impulse-response (IIR) filter with two poles described by the equation $y(n) = x(n) + \beta_1 y(n-1) + \beta_2 y(n-2)$ (b).



the input is sinusoidal, the steady-state output is sinusoidal with the same frequency. The amplitude and phase of the frequencies are determined by the system. That is why this transformation is called a filter.

Subtractive Synthesis

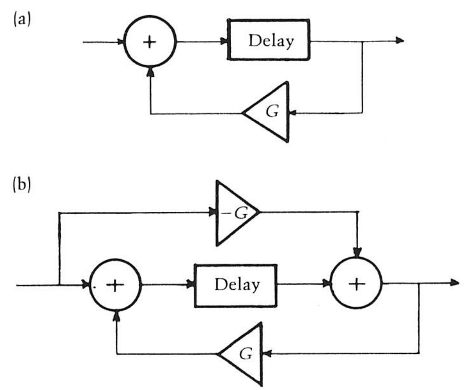
Sound produced by filtering a complex waveform is called, sometimes inappropriately, *subtractive synthesis*. First, a periodic or aleatoric signal rich in harmonics is generated by the previously examined techniques or others. This signal must contain energy in all frequencies required in the output sound. Second, one or more filters are used to alter selectively the specific frequency components. The undesired components are attenuated (subtracted) and others are eventually amplified. When the filter coefficients change, the frequency response changes, too. Thus, it is possible to vary characteristics of the output sound.

In modular diagrams, filters are usually represented by rectangles and the difference equation or the transfer function is given as a label near the rectangle. Two examples of simple digital filters, showing their internal structure, are shown in Fig. 6. The first filter (Fig. 6a) has a finite-impulse response (FIR). This structure is useful to produce transmission zeros: that is, it can nullify some frequencies that depend on α_1, α_2 values and on the sampling rate. The second filter (Fig. 6b) is recursive, or has an infinite-impulse response (IIR). Feedback in the structure amplifies certain frequencies, that is, produces transmission poles. When used as bandpass filter, in general terms, the coefficient β_1 controls the center frequency and the coefficient β_2 the bandwidth.

One of the most attractive aspects of digital filtering is that it is analogous to the functioning of many acoustic musical instruments. Indeed, instrument physics can be used as a model for synthesis. For example, in the brasses and woodwind instruments, the lips or vibrating reed generate a periodic signal rich in harmonics. The various cavities and the shape of the instrument act as resonators, enhancing some spectral components and attenuating others. In the human voice, the excitation signals are periodic pulses of the glottis (in the case of voiced sounds) or white noise (in the case of unvoiced sounds—for example, the consonants *s* and *z*). The throat, the mouth, and the nose are the filtering cavities, and their dimensions vary in time. Their great variability makes the human voice the most rich and interesting musical instrument.

Today, subtractive synthesis is the standard means of speech synthesis. An analysis procedure, called linear predictive coding (LPC), allows us to obtain

Fig. 7. Elementary filters used in reverberators. Comb filters (a). All-pass filter (b).



the pitch and the coefficients of a recursive (poles only) filter (see Cann's [1979–1980] tutorial and Moorer's paper [1979*a*]). These data can be utilized to synthesize the sound directly or following modification. For example, speech can be accelerated or slowed down, and pitch can be varied. An instrument or orchestral sound can be used as input to the filter, producing the effect of a "talking orchestra."

Interesting possibilities for *musique concrète* sound processing arise. Not only simple filtering of sounds is possible, but the modification of their most intrinsic characteristics is also made possible by varying the parameters of the deduced sound-production model.

Generally, LPC is relatively difficult to use. Intuitively, the filter characteristics depend on the position of the zeros and the poles in the transfer function. These characteristics are affected in a complex and nonintuitive way by the filter coefficients. In some simple cases, approximate formulas give the coefficients as functions of significant parameters, that is, center frequency and bandwidth, or cutoff frequency and slope. The filters can be used in series or in parallel. In the most complex cases, a precise analysis is obtained by using specific programs for digital filter design and analysis. Such digital filters can be very stable and precise, but only at the cost of a large amount of calculation. Simple linear digital networks can also be used as oscillators (Tempelaars 1982) by applying a pulse sequence to the input and choosing an impulse response equal to the signal function to be generated.

Reverberation

One application of digital filters is sound reverberation. An acoustic environment can be simulated by distributing sound among different loudspeakers and by adjusting the ratio between direct and reverberated sound (Chowning 1971). Most of the studio reverberators sold today use digital technology.

The two elementary filters used in reverberation are shown in Fig. 7. The first filter is called a *comb filter*; in it, the signal is delayed a certain number of samples, attenuated, and added to the input. An exponentially decaying, repeated echo is so obtained. The frequency response is characterized by equispaced peaks—hence this filter's name. The peaks' amplitude increases as G approaches 1.

The second filter is called an *all-pass filter*, since the frequency response is flat and there is only a phase shift. The input signal is attenuated and subtracted from the delayed signal so that the feedback effect is compensated and the echoes are maintained. The all-pass property is valid only in the steady state with stationary sounds, not in transient states. Thus, it has a well-defined sound quality that a skilled listener can easily distinguish.

Reverberators are built combining some of these filters (Moorer 1979*b*). Distinguishable signal repetitions should not occur in them, since the reverberated result should consist of a diffused sound. The delay time of each elementary filter has to be chosen very carefully. Sometimes a nonrecursive echo generator is added to produce the first aperiodic echoes, which are the main perceptual determinants of the characteristics of the room.

Nonlinear Techniques

In addition to linear transformations, which are used in other fields and have a rather developed theory, nonlinear transformations are used more and more commonly in musical applications. They derive mainly from electrical communication theory, and they have proved to be promising and effective. One use of nonlinear synthesis is in the large amount of computer music generated by frequency modulation (FM) synthesis (Chowning 1973).

In the classic case, nonlinear techniques use simple sinusoids as input signals. The output is composed of many sinusoids, whose frequency and amplitude depend mostly on the input ones.

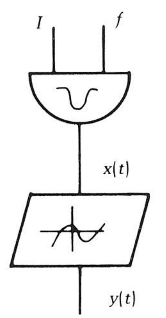
Two main types of nonlinear techniques can be distinguished, waveshaping and modulation. In *waveshaping*, one input is shaped by a function depending only on the input value in that instant. In *modulation* (with two or more inputs), a simple parameter of one signal, called the carrier, is varied according to the behavior of another signal, called the modulator. In electrical communications (e.g., radio) the spectra of the signals are clearly distinguished and therefore easily separable. The originality in computer music application is the utilization of signals in the same frequency range. Thus, the two signals interact in a complex way, and simple input variation affects all the resultant components.

Often, the input amplitudes are varied by multiplying them by a constant or time-dependent parameter I, called the modulation index. Thus, acting only on one parameter, the sound characteristics are substantially varied. Dynamic and variable spectra are easily obtainable. In additive synthesis, similar variations require a much larger amount of data.

Waveshaping

A linear filter can change the amplitude and phase of a sinusoid, but not its waveform, whereas the aim of waveshaping is to change the waveform. The distortion of a signal heard from a nonlinear amplifier is common. The output from a nonlinear amplifier of a sinusoidal signal is a signal with the same period, but with a different waveform. The various harmonics are present, and their amplitude depends on the input and on the distortion. In stereo systems, these distortions are usually avoided,

Fig. 8. Waveshaping.



while waveshaping (Arfib 1979; Le Brun 1979; Roads 1979) exploits them to generate periodic sounds, rich in harmonics, from a simple sinusoid.

The function $F(x)$, describing distortion, is called the shaping function, and it associates with each input value the corresponding output value independent of time. If the input is $x(t) = \cos(2\pi ft)$, the output is

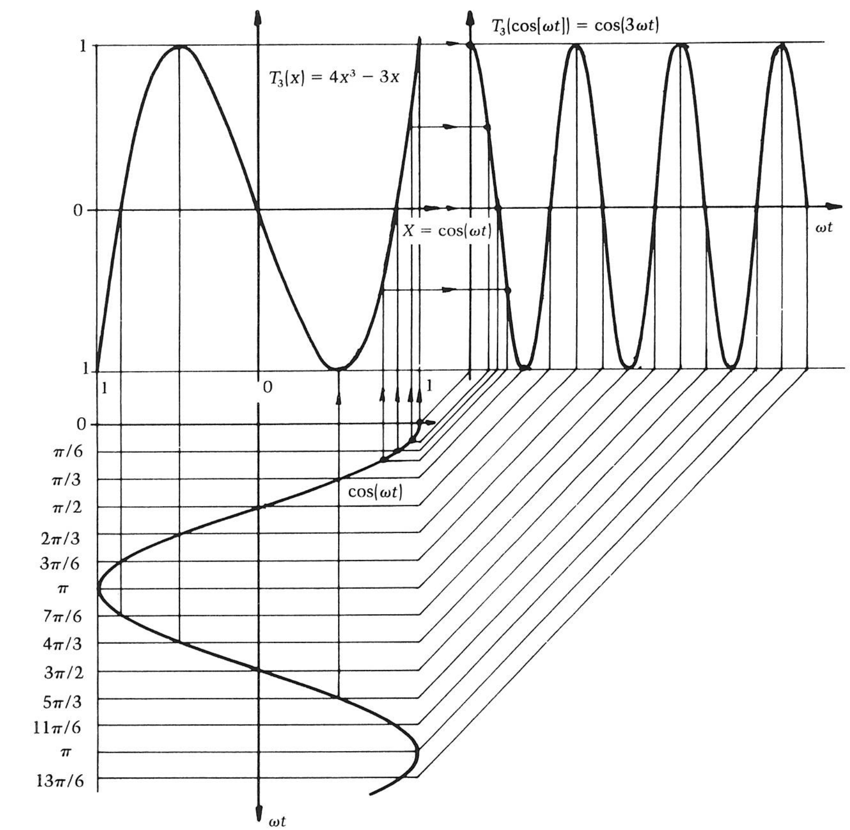
$$s(t) = F(x(t)) = F(\cos(2\pi ft)).$$

In analog synthesis, it is difficult to have an amplifier with a precise and variable distortion characteristic. In digital synthesis, this technique is extremely easy to implement (Fig. 8). As in the case of the oscillator, the shaping function can be previously computed and stored in a table. All that is necessary is to look up the proper value from the table.

Generally, if $F(x) = F_1(x) + F_2(x)$, the distortion produced by F is equal to the sum of those produced by F_1 and F_2 separately. Usually, the shaping produces infinite harmonics. But when a polynomial of degree N is chosen as shaping function, only the first N harmonics are present. Thus, foldover is easily avoided. Arfib and Le Brun deal extensively with the mathematical relations among the coefficients d_i of the shaping polynomial and the amplitudes h_i of harmonics generated when the amplitude I of the cosinusoidal input varies.

The shaping function, producing the jth harmonic, is the Chebychev polynomial $T_j(x)$ of degree j (Fig. 9). Thus, to obtain the various harmonics of

Fig. 9. Chebychev polynomial of degree K used as shaping function produces only the Kth harmonic. In the figure, K = 3.



amplitude h_i, it is sufficient to add the correspondent Chebychev polynomials, each multiplied by h_i:

$$F(x) = \sum_{i=0}^{N} h_i T_i(x) = \sum_{i=0}^{N} d_i x^i.$$

From these relations, it follows that the even harmonics are completely separate from odd ones. An even polynomial coefficient affects only the even harmonics. In the same way, odd coefficients affect only the odd harmonics. For example, the coefficient of x^7 affects only the first, third, fifth and seventh harmonics. Moreover, an even (or odd) harmonic of order n is affected only by the even (or odd) coefficients of order n and greater than n. For example, the seventh harmonic is affected by the odd coefficients from the seventh up to the degree of the polynomial.

When the input amplitude I varies, the distortion and the output spectrum vary. This is similar to an expansion or contraction of the function, since greater or smaller range of the function is employed. From a mathematical point of view, the amplitude variation corresponds to the multiplication of each polynomial coefficient d_i by I^i. The amplitudes of the even or odd harmonics depend on I according to the even (or odd) polynomials, which contain the terms from the harmonic order up to the polynomial degree.

If the spectrum is rather smooth, the number of significant harmonics increases with the index. Thus, a typical characteristic of real instruments is reproduced, in that amplitude and spectrum are correlated. The amplitude and loudness of the output vary with the input amplitude. In simple cases, this effect can be compensated for by multiplying the output by a suitable normalization function. But in musical applications, the amplitude of the signal is rarely constant, and it is multiplied by an envelope. Normalization can be avoided by combining it with the amplitude envelope in experimental or intuitive ways after considering the normalization function.

It is also advisable to choose the even (or odd) polynomial coefficients with alternating signs, that is, according to the following model: + + − − + + − −. It is also advisable that the h_i amplitude not decrease abruptly, sharply limiting the band. Otherwise, a spectrum would result that varied very irregularly with I.

Dynamic spectral behavior cannot be easily anticipated from the coefficients or from the static spectrum. Moreover, the same (absolute-value) spectrum can be produced by many polynomials with different dynamic behaviors (Forin 1982). With waveshaping, listening and graphic considerations have more relevance than purely mathematical formulations.

Another dynamic variation of waveshaping that is easy to implement occurs when a constant is added to the input; the shaping function shifts horizontally. Even in this case, the spectrum varies. The signal is periodic, with the same number of harmonics. But in this case, the harmonic behavior depends on both the even and the odd coefficients.

Generalizations of waveshaping technique are possible. Reinhard (1981) studied the relations that produce the partials generated by the polynomial distortion of two cosine waves of frequency f_1 and f_2. All the components of frequency $|K f_1 \pm j f_2|$ with $|K + j| \leq N$, where N is the polynomial degree, are present.

Shaping functions that are not polynomial can be used if the spectra produced by them are almost band limited. Of particular interest is the use of trigonometric and exponential functions (Moorer 1977) and of those where the input also appears in the denominator (Winham and Steiglitz 1970; Moorer 1976; Lehmann and Brown 1976; De Poli 1981).

Due to the wide spectral variation induced by only one parameter (amplitude or shift), waveshaping is particularly convenient in musical applications, especially in combination with multiplicative synthesis. Moreover, it is suitable for modeling the sound production of some acoustic instruments (Beauchamp 1979, 1982). There is a large and not intuitive problem in choosing the coefficients, however, and further research is required.

Multiplicative Synthesis (Ring Modulation)

The simplest nonlinear transformation consists of the multiplication of two signals. In analog synthesizers, it is called ring modulation (RM). Sometimes it is also called amplitude modulation (AM), but the two differ, especially in their realization.

With two inputs $x_1(t)$ and $x_2(t)$, the output is $s(t) = x_1(t) \cdot x_2(t)$. Obviously, when the inputs interchange, the result does not vary. The resulting spectrum is obtained from the convolution of the two signals' spectra. Usually, one of the two signals, called the carrier, is sinusoidal; the result is not too complex and noisy.

When x_1 is the sinusoidal carrier of frequency f_1, and x_2 (modulator) is sinusoidal with frequency f_2, from $\cos(\alpha) \cdot \cos(\beta) = \frac{1}{2}[\cos(\alpha + \beta) + \cos(\alpha - \beta)]$, the output consists of two sinusoidal partials of frequency $f_1 + f_2$ and $f_1 - f_2$. The phases of the output are also the sum and the difference of the phases of the two inputs. For example, if x_1 and x_2 frequencies are 400 Hz and 100 Hz, the output has two partials of frequency 500 Hz and 300 Hz.

Negative frequencies may occur, for example, when f_1 = 100 Hz and f_2 = 400 Hz. This often happens in modulations (foldunder) and can be explained by the trigonometric relation $\cos(\alpha) = \cos(-\alpha)$, from which $\cos(2\pi ft + \phi) = \cos(\pi|-f|t - \phi)$. The alteration of the frequency sign only changes the sign of the phase with respect to the cosine. In particular, a cosine signal is unaffected, while a sine wave changes its sign. In the interpretation of the results, only absolute frequency values have to be considered. Usually, the phase is not significant, as the ear is not terribly sensitive to it. But the phase has to be taken into account while summing the amplitude of components of identical frequencies.

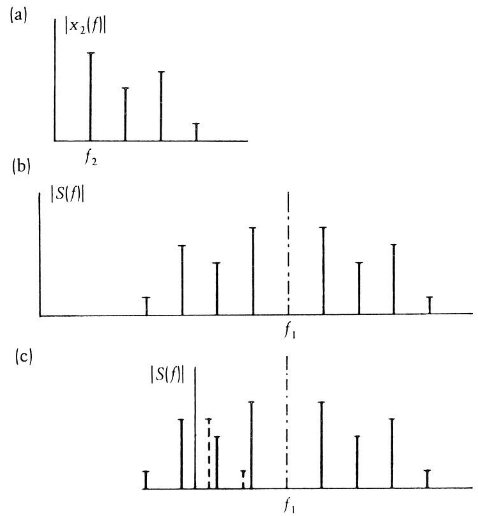
In multiplicative synthesis, usually x_2 is periodic with frequency f_2. The multiplication causes every harmonic spectral line of frequency $K \cdot f_2$ in the original signal to be replaced by two spectral lines (called sidebands) of frequency $f_1 + K f_2$ and $f_1 - K f_2$. The resulting spectrum has components of frequency $|f_1 \pm K f_2|$, where K is equal to the order of the different harmonics in x_2 (Fig. 10).

Thus two sidebands, symmetric with respect to the carrier, occur. When f_1 is less than the greatest frequency in x_2, then the negative frequencies fold around zero, as discussed above.

The possibility of shifting the spectrum is very intriguing in musical applications. From simple components, harmonic and inharmonic sounds can be created, and various harmonic relations among the partials can be established. If x_2 is a signal with spectrum X_2, the signal obtained from its multiplication with a sinusoid of frequency f_1 has two sidebands symmetric with respect to f_1 and shaped like X_2.

A periodic signal x_1 can be expanded in Fourier series. Each x_1 partial will have sidebands of amplitude proportional to its own. If f_1 is less than the bandwidth of x_2, then the sidebands overlap with eventual component superimposition. In this case, the phases have to be taken in account while summing. Dashow (1978, 1980) describes some generalization of this technique and employs the generated spectra for particular "harmonizations" of pitches specified by the composer.

Fig. 10. Multiplicative synthesis. Spectrum of a periodic signal X_2 with four harmonics (a). Resulting spectrum when d_2 is multiplied by a sinusoid of frequency f_1 greater than its bandwidth ($f_1 = 7f_2$) (b). Resulting spectrum when x_2 is multiplied by a sinusoid of frequency inferior to its bandwidth ($f_1 = 26f_2$) (c). The components deriving from the folding of negative frequencies are shown as dashed lines.



Amplitude Modulation

In RM, the carrier does not appear in the spectrum created by the product of a sinusoidal carrier with another signal, except when the modulator has a direct current (dc) component. In carrying out the modulation in AM (Fig. 11), the carrier is present in the output, with an amplitude independent of the sidebands. The formula for AM is as follows:

$$s(t) = x_1(t) \cdot (K + x_2(t)).$$

The result is RM with carrier added. When the carrier is sinusoidal and the modulator is periodic, the spectrum is composed of partials of frequency $|f_1 \pm K f_2|$, with $K = 0, 1, \ldots$. It is useful to distin-

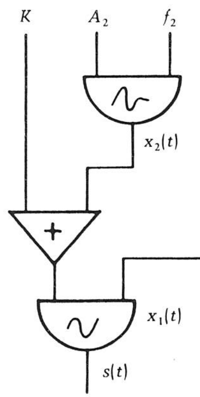


Fig. 11. Amplitude modulation.

guish between the two modulations because they have different realization schemes.

Spectra of Type $|f_1 \pm K f_2|$

The following considerations are valid for all spectra whose components are of type $|f_1 \pm K f_2|$, with $K = 0, 1, \ldots$. The spectrum is characterized by the ratio f_1/f_2. (This is often referred to as the carrier-to-modulator [c:m] ratio.) When this ratio is rational, it can be expressed as an irreducible fraction $f_1/f_2 = N_1/N_2$, with N_1 and N_2 as integers that are prime between themselves. In this case, the resulting sound is harmonic, since the various components are a multiple of a fundamental according to integer factors. The fundamental frequency is

$$f_0 = \frac{f_1}{N_1} = \frac{f_2}{N_2},$$

and the carrier coincides with the N_1th harmonic.

If $N_2 = 1$, all the harmonics are present and the sideband components coincide. If $N_2 = 2$, only odd harmonics are present and the sidebands superimpose. If $N_2 = 3$, the harmonics that are multiples of 3 are missing. The c:m ratio is also an index of the harmonicity of the spectrum. The sound is more "harmonious" intuitively when the N_1/N_2 ratio is simple and formally when the $N_1 \cdot N_2$ product is smaller.

The ratios can be grouped in families (Truax 1977). All ratios of the type $|f_1 \pm K f_2|/f_2$ can produce the same components that f_1/f_2 produces. Only the partial coinciding with the carrier (f_1) changes. For example, the ratios 2/3, 5/3, 1/3, 4/3, 7/3 and so on all belong to the same family. Only the harmonics that are multiples of 3 are missing (see $N_2 = 3$), and the carrier is respectively the second, fifth, first, fourth, seventh, and so on harmonic.

The ratio that distinguishes a family is defined in normal form when it is $\leq 1/2$. In the previous example, it is 1/3. Each family is characterized by a ratio in normal form. Similar spectra can be produced using ratios from the same family. Different spectra are obtained by sounds of different families.

When the f_1/f_2 ratio is irrational, the resulting sound is aperiodic and hence, inharmonic. Of particular interest is the case of an f_1/f_2 ratio approximating a simple value, that is,

$$f_1/f_2 = N_1/N_2 + \varepsilon.$$

Here the sound is no longer rigorously periodic. The fundamental frequency f_0 is still f_2/N_2, and the harmonics are shifted from their exact values by $\pm\varepsilon \cdot f_2$. When N_2 is equal to 1 or 2, the positive and negative components are not superimposed but beat with a frequency of $2\varepsilon \cdot f_2$. Hence, a small shift of the carrier does not change the pitch, even if it slightly spreads the partials and makes the sound more lively. But the same shift of the modulating frequency f_2 changes the sound's pitch.

Frequency and Phase Modulation

Another type of modulation, suggested by Chowning (1973), has become one of the most widely used synthesis techniques. In general, it consists of *angle modulation* and it can be realized both as phase modulation (ϕM) or as FM. This technique does not derive from models of production of physical sounds, but only from the mathematical properties of a formula. It has some of the advantages of waveshaping and RM, and it avoids some of their drawbacks.

The technique consists of the modulation of the instantaneous phase or frequency of a sinusoidal carrier according to the behavior of another signal (modulator), which is usually sinusoidal. It can be expressed as follows:

$$s(t) = \sin(2\pi f_c t + I \sin(2\pi f_m t)) = \qquad (1)$$
$$= \sum_{K=-\infty}^{+\infty} J_K (I) \sin[2\pi (f_c + K f_m) t].$$

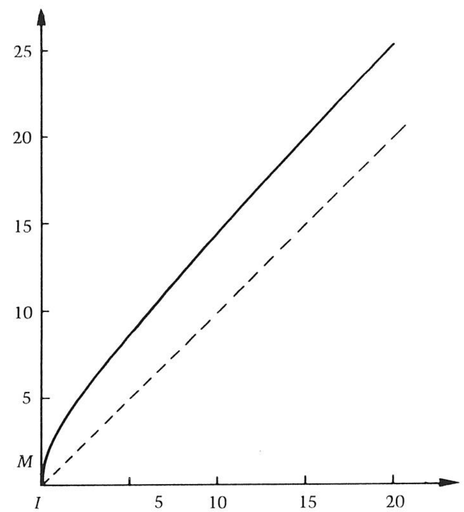
The resulting spectrum is of the type $|f_c \pm K f_m|$. All the spectral considerations discussed previously are applicable, particularly those regarding negative frequency, foldunder, f_c/f_m ratios, and harmonic and inharmonic sounds.

The amplitude of each Kth side component of the FM technique is given by the Bessel function of Kth order computed in I. To plot the spectrum, a table of Bessel functions has to be referenced to obtain the amplitudes of the carrier and of the side frequencies in the upper sideband. The odd-order side frequencies in the lower sideband have signs opposite to those in the upper one, and the even-order side frequencies have the same sign. The negative frequencies, being sine waves, are folded, changing the sign. When superimposition occurs, the amplitudes are added algebraically.

When I (called the modulation index) varies, the amplitude of each component varies as well. Thus, dynamic spectra can be obtained simply by varying this index. Each component varies its amplitude by following the corresponding Bessel function. A Bessel function can be asymptotically approximated by a damped sinuosoid. So when the index varies, some components increase and others decrease, all without sharp variations.

In Eq. (1), the sum includes infinite terms, so theoretically the signal bandwidth is not limited. But, practically, it is limited. In the Bessel function's behavior, only a few low-order functions are significant for small index values. When the index increases, the number and the order of the significant functions increase. For a given index, the side amplitudes oscillate with gradually increasing amplitude and slowly increasing period all the way from the origin to a maximum, and then decrease rapidly toward zero. The maximum is reached for a value slightly below the modulation index.

Fig. 12. The number of significant sidebands in FM.



Usually, in the bandwidth definition of the FM signal, all side frequencies of amplitude greater than 1/100 of the nonmodulated signal are considered. The number M of the significant sidebands is

$$M = I + 2.4 \cdot I^{0.27}.$$

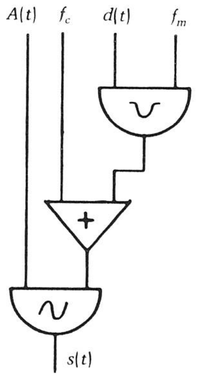
(See Fig. 12.) Often, as a rule of thumb, it is roughly considered as

$$M = I + 1.$$

In Eq. (1), the sum can be performed for K from $-M$ to $+M$. For a harmonic sound, that is, when the ratio $f_c/f_m = N_1/N_2$ is simple, the maximum order of significant harmonics is $N_1 + M \cdot N_2$.

For wide index variations, the sounds produced are characteristic of the FM technique. A typical timbre of FM sound is easily recognizable and thus well defined. This does not happen for small index variations or for compound carriers or modulators. Frequency modulation synthesis has another prop-

Fig. 13. Frequency modulation.



erty that is very important in musical applications: the maximum amplitude and the signal power do not vary with the index I. Unlike the situation in waveshaping, normalization of the output is not necessary.

Let us now examine the difference between ϕM and FM. Phase modulation is defined as follows:

$$s(t) = \sin(2M f_c t + \phi(t)),$$

and it corresponds to Eq. (1) if the modulating signal is $\phi(t) = I \sin(2\pi f_m t)$.

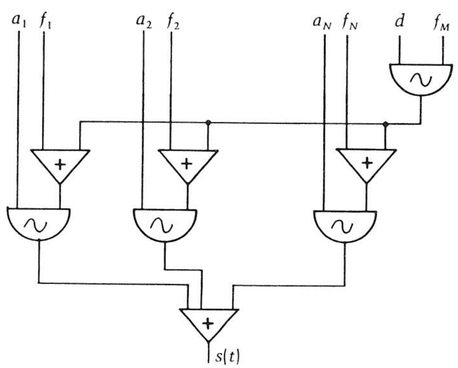
Frequency modulation occurs when the instantaneous frequency varies around the carrier value according to the behavior of the modulating wave. For a signal $s(t) = \sin(\psi(t))$, the instantaneous frequency is $f_i = (1/2\pi)(d\psi(t)/dt)$. Thus, the instantaneous frequency of the signal in Eq. (1) is as follows:

$$f_i = f_c + I f_m \cos(2\pi f_m t).$$

The frequency varies around f_c with a maximum deviation $d = I \cdot f_m$. Thus, with a modulating wave $I \cdot f_m \cos(2\pi f_m t)$, an FM equivalent to ϕM is obtained. Both phase and frequency modulations are special cases of angle modulation.

In sound synthesis programs, frequency-driven oscillators are provided. The integration involved in calculating the instantaneous phase is therefore

Fig. 14. Frequency modulation with N carriers modulated by the same oscillator.



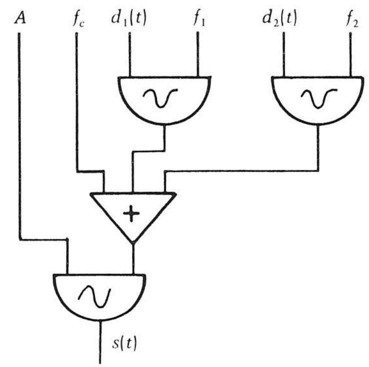
computed automatically. Frequency modulation is normally implemented as in Fig. 13. A change of the phase between the carrier and the modulating wave in Eq. (1) only changes the reciprocal phase of the partials. If components superimpose, their total amplitude changes, and a direct-current component may appear. The next sections examine some useful extensions of the basic algorithm.

Nonsinusoidal Carrier

Here we consider a periodic nonsinusoidal carrier. The result of its modulation is the modulation of each of its harmonics by the same wave. Sidebands of amplitude proportional to each harmonic will be present around the carrier. The result is a spectrum with components of frequency $|n \cdot f_c \pm K \cdot f_m|$, with $K = 0, \ldots, M$ and $n = 1, \ldots, N$, when N is the number of significant harmonics. The maximum frequency present is $N \cdot f_c + M \cdot f_m$. In general, there may be various independent carriers modulated by the same wave (Fig. 14) or by different modulating signals. This is like additive synthesis, only instead of sinusoidal addends, more complex addends are used. For example, harmonic sounds can be generated by controlling the various spectral ranges with a few significant and independent parameters. Sounds of the same "family" are possible.

The frequency of each carrier determines the

Fig. 15. Frequency modulation with two modulators.



location of the formant position, the amplitude determines its energy, and the modulation index specifies its bandwidth. Chowning (1981) demonstrated these facilities in synthesis of the singing voice of a soprano.

Compound Modulation

Let us examine the case of a modulation composed of two sinusoids (Fig. 15), each with its own modulation index, applied to a sinusoidal carrier. The formula for two-sine-wave ϕM (Le Brun 1977) is as follows:

$$s(t) = \sin((2\pi f_c t + I_1 \sin[2\pi f_1 t] + I_2 \sin[2\pi f_2 t]))$$

$$= \Sigma_K \Sigma_n J_K(I_1) \cdot J_n(I_2) \cdot \sin((2\pi(f_c + K f_1 + n f_2)t)).$$

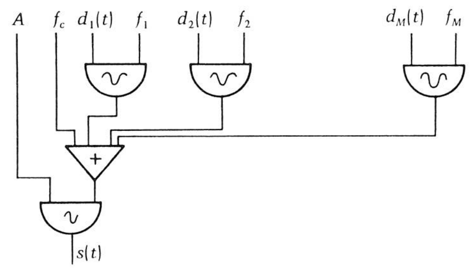
The same result can be obtained with FM using as modulating signal the following expression:

$$I_1 f_1 \cos(2\pi f_1 t) + I_2 f_2 \cos(2\pi f_2 t).$$

The resulting spectrum is much more complex than in the one-modulator case. All the components of frequency $|f_c \pm K f_1 \pm n f_2|$ are present, and their amplitude is $J_K(I_1) \cdot J_n(I_2)$.

To interpret the effect, let us consider $f_1 > f_2$. If only f_1 were present, the resulting spectrum would have a certain number of components of amplitude $J_K(I_1)$ and frequency $f_c \pm K f_1$. When the modulator

Fig. 16. Frequency modulation with N modulators.



f_2 is applied, these components become carriers, with sidebands produced by f_2. The resulting bandwidth is approximately equal to the sum of the two bandwidths.

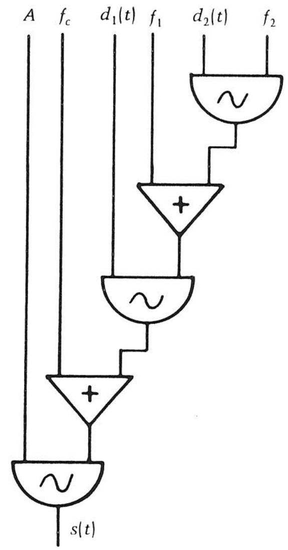
If the frequencies have simple ratios, the spectrum is of the type $|f_c \pm K f_m|$, where now f_m is the greatest common divisor of f_1 and f_2. For example, with $f_c = 700$ Hz, $f_1 = 300$ Hz, and $f_2 = 200$ Hz, the components are $|700 \pm K \cdot 100|$. Thus, by choosing f_1 and f_2 multiples of f_m, sounds belonging to the same family as a simple modulation, but with a more complex spectral structure, can be generated. In general, if the modulating signal is composed of N sinusoids (Fig. 16), the following relations hold:

$$s(t) = \sin\left(\left|2\pi f_c t + \sum_{1}^{N} I \sin[2\pi f_s t]\right|\right)$$

$$= \Sigma_{K_s} \prod_{s=1}^{N} J_{K_s}(I_s) \cdot \sin|2\pi(f_c t + \Sigma K_s f_s)t|.$$

Thus, all the components of frequency $|f_c \pm K_1 f_1 \pm \ldots \pm K_N f_N|$, with amplitudes given by the product of N Bessel functions, are obtained. A very complex spectrum results. If the relations among the frequencies f_s are simple, that is, if the modulating wave is periodic, then the spectrum is of the type $|f_c \pm K f_m|$, where f_m is the greatest common divisor among the modulating components. Otherwise, the sonorities are definitely inharmonic and particularly noisy for high indexes.

Fig. 17. Nested FM.



Nested or Complex Modulation

Let us examine the case of a sinusoidal modulator that is phase modulated by another sinusoid. The signal is defined as follows:

$$s(t) = \sin(|2\pi f_c t + I \sin[2\pi f_1 t + I_2 \sin\{2\pi f_2 t\}]|)$$
$$= \Sigma_K J_K(I_1)\sin(|2\pi[f_c + K f_1]t + K I_2 \sin[2\pi f_2 t]|)$$
$$= \Sigma_K J_K(I_1) \cdot J_n(KI_2)\sin(2\pi[f_c + K_1 f_1 + n f_2]t).$$

The result can be interpreted as if each partial produced by the modulator f_1 were modulated in its turn by f_2 with modulation index $K I_2$. Thus, all the partials of frequency $|f_c \pm K f_1 \pm n f_2|$, with approximately $0 \leq K \leq I_1$, $0 \leq n \leq I_1 \cdot I_2$, are present. The maximum frequency is $f_c + I_1(f_1 + I_2 f_2)$.

The structure of the spectrum is similar to that produced by the two-sinusoid modulation, but with a larger bandwidth. Even where f_m is the greatest common divisor between f_1 and f_2, the spectrum is of the type $|f_c \pm K f_m|$.

In the equivalent realization by FM (Fig. 17), the spectrum is of the same type, but with slightly different amplitudes. A direct-current component in the resulting modulating wave added to the carrier is avoided by choosing a sine wave modulated by a cosine wave.

This technique is made more interesting by an algorithm suggested by Justice (1979), which enables an analysis of a sound according to this model, with the frequency and the index behavior of two or more nested modulators being deducible.

Other Two-Input, Nonlinear Transformations

Mitsuhashi (1980) proposed a more complex two-input, nonlinear transformation, in which the instantaneous phase and amplitude of an approximately sinusoidal signal are simultaneously varied. In another paper, Mitsuhashi (1982c) generalized this technique while discussing some criteria in choosing the two-input, nonlinear function and suggesting two examples. The function is time independent, bidimensional, and considered periodic outside the definition field. Thus, it can be implemented with a two-dimensional table, with analogy to an oscillator. This technique appears very interesting, even if it seems to be difficult to find a simple expression that bounds significant parameters of the resulting spectrum to the input and function characteristics. Another promising modulation technique is linear sweep synthesis, recently suggested by Rozenberg (1982).

Conclusion

As a consequence of progress in digital hardware and software, the initial antithesis between computing efficiency and timbral richness is lessening. Digital sound quality largely depends on the amount of introduced or controlled detail; excessive simplifications lead often to trivial results. It follows that increased computing power can generate more sophisticated results.

A musically interesting sound can be obtained in two ways. The first consists of the utilization of more complex techniques or of the combination of many of the techniques described here. Many linear and nonlinear transformations are possible. Most of the parameters do not have to be constant and can be varied by control functions and random signals.

The other synthesis approach consists of the superimposition of many simple sounds produced by basic techniques. The evolution of the individual sounds is not complex, and the richness of the result essentially depends on their combination. In this approach, the parameters of many elementary sounds have to be given. Specific programs are often used to define these parameters.

Sound evolution can be regulated either by control functions in the synthesis or by programs computing the parameters for the synthesis. In any case, many details of the sound have to be accurately controlled. Their coherence both within the sound and in the context of adjacent and simultaneous notes has to be guaranteed. The relations among sounds can be more easily highlighted when they are reflected not only in macroscopic parameter variations but also in internal structure.

The extensive utilization of a single technique reveals its peculiar characteristics. This derives from the finite repertoire of obtainable sounds and, more specifically, from the more easily producible dynamic variations associated with it. Thus, it is wise to use different techniques, the better to exploit their different potential. Moreover, the musician must study and experiment with a technique. This is essential in order to determine all its characteristics and to acquire a feeling for the parameter choices necessary for nontrivial use. In any case, a synthesis technique is simply a tool to produce sound, and sound is not yet music.

References

Arfib, D. 1979. "Digital Synthesis of Complex Spectra by Means of Multiplication of Nonlinear Distorted Sine Waves." *Journal of the Audio Engineering Society* 27(10):757–768.

Bass, S. C., and T. W. Goeddel. 1981. "The Efficient Digital Implementation of Subtractive Music Synthesis." *IEEE Micro* 1(3):24–37.

Beauchamp, J. W. 1975. "Analysis and Synthesis of Cornet Tones Using Non Linear Interharmonic Relationships." *Journal of the Audio Engineering Society* 23(10):778–795.

Beauchamp, J. W. 1979. "Brass Tone Synthesis by Spectrum Evolution Matching with Nonlinear Functions." *Computer Music Journal* 3(2):35–43.

Beauchamp, J. W. 1982. "Synthesis by Spectral Amplitude and 'Brightness' Matching of Analyzed Musical Instrumental Tones." *Journal of the Audio Engineering Society* 30(6):396–406.

Cann, R. 1979–1980. "An Analysis Synthesis Tutorial." Part 1, *Computer Music Journal* 3(3):6–11; Part 2, *Computer Music Journal* 3(4):9–13; Part 3, *Computer Music Journal* 4(1):36–42.

Cerruti, R., and G. Rodeghiero. 1983. "Comments on 'Musical' Sound Synthesis by Forward Differences." *Journal of the Audio Engineering Society* 31(6).

Charbonneau, G. 1981. "Three Types of Data Reduction." *Computer Music Journal* 5(2):10–19.

Chowning, J. M. 1971. "The Simulation of Moving Sound Sources." *Journal of the Audio Engineering Society* 19(1):2–6. (Reprinted in *Computer Music Journal* 1(3):48–52, 1977.)

Chowning, J. M. 1973. "The Synthesis of Complex Audio Spectra by Means of Frequency Modulation." *Journal of the Audio Engineering Society* 21(7):526–534. (Reprinted in *Computer Music Journal* 1(2):46–54, 1977.)

Chowning, J. M. 1981. "Computer Synthesis of the Singing Voice." In *Sound Generation in Winds Strings Computers*. Stockholm: *KTH Skriftserie* 29, pp. 4–13.

Dashow, J. 1978. "Three Methods for the Digital Synthesis of Chordal Structure with Non-Harmonic Partials." *Interface* 7(2/3):69–94.

Dashow, J. 1980. "Spectra as Chords." *Computer Music Journal* 4(1):43–52.

De Poli, G. 1981. "Sintesi di suoni mediante funzione distorcente con poli complessi coniugati." *Atti del IV Colloquio di Informatica Musicale* 1, Pisa, pp. 103–130.

De Poli, E., and G. De Poli. 1979. "Identificazione di parametri di un oscillatore VOSIM a partire da una descrizione spettrale." *Atti del III Colloquio di Informatica Musicale*, Pisa, pp. 161–177.

Forin, A. 1982. "Spettri dinamici prodotti mediante distorsione con polinomi equivalenti in un punto." *Bollettino LIMB* 2:62–76.

Grey, J. M., and J. A. Moorer. 1977. "Perceptual Evaluation of Synthesized Musical Instrument Tones." *Journal of the Acoustical Society of America* 62:434–462.

Justice, J. M. 1979. "Analytic Signal Processing in Music Computation." *IEEE Transactions on Acoustics, Speech and Signal Processing* (ASSP) 27(6):670–684.

Kaegi, W. 1973. "A Minimum Description of the Linguistic Sign Repertoire (part 1)." *Interface* 2:141–156.
Kaegi, W. 1974. "A Minimum Description of the Linguistic Sign Repertoire (part 2)." *Interface* 3:132–158.
Kaegi, W., and S. Tempelaars. 1978. "VOSIM—A New Sound Synthesis System." *Journal of the Audio Engineering Society* 26(6):418–424.
Le Brun, M. 1977. "A Derivation of the Spectrum of FM with Complex Modulating Wave." *Computer Music Journal* 1(4):51–52.
Le Brun, M. 1979. "Digital Waveshaping Synthesis." *Journal of the Audio Engineering Society* 27(4):250–265.
Lehmann, R., and F. Brown. 1976. "Synthèse rapide des sons musicaux." *Revue d'Acoustique* 38:211–215.
Lorrain, D. 1980. "A Panoply of Stochastic 'Cannons.'" *Computer Music Journal* 4(1):53–81.
Mailliard, R. 1976. "Les distorsions de Music V." *Cahiers récherche/musique* 3:207–246.
Mathews, M. V. 1969. *The Technology of Computer Music.* Cambridge, Massachusetts: MIT Press.
Mitsuhashi, Y. 1980. "Waveshape Parameter Modulation in Producing Complex Audio Spectra." *Journal of the Audio Engineering Society* 28(12):879–895.
Mitsuhashi, Y. 1982*a*. "Musical Sound Synthesis by Forward Differences." *Journal of the Audio Engineering Society* 30(1/2):2–9.
Mitsuhashi, Y. 1982*b*. "Piecewise Interpolation Technique for Audio Signal Synthesis." *Journal of the Audio Engineering Society* 30(4):192–202.
Mitsuhashi, Y. 1982*c*. "Audio Signal Synthesis by Functions of Two Variables." *Journal of the Audio Engineering Society* 30(10):701–706.
Moorer, J. A. 1976. "The Synthesis of Complex Audio Spectra by Means of Discrete Summation Formulae." *Journal of the Audio Engineering Society* 24(9):717–727.
Moorer, J. A. 1977. "Signal Processing Aspects of Computer Music: A Survey." *Proceedings of the IEEE* 65(8):1108–1132. (Reprinted in *Computer Music Journal* 1(1):4–37, 1977.)
Moorer, J. A. 1979*a*. "The Use of Linear Prediction of Speech in Computer Music Applications." *Journal of the Audio Engineering Society* 27(3):134–140.
Moorer, J. A. 1979*b*. "About This Reverberation Business." *Computer Music Journal* 3(2):13–28.
Reinhard, P. 1981. "Distorsione non lineare della somma di due cosinusoidi: analisi dello spettro tramite matrici." *Atti del IV Colloquio di Informatica Musicale,* Pisa, pp. 160–183.
Risset, J.-C. 1969. "An Introductory Catalog of Computer Synthesized Sounds." Murray Hill, New Jersey: Bell Laboratories.
Risset, J.-C., and M. V. Mathews. 1969. "Analysis of Musical Instrument Tones." *Physics Today* 22(2):23–30.
Roads, C. 1978. "Automated Granular Synthesis of Sounds." *Computer Music Journal* 2(2):61–62. Revised and updated version forthcoming in C. Roads and J. Strawn, eds., *Foundations of Computer Music.* Cambridge, Massachusetts: MIT Press.
Roads, C. 1979. "A Tutorial on Non-linear Distortion or Waveshaping Synthesis." *Computer Music Journal* 3(2):21–34.
Rodet, X. 1980. "Time Domain Formant Wave-Function Synthesis." In *Spoken Language Generation and Understanding*, ed. J. G. Simon. Dordrecht: D. Reidel.
Rozenberg, M. 1979. "Microcomputer-controlled Sound Processing Using Walsh Functions." *Computer Music Journal* 3(1):42–47.
Rozenberg, M. 1982. "Linear Sweep Synthesis." *Computer Music Journal* 6(3):65–71.
Tempelaars, S. 1976. "The VOSIM Signal Spectrum." *Interface* 6:81–86.
Tempelaars, S. 1982. "Linear Digital Oscillators." *Interface* 11(2):109–130.
Truax, B. 1977. "Organizational Techniques for C:M Ratios in Frequency Modulation." *Computer Music Journal* 1(4):39–45.
Winham, G., and K. Steiglitz. 1970. "Input Generators for Digital Sound Synthesis" (Part 2). *Journal of the Acoustical Society of America* 47(2):665–666.
Xenakis, I. 1971. *Formalized Music.* Bloomington, Indiana: Indiana University Press.

Gottfried Michael Koenig
Institute of Sonology
Plompetorengracht 14-16
3512 CD Utrecht, The Netherlands

Aesthetic Integration of Computer-Composed Scores

Introduction

I have been involved with the algorithmic description of composition processes for about 30 years. My interest started with the analysis of music composed in a strict style (like Bach's or Webern's), and compositions of my own in which the material was to be organized according to a plan. Instrumental music in this category consisted of several pieces for chamber ensemble and orchestra (1952–1955), piano pieces (1957), a wind quintet (1958–1959), and a string quartet (1959).

My interest in algorithmic description was reinforced by my involvement with electronic music, which, owing to new kinds of production techniques using electroacoustic devices, required more intensive planning. This period is documented by my *Klangfiguren I* and *II* (1955–1956), *Essay* (1957–1958), and *Terminus I* (1962).

Eventually my attention focused on algorithmic description itself when, instead of obeying compositional rules, I started using a computer to carry them out. This led to my computer programs Project 1 (first version 1964–1966), Project 2 (first version 1965–1969), and my sound synthesis program SSP (1972–1979).

I was able to make certain observations, which I should like to mention in this outline, from the following computer-aided compositions: *Version 1* for 14 instruments (1965–1966), *Version 3* for 9 instruments (1967), *Segments 1–7* for piano (1982), *Segments 99–105* for violin and piano (1982), and *3 ASKO Pieces* for chamber orchestra (1982). I composed all these pieces with Project 1.

At present I am working, albeit sporadically, on another program (Project 3), which I call a *mentor system*. With it, I will carry on the work started with Project 1 and Project 2.

Modes of Production, Division of Labor

While I was occupied with the algorithmic description of compositional processes, I became increasingly aware of how much the planning and execution of a compositional idea depend on, or are influenced by, the mode of production (solo instrument, ensemble, improvisation, score, tape, synthesizer). By *mode of production*, I mean the technology and social circumstances of sound production. Social circumstances include the function of music (e.g., mythical and ritual, to the glory of God, court splendor, bourgeois status symbol, social contacts, amusement, the reinforcement of the communal spirit, education, mood setting, background, documentation of individuality by musical means), but they also include the social division of labor (solo performance, group performance, ensemble, large orchestra, music making in the home, such automata as music boxes and street organs, electronic sound production, loudspeaker performance). Production modes also include technique in a stricter sense, in the form of the technique of playing instruments, improvised or "from the score"; the extension of the body by complicated tools, such as the piano or organ; the use of electric current (electric and electronic music) and digital circuitry (computer processes). Increasing change, both in technology and in the social division of labor for purposes of musical sound production, gives rise to the question of how a composer reacts to the different variants of the production mode. I have not done any sociological or psychological research on this subject, but as a composer I have felt a need to react "functionally." This means that players

This essay is dedicated to Otto Laske.

Computer Music Journal, Volume 7, No. 4,
Winter, 1983, 0148-9267/83/040027-06 $04.00/0,