Analog Circuits for Sound Animation*

BERNARD A. HUTCHINS, JR.

Electronotes, Ithaca, NY 14850, USA

One important practice of analog music synthesis is the achievement of a harmonically rich and dynamically changing sound, the so-called "fat sound." This sound is inherently more interesting to the ear than are standard periodic waveforms, and as a consequence, use of an animated sound instead of a periodic waveform enhances any final result, often quite dramatically. A standard practice is to achieve animation through the use of several voltage-controlled oscillators in parallel, with small tuning errors resulting in a rich, slow beating effect. This method is very expensive, however. This report is concerned with devices which use less hardware, and which produce effects similar to those achieved with parallel oscillators. In addition, new types of animation are discussed.

0 INTRODUCTION

The term "electronic sound" has been used over the years (in a largely pejorative manner) to describe many of the sounds from electronic music synthesizers. While it is quite natural to call sounds from electronic equipment "electronic," the fact that the term exists in the general listening public is an indication of the strength of the characteristics that manifest the sounds so termed. In general the term reflects the extreme regularity of the tones in their steady state, as compared to the subtle variations in the steady state of tones produced in traditional acoustic musical instruments. While the imitation of acoustically produced sounds need not be the goal of the electronic synthesizer designer, the ability to do so is important in demonstrating a certain level of sophistication in the electronics. It also indicates some level of isomorphism with acoustical instruments and their interaction with the listener's hearing mechanism.

A standard response to the problem of extreme regularity has been to try to break it up by employing random signals either in a mix or to control parameters of the regular signal. However, the "ear brain" is extremely good at dissociating the signal and noise components of such an effort, rather than recognizing a merger of the two into something new. As a result, simple schemes employing random elements do not work out in general.

Modulation methods of various types enjoy a long and successful tradition in electronic music. These are useful in enriching an otherwise perfectly periodic sound, either by adding nonharmonic components or by permitting the dynamic variation of harmonic components, and are thus more suited to the generation of unusual timbres and transients than steady-state sounds. Thus the achievement of an electronically generated sound useful for sustaining the ear's interest, even during long-duration tones, has been an elusive goal for the instrument designer. This rich tone is the so-called "fat sound" and is characterized by impressions of "ensemble" or "evolution." One successful (though uneconomical) method of achieving this has been to run a number of voltage-controlled oscillators (VCO) in parallel (see Fig. 1), allowing small tuning errors to achieve a beating ensemble effect.

In this paper various methods of achieving what we will call an "animated" sound are discussed. These include methods of achieving ensemble types of effects (through parallel processors) and methods of achieving evolution of the tonal characteristics, or both. After developing a basic model for the receiver of animated sounds, examples are given, and the possibility of achieving theoretical (or at least empirical) design procedures are discussed.

* Presented at the 67th Convention of the Audio Engineering Society, New York, 1980 October 31–November 3.

  0004-7554/81/110814-07$00.75

1 A WORKING MODEL FOR THE RECEPTION OF MUSICAL TONES

In order to understand animation, it is useful to employ here an oversimplified model of the receiver involved: the ear, the brain, and the so-called "ear brain." We shall assume that this receiver is presented with an acoustical stimulus, a proposed musical tone, but not in the context of any musical composition. We assume that the receiver attempts to deal with this in two ways: by classification and by analysis, not necessarily in that order, or independently. Classification will be thought of as a brain activity, while analysis will be considered an ear-brain activity (which includes frequency analysis, a known capability of the ear brain, but is not limited to that). Analysis (or partial analysis) is probably part of the full classification process, although we will consider these processes to be independent in their end results.

Our concern shall be with how these two supposed processes of analysis and classification are involved in the holding of the attention of the listener. We are thus looking for some inherent interest in the sound prior to its association into a clearly musical usage. It will turn out that classification may be completed, and attention may still be held, as long as analysis continues. Analysis should not terminate, nor should the analysis system be overloaded or underloaded, or the receiver may drop attention. It is the purpose of an animator to defy exact classification if possible, and to provide variations as the sound progresses so that analysis may not be terminated.

It will be useful here to look at an example from analogy. Suppose that you receive a package in the mail. Its very appearance will draw your attention. If it is some curious object which you cannot recognize, your failure to classify it will hold your attention as long as you can continue to analyze it at a reasonable mental level. On the other hand, you may be easily able to classify it—a book, let us say. If it is a book, you will continue to analyze it by reading it. It will hold your attention best if it is consistent with your intellectual level (not overloading or underloading) and if it has a comfortable flow rate of information.

Using this type of model, we can better understand our processing of sounds for their inherent musical interest. The sound of a violin, for example, may be rapidly classified (it is a violin or a string sound), but analysis of the tone will not terminate. An electronic oscillator is also rapidly classified (it is an oscillator, an electronic sound), and analysis is also rapid in such a case due to the regularity of the waveform, so interest is dropped. White noise is an interesting case, because analysis is not completed, just overloaded, while at the same time classification is completed (it is noise), so interest is dropped. Finally, we can use the model to understand the success of the use of short-duration tones in synthesized, rapidly moving musical passages. In these short tones, the rate of appearance is fast enough that the analysis process, and probably the classification one as well, do not terminate.

We thus look to animators mainly for long-duration tones, as a means of making exact classification more difficult, while at the same time making the analysis process continue indefinitely. At the same time we must make the level of animation such that analysis is neither overloaded nor underloaded, either with respect to complexity or with respect to progression in time.

2 ANIMATION THROUGH PARALLEL PROCESSORS

As stated above, one approach to a "fat sound," which has been used, is to employ a number of VCOs in parallel and sum the outputs (Fig. 1). Each of the VCOs is tuned to very nearly the same frequency, and all are connected to the same control voltage, so in theory they track exactly. However, small tuning errors are unavoidable, or the user may have displaced the manual controls slightly from a best effort to achieve a desired beating effect. The technique is often used, but is somewhat unpredictable since it relies on a generally unknown tuning error, and it is uneconomical, because quality VCOs are expensive. Also, paralleling VCOs makes them unavailable for other purposes in a synthesis patch.

By way of imitation of the VCO method, a general view of a parallel-processor-type animator emerges as suggested in Fig. 2. Here we have one VCO, and it is assumed that the processors $P_1, P_2, \ldots, P_n$ are simple

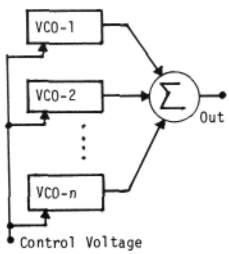

Fig. 1. Parallel VCOs.

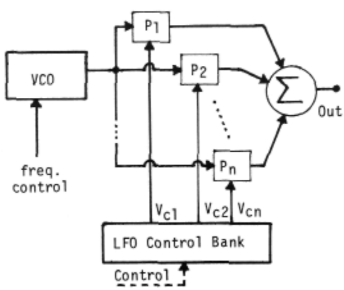

Fig. 2. General parallel-processing animator.

devices so that their circuitry, plus that of their portion of the low-frequency oscillator (LFO) control bank, is less involved than that of a corresponding VCO. The control bank is generally a set of LFOs, each of which has its own frequency in the general range of 1/100 to 1 Hz. In some cases it is desirable to have the entire bank controlled as well, and the dashed line below the bank in Fig. 2 is intended to suggest a voltage-control line to the bank. The number of processors varies from a minimum of perhaps three or four up to about eight.

The exact nature of the parallel-processor technique is not well understood, but one effect is fairly clear. If we use only one processor, the ear is very good at picking up its cyclic variation and/or following the change in the sound quality if a noncyclic control is used in place of the LFO. The ear is also good at following the patterns in two or three processors, but when we get to five or more, the patterns are pretty well disguised in a more general mix, resulting in a full animated sound, but still not sounding random. There is also evidence that if too many processors are used, the result becomes too homogeneous, and the apparent level of animation decreases. In one case (the MPWA circuit to be discussed shortly), a change from 8 to 16 processors resulted in a less satisfactory animated sound.

The use of phase or time delay as a processing means in a scheme as suggested in Fig. 2 is a more or less standard technique for achieving a "stringlike" ensemble effect. However, there are many other types of processor that are available. Here an algebraic phase shift unit for sawtooth waves and a moving filter bank in imitation of the vocal tract will be discussed individually.

2.1 Multiphase Waveform Animator (MPWA) [1]

The MPWA is based on an algebraic phase shift of a sawtooth wave by the scheme illustrated in Fig. 3, with Fig. 4 showing a practical circuit for achieving this shift. The input to the shifter is a sawtooth from a VCO, while the output is a sawtooth that is shifted in time from the original by an amount that depends on the control voltage V_c. Thus the shift is already under voltage control. If a number (typically eight) of these units are used in parallel in the manner of Fig. 2, a rich animated bright sound results. (Note that the amplifier A1 in Fig. 4 can be used to provide the inverted sawtooth for all the shifter units used, so additional shifters require only two operational amplifiers each.)

A typical LFO for the control bank is shown in Fig. 5, while a voltage-controlled LFO (VCLFO) using the CA3080 operational transconductance amplifier (OTA) as a control element is shown in Fig. 6. These same types of LFO circuits can be used with other types of animators. In fact it is reasonable to consider a bank of VCLFOs as a subunit, perhaps driving several types of animators, thus avoiding the need to produce LFOs for each different type. In addition, sums of the voltages from LFO outputs form complex patterns that can also be used for control. By using summing, the number of different (although correlated) control voltages which can be obtained from an LFO bank is greatly increased.

The waveform of the MPWA output is very rich, sounding at times like an orchestra tuning up, where different instruments and overtones are favored for brief moments. Substantial amplitude variations are also evident by ear and particularly when the waveform is viewed on a oscilloscope. The change of the relative amplitude of harmonics can be understood in terms of enhancement and cancellation of like frequencies as the

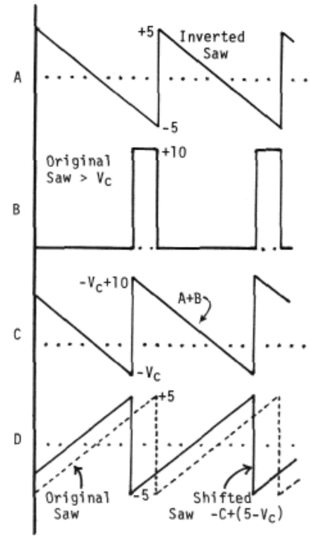

Fig. 3. Waveforms of algebraic sawtooth shifter.

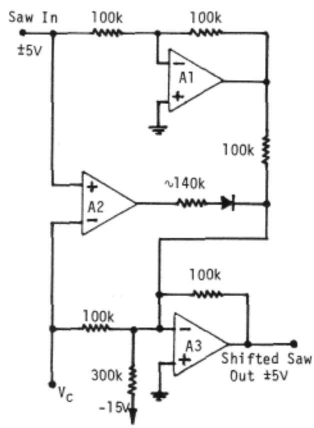

Fig. 4. Practical circuit for algebraic sawtooth shifter.

sawtooth waves shift relative to each other. While designed for sawtooth waveforms, the circuitry responds to other waveforms, even though it does not shift them without changing their shapes. In general the use of another input waveform causes the output to take on more of the character of the input, relative to the sawtooth case. For example, a sine-wave input is more mellow sounding than the sawtooth case.

2.2 Vocal-Effects Waveform Animator (VEWA) [2]

A second type of parallel-processing animator is the VEWA. The VEWA was designed with three (variable Q) band-pass filters (Fig. 7), each set to cover the range of the three principal formants of the human vocal tract for vowel sounds. However, the design principles can easily be applied to other types of moving filter banks. This type of animator causes the spectrum at the output to evolve in a manner more obvious than the MPWA. VCLFOs are suggested for this design. The animator can be used in several ways.

When the filters are set for low Q and the VCLFO bank is set for slow sweeping, a general type of subtle animation results, suitable for many input signals. With a somewhat higher Q, and slow sweep speed, white noise can be filtered for an excellent wind effect. If we raise the Q, use a moderate sweep speed, and feed a summed LFO control back to a driving VCO, vocal-like effects of changing pitch are achieved, the original design goal. Finally at the highest Q, a pulse input will cause the three filters to ring, and as they ring, their pitch will change slightly in response to their LFO control. This technique is an interesting special effect, and can also be used to add animated fine structure to a waveform. Fig. 8 shows a typical filter, which is a voltage-controlled state-variable design, with the upper CA3080 OTA serving as a voltage-controlled Q stage, while at the same time maintaining a constant peak response.

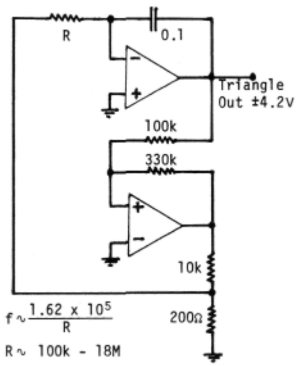

Fig. 5. Practical LFO.

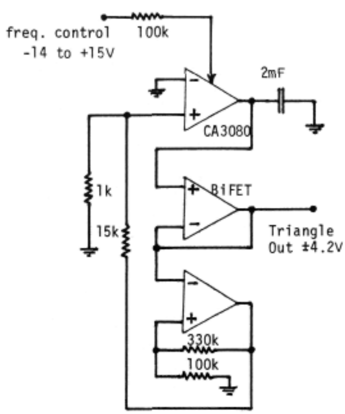

Fig. 6. Practical VCLFO.

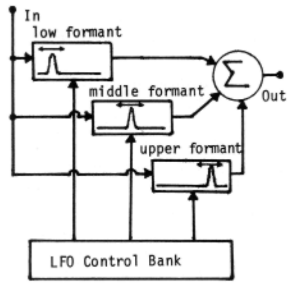

Fig. 7. Vocal effects waveform animator.

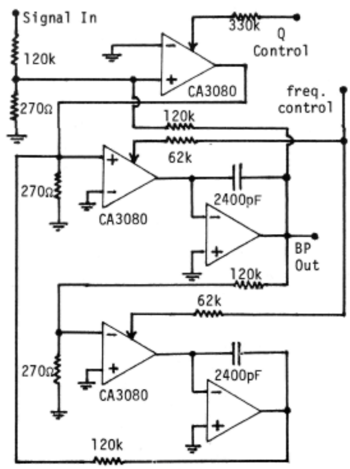

Fig. 8. Voltage-controlled band-pass filter.

3 ANIMATION THROUGH EVOLUTIONARY GENERATORS

The parallel-processing technique discussed is one method of achieving a constantly changing waveform by breaking it apart, altering it, and reassembling it in this case. Another process is that where a waveform is created in an evolving manner. This is a quite reasonable approach if we consider that in traditional acoustic instruments consecutive waveforms are similar, but have features that show evolving trends. Thus we might start with a certain waveform, select one feature of it, and gradually change this feature to another feature as different cycles are outputted. At the same time other features may be changing as well.

An example of this process is the binary tone wheel [3] which we can consider to be a shift register connected back on itself so that a binary sequence on it will be recycled over and over (Fig. 9). Each time the sequence cycles, one pitch period is in general completed. If we change the sequence by changing any one bit, the timbre of the tone will change. The remaining task is to make this transition smooth and subtle rather than abrupt.

A first thought might be that if we have many bits in the sequence, the change of any one bit would be minor, and that single bit changes could be made as needed and relatively frequently to achieve a pleasing level of evolution. It seems that this is not the case, however. Even in a sequence as long as 48 bits, the change of any one bit is abrupt and may have an annoying "ping" associated with it.

A second thought would then be not to make this sudden transition from one binary state to the other, but rather to go through a number of intermediate voltage steps. For example, if the binary levels are 0 and +15 V, we might have a +15-V to 0 transition by a series of levels like +14, +13, +12, . . . appearing during consecutive cycles until 0 is reached (Fig. 11). This technique does in fact work, although careful consideration must be given to the way in which it is carried out.

One approach is to connect additional shift registers to the output of the tone wheel in the manner of a transversal filter (Fig. 10). For example, the tone wheel could be 8 bits long, and a number of additional 8-bit shift registers could be added on. Shift registers such as the 74C164 were used in our experiments for these functions. Each of the transversal filter stages would have its taps weighted in the same manner, with the final output being the sum of all stages. (The taps should not be equally weighted the same for all eight taps on any stage, as this would just read out the dc level of the sequence, but other weightings, such as binary, could be used.) When the tone wheel is not changing, the same sequence will rapidly propagate into all filter stages, and the output is the same as it would be for just one stage (disregarding amplitude). However, if the tone wheel does change, this change will appear first in stage 1, then in stage 2, and so on. When the change first moves into stage 1, its own taps will feel the effect, but the corresponding taps of the other stages down the line will be unchanged from their original sequence values, so the change will be only one part in n, where n is the number of stages. The effect at the output increases until it is complete when the changed sequence reaches the last stage. For example, if there is only one tap per filter stage, the output will evolve as suggested in Fig. 11.

Test setups with as few as eight transversal filter stages are effective in removing the "ping," but are not otherwise very effective in removing the abrupt feeling of the change. Also, as many as 30 filter stages were tried and found to be lacking at least in some respects. While 30 stages certainly remove the "ping" and the abruptness, subjectively we would prefer that the transition time were on the same order as the occurrence time of transitions. That is, there should be some smoothing going on all the time. Either an old change should be working out, or a new one should be starting, on the average at least. Thus 30 stages meet this requirement only at low frequencies and at high rates of

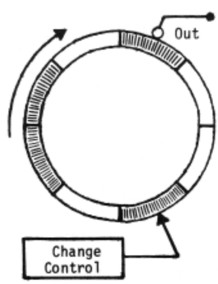

Fig. 9. 8-bit tone wheel.

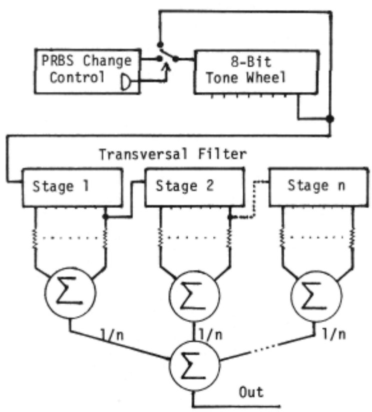

Fig. 10. Tone wheel with transversal filter.

change initiation. For a 100-Hz tone (a low frequency), changes should be initiated at about three per second (a fairly fast rate of animation). For higher frequencies and/or slower change initiation rates, some sensation of dead time is apparent, although not always bothersome. It is clear, however, that more stages than 30 would be useful if it were practical.

Fortunately, in the case where only one tap per filter stage is used, a simpler device equivalent to many filter stages is possible [4]. This device is suggested by Fig. 12, which shows what can be considered to be a fixed-sequence wheel which is read out by a rotating commutator. Between each wheel tap and the output commutator tap point a lag unit L has been placed. The lag unit might be an R-C low-pass or an integrator. In the center of the fixed-sequence wheel is the usual change unit, which may change any tap. If it does change one, the transition appears gradually (and under the control of an adjustable lag time) at the commutator tap point. Thus the outside commutating wheel can be expected to pass over this contact point many times during the transition. The result (if an integrator is used) will be similar to that in Fig. 11, and if an R-C low-pass is used, exponential level changes will be outputted. Thus the structure is the same as the single tap per stage transversal filter, except that a very large number of effective stages is obtained with vastly simpler hardware.

So far nothing has been said about the means by which changes are caused on the tone wheel. Two approaches are possible, one involving periodic forced changes, and a second involving probabilities of change at random intervals under the control of a random (pseudo-random) generator. The periodic change approach results in interesting sounds, but the regularity of changes restores an unnaturalness to the final waveform.

If a pseudorandom binary-sequence (PRBS) generator is employed, it can be used to initiate certain events according to the state of selected taps on the shift register body of the generator [5]. This process is achieved if it is realized that each stage of the sequence has, at any clocking interval, a 50:50 chance of being high (or low), just as a coin has a 50:50 chance of being a head (or tail). If we logically AND a number of stages on the generator, the probability of success changes from 1/2 to 1/4 to 1/8, and so on, dividing by two with each additional AND tap, until a relatively low probability of change, relative to the clock rate, is achieved. Thus the times at which changes begin can be controlled on the average over a wide range (by factors of 2) and can even be put under voltage control.

It might be supposed that this sort of approach would lead to sounds evolving in different directions giving constantly detectable new sounds. This is not the case. The evolutionary trend is clear, but most of the sequences that are possible on the tone wheel have somewhat similar spectra, so the effect is more one of a shifting harmonic emphasis, and less one of changing from one distinct timbre to another.

4 SOME ELEMENTS OF ANIMATOR DESIGN PROCEDURE

Each successful animator that is designed, constructed, and found to be useful provides a "data point" for our understanding of the general design principles involved. We can perhaps say that we need a continuous, gradual, overlapping series of changes occurring in the spectrum and the amplitude. This can be approached with parallel processors where roughly eight processors are used, each with its own control cycle from an LFO or from summed LFOs. Another method is to produce a waveform generator that evolves, with full evolution of any one trend occurring on a time scale roughly the same as that between initiations of transitions.

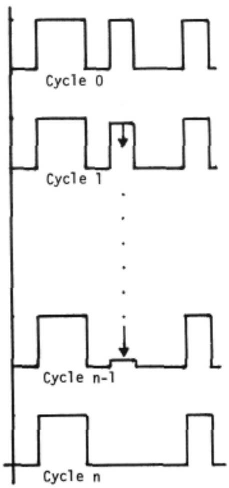

Fig. 11. Gradual change of sequence from *01101001* to 01100001.

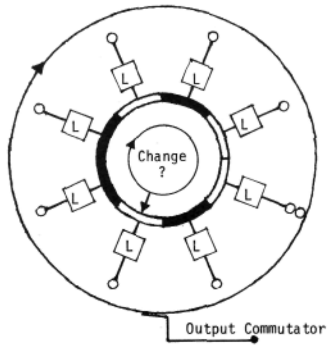

Fig. 12. Commutating tone wheel with lag units.

Unfortunately a theory as to why certain combinations and methods are successful is not at hand, and an empirical approach at present is the only one available to us. Some possible directions from which a theory could originate can be suggested, however. Information theory might be used to determine the rate of information flow from traditional instruments and from electronic devices, in which cases an optimum region of favorable transition rates might be determined, and animators could be adjusted in complexity and rate to match this region. Another possible area of interest begins with the findings of Voss and Clarke [6] that the fluctuations of the parameters which characterize sounds that interest the ear (speech and music) have a $1/f$ type spectrum. It would thus be of interest to isolate parameters from animator outputs and then perform a spectral analysis on them to see if they might be $1/f$.

5 CONCLUSIONS

The area of animator design is very rich with possibilities. Some generally simple and useful techniques have been identified and put into practice. In addition to the refinement of known techniques, numerous other methods have yet to be tested at all. In some cases testing by digital simulation would be helpful. A general theory of some sort of animator test also needs to be developed.

6 REFERENCES

[1] B. Hutchins, "A Sawtooth-Driven Multi-Phase Waveform Animator," *Electronotes*, vol. 10, pp. 3–11 (1978 Mar.).

[2] B. Hutchins, "A Vocal-Effects Waveform Animator," *Electronotes*, vol. 11, pp. 2–10 (1979 June).

[3] B. Hutchins, "A Pseudo-Random Tone Wheel Animator," *Electronotes*, vol. 11, pp. 3–21 (1979 Oct.).

[4] L. Ludwig, private communication, 1980 Aug.

[5] Y. Neuvo and W. H. Ku, "Analysis and Digital Realization of a Pseudorandom Gaussian and Impulsive Noise Generator," *IEEE Trans. Commun.*, vol. COM-23, pp. 849–858 (1975 Sept.).

[6] R. Voss and J. Clarke, "$1/f$ Noise in Music: Music from $1/f$ Noise," *J. Acoust. Soc. Am.*, vol. 63, p. 258 (1978 Jan.).

THE AUTHOR

Bernard A. Hutchins, Jr. received his B.S. degree in engineering physics from Cornell University in 1967. He has been associated with the School of Electrical Engineering, Cornell University, since 1974 with both research and teaching responsibilities. Mr. Hutchins is also the editor and publisher of *Electronotes: Newsletter of the Musical Engineering Group*. In this connection, he has pursued his interest in the analysis, synthesis, and processing of audio and musical sounds. He is a member of Audio Engineering Society.