Computing Science Technical Report #10

COMPUTER CONTROLLED ANALOG SYNTHESIZERS

F. Richard Moore

May, 1973

Computer Controlled Analog Synthesizers

by

F. Richard Moore
Bell Laboratories
Murray Hill, New Jersey

A colleague of mine at Bell Laboratories recently made the comment: "Computers - ah yes! Those are those marvelous machines capable of making nearly one million mistakes per second!" It may turn out that one of the saving graces of human beings is their inability to make nearly so many mistakes per second as computers. When we first considered the problem of how to make the "ideal" computer music system, it became apparent that we needed to inject as much as possible of the human element into what the computer was to do. We needed to harness the computer's speed in such a way as to allow a human being to interact with it on his own time scale - in terms of seconds and minutes instead of microseconds. But how could we allow a human being - a musician who wants to compose music - to interact with a computer so as not to simply limit what the computer does to whatever the human can already do? Indeed, what is to be gained by using computers to make music at all? In order to answer these questions, we have to rethink some of our basic ideas about music, and about musicians.

*** Figure One ***

When a musician sits and improvises at an instrument, he is essentially composing music in real time. This means he is thinking very quickly about various possible sound combinations, he is evaluating each possibility as best he can in the time allotted, and he is executing the various mechanical motions necessary to produce the music he wants. Because of the considerable difficulty involved in doing all these things at once, musicians have traditionally tended to specialize in either the inception/evaluation process, or in the execution or "realization" process; that is, we have composers and performers.

*** Figure Two ***

"IMPROVISING" CASE:



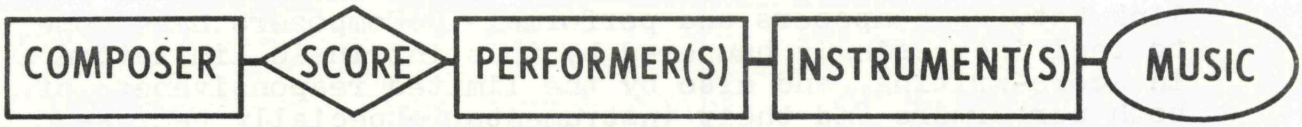
MUSIC

Figure 1. MUSIC - Improvising Case

"IMPROVISING" CASE:



"TRADITIONAL" CASE:
(IDEALIZED)



MUSIC

Figure 2. MUSIC - Improvising and Traditional (Idealized) Cases

It's fairly easy to see the advantages of this rather artificial subdivision: composers, freed of the necessity for quick thinking, can be more thorough in the inspection of musical possibilities, while performers are free to concentrate on perfecting their ambidextral virtuosity. One important consequence of this traditional arrangement has been the extensive development of a means for communication between these two groups in the form of symbolic messages. When these messages are written down on pieces of paper, they constitute not only the embodiment of a communication channel, but also an all-important permanent record of musical thought which can be transmitted through both space and time. Not only that, but these messages allow a single person, a composer, to control the actions of a large group of musicians playing in concert, as in the case of a symphony orchestra. We call such messages musical scores.

The disadvantages of this dichotomous arrangement are perhaps also well-known, but in the past we have more or less disregarded them.

*** Figure Three ***

Communication of musical thought via scores has always been far from perfect, and composers and performers have always had to take this into account whenever they interpret their own tasks. Fortunately, most performers usually have enough insight into the art of composition that they can make proper judgments about something which the score leaves ambiguous, thereby enhancing the musical result. But this so-called "art of interpretation" is really just an attempt to compensate for the deficiencies in the communication link between composers and performers. Composers have come to feel increasingly hampered by this intrinsic limitation in score-writing, and also by the limited responsiveness of both performers and their instruments, especially because alternatives to this composer/performer schema have recently presented themselves.

In order to fully appreciate the implications of these alternatives, we will have to do some careful thinking about music as it has been. For example the score, or the "symbolic music" as we have called it, has come to be commonly confounded with the music itself. There seems to be a deeply-rooted belief in the identity of music and its notation, as we can see in the rather contumelious attitude of so-called serious musicians, who read and write their music, towards the improvising jazz players, who don't seem to use music at all! It is crucial to emphasize that by music we should mean sound, and not the symbolic language used as a link in the composer/performer schema. Another

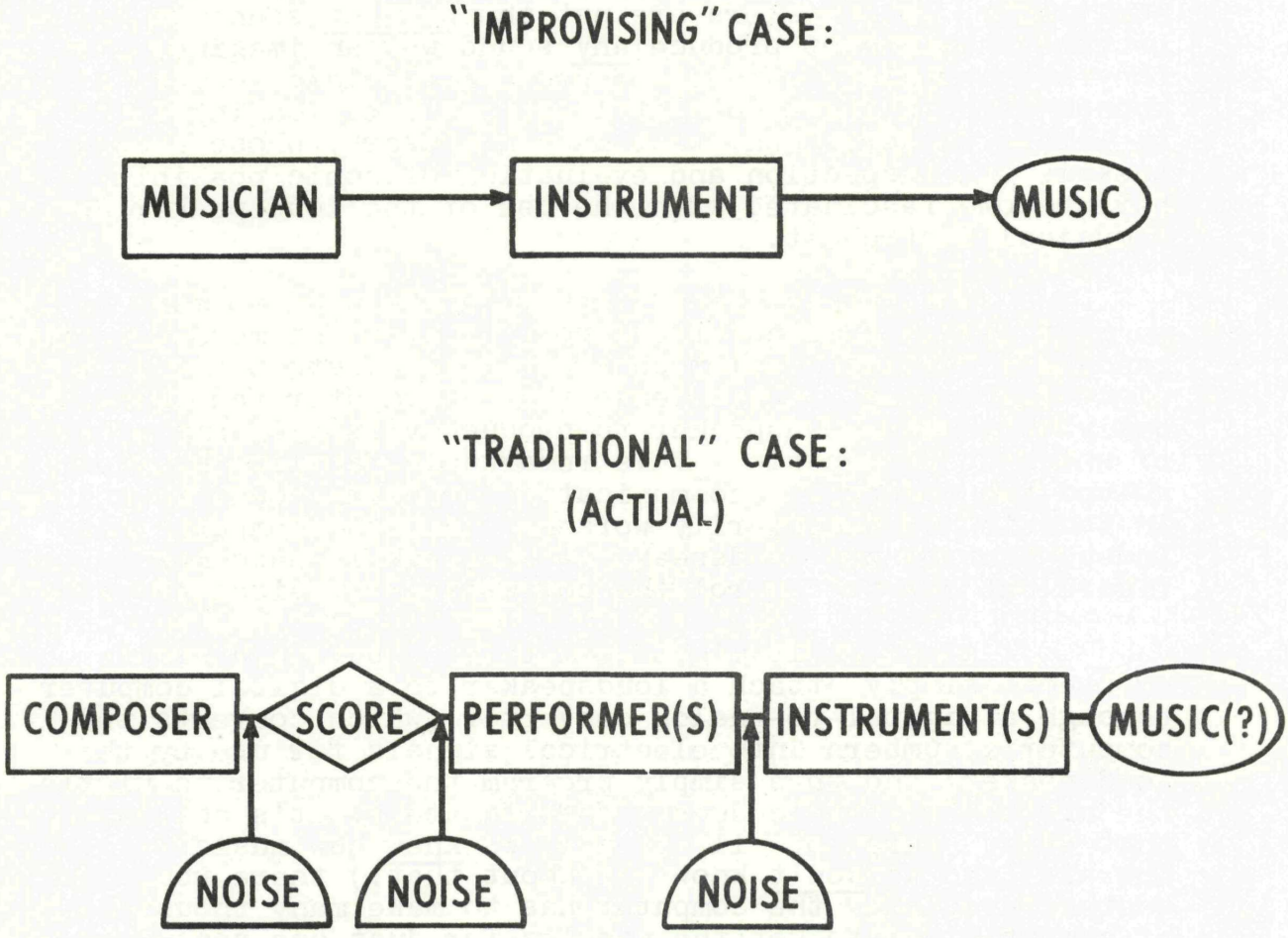


Figure 3. MUSIC - Improvising and Traditional (Actual) Cases

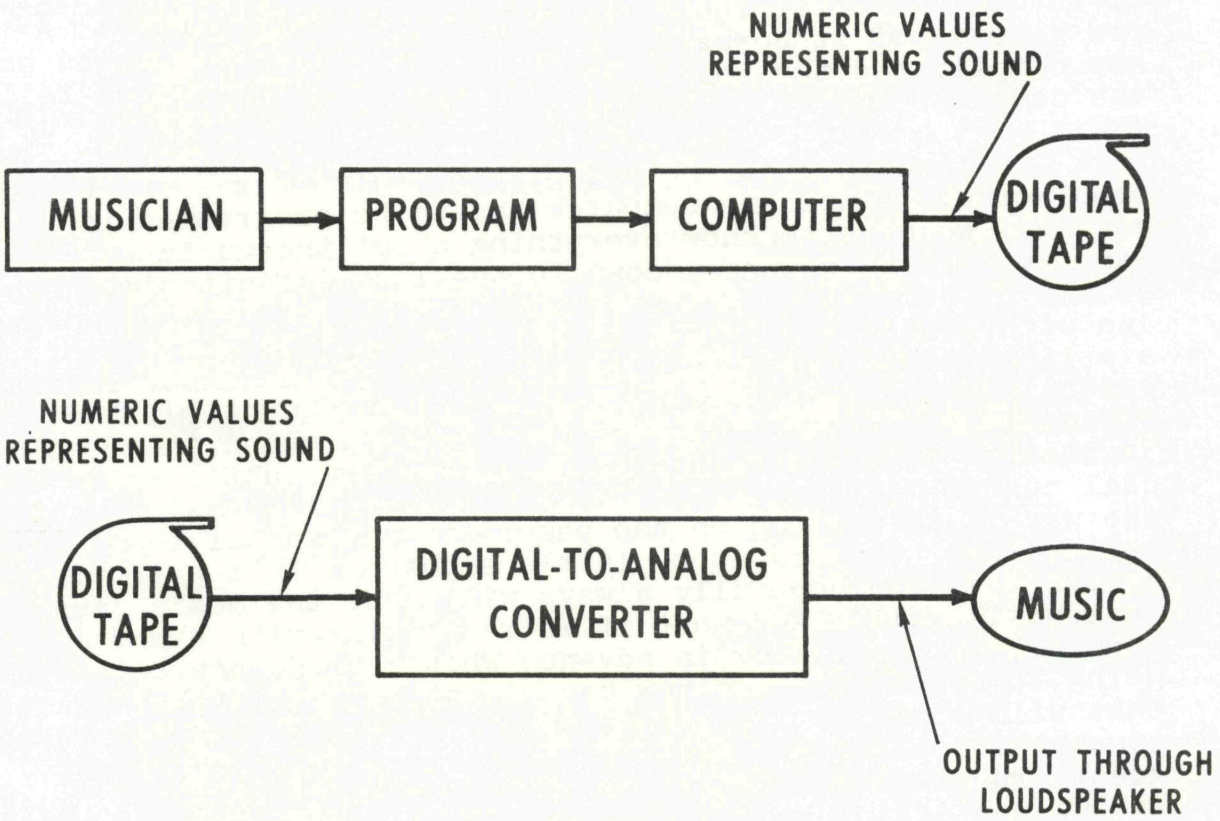
more subtle misconception is the tendency to identify music with those particular mechanical contrivances we call musical instruments. It should be obvious that not every sound which can come from a piano or a violin is music, yet it is still an exotic idea to many people that music may come from other sources than winds, percussion, strings, or voices.

In particular, there has come into existence a much more generalized sound source which is capable not only of closely approximating the sounds of all other instruments, but a host of new ones as well: the loudspeaker. Admittedly, loudspeakers are not the perfect sound source because they cannot produce any sound we can imagine. But they are certainly capable of producing a large number of sounds which we have not, at least not yet, imagined. It is little wonder, then, that many composers, whose avowed task is the inspection and evaluation of sonic possibilities, have become fascinated with the use of the loudspeaker as a musical instrument.

But before a loudspeaker can be used to make music, we must devise a means of controlling it. The most powerful candidate for a control mechanism which exists today is the digital computer, in the sense that a computer can theoretically cause the loudspeaker to produce virtually any sound of which it is capable. This remarkable facility has not solved the composer's communication problem, however, and in fact it has temporarily worsened it: the computer can indeed control the loudspeaker - the task it hand now is to figure out how to control the computer for musical purposes.

At first it might seem to be a pure and simple problem: simply attach a loudspeaker to a digital computer through a digital-to-analog converter (which converts the computer's numbers into electrical signals for use by the loudspeaker) and then simply program the computer to wiggle the air the way music does. Unfortunately, it's not quite that easy. First of all, even if we knew how music wiggled the air (and we don't know all about that!) there would be another problem: the computer has to make many thousands of computations to get the wiggles for just one second of sound. Everyone is used to hearing about how fast computers can operate, so it might come as some surprise to realize that computers aren't nearly fast enough to compute most musical sounds in real time. To make all the calculations necessary to imitate the sound of, say, a string quartet, computers would have to go hundreds, even thousands of times faster than any of those presently available. The wiggles are just that complicated.

*** Figure Four ***



MUSIC V SYSTEM

Figure 4. MUSIC V System

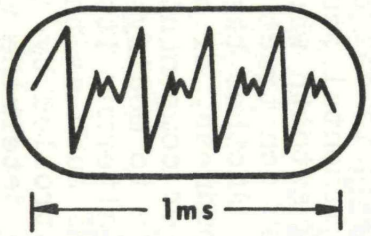
So what we have done in the MUSIC V program, for example, is to let the computer compute as fast as it can to put its numbers on a digital tape. After all the numbers are there, perhaps after 10 minutes or several hours of computing, we can play back the numbers in real time through the digital-to-analog converter and loudspeaker to hear the music. A few seconds of it perhaps, or a few minutes - it depends on how complicated the sounds are, and how good we are at telling the computer how to make them.

And there's another problem which turns out to be even worse: since the computer has to be programmed in advance, we have to know everything about the music we want to make in advance, down to the last infinitesimal detail of each wiggle. This might be fine if we were working with sounds we already knew a lot about, like that of a string quartet, but even then something, an almost ineffable something, would be missing. This missing quantity, the X-factor that makes music "come alive" and sound interesting instead of dull and mechanical, is the very subtle thing call nuance. It turns out that if we play the music exactly the way it is written on the page, without any of the natural inflections, tiny rubatos, little emphases, that performers always have and hopefully always will use, the music sounds just plain dull. And what if we're working with brand new sounds? We can't know in advance what a brand new sound is going to sound like, so how can we put in the little nuances that will optimize its effect on our ears? We have to resort to what is called the scientific method and make our best guess, run it through the computer, see what comes out, guess again, run again, and so on, until what comes out the last time either sounds wonderful or uses up our last penny's worth of computer time. And each trial might take several hours, even days to accomplish. Compared with sitting at a piano, it is very difficult to compose new music in this way.

Are we stuck? In a way we are, for there is no other way to control a loudspeaker that is nearly so flexible and rich with possibilities. But there is another approach which, as we shall see, very neatly solves the problem of how to get performance nuance back into the music and at the same time take advantage of computer control of many of the aspects of sound. Let's take yet another look at the nature of music to see how to do it.
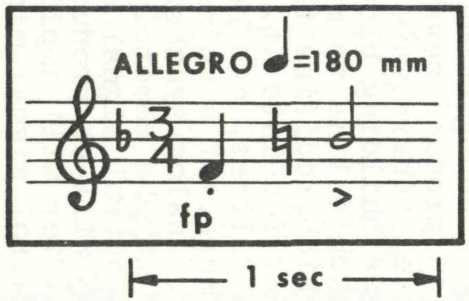
*** Figure Five ***

1. FASTEST VARIATIONS (~10^2 TO 10^4 CHANGES PER SECOND)



EXAMPLE: ACOUSTIC WAVEFORM (SOUND ITSELF)

2. MEDIUM-SPEED VARIATIONS (~10^0 TO 10^2 CHANGES PER SECOND)



(PSYCHOLOGICAL OR PERCEPTUAL CHANGES)

EXAMPLES: PITCH
LOUDNESS
DURATION
TIMBRE (TONE-QUALITY)
SOURCE OR DIRECTION

3. SLOWEST VARIATIONS (LESS THAN 1 CHANGE PER SECOND)

(GROSS MUSICAL CHANGES)

EXAMPLES: ACCELLERANO
CRESCENDO

"LEVELS" OF TEMPORAL VARIATIONS

Figure 5. Levels of Temporal Variation in Music

Music might be thought of as the art of temporal variation on many levels. The fastest variations are those of the wiggles we discussed earlier. These are of course the pressure variations in the air, and represent the actual acoustic waveform of the sound itself. Musical sounds vibrate from about 20 to 20,000 times each second. So there are on the order of 10^2 to 10^4 wiggles, or cycles of air pressure changes, per second.

The next-to-fastest variations change on the order of only 10^0 to 10^2 times per second. These are what might be called the psychological or the perceptual characteristics of musical sounds. They are pitch, loudness, duration, timbre, and the direction from which the sound emanates.

There is also an abstract category of very slow time variations which change even less rapidly than once per second, such as tempo, crescendi, and whole compositions, but we won't consider these any further here.

If we were to use the computer to control only the medium-speed variations like pitch and loudness, the computer could easily keep up with what is going on at that level in most music. The fast wiggles of the acoustic waveform, then, would have to be produced by something outside the computer which the computer can control. Fortunately, just such a thing exists today: the analog music synthesizer such as the Moog, or the ARP.

*** Figure Six ***

An analog music synthesizer is made up of several electronic devices which either produce or modify sound in some way. The basic sound-generating unit is called an oscillator, which can produce several elemental sounds, such as the sine wave, square wave, or sawtooth wave. These sounds can be used as building blocks to construct nearly any other sound we would like to make. Another basic unit is the amplifier, which modifies the loudness, or amplitude, of any sound which is fed into it. Other components of the synthesizer, called filters, can be used to modify the tone-quality, or timbre, of a sound by altering its spectral envelope. In modern analog synthesizers such as the Moog, each of these units may be remotely controlled by a voltage signal. A voltage control signal could determine what pitch, or frequency, will be produced by an oscillator, or by how much an amplifier will amplify the sound being fed into it. Such a control signal might be generated by a piano-like keyboard which produces differing voltages according to which keys are depressed. we can even use the output signal of one unit, such as an oscillator, to



Figure 6. An Analog Music Synthesizer

control the action of another unit, such as an amplifier, which would result in the amplitude modulation of the sound being fed into the amplifier. Complex and interesting sounds can be made by utilizing the virtually unlimited number of ways of interconnecting these modules to each other - this is how the synthesizer is normally used. But we could also produce the control signals with a digital computer.

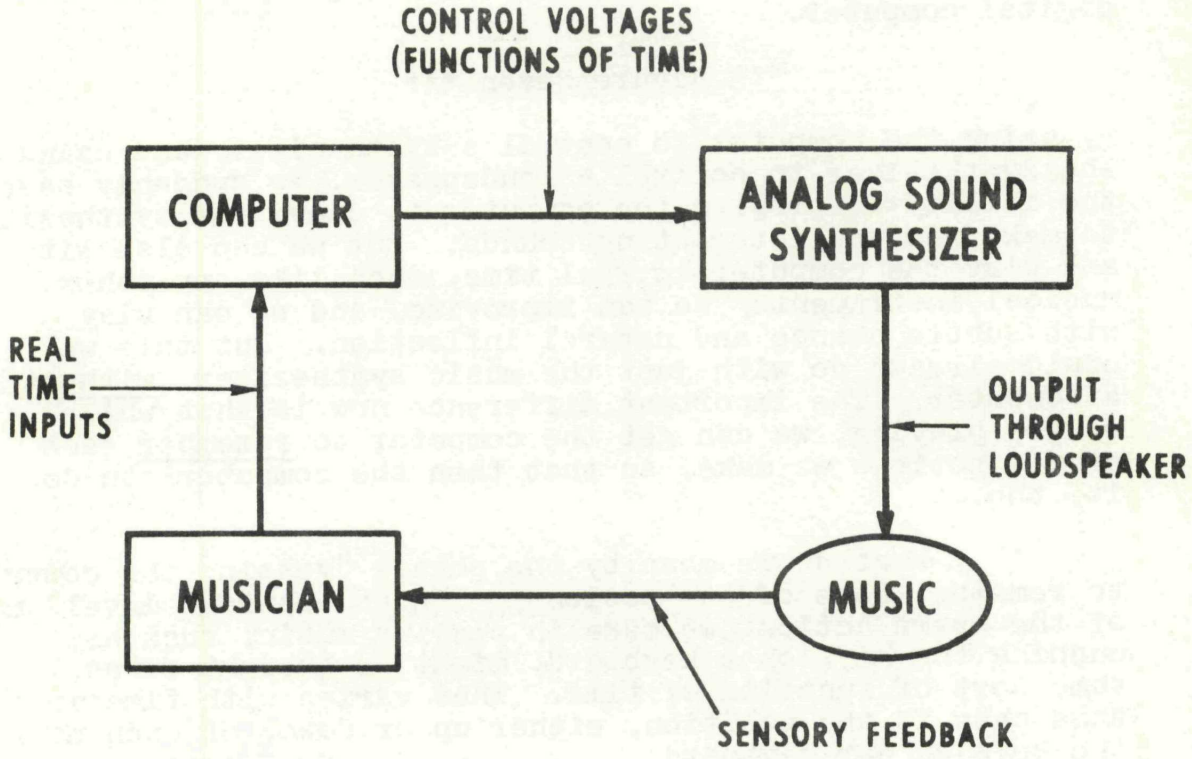
*** Figure Seven ***

By using the computer to control a synthesizer, and using the synthesizer to control a loudspeaker, we suddenly have the ability to program the computer to cause the synthesizer to make new and interesting sounds. But we can also sit and play the computer in real time, much like any other musical instrument. We can improvise, and we can play with subtle nuance and natural inflection. But this we could already do with just the music synthesizer, without a computer. The important difference now is that while we are playing, we can get the computer to remember each of the actions we make, so that then the computer can do it, too.

What do we mean by the phrase "getting the computer to remember each of our actions?" On an abstract level, many of the overt actions we make in playing music, such as pushing the keys on a keyboard, might be thought of as some sort of function of time. What varies with time in this case is the position, either up or down, of each of the keys on our keyboard.

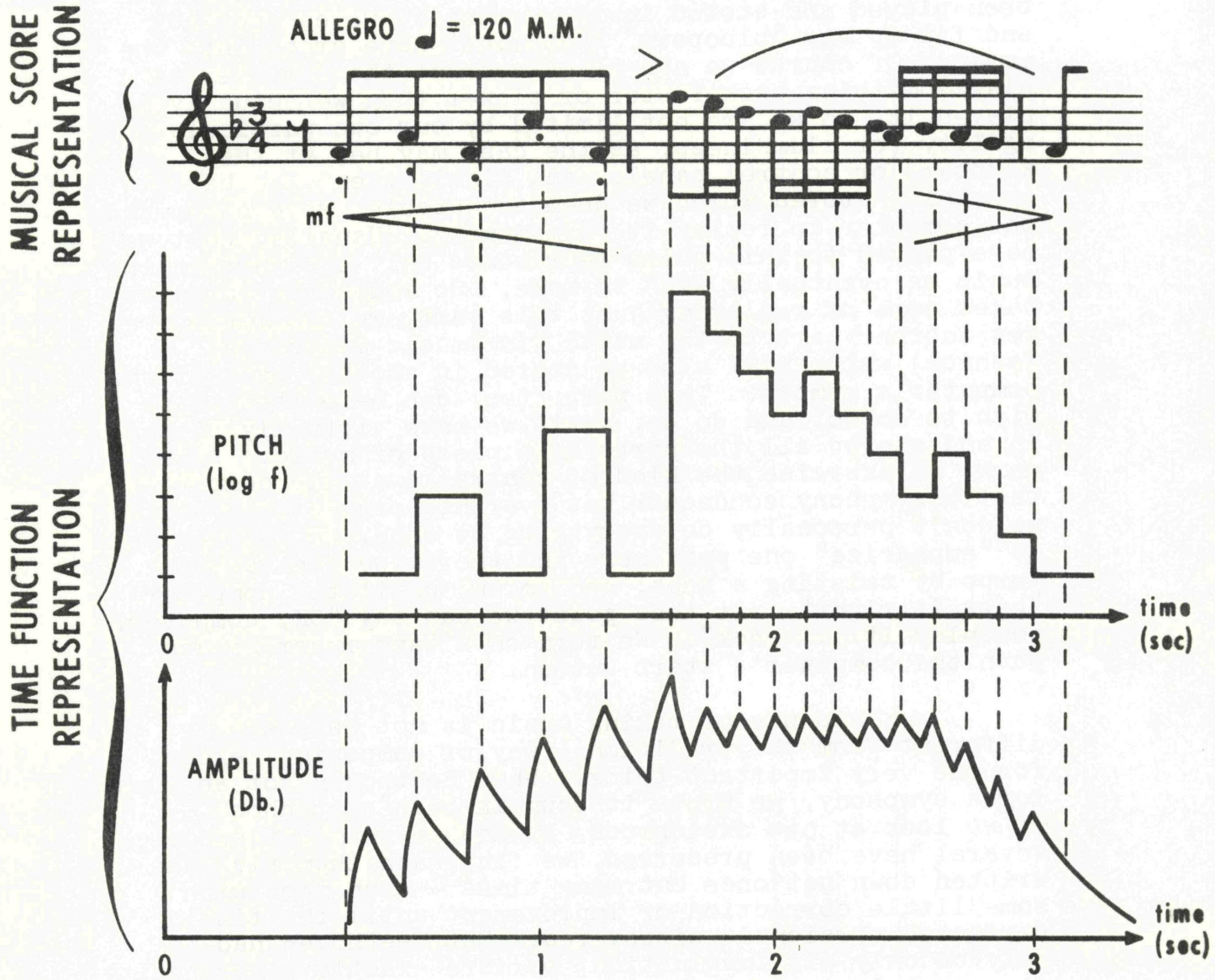
*** Figure Eight ***

In electronic music we often use knobs and dials, so another function of time might describe the setting of a knob or dial at any particular time throughout a piece of music. What the computer does is to look at the position of a key or a knob that we are playing on many times a second, at regular intervals, and to store a number in its memory that represents each particular setting. This process is called sampling. The maximum number of changes the computer can detect per second will be equal to the number of samples it takes per second. The number of samples per second is called the sampling rate. It turns out that for most any overt action that can be made by a human being, if we sample it at a rate of about 100 times a second, the samples will very adequately represent what the human did. That is, we can imitate, or recreate his actions merely by retrieving the sample values from the computer's memory and applying them to the controlled device, which is the analog music synthesizer. But that isn't all we can do with the stored



GROOVE SYSTEM

Figure 7. GROOVE System



MUSIC REPRESENTATIONS

Figure 8. Music Representation - Score and Time Functions

samples. Once we have some samples stored in memory, we can also instruct the computer to make large or small changes in them. This means that we suddenly have the improvising musician's dream: once an improvisation has been played and stored in the computer, we can go back and fix up any "bloopers" that might have crept into the music. Of course we can also instruct the computer to play something back faster or slower than we actually played it, so we are not limited by our own particular agility (or clumsiness, as the case may be) at the keyboard or control panel. And finally, once we have something stored which we deem acceptable, we can command the computer to replay it from memory. Usually, what we have played to this point represents only one part in the music we eventually wish to make. So while the computer takes care of replaying just this part, we can improvise yet another part to the music (in "music-minus-one" fashion) which will also be stored in another part of the computer's memory. This part, too, can be edited if we wish to do so, and so on, until we have virtually played in and stored all the parts of a piece of music. As we work, we exercise the kind of control over the computer that a symphony conductor has over his orchestra: while we don't personally do everything at once, we can decide to "emphasize" one part here and there, we can vary the tempo by twisting a knob, and so on, until the computer "gets it right". At that point we can say "OK, computer, now play it all back". We turn on a tape recorder and push the computer's start button.

This way of making music is not essentially different from the traditional way of composing, except for one very important thing. When Beethoven got an idea for a symphony, he wrote it down in his "sketchbook". If we look at his sketchbook, as indeed we might since several have been preserved, we find each thematic idea written down not once but many times - each time with some little correction or improvement until the final, perfected version is discovered. But Beethoven had to rely on only his imagination, or more exactly his capacity for aural imagery, to decide what changes to make. Using the computer, we are working directly with the sound itself as we decide what to do about it. We can actually hear the effects of any alterations as we make them, which is not only more efficient, but it is absolutely essential if we are working with brand new kinds of sounds. We might call this process "sound-sculpting", since we are working with the sound now in much the same way that a sculptor works in his plastic medium - each of our little changes can immediately be observed and evaluated with relation to the whole composition.

Such a system as I have described here exists now at Bell Laboratories. We call it the GROOVE system. GROOVE is of course an acronym. It stands for Generated Real-time Operations On Voltaged-controlled Equipment. It has this rather general name because in principle, the system could be used to control any time-varying process that a human normally might control, such as driving a car, operating a machine tool, or even running a factory. By watching the controlled process, an operator could cause the computer to turn out the ultimate widget, a working automobile, or, if he listens instead of watches, a piece of music. The key here is the feedback to the operator gained by watching or listening, coupled with the computer's memory and its ability to operate in real time, as the person does.

All this sounds well and good - what is wrong with it? Why use the MUSIC V program at all, if GROOVE is so much better at producing lively music? There is in fact nothing wrong with the computer part of it - the control principles embodied here are flexible and powerful. What is wrong now is that the sounds which may be made with analog synthesizers are far more limited in range of qualities than those which might be made with MUSIC V and digital synthesis. First, we have only a limited number of modules in any particular music synthesizer. Also, analog electronic equipment tends to drift as it warms up, like a small AM radio does, and it is extremely difficult to get analog oscillators to play in tune. Because of these inaccuracies, it is virtually impossible to achieve some of the more subtle effects - like the ever-rising or ever-decending Risset's tones - that can be digitally synthesized.

At the present time, we are working on solving this problem in several different ways - two of them are of interest to us here. First, analog equipment has the advantage of being pretty well understood, since it has been around a long time. If we decide that what we need to improve the quality of some sound is another filter, or a ring modulator, it's not too hard to simply whip one up at the work bench and add it to the analog synthesizer. But we still have to put up with the electronic foibles of such devices, so the problem of getting better sounds is not fundamentally solved by adding more analog equipment.

The second solution is far more experimental but also far more promising: we can construct small, special-purpose digital circuits which perform many of the functions of the analog music synthesizer. These digital circuits, which are in effect small, special-purpose computers, have the advantage that they are perfectly accurate, never drift,

and are fairly inexpensive to make. And they can be built to have the precision necessary to perform any of the tasks which might be done by a general-purpose computer using MUSIC V, only in real time. A combination of a large, slow, general-purpose computer with the GROOVE system, controlling several small, high-speed, specialized digital circuits, is clearly the way in which loudspeakers can best be controlled in the future. Such a system could produce speech, music, and indeed, any sound of which the loudspeaker is capable. And the control system itself will also become more sophisticated.

In 1932, conductor Leopold Stokowski made the following prediction: "One can see coming ahead a time when a musician who is creator can create directly into tone not on paper." With the GROOVE system, which now exists, and with its descendants, nearly 100 years after the invention of the loudspeaker, we are nearing the realization of this goal.