- NJR
- Return to LWO

Electronics

xc f 571-17

9/17/83

Entertainment technology

Electronically synthesized music

Whether deemed an imitation or an art form of its own, electronic music is competing with conventional instruments

During the Academy Awards presentations in 1982, millions of TV viewers heard the theme music from the motion picture *Chariots of Fire*, as had millions of others in movie theaters around the country. Few, however, realized they were not listening to music produced by the conventional orchestra of woodwinds, strings, and brass. Rather the sound track for this film, which won the Academy Award that year, was produced mainly by electronic synthesizers, which created sounds closely resembling those produced by conventional instruments.

In the past few years, more and more of the music heard in TV commercials and movie sound tracks has come from electronic synthesis. Not only are producers and advertising companies turning to synthesizers as a way of producing music more cheaply and easily than with live orchestras, but composers of popular music are also using synthesizers to hear what finished compositions sound like, and performers are employing them in live acts.

At present synthesizers cannot accurately imitate all common musical instruments (pianos, for one, are still too difficult), nor can they mimic the full range of sound that a skilled performer can obtain from any single instrument. But they can achieve sufficient similarity for much popular and commercial music and can offer composers and sound editors capabilities not fully possible with more conventional approaches.

Waveforms produced by circuitry

Electronically synthesized music is any music with waveforms that have been generated by electronic circuitry rather than by the mechanical oscillators in conventional instruments. Electronic waveforms can be produced from scratch, built up out of individually defined sequences of pulses or through the combination of analog oscillator circuits, or they can be produced by modifying digitized sounds previously produced by conventional instruments or by natural events—thunder, bird songs, and so on.

Electrically generated music dates back three quarters of a century to the Telharmonium, an instrument developed by Thaddeus Cahill. Its sounds, generated by over 100 alternators, were controlled from a keyboard and fed to subscribers' speakers over leased telephone lines. However, synthesized music did not become at all common until the late 1960s, when the Moog synthesizer was developed. This synthesizer, using a combination of analog circuits such as oscillators and filters, could be manipulated by an array of knobs and switches. It produced sounds that did not resemble those of conventional instruments, and a number of pop bands made use of it.

But the Moog and similar analog synthesizers have serious limitations. Knob settings and module connections for the production of a given sound must be made manually, and, in general, the synthesizers can produce only one note at a time. They are thus far less flexible than conventional instruments.

These limitations were overcome, at least in part, with the development of digital synthesizers in the mid-1970s. Digital synthesizers, first marketed on a significant scale in 1979, allowed users to reprogram waveforms rapidly and to play more than one note at a time. Sufficiently complex waveforms allowed the mimicking of a number of conventional instruments.

Eric J. Lerner Contributing Editor

Limited supply at high prices

At present, two small companies have split up most of the digital-synthesizer market between them: New England Digital Corp. of White River Junction, Vt., with its Synclavier II (which is a digital-analog hybrid) and Fairlight Instruments Pty. Ltd. of Sydney, Australia, with its Fairlight Computer Musical Instrument (CMI). The market remains limited as yet—only 200 Synclaviers and 80 CMIs have been sold worldwide, but Synclavier reports a 30-percent annual growth rate in sales. One factor holding back sales is price: the Fairlight CMI sells for $27 750 and the Synclavier II for $14 150. Though these synthesizers are currently the most popular for professional performances, there are also many less expensive synthesizers with progressively more limited flexibilities, extending down to hobbyist systems costing several hundred dollars.

The new digital synthesizers are having their greatest impact in music for TV commercials and TV and movie sound tracks. In addition to the *Chariots of Fire* sound track, digital synthesizers have been used in the sound tracks for the TV movie *Word of Honor*, the series *The Power of Matthew Starr*, and the movies *The Wall* and *Cat People*. Many TV commercials now use synthesizers for all or part of the sound track, including those for Coca Cola, Chevrolet, DuPont, and Rockwell International.

The advantages of synthesized sound over conventional instruments for these applications are precision, convenience, and cost, according to consultants who use the devices. In TV commercials, for example, music timing can be matched perfectly with the image, down to an individual frame if necessary. Every aspect of the sound can be controlled directly by the composer, and the composer can try out different arrangements in quick succession without wasting valuable rehearsal time with a live band or orchestra.

The substitution of instruments, for example, can generally be accomplished with a few minutes of reprogramming. In addition, a single performer, often the composer, can play a variety of simulated instruments by reducing them all to a single keyboard. Simulated violins, trombones, and clarinets can all be played in sequence by one person on a synthesizer and then combined to produce the sound of a full orchestra.

Some of the same features have also attracted composers like Pierre Boulez, former conductor of the New York Philharmonic. At a conference on "The Composer and the Computer," held in Paris in February 1981, Mr. Boulez explained that digital synthesizers could be used as a form of musical calculator, allowing the composer to transform a piece rapidly in different ways. Composers have also used synthesizers to produce sounds that are beyond the reach of human performers, such as notes played extremely rapidly.

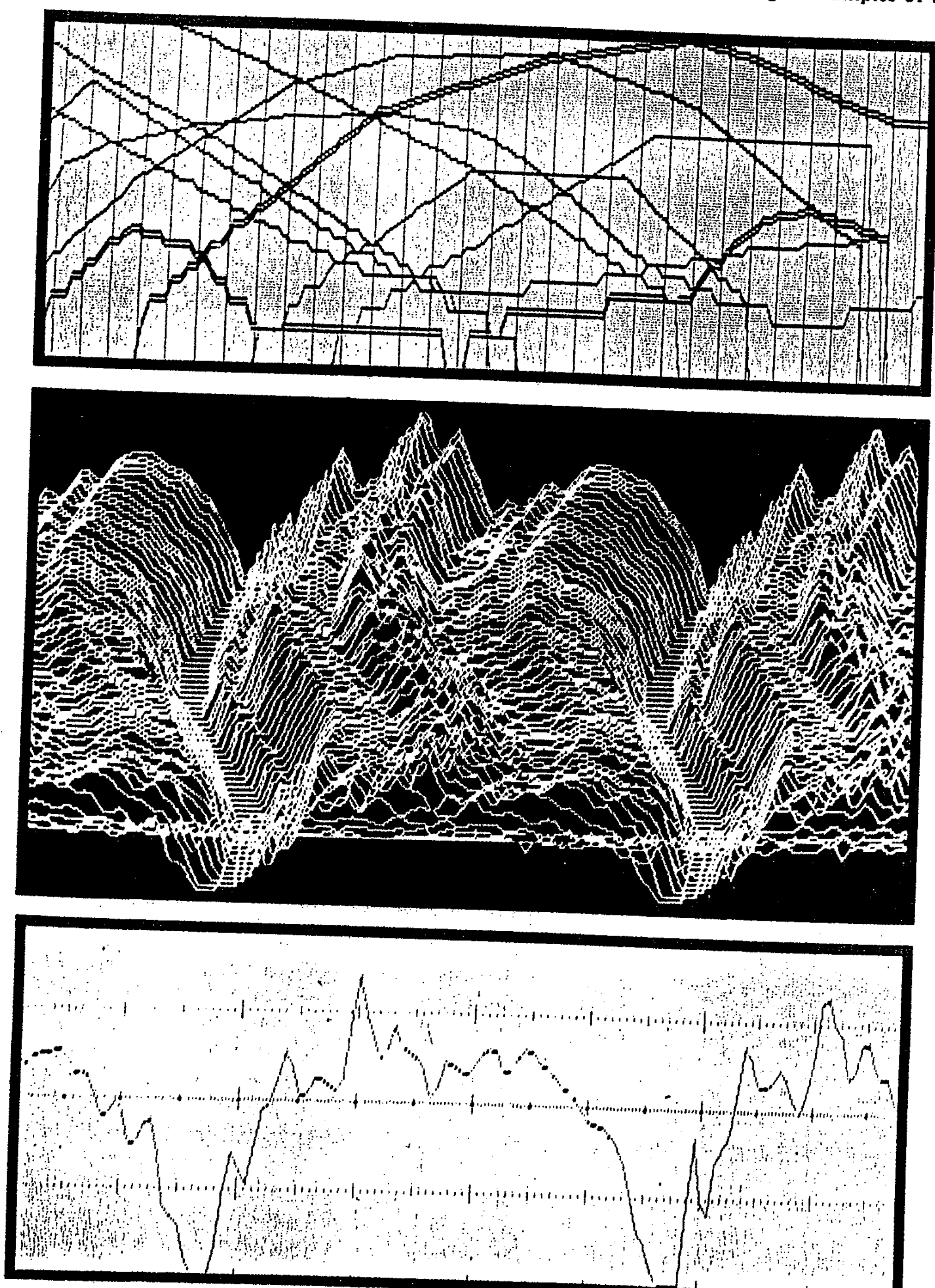
Finally, popular music performers—among them Oscar Peterson, John Appleton, Gerry Rafferty, and Jimmi Destri—have used the synthesizers in their recordings. One of the main appeals to performers is the ability to obtain a wide range of instrumental sounds from a keyboard.

Digital synthesizers are commonly used today in combination with conventional instruments. A piece may be composed on a synthesizer and then played for a final recording with all conventional instruments, or a portion of the band or orchestra—the rhythm or strings section, for example—may be replaced by a synthesizer.

Sounds by the numbers

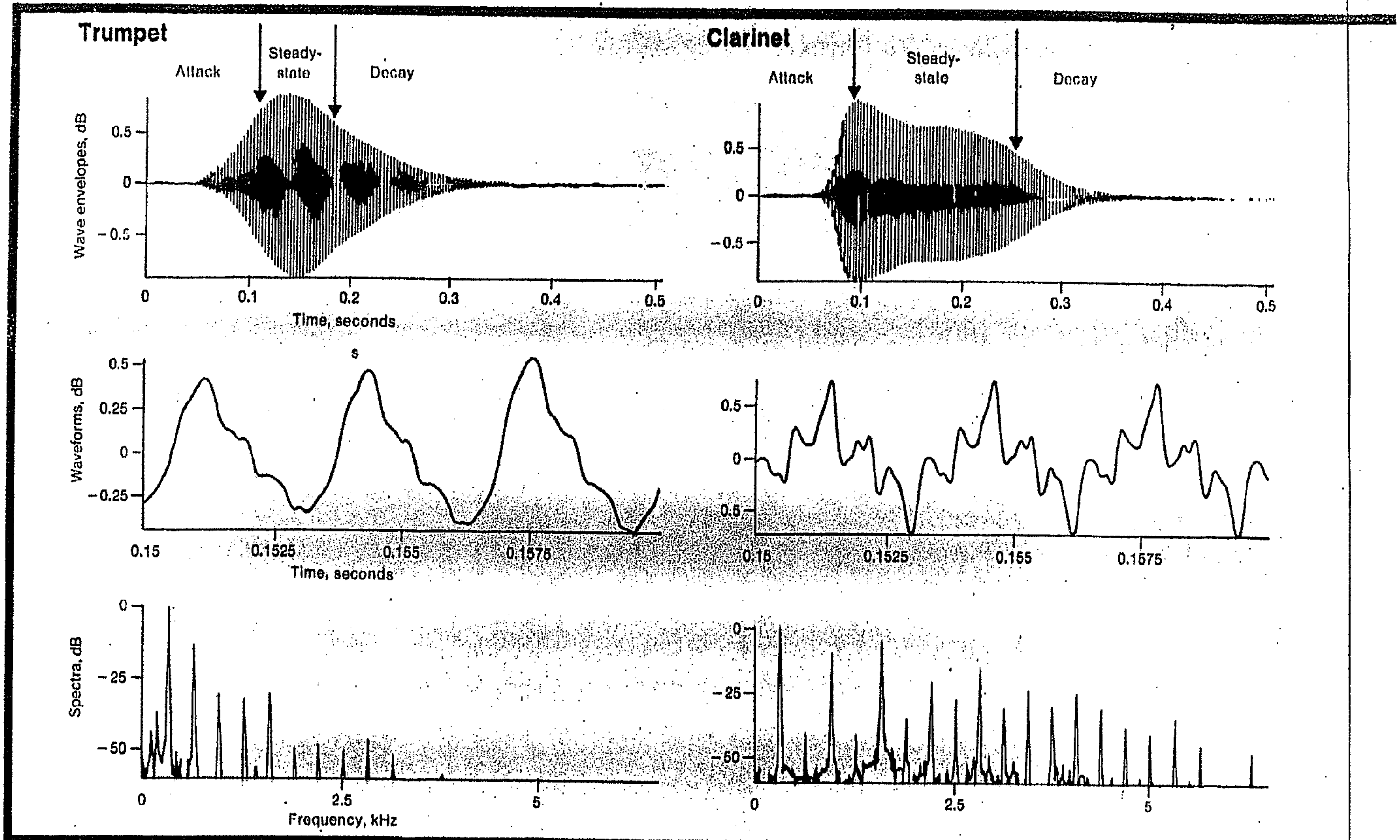
To understand the approaches used in the digital synthesis of music, it is useful first to review what conventional musical sounds are and how they are generally produced. Musical instruments produce sounds by activating some form of mechanical oscillator—a column of air, in the case of woodwinds and brass; a stretched piece of cat gut, in the case of strings; or steel wire, in the case of the piano, to name a few types. This primary oscillator sets other parts of the instrument vibrating. The strings of the violin start the wooden body of the instrument vibrating, setting up in turn vibrations in the air inside the violin; the piano wire that is hit with a piano hammer sets other wires vibrating and sets up vibrations in the wooden body of the piano [Fig. 1].

Together these vibrations produce the instrumental sound. The pitch of the sound is generally defined by the lowest-frequency vibration present. But the timbre of the sound, which tells the ear which instrument produced the sound, is formed by the higher-frequency components. Some of these components are integral multiples of the fundamental frequency—true har-



Keith Lovett Assoc. (photos)

A leading digital synthesizer, the Fairlight Computer Musical Instrument (CMI), allows the user to sketch with a light pen the time variations of up to 32 harmonic components of a sound (A). The CMI then displays the evolution of the resulting waveform with time (B). Individual segments of the waveform can be excerpted and modified with a light pen (C). As with the Synclavier II, the bestseller among digital synthesizers, the user of the CMI can mix several waveforms to produce a more complex sound. (These CMI displays were photographed in the company's New York showroom.)



monics. A flexible string, as in a violin, sets up overtones that are exact integral multiples of the fundamental. Other overtones are not exact multiples and are therefore called inharmonic. A piano string, which is quite stiff, sets up higher overtones that are sharper—higher in frequency—than true harmonics. The resonating body of the instrument then acts as a filter, selectively reducing or enhancing the overtones created by the primary oscillators. To complicate matters, the overtones do not remain constant but vary over the duration of a note.

All of these characteristics of the tone produced can be modified by the way the instrument is played. A violinist, for example, can introduce a vibrato (a rapid rising and falling of the volume or pitch of the note) through his fingering of the instrument and can greatly alter the relative weight of overtones through his control of the bow. Such modifications are termed articulations.

From an information-processing standpoint, the process of using instruments to produce music from a score is similar to that of turning written words into speech. In both cases the initial input has a relatively low information content. A score with even a dozen different parts defines a maximum of little more than 100 notes per second, requiring a total input rate of 500 bits per second. But the output in both speech and music contains far more information—frequencies ranging to 20 kilohertz, defined to a precision of 16 bits, or a total of up to 500 000 bits per second. In music the additional information is added both by the musicians, who define the amplitude and articulation of each note on the score, and by the instruments themselves.

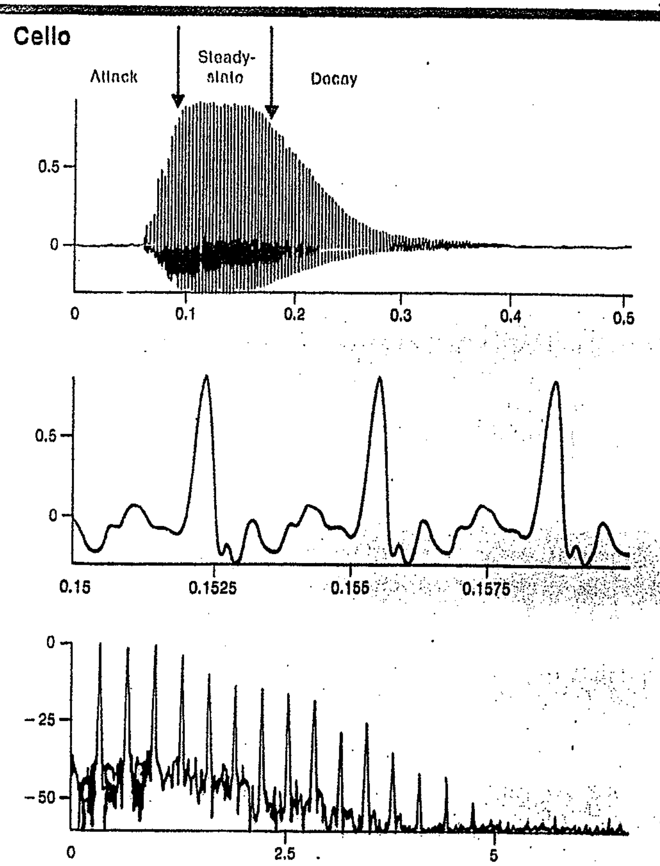
The task of digital musical synthesis is to bridge the gap artificially between the 500-b/s input and the 500 000-b/s output, just as the task of speech synthesis is to bridge the gap between a written input and spoken output. But the problem in musical synthesis is the enormous complexity of the waveforms. A synthesizer represents these waveforms as sums, differences, or products of simpler forms or by more complex computations. The number of different frequency components in instrumental music is far too high for each overtone to be produced separately. The trick in music synthesis, therefore, is to arrive at a suitable approximation that uses a reasonable number of independently produced waveforms but that nonetheless sounds like the instruments mimicked.

Additive synthesis: a simple approach

Different approaches are being used to perform this approximation. The simplest conceptually is that of additive synthesis, developed initially in the 1950s and 1960s and elaborated by James A. Moorer, then a music professor at Stanford University and now with Lucasfilms in San Rafael, Calif., and others. In additive synthesis a musical tone is taken to be the sum of a relatively small set of pure harmonic overtones, each of which varies in amplitude and frequency over time.

Suppose a composer wants to mimic the sound of a trumpet. He begins by playing a single note—say, E above middle C—from a real trumpet and recording and digitizing it. The exact fundamental frequency is then determined electronically, and, using that frequency, a circuit extracts narrow bands of signal spectra centered on each harmonic. The precise frequency and amplitude of each harmonic component is then sampled repeatedly at a relatively high sampling rate (above 20 kHz). The net output is a set of curves showing the variation over time of each harmonic component [Fig. 2].

In using this data for music synthesis, the synthesizer approximates each curve of harmonic variation by a series of straight lines and then stores the approximation. When the appropriate



[1] Synthesis of sounds resembling musical instrument notes is difficult because the instruments produce complex waveforms that vary with time. Here are the wave envelopes, waveforms, and spectra of notes played on three instruments: trumpet, clarinet, and cello. The spectra are taken from an 80-millisecond segment near the middle of the note. Note that the cello has relatively few harmonics and almost no energy between the harmonic peaks. These characteristics make cellos and other string instruments easier to simulate. The trumpet and clarinet have more significant harmonics, and the trumpet has a considerable amount of energy in the inharmonic regions of the spectrum. The wave envelope of a note is often divided into an initial rise (or attack), a peak, a sustain (or steady-state) period, and final decay.

key on a keyboard is hit, the stored data are used to control a set of oscillators whose frequency and amplitude are modified according to the harmonic curves. To produce a complete set of trumpet notes, the process is simply repeated through the entire scale, or the waveform is reproduced with a higher fundamental.

An alternative approach is to treat the harmonic curves produced in the analysis phases as raw material and to modify the curves by hand to produce a better likeness or some desired effect.

Additive synthesis can produce some strikingly lifelike tones, but it has a number of inherent limitations. One is that the reliance on an accurately determined pitch eliminates mimicking of notes with rapid pitch variation, or vibrato. Moreover, the method is not particularly suited to instruments with strong inharmonic overtones, such as drums or even the piano.

The FM and other approaches

A second general approach is based on frequency modulation. In this method, first investigated by John Chowning (also of the Stanford University music department), a carrier frequency—the fundamental—is modulated by another frequency. This produces a frequency spectrum with peaks at the carrier and in sidebands spaced at integral multiples of the modulation frequency, about the carrier. If the carrier and modulation frequency are set equal, the result is a series of harmonics. However, inharmonic tones can also be produced by setting the modulation frequency equal to some irrational number multiplied by the carrier frequency.

Both the amplitude of the carrier and that of the modulating frequency can be made functions of time and allowed to vary according to the attack and decay characteristics of the desired tone.

Since there is no easy way to analyze natural tones by frequency modulation, FM synthesis is mainly a matter of empirically trying various modulation frequencies and time-varying modulation indexes. A key limitation of simple FM tones is that the resulting spectrum tends to be rectangular, with a fairly sharp high-frequency cutoff, unlike the gradual, high-frequency decline of instrumental spectra. In practice this limitation is partly overcome by having the synthesizer sum a number of FM tones with successively higher modulation frequencies to produce the necessary spectrum [Fig. 3].

Many other possible schemes are based on the same principles as frequency modulation. Mathematically, frequency modulation results when one sine wave is taken as the argument for another sine-wave function.

Many potentially useful synthesizers can be created if other functions are substituted for the sine waves. For example, Yasuhiro Mitsuhashi has proposed a technique based on parabolic functions, which result in more gradual high-frequency cutoffs. Combinations of exponential and sinusoidal functions can also be used. Other techniques are subtractive synthesis, in which an initial excitation sound is digitally filtered, and waveshaping, the digitally controlled distortion of a given waveform.

Finally, programmers can achieve a form of synthesis simply by recording the individual notes of an instrument and modifying them in various ways—for instance, by changing their pitches. But one problem with this approach is that simply playing a waveform faster changes its timbre and length as well as its pitch, so that notes played at different pitches generally have to be "touched up" in some way to resemble conventional instrumental tones. However, an advantage of this system is that it can turn any sound, whether initially musical or not, into the basis for a keyboard scale, and thus can be used to produce special musical effects, such as bird chirping.

The Synclavier: a combination of techniques

An idea of how these principles of synthesis are applied can be obtained from looking at the two most popular digital synthesizers, the Synclavier and the Fairlight CMI. The Synclavier uses a combination of additive synthesis and frequency-modulation techniques.

The first step in synthesizing a sound with the Synclavier is to enter a fundamental frequency and the relative amplitudes of up to 24 exact harmonics. The values of the harmonics can be obtained from a recorded instrumental tone through a spectral analyzer called Sample to Disk, which comes as an option. A volume envelope, defining the way the amplitude of the note varies with time, can then be entered. The volume envelope is specified by the attack rise time, decay times, and sustain times, as well as by relative peak and sustain amplitudes.

This first step produces a tone with a timbre that does not vary over time. A frequency-modulation signal, a sine wave, can then be superimposed on the fundamental frequency plus its harmonics. The modulation signal's amplitude is controlled by a second time-dependent envelope. Since the modulation creates new harmonics with an amplitude that varies with time in a different way than that of the initial sets; the resulting timbre also varies with time, giving a more lifelike sound. Since the ratio of the modulation frequency to the fundamental can be varied at will, inharmonic frequencies can also be added.

Finally, the user can modify the fundamental slightly and repeat the same process with a new set of harmonics and a new frequency modulation.

Each pair of oscillators—fundamental and modulator—is collectively termed a "voice." Up to four voices can be tied to a single key to produce a note of one synthesized instrument. Each Synclavier, depending on the model, can work with 8 to 32 voices, with the 16- and 32-voice models being the most popular.

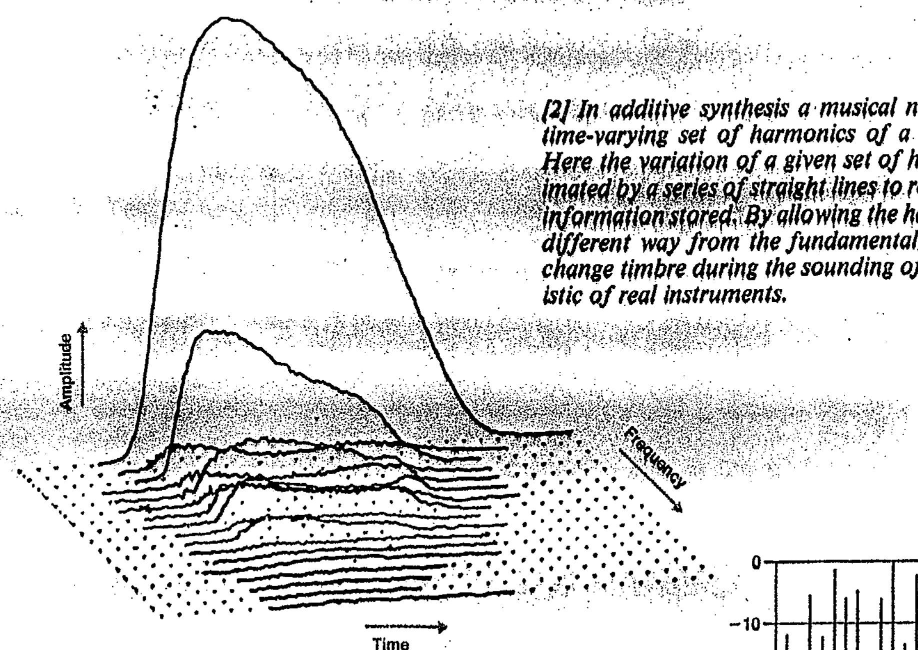
The hardware of this system is reasonably straightforward. Each group of eight voices is on a module with five circuit boards, while overall control is supplied by a minicomputer designed by New England Digital Corp. The minicomputer's own software is programmed in XPL, a subset of PL-1, rather than in assembly language—a feature that has reduced the manufacturer's software-development costs. The minicomputer takes the data entered by the user and calculates from the harmonic ratios a complex waveform for each voice used. The waveforms of the appropriate voice modules are then sent to a memory and [illegible]ed.

When a key is struck, a signal is sent, activating the appropriate voices and giving them the fundamental frequency assigned to the key. In each voice module a signal from a crystal oscillator is split into a carrier and modulator channel. The modulator channel, using a sine wave lookup table, produces a sine wave of the proper frequency and scales it to the proper amplitude by multiplying it by the output of the modulator envelope interpolator. The modulator sine wave is then added to the phase angle signal from the carrier channel, and the resulting phase angle is used to generate the final waveform from the shape stored in memory. The overall volume is determined from the volume envelope interpolator, and the waveform is converted into an analog output.

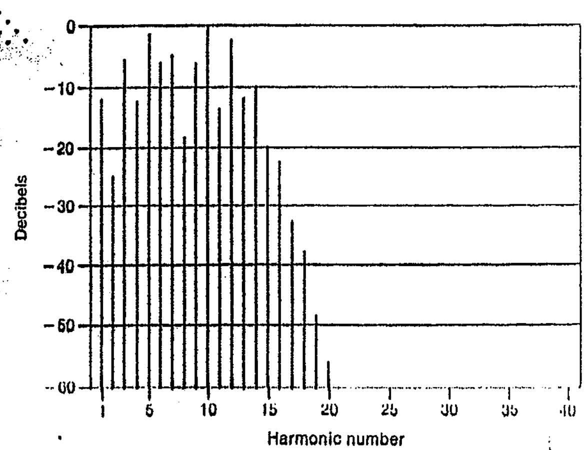
Software implemented by the minicomputer creates additional effects, such as pitch glide—in which one note glides smoothly into a second—or vibrato.

The input devices of the Synclavier are designed to be used by performers familiar with analog instruments. In creating the voices, the user pushes the button describing the variable desired, such as the number of the harmonic to be entered, and then turns a knob to increase or decrease the value from starting point. The user enters the actual music to be played by the voices by performing it on a pianolike keyboard.

The Synclavier records its own output encoded digitally on floppy disks. Up to 16 "tracks" can be recorded in a synchronized fashion. A performer might, for example, set up a 32-voice



[2] In additive synthesis a musical note is modeled as a time-varying set of harmonics of a given fundamental. Here the variation of a given set of harmonics is approximated by a series of straight lines to reduce the amount of information stored. By allowing the harmonics to vary in a different way from the fundamental, this technique can change timbre during the sounding of a note, a characteristic of real instruments.



[3] In FM synthesis, modulation is imposed on the carrier, producing a series of higher harmonics that are integral multiples of the modulation frequency. Simple FM leads to a relatively sharp high-frequency cutoff, as seen here, in contrast with the gradual cutoffs of real instruments. However, if modulation is applied to a complex waveform, consisting of a fundamental and harmonics, a more gradual high-frequency decline, and a more lifelike sound, occur.

Synclavier to play an instrument that has four voices assigned to each key. Up to eight keys could be played at once. A second instrument could then be recorded, and the two superimposed on tape. By this means, the sound of an entire 16-piece band could be built up, even though only one performer was doing the actual playing. With multiple-track tape recording, even more complex superimpositions can be created.

While the Synclavier is basically designed as a performer's instrument, with music entered through the keyboard as it is played, an optional terminal can enter music through a typewriterlike keyboard in a special music language called Script. Another option allows the user to print out a performance at the keyboard in conventional musical notation.

The Fairlight CMI: operation in two modes

Synclavier's principal competitor is the Fairlight CMI. It uses two modes to produce sound, one based on naturally produced samples. In the sample mode a note from an instrument at a given pitch is recorded and digitized, with the sampling rate varying according to the pitch of the note. (Since the amount of memory available is fixed, higher pitched notes will be sampled for a shorter duration.) The same sound is then played back at different pitches when the keys on the keyboard are touched. In effect, the waveform is simply accelerated or slowed according to the key touched.

The second mode is based on additive synthesis. Using a light pen, the user draws on a screen the time-varying amplitudes of each of 32 harmonic overtones. From these the Fairlight CMI calculates a waveform that varies over the duration of the note. A three-dimensional picture of the waveform is displayed on the screen, and any waveform for a given segment of time can also be displayed [see photos, p. 47]. (Each note is divided into 128 time segments.) The user can modify the individual waveforms on the screen with the light pen.

Up to eight synthesized or natural sounds can be merged and played simultaneously when one key is depressed. In effect, the synthesized mode plays the same role in the Fairlight unit as a single partial timbre does in the Synclavier—except that in the Fairlight only exact harmonics are produced, and each harmonic can be individually varied. The programmer can add inharmonic tones by producing another voice based on a different fundamental (as in the Synclavier), by mixing in natural sounds, or by modifying the resulting waveforms.

While the Synclavier synthesizer is designed mainly for performers, the Fairlight leans more toward the composer. Scores can be typed in, with notes designated in conventional format by letter. Music can also be entered directly through a piano-style keyboard.

Both the Synclavier and the Fairlight units give remarkably lifelike imitations of strings, horns, accordians, organs, and bells. But both have inherent limitations in producing complex sounds. For one thing, both instruments rely on a simple scaling to change the pitch of notes, while in reality, when the pitch of a note produced by a real instrument doubles (a one-octave shift), the overtones do not double but undergo much more complicated changes.

Equally important, both instruments have fairly strict limits on the total number of overtones that can be included. This presents a problem with synthesizing the piano in particular. For many of the 88 keys of the piano, there are more than two or three strings tuned to slightly different frequencies. When a key is struck, the strings of that key first vibrate at these different frequencies, then pull each other together toward a single tone. In addition, all the other hundreds of strings vibrate somewhat in sympathy when the damper is lifted. Each string has its own overtones, and, as pointed out earlier, the higher ones are not exact harmonics of the lower ones for a given string. So, an accurate rendition of a piano would need hundreds of overtones. Also, the piano note does not decay at an even rate, as do synthesized tones, but decays first faster and then slower.

As a result, the piano tones offered on Synclavier's demonstration recording sound much harsher than the real thing, because many overtones, especially the inharmonic ones, are omitted.

Finally, complicated articulations, in which the timbre of an instrument changes rapidly, as in some classical violin pieces or a jazz trumpet solo, are difficult to program on existing synthesizers simply because so many rapid changes take place in such a short time.

Despite these limitations, synthesizers are being used more and more to replace conventional instruments. For most commercial applications, the music simply does not call for anywhere near the full potential of an orchestra or the full talents of musicians, and the limitations of synthesized music are simply not relevant.

In fact, synthesizer use has become widespread enough to cause some concern among musicians that electronic music may mean fewer jobs for them. But Dick Hyman, a spokesman for the Federation of Musicians in New York, said that existing contracts offer sufficient protection, since they guarantee that if a musician uses a synthesizer to play several instrumental lines, the musician will be paid the same as if several individual musicians are playing the same music. There is thus as yet no direct financial gain to an employer of union musicians to substitute a single synthesizer for many conventional instruments.

New sounds being developed

As commercial digital synthesizers become more common, research is continuing at a number of universities and research laboratories into ways of improving the quality of synthesized music. Major centers of research include Stanford University, Palo Alto, Calif.; Bell Laboratories in Murray Hill, N.J.; the University of Michigan in Ann Arbor; the Massachusetts Institute of Technology in Cambridge; and the University of Bradford in England. In most of this work the trend is to integrate more and more elaborate models of tones into synthesizers. For example, some researchers are investigating ways of systematically changing timbre with pitch to imitate more accurately the resonances of real instruments.

At some point, Lucasfilms may use electronically synthesized music or special effects in its *Star Wars* series of movies. But the next installment, *The Return of the Jedi*, will have a score performed by the London Philharmonic Orchestra. The sound of a full symphony orchestra is still well beyond the capabilities of digitally synthesized music.

To probe further

Two excellent introductions to the basic techniques of digital synthesis are provided by James A. Moorer in "Signal Processing Aspects of Computer Music: A Survey," *Proceedings of the IEEE*, Vol. 65, no. 8, August 1977, pp. 1108–37, and by Harold G. Alles in "Music Synthesis Using Real Time Digital Techniques," *Proceedings of the IEEE*, Vol. 68, no. 4, April 1980, pp. 436–49. The underlying principles have not changed since these papers were written.

More current information on digital music synthesis can be obtained from the *Computer Music Journal*, published by the MIT Press. *Electronotes*, a newsletter of the Musical Engineering Group at Ithaca, N.Y., also gives current information. The May 1981 issue contains a thorough bibliography on the subject. ◆