THE TECHNOLOGY OF SOUND SYNTHESIS

Neale Hancock

Through the psychedelic haze of the 1960s there emerged a new generation of musical instrument: an instrument not limited to one typical sound, but able to copy or synthesise existing sounds and generate sounds never heard before.

OVER THE 21 years since synthesisers have been with us their appearance has changed noticeably. The bulky boxes covered in masses of knobs have given way to slim featureless units with only one or two sliders and display. But in many cases the method used to generate sound is very similar to that which Dr Robert Moog used in his first synthesiser in 1965. The main differences lie in changes made to how the instrument is controlled, its ability to store sounds and how the information is displayed. The result is quite often a box full of old technology behind a high tech front panel which is impossible to program. But synthesis techniques have been developed that do away with the tried and true building blocks and bring better sound at a lower cost.

Traditional synthesis techniques

The method of sound generation created by the good Dr Moog was based around the use of voltage controlled oscillators (VCOs), low frequency oscillators (LFOs), voltage controlled filters (VCFs), voltage controlled amplifiers (VCAs), a mixer, and envelope generators. Whilst oscillators, filters, amplifiers and envelope generators have been around for a long time, it was the ability to control them via a control voltage (CV) which made them useful in the synthesis of sound.

Each of these building blocks needs to have at least two of its parameters variable, for instance, a filter can have its breakpoint and sharpness (resonance) changed, thus it requires at least two control voltages. Since a knob or slider is used to change each parameter, it is no wonder that many synthesisers appear as a multiplicity of sliders and knobs!

Figure 1 shows a block diagram of a synthesiser which uses voltage controlled oscillators, filters and amplifiers. The control voltages for these building blocks come from the front panel knobs, the low frequency oscillator, the envelope generator, the musical keyboard and controllers. The technique which utilizes these building blocks is called subtractive synthesis.

The VCO takes control voltages from many different sources, namely, the musical keyboard, the envelope generator and the low frequency oscillator. When the VCO is controlled by the keyboard, its pitch is shifted to match the musical scale. Controlling the VCO with an LFO results in a tremlo effect whilst controlling it with the envelope generator results in a pitch sweep. All three control voltages are mixed via a summing amplifier to allow them all to control the VCO at the same time.

The VCO can generate triangular, saw tooth, square and pulse waves, all of which have a different harmonic content (see Figure 2). Each of these waveforms has a characteristic sound, and should be viewed as the raw material for making new sounds. For instance, a triangular wave is useful when generating flute-like sounds, a sawtooth wave is useful for brass sounds, square waves are useful for wood-wind or bass sounds and pulse waves are good for harpsichord or piano sounds. The waveshapes correspond to these typical sounds because their harmonic content is similar to that of the corresponding sounds.

The pulse wave is different from the other

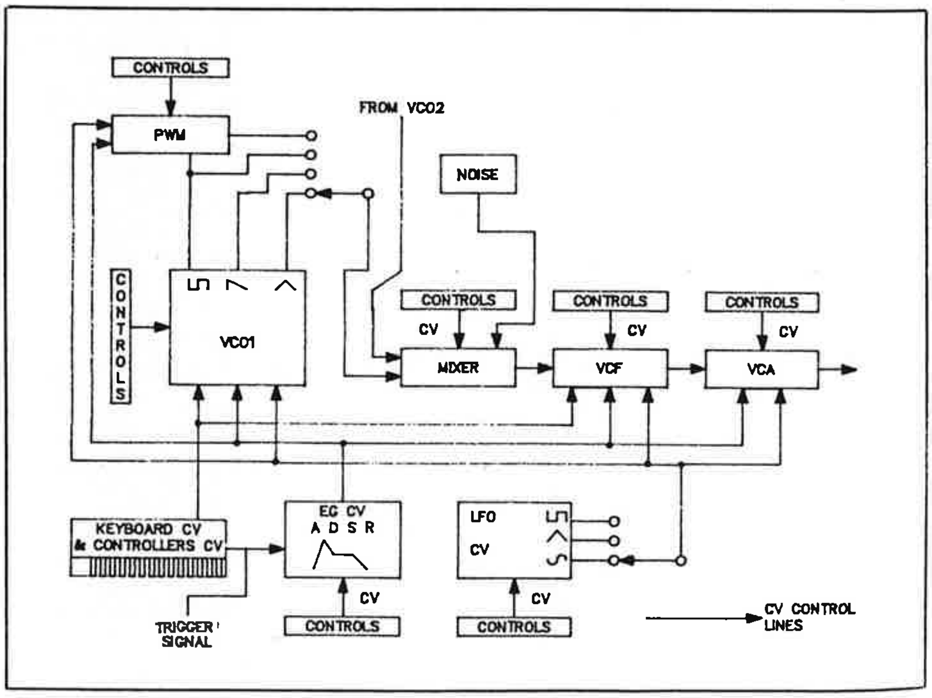
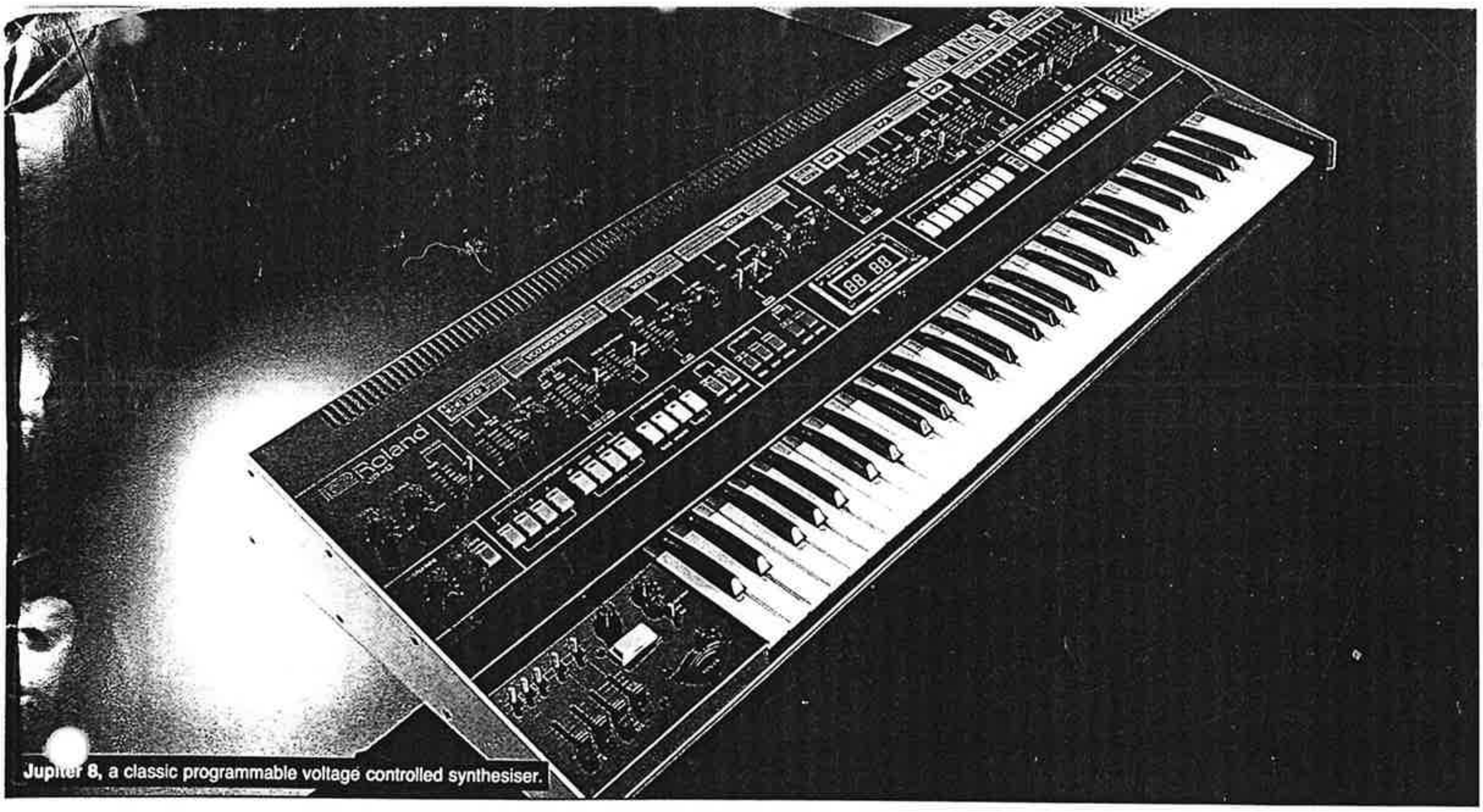


Figure 1. Traditional synthesiser a la Moog.



Jupiter 8, a classic programmable voltage controlled synthesiser.

VCO waveshapes because it can have its pulse width controlled by the LFO and the envelope generator (pulse width modulation). Varying the pulse width creates a rich and variable harmonic content.

The VCF is used to modify the basic sound of the waveshapes generated by the VCO by reducing or increasing the harmonic content. This control over harmonic content is achieved by changing the resonance or the breakpoint of the VCF. The VCF's breakpoint (in some cases the resonance as well) can be changed via control voltages from the keyboard, the LFO and the envelope generator. Since a summing amplifier is used, all three control voltages can be used to control the VCF at the same time. Controlling the breakpoint of the VCF with a control voltage from the envelope generator changes the harmonic content while the note is being played. Controlling the breakpoint of the VCF via the keyboard control voltage enables the VCF to track the VCO.

The envelope generator (EG) typically has an attack time, a decay time, a sustain level and a release time. The envelope generator is sometimes referred to as an ADSR (attack, decay, sustain, release). The duration of the note can be set by using the EG to control the voltage controlled amplifier (VCA). The EG can be set up differently for different instrument types, for instance, plucked instruments would have a short attack time, short decay, no sustain and a long release. Vibrato is obtained by controlling the VCA with the LFO.

Later models of synthesiser (see Figure 5) have their oscillators, filters and amplifiers controlled by digital signals instead of control voltages. As these synthesisers are velocity sensitive, a digital control signal corresponding to playing velocity (how hard a key is pressed) results. When the keyboard velocity controls the amplifier (DCA), the note is louder when a key is pressed hard and quieter when the key is pressed softly. Harmonic content can also be changed by controlling the filter (DCF) with the keyboard velocity.

History

The early synthesisers had VCOs, VCFs, VCAs, envelope generators, etc, all in independent modules. Each module could be patched to the next with leads and the level of each control voltage was set by a slider or a knob. This design concept was the most flexible since patching allowed the output from any module to be fed into the input of any other module. Whilst this inherent flexibility permitted a vast array of sounds to be created, the trade off was that it was only done with difficulty. Most serious musicians want to spend their time playing music, not creating sounds, although there are those amongst us who derive as much pleasure from creating an excellent sound, as we do from using it!

In the next generation of synthesiser, the patch leads were replaced by rotary switches. Whilst this made the synthesiser easier to use, it also meant that inputs and outputs were hard wired thus limiting the flexibility of the synthesiser. One other disadvantage of hard wired synthesisers is that they do not allow external signals (eg, from a microphone or guitar) to modulate the signal or be inserted into signal path at any point.

After synthesisers had been made easier to program, the keyboard players of the world wanted to store and recall sounds and play them polyphonically. The alternative was to use a synthesiser with selectable preset sounds. This, however, limited the variety of sounds. ▶

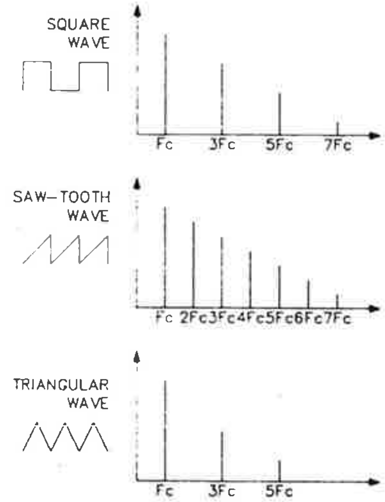


Figure 2. Harmonic content of various VCO outputs.

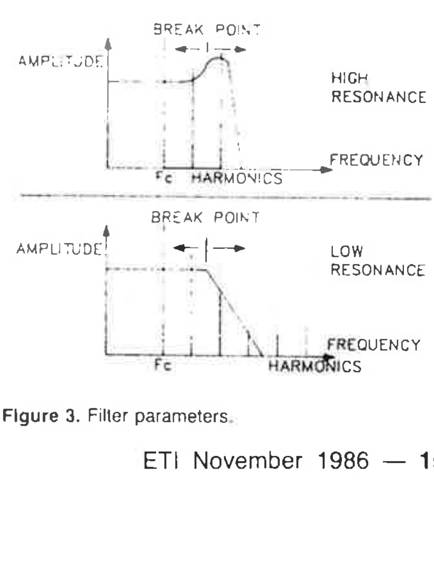


Figure 3. Filter parameters.

String and brass ensembles were polyphonic versions of these preset synthesisers minus the facility to create and store new sounds. The first polyphonic synthesisers were programmable, which meant that they *could* store and recall sounds. A block diagram of such a synthesiser appears in Figure 4.

For a sound to be stored, the settings of the oscillators, mixer, filter and amplifier were stored in random access memory (RAM). Thus the corresponding control voltage of each knob and slider was digitized by an anologue-to-digital converter (ADC) and stored in memory. A previously generated sound could be recalled by reading the RAM and the digital signal sent to a corresponding digital-to-analogue converter to set the control voltage. The control voltage connected to the VCOs, VCFs, VCAs, etc, to re-generate the sound.

The next evolutionary step led to the present generation of synthesiser by making the oscillators, filters, amplifiers, etc, digitally controlled (see Figure 5). Thus VCOs became DCOs, VCFs became DCFs, VCAs became DCAs and EGs became DEGs. Direct digital control over parameters removes the need for converting control voltages into digital signals thus simplifying the internal workings of the synthesiser and reducing cost. Digital control has the added advantage of being more precise and stable than voltage control and in the case of digitally controlled oscillators (DCOs) their inherent stability enables synthesisers to say in tune for longer periods of time.

Synthesisers that are digitally controlled are often referred to as digital synthesisers but don't be fooled into thinking that you are getting some new form of sound generating instrument. In most cases (with some notable exceptions that I will mention later), digital synthesisers use the same old synthesiser building blocks but control them with digital signals instead of control voltages. Some so-called digital synthesisers are just voltage controlled synthesisers with a DCO instead of a VCO (and only one DCO in some cases). Don't get me wrong, digitally controlled synths can sound fantastic (just listen to the latest Oberheims). This is the case when manufacturers implement more than one DCO, DCF, DCA, and envelope generator. Also the quality of the DCF can make a huge difference.

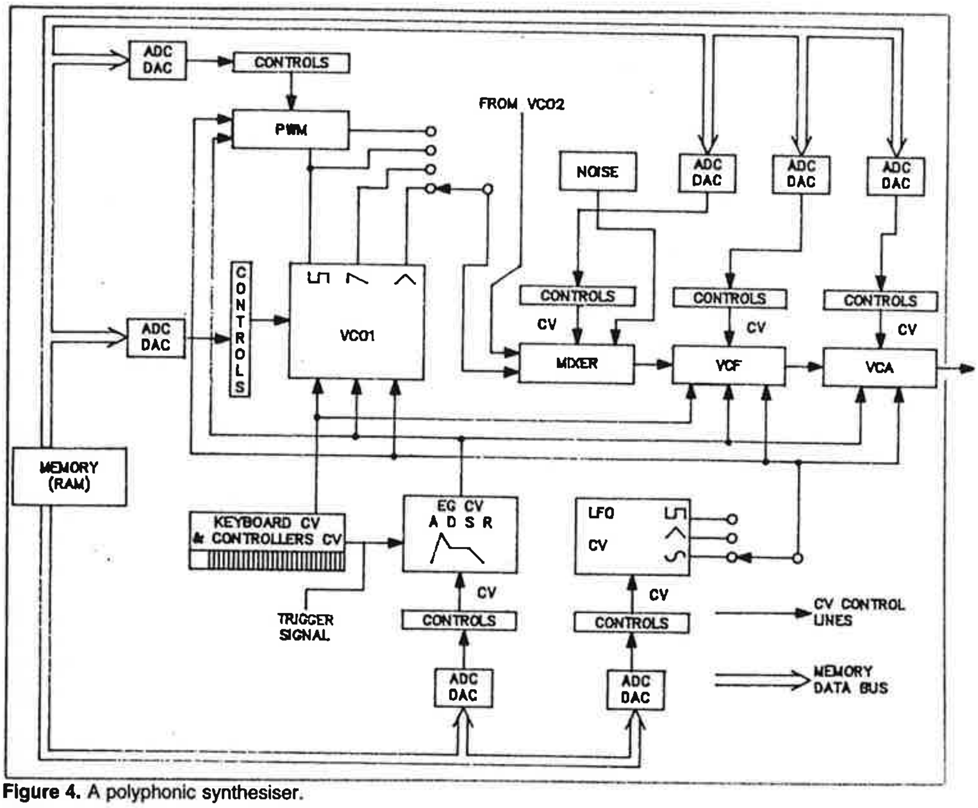


Figure 4. A polyphonic synthesiser.

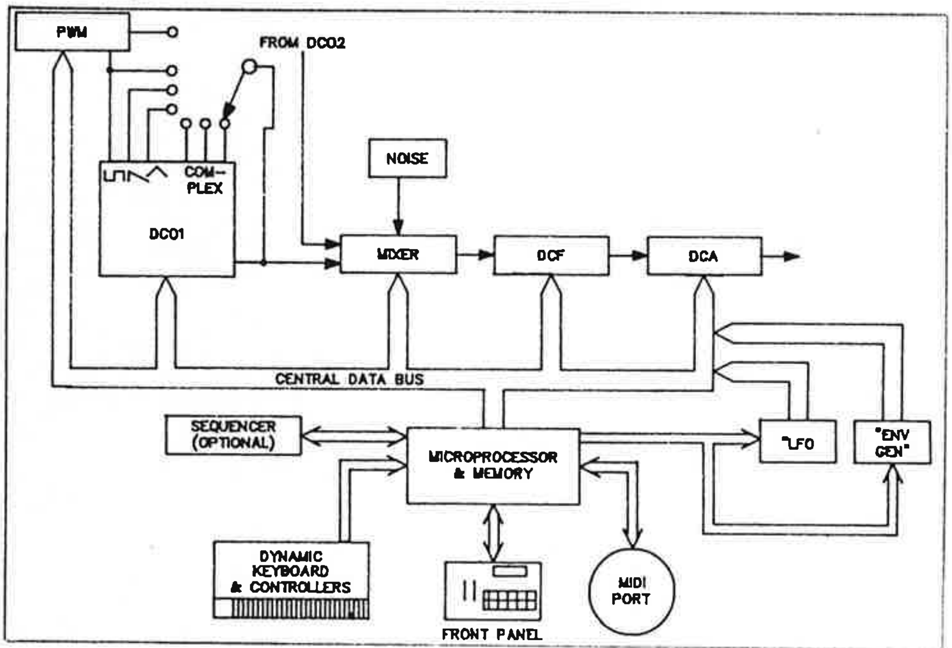


Figure 5. A typical modern synthesiser.

Digitally controlled synthesisers use one control and an array of buttons to modify all the parameters. Instead of a separate control to change each parameter, a particular button selects each parameter and one slider knob controls the value of the lot. The values of these settings are then stored in memory. This way of changing parameters enables synthesisers to be much lower in price largely due to the lower component and manufacturing costs involved. However, the saving in cost is lost in user friendliness, since a panel controlled this way requires each parameter to be called up individually before it can be viewed or controlled.

The ability to control a synthesiser with digital signals also permits it to be controlled remotely, provided there is a way to transfer the digital signals. Remote control allows one synthesiser to be played from the keyboard of another synthesiser or from a computer. To allow the playing information (represented by digital signals) to be transferred, a digital communications bus was developed, this digital communications bus was called the Musical Instrument Digital Interface or MIDI, (see our October issue for all the ins and outs of MIDI). To allow instruments made by different manufacturers to control each other (to communicate), the playing data transferred by the bus was made standard.

Some manufacturers of digital synthesisers have broken away from the four traditional oscillator waveshapes (pulse, square, ramp and triangular) and use new waveshapes (called complex waveforms) to generate sounds. Some manufacturers have even been daring enough to get rid of the traditional synthesiser building blocks and have invented▶

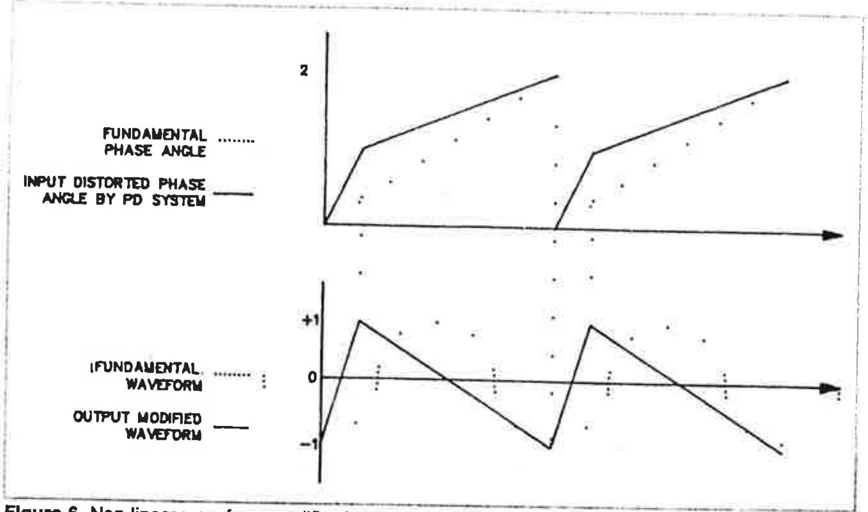


Figure 6. Non-linear waveform modification.

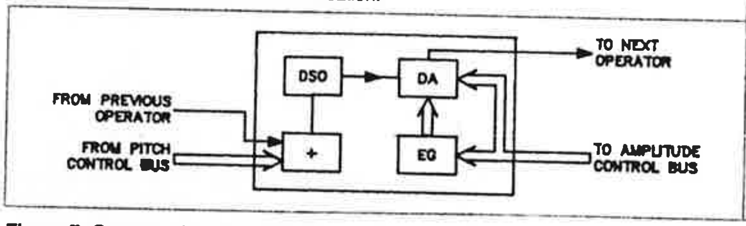


Figure 7. Operators in FM synthesis.

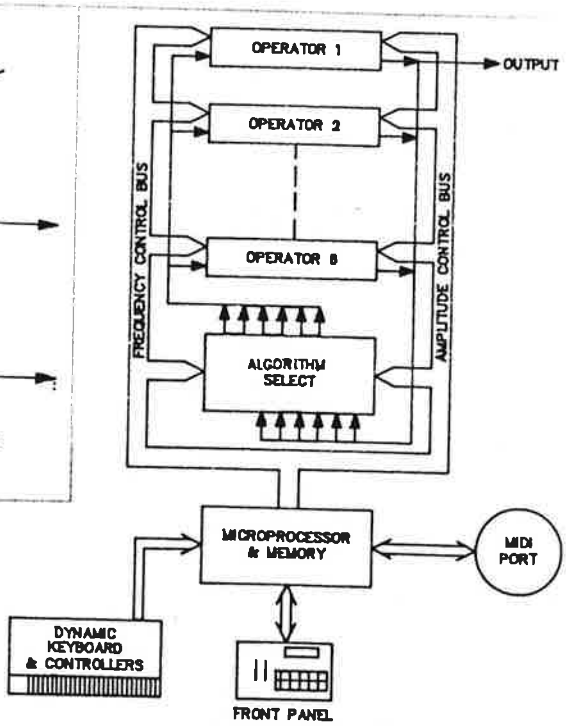


Figure 8. FM synthesis block diagram.

totally new ways of synthesising sound as well as totally new sounds.

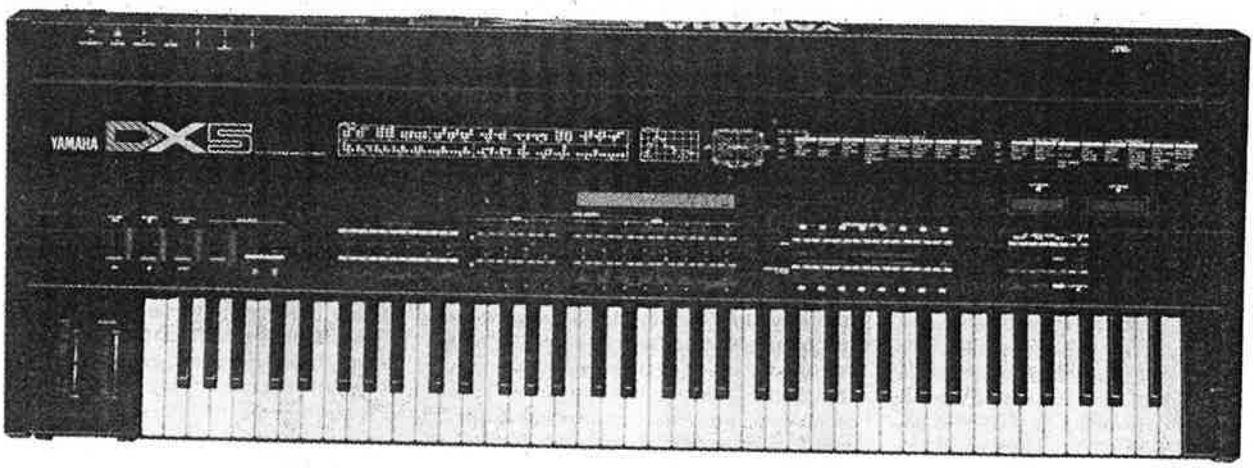
Surprisingly, synthesisers employing new techniques are no more expensive than synthesisers using old techniques. So "where is the rub?" you ask. Well, the problem lies with getting people to change their perception of sound synthesis and adapt their thinking to create sounds in a different way. Casio has succeeded in using a new synthesis method called phase distortion, and made an economical and easy to use synthesiser. Its phase distortion synthesisers use similar building blocks to traditional synthesisers but use complex waveforms and a digitally controlled wave (DCW) converter in the place of a filter. This method of synthesis apparently uses a DCW to modify the phase of an incoming waveform in a non-linear way (see Figure 6) to change its timbre. The Casio synthesisers are extremely easy to create sounds with, largely due to their conveniently laid out front panels.

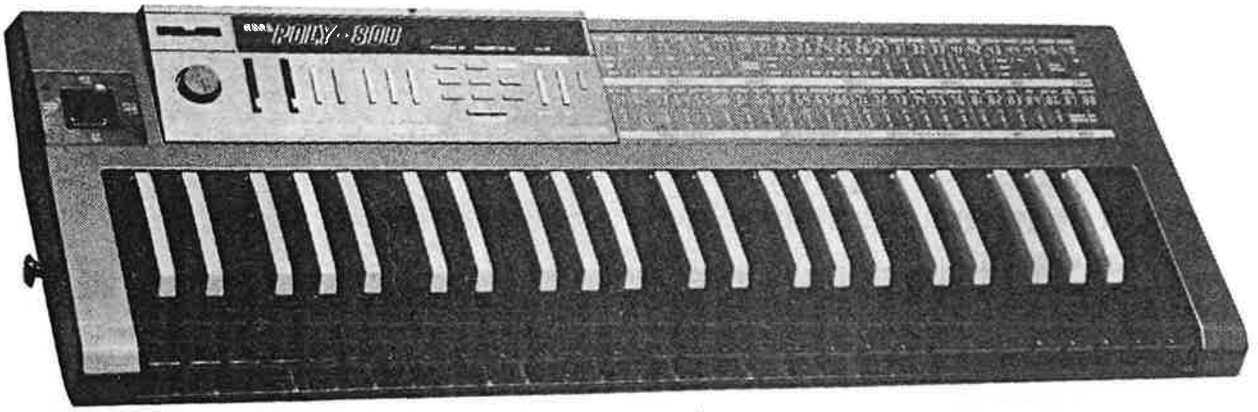
Sequential Circuits has implemented complex waveforms in its Prophet VS synthesiser. This synthesiser allows four complex waveforms to be mixed to create the basic sound source. This sound source is then filtered and modulated by envelopes and LFOs. Combining four complex waveforms is similar to using four DCOs to create a sound (most digital synthesisers only have two). The difference is that complex waveforms give a different sound. Sequential Circuits calls this synthesis method vector synthesis, and it enables complex waveforms to be easily mixed using a joystick. The synthesiser combines excellent sound generating capacity with user friendliness, but at a cost.

FM synthesis

In the early 1980s a breakthrough occurred in the synthesis of sound that was as innovative as Moog's use of voltage controlled oscillators, filters and amplifiers was in the 1960s. The breakthrough was FM (frequency modulation) synthesis. This form of synthesis was a total departure from traditional methods and since it was implemented on VLSI circuits, the FM synthesisers were not only low in price, but were also very reliable.

When the first of Yamaha's FM synthesisers (the DX7 and DX9) hit the market, they were generally capable of making sounds far superior to any synthesiser in the same price range. But since these synthesisers used a different method to generate sound, they were criticized as being difficult to use. In practice, one only had to be slightly open minded and spend a few hours getting familiar with it, and the FM synthesiser was no more difficult to use than any other digital synthesiser. It is not surprising that within





a year the DX7 became the industry standard in synthesisers.

A feature of FM synthesisers which makes them unique is that they do not contain filters. This means that the sounds can be synthesised with software.

FM synthesisers create sounds using four to six operators (see Figure 7), which can be arranged to create different algorithms. Each operator consists of a digital sine oscillator (DSO) and a digital amplifier (DA), both of which are implemented in software. The envelope generator used to control each operator is implemented in software too, therefore the diagram is purely representative. Sounds are generated by digitally controlling the frequency and amplitude of the operators via the DSOs and the DAs. Figure 8 shows a block diagram of an FM synthesiser.

The individual operators are arranged so that the output from one operator (operator 1) can frequency modulate the next (operator 2). Harmonics are generated when the sine wave from operator 1 modulates operator 2. As the output from operator 1 increases in amplitude more harmonics are generated; Figure 9 illustrates this effect. Since the output level from an operator controls the amount of harmonics generated, a filter is not required to control these harmonics.

A group of operators can be arranged in many different combinations to form different algorithms by switching the outputs from some operators to the inputs of others. The amplitude of each of the operators (and thus the harmonic content of the sound) can be controlled by the envelope generators, the playing velocity, and the front panel. Controlling the harmonic content with playing velocity allows the FM synthesiser to copy the dynamics of an instrument with remarkable realism.

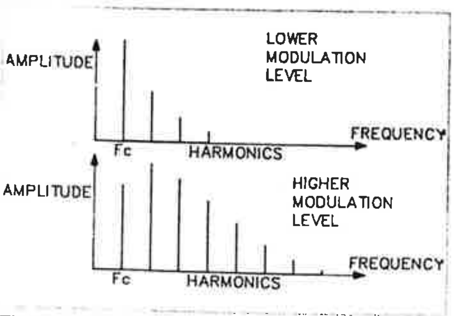


Figure 9. Amplitude causes more harmonics.

Sampling v structured adaptive synthesis

With the advent of sampling (a method where sounds are digitally recorded and musically scaled) any sound, musical or not, can be played from a musical keyboard. This method of sound generation is not actually synthesis. It uses existing sounds and modifies them, whereas synthesisers use oscillators, filters or software, to create sounds.

Sampling has some major drawbacks due to the massive amounts of memory required. For instance, one second of high quality sound (compact disc quality) sampled at a 44.1 kHz rate using 16-bit analogue-to-digital conversion would require 88.2 Kbytes of memory. To perfectly sample a grand piano (88 sample points, 2 second sample time and 50 levels of dynamics, at least) one would require a staggering 755 megabytes of memory!! (Separate samples are required for each level of dynamics because the harmonic content of a sound changes with dynamics.) However, there is an alternative to sampling when recreating real sounds. This alternative is Roland's latest contribution to the world of music technology and is called SAS, structured adaptive synthesis.

The Roland RD 1000 keyboard and MKS-20 rack module both use the SAS technique to create extremely real piano-type sounds. This method of synthesis also allows better dynamic control of harmonics than FM. Structured adaptive synthesis method implements an algorithm to generate the sound instead of using masses of memory. To get this algorithm, the 88 keys of a grand piano are sampled at many different levels of dynamics. All these samples are then fed into a powerful computer, and using some heavy duty signal processing software the relationship between the harmonic content of each of the samples is then worked out. This relationship is transformed into an algorithm, which is then implemented on an integrated circuit.

The SAS method allows 128 levels of dynamics, (some form of interpolation is probably used to get this many levels) whilst even the most expensive sampler can only give a few. About 2 gigabytes (2000 megabytes) of memory would be required to get the equivalent sound quality with 16-bit PCM sampling. All the sonic characteristics of a piano are stored in the SAS algorithm, so when note and velocity information is sent to it from a velocity sensitive keyboard the corresponding audio signal (sound) is output by the algorithm.

Unfortunately SAS requires an algorithm to be worked out for each sound, thus sounds are dependent on what Roland creates. The other drawback is that Roland cannot at present synthesise sustained sounds (such as strings and brass) using SAS. An SAS synthesiser with a full range of sounds would pose a threat to samplers but would not be of any use to people who want to sample their own sounds.

The future

Ideally, a combination of SAS with sampling would be an excellent way to use sounds, as it would give excellent dynamics and user sampling and not require half a ton of memory. To implement this technique a sampled waveform could be fed into a signal processor and have its harmonics modified by an algorithm. Different typical algorithms could be used for typical types of sound, for instance string sounds could have a typical algorithm, as could brass and wind instruments. One could sample a trumpet, then dynamically control its harmonics via a trumpet algorithm, or for more bizzare combinations, sample a hand clap and dynamically control its harmonic content via a piano algorithm. At present it is not feasible to use real time signal processing to control sound in this manner, however, this is probably due to economic and ergonomic reasons.

As to the future I consider that complex waveforms, FM and other new methods of synthesis have the best potential for creating new and realistic sounds. However, a better means of displaying and editing parameters needs to be developed.

The traditional waveforms and filters have been with us for over 20 years, and even though they are now digitally controlled it seems unlikely that many more new sounds can be squeezed from synthesisers designed this way. I suspect that these traditional methods are still with us largely due to people's familiarity with them. As people become accustomed to the new methods of sound synthesis the traditional VCO . . . sorry, DCO, DCF and DCA type synthesisers will be finally out-evolved. ●