Richard Brice looks at the evolution of music synthesis techniques from the Moog to wave tables.

Electronics and music

The analogue music synthesiser owes its genesis to Robert Moog. He invented the first commercial unit, which gave artistic inspiration to composers working in the early post second world war electronic music studios; composers of the stature of Stockhausen, Eimert and Berio. An early modular Moog synthesiser is illustrated in **Fig. 1**.

All analogue synthesisers contain certain cardinal circuit blocks each, originally, more at home in a laboratory than in a music studio. These are described later.

What Moog did was to take everything that the musicians had found useful at that time and build it all in a neat form so that all the various components interfaced properly.

Although many of the first customers were experimental musicians, it wasn't long before advertising agencies latched onto Moog equipment. The synthesiser turned out to be the perfect way to bridge the gap between sound effects and music.

Voltage-controlled oscillator

Fundamental to the whole concept of analogue sound synthesis is the voltage-controlled oscillator. This may be controlled by a switched potentiometer, perhaps arranged like a conventional musical keyboard. Alternatively, it may be controlled by a constantly variable voltage, thereby providing a sound source with endless portamento like the Ondes Martenot and the Theremin.

Control voltage for the VCO may also be controlled by the output of another oscillator resulting in one waveform frequency-modulated by another. And perhaps this resultant waveform might be made to modulate a further source! By this means, the generation of very rich waveforms is possible.

Design of voltage-controlled oscillator for audio synthesis applications is not altogether straightforward . This is because the oscillator must be made to swing over the entire audible range – a frequency range of some eleven octaves.

Often, the oscillator is a sawtooth generator type like that illustrated in **Fig. 2**. Note that the rate at which the integration capacitor charges in the feedback loop of the op-amp is variable by means of the adjustable current source. The circuit must itself generate the ramp termination pulse.

The self-generation of the termination pulse occurs due to the action of the comparator circuit, which has a pre-set negative voltage on its positive input terminal.

Once the ramp circuit output voltage – which is shown supplied to the comparator's inverting input – has reached this threshold, the comparator changes state. This closes the electronic switch shown connected across the integration capacitor. The charge-integrating capacitor is thereby shorted and the ramp terminates, allowing the whole process to start once again.

It's worth pointing out that there is nothing to stop an external pulse being sent to this oscillator in the manner shown in Fig. 2. This is often done in commercial synthesisers, a technique known as 'synching'. By setting the natural oscillation of one oscillator to a different frequency from that of its externally supplied synching pulse, some very complex waveforms are obtainable.

One major complication with voltage control for synthesisers is due to the nature of the relationship between control voltage and frequency.

From a technical point of view, the easiest control law to generate is linear, or V/F=k, where k is a constant. But from a musical standpoint, a law that relates a constant change in pitch (frequency) to a constant change in control voltage is far better. This is a logarithmic law and considerable complication exists within most analogue synthesisers to alter the control law of the VCO to that suitable for musical applications.

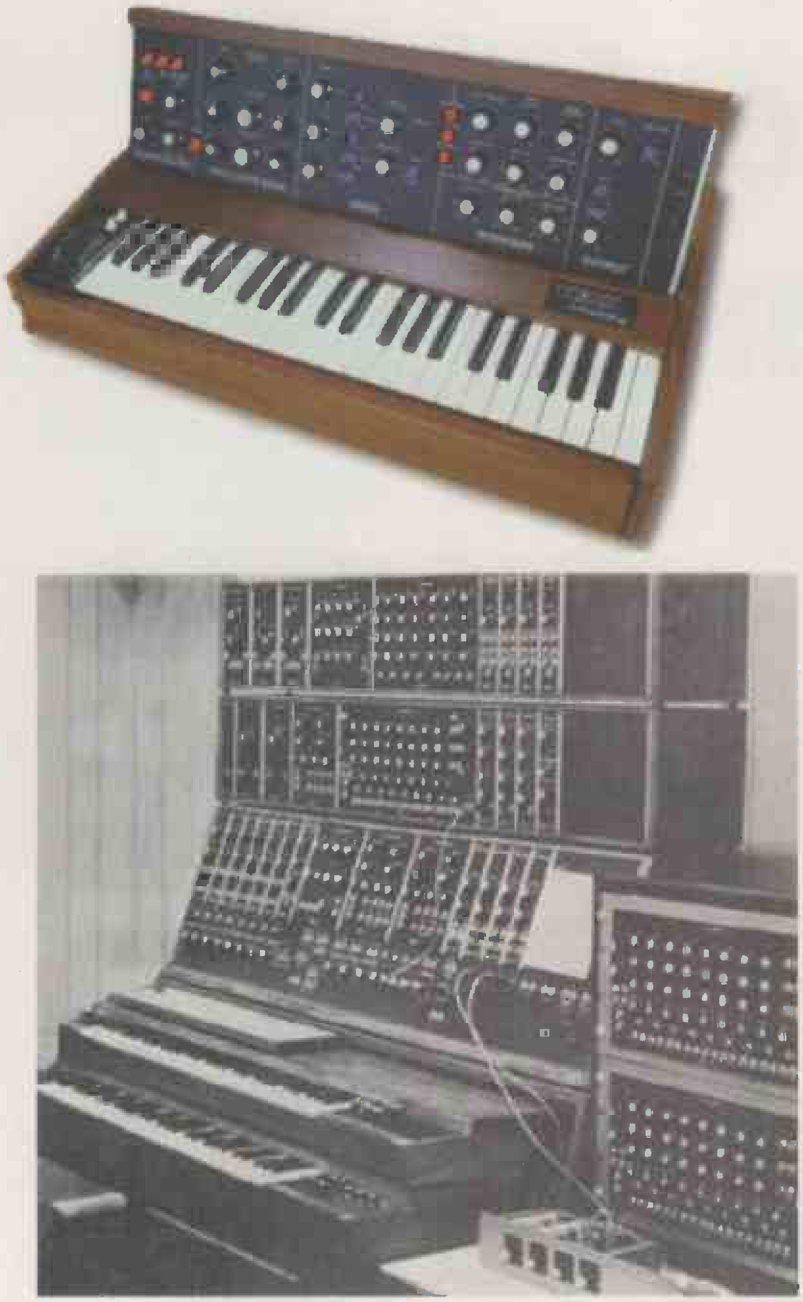

Fig. 1. Moog took al the electronic building blocks that musicians were using at the time and incorporated them into one instrument. Shown are an early Moog and the Minimoog.

Voltage controlled filters

A voltage-controlled filter, or VCF, is a frequency selective circuit that may be made to alter its cut-off frequency under the control of an externally applied voltage. The most usual type in synthesiser applications is the voltage-controlled low-pass filter, which is the most useful in musical applications.

A simplified schematic of a VCF is given in **Fig. 3**. This unusual circuit operates like this: The cut-off frequency is programmable by means of the current sink 'tail' which may be made to vary its sink current as in the manner of a normal current mirror. This current divides between the two cascode pairs and into the collector loads of Tr_3 and Tr_4; themselves another cascode pair.

At very low sink currents, the value of the collector loads Tr_1 and Tr_2 will be relatively high. This is because the output impedance from an emitter follower – which is what these loads are – is inversely proportional to emitter current.

Similarly, the transconductance of the differential pair will be low too. The gain of the stage will therefore be the product of the low-ish transconductance of the pair, multiplied by the relatively high impedance of Tr_1 and Tr_2 collector loads.

At high tail current, these conditions alter so that the transconductance of the differential cascode pair will be high, but the impedance of the collector loads – from which the signal is taken differentially – will be low. The overall low-frequency gain of the circuit will thereby remain constant; irrespective of changes in tail current.

What will alter however will be the available slew-rate of the amplifier which will be severely limited by the ability to charge and discharge C_1 and C_2 at low standing currents. At high standing currents, this situation will improve, thereby increasing the bandwidth of the circuit.

Sometimes practical circuits repeat the cascode structure of this circuit many times to increase the number of poles in the filter, often earning the circuit the name 'ladder filter'.

An important further function of the circuit is its ability to feed a proportion of the signal back to the other side of the differential pair. Note that this is not negative feedback but positive feedback. This has the effect of increasing the Q of the circuit – especially near the turnover frequency. It is therefore possible to produce a range of responses like those shown in **Fig. 3**. These offer a gamut of musically-expressive possibilities by permitting the possibility of imprinting high-Q formats on the fundamental wave.

Sometimes this Q control allows the possibility to produce instability at extremes of the control's range, thus turning the VCF into another, somewhat unpredictable, VCO.

Envelope generation

Real musical sounds have a dynamic envelope. Turning a primitive synthesised musical waveform into a musical sound involves imprinting attack, sustain and decay envelopes onto the fundamental sound source. This manipulation requires a controlled multiplication function. The attack is the speed with which a signal is multiplied from zero – i.e. silence – to a constant sustain level.

The rate at which the sustain level decreases back to zero once the keyboard key is released is the decay, or release period. The multiplication function is performed by a voltage-controlled amplifier or VCA.

Voltage-controlled amplification

Analogue multiplication techniques involve the use of current 'steering' via two alternative circuits to achieve such a multiplication. The circuit in **Fig. 4** demonstrates the general principle.

Essentially the audio signal at the base of the lower transistor is turned into a current in the collector circuit of the same transistor by the transistor's transconductance mechanism. Notice that a resistor is included in the emitter circuit to linearise this current. This collector current divides into two circuits, through Tr_1 and Tr_2, the ratio of the current being dependent on the voltage on the base of Tr_1. If this voltage is higher than the signal on Tr_2 base (V_k), current will flow predominantly through Tr_1 and appear as a voltage signal on R_r. If it is lower than the signal on Tr_2 base, current will flow predominantly in Tr_2.

By altering the value of the voltage on the base of Tr_1, a variable proportion of the original signal voltage (suitably inverted) can be recovered across R_r.

Attack-sustain-release generator

The controlling signal is derived from an envelope generation circuit sometimes known as an attack-sustain-release generator, or ASR, and an example is illustrated in Fig. 5.

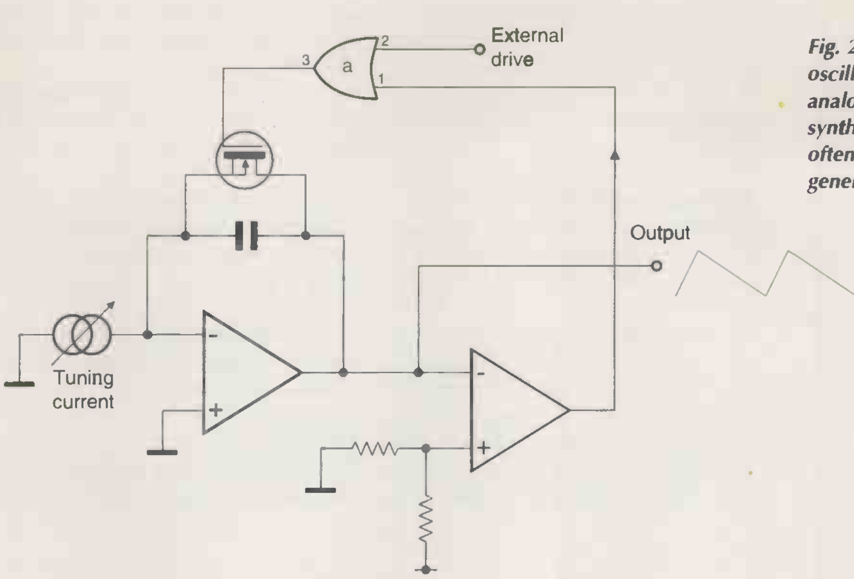

Fig. 2. The oscillator used in analogue synthesisers is often a sawtooth generator.

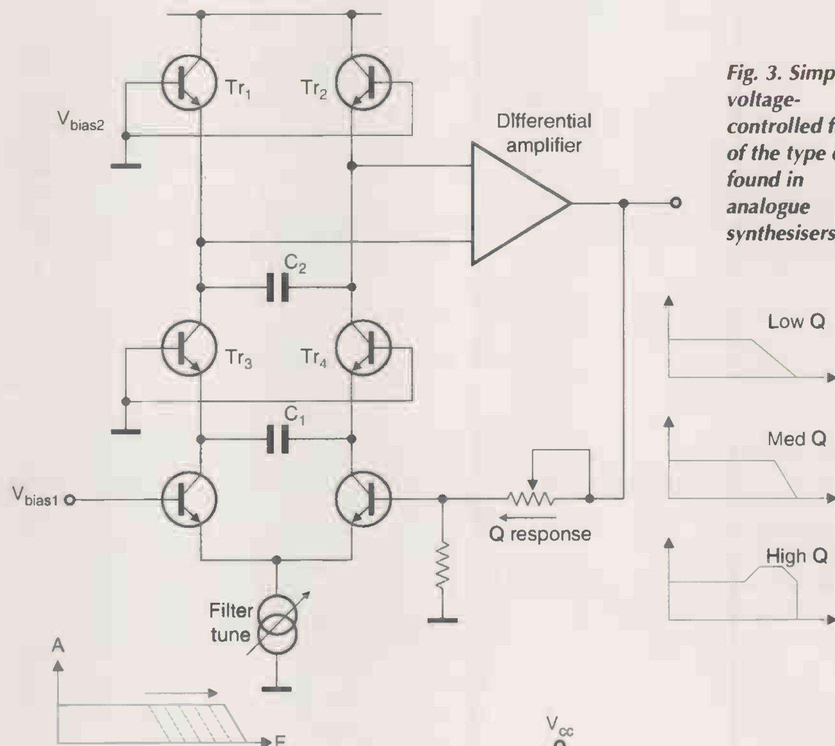

Fig. 3. Simplified voltage-controlled filter of the type often found in analogue synthesisers.

When the key closes the control voltage fed to the VCA rises at a rate predetermined by the setting of VR_1 and its interaction with C_1. Ultimately, the control voltage rises to the value set on VR_2 which determines the sustain level. Finally, once the key is released, the control voltage will fall as determined by the setting of VR_3 and C_1.

Low-frequency oscillator

The low-frequency oscillator is often a ramp generator, or sometimes Wien bridge type. External voltage control is seldom provided; instead the function of this oscillator is to control either VCA or VCO in order to pro-

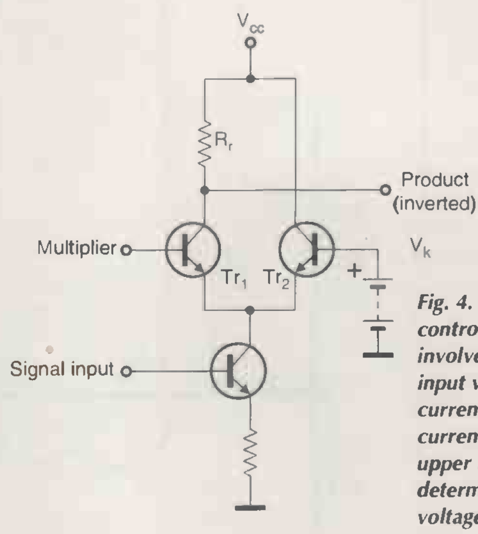

Fig. 4. Voltage-controlled amplifier involves turning the input voltage into current. The ratio of current in the two upper transistors is determined by the voltage at the base of Tr_1

vide tremolo or vibrato respectively.

Alternatively the LFO is often used to control a VCF, in order to obtain a protean musical timbre – one of the musical hallmarks of analogue synthesis.

Analogue noise generators

Analogue noise generators are relatively simple; often as no more elaborate than a resistor followed by a high-gain amplifier.

Another source of noise is a reverse-biased diode, followed by a high gain amplifier. This is a good generator of noise because the noise in a diode is relatively high. This is due to the avalanche effect as high-velocity electrons, entering the n-region from the p-region, liberate others in the valence bands of the n-region – a process that is inherently noisy.

Furthermore it side-steps a problem with a resistor circuit. In order to get a high enough noise voltage, a very high value of resistor must be used. This invites the possibility of other signals being electrostatically or electromagnetically coupled into the circuit; hum being especially problematic in this respect.

The diode, on the other hand, generates a comparatively high noise voltage across a low impedance.

Patching

In order to produce a practical, usable musical sound from an analogue synthesiser, each of the circuit blocks mentioned above must interconnect.

A simple patch – the name used for a particular interconnection scheme – is illustrated in **Fig. 6**. Notice that the keyboard controls the VCO to generate a pitch; the VCO is followed by the VCF to impart a character onto the basic sound source, and this is driven by the output of the LFO to create a changing formant; and the output of these modules is passed to the VCA block where the trigger-signal from the keyboard controls the ASR generator.

Figure 6 illustrates but one possible patch. Remember, the power of the analogue synthesiser lies in its ability to cause the various sound and noise sources housed within it to interact.

Commercial synthesisers differ greatly in the flexibility they present to the user; in terms of being able to route the various signals and thereby have control over the pattern of possible interactions.

Some synthesisers offer very limited re-patching, others permit virtually unlimited flexibility, so that practically any signal may be used as a parameter to control another process. Electrically this switching function may be provided by hardware switches, by electronic switches or even by a traditional telephone-style jack-field; as illustrated in Fig. 1.

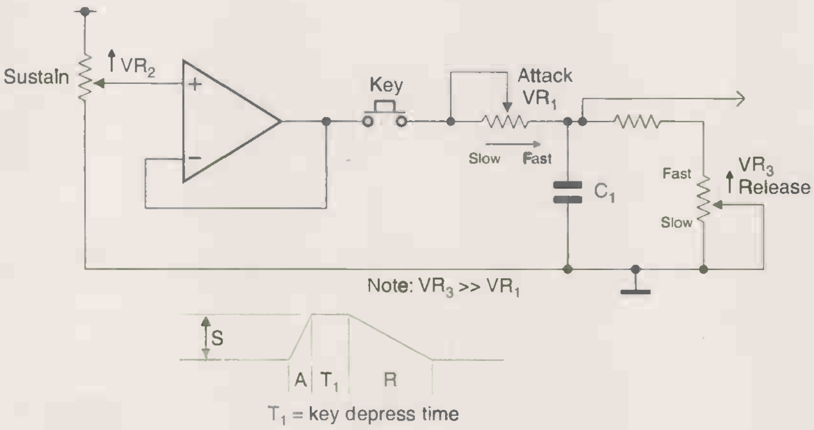

Fig. 5. To produce realistic music synthesis, an attack-sustain-release generator is needed to shape the note.

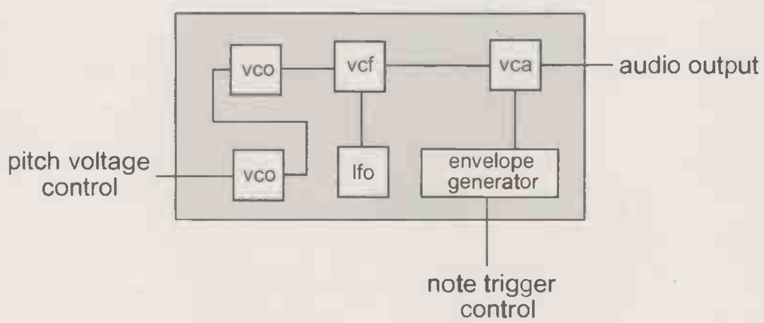

Fig. 6. How the circuit described in previous illustrations interconnect – known as a 'patch'.

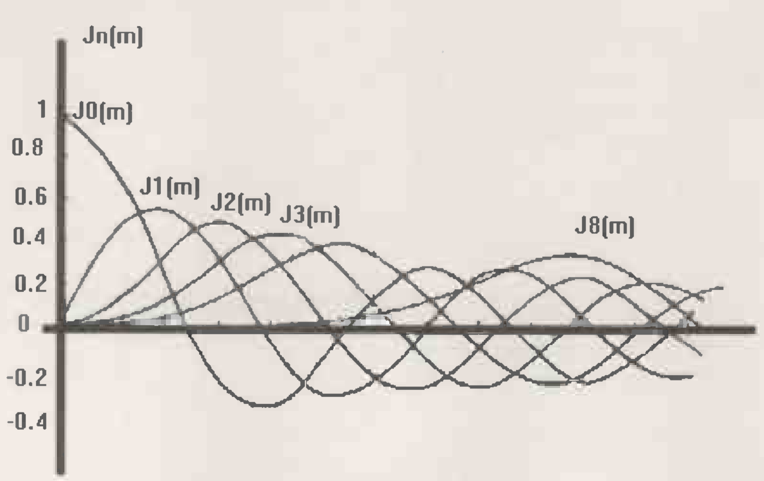

Fig. 7. A collection of first-order Bessel functions.

Digital synthesis techniques

One of the lessons learnt from analogue synthesis – and especially from controlling one oscillator from another – was that extremely complex timbres may be generated by relatively simple means.

In 1967 Dr John Chowning – a musician who had studied with Nadia Boulanger in Paris and subsequently set up a music composition programme at Stanford University – realised the possibilities of using frequency modulation to imitate complex musical sounds.

The complex nature of fm sidebands seemed to suggest that here might be a technique where complex musical tone-structures could be built up using one or two oscillators. Chowning was right, but he had to wait for the advent of digital techniques to guarantee the necessary stability, predictability and repeatability required for this synthesis technique.

Frequency modulation in detail

The sideband frequencies produced around a frequency-modulated carrier are related not only to the deviation as a proportion of the modulating wave frequency – the so-called modulation index – but to all the harmonics of the modulating frequency as well.

The structure of the resulting sidebands issuing from all these variables are determinable using mathematical relationships known as Bessel functions.

A collection of first-order Bessel functions is illustrated graphically in **Fig. 7**. These illustrate the harmonic content of the fm process. The abscissa value represents the modulation index – which is equal to the frequency deviation divided by the modulation frequency. The ordinate represents the amplitude value of the carrier J0 and the first eight harmonics, J1 to J8 respectively.

Intuitively, you can see that when the modulation index is zero – i.e. no modulation – all the energy in the resulting output is concentrated in the carrier: That is, J0 is unity and all the other functions J1 through 7 are zero.

As the modulation index increases slightly, J0 declines and J1 climbs. Note that J2, 3, 4, etc., climb more slowly. This illustrates why low modulation index fm has a spectrum similar to that of amplitude modulation, with only first-order sidebands. It also illustrates that a signal modulated with a very high modulation index will have a very rich spectrum indeed.

At this stage, you may find yourself asking, "OK, I can see frequency modulation produces complex harmonically related structures around a carrier; but doesn't that mean that the carrier would have to be different for each note of the keyboard? And what happens to the lower sidebands since these don't normally exist in the case of a musical sound?"

A carrier at zero hertz?

The answer to the mystery is that the carrier used in fm synthesis is zero hertz! It is the modulating signal that determines the fundamental frequency, and the depth of the modulation, or modulation index, which may be manipulated to produce the harmonic structure above the fundamental.

Look at the figure of Bessel functions and imagine that a modulating frequency is employed that relates to the note middle C; 261Hz. Suppose a deviation is chosen of 261Hz, that is with a modulation index of 1.

From the curves in Fig. 7, it's clear that at m=1, J1 equals about 0.45 and that J2 equals about 0.1, all the other harmonics still remaining very small at this low index. The resulting sound – suitably synthesised, amplified and transduced – would be a note at middle C with a second harmonic content of about 22%. The zero-hertz carrier would, of course, effect no audible contribution.

This might make a suitable starting point for a flute sound. Now imagine a much greater modulation index being used: m=3, for example. Reading off from the curves, it shows that in this case; J1=0.34, J2=0.49, J3=0.31 and J4=0.13. This would obviously create a much richer musical timbre.

If the index of modulating signal is changed over time, it is possible to create musical sounds with extremely rich musical transients, full of spectral energy which then segue into the relatively simple on-going motion of the resulting note. In other words, it is possible to create synthetic sounds just like real musical sounds, where a 'splash' of harmonics is created as the hammer hits the piano string, or the nail plucks the string of a guitar, or as the first breathy blast of air excites the air inside a flute. All of these decay quite rapidly into a relatively simple on-going motion.

Each of these effects may be synthesised by generating a modulating signal that initiates the carrier modulating with a large deviation, creating a rich transient part to the sound. This then decays to a level where it causes only relatively small deviation, generating a relatively pure on-going sound.

A complication arises as a result of the choice of the zero frequency carrier. Just because the carrier is zero hertz, it doesn't stop there being sidebands in the negative frequency region. These 'fold back' into the positive frequency region and destructively interfere or constructively reinforce the sidebands present there already. These too have to be taken account of in generating fm synthesis algorithms.

But there are other more interesting uses for the lower sidebands and these embrace the use of a very low frequency carrier instead of one of zero frequency. When a low frequency carrier is used, the negative frequency sidebands fold back into the positive region to be interleaved with the positive frequency sidebands. In this manner, even more complicated timbral structures may be built up.

Fig. 8. Yamaha's DX7 – the first synthesiser to use frequency-modulation techniques.

The DX7

FM synthesis techniques were first used commercially in the Yamaha *DX7* keyboard, **Fig. 8**. At the time, it caused a sensation in synthetic music.

So successful was the method and so excellent Yamaha's implementation, that the 'sound' of the fm synthesis *DX7* dominated the popular music of the eighties. Artists who used it include Talking Heads, Brian Eno, Cabaret-Voltaire, D:Ream, Front 242, Depeche Mode, U2, A-Ha, The Cure and Enya.

Frequency modulation is remarkable in that it represents the high-point of pure, electronic sound generation. It still remains the sound technique employed in cheaper PC sound cards.

Sampling

Digital sampling systems rely on storing high-quality, digital recordings of real sounds and replaying these on demand.

The tough problem in sampling is the amount of memory it requires. Sampling is well suited to repetitive sounds like drums and other percussion instruments. This is because the sample is mostly made up of a transient followed by a relatively short on-going sustain period. As such, it may be used over and over again so that an entire drum track could be built from as few as half-a-dozen short samples.

Problems arise when long, sustained notes are required; like the sounds generated by the orchestral strings. The memory required to store long sustained notes would be impossibly large.

Pure sampled-synthesis systems rely on 'looping' to overcome the limitation of a non-infinite memory availability.

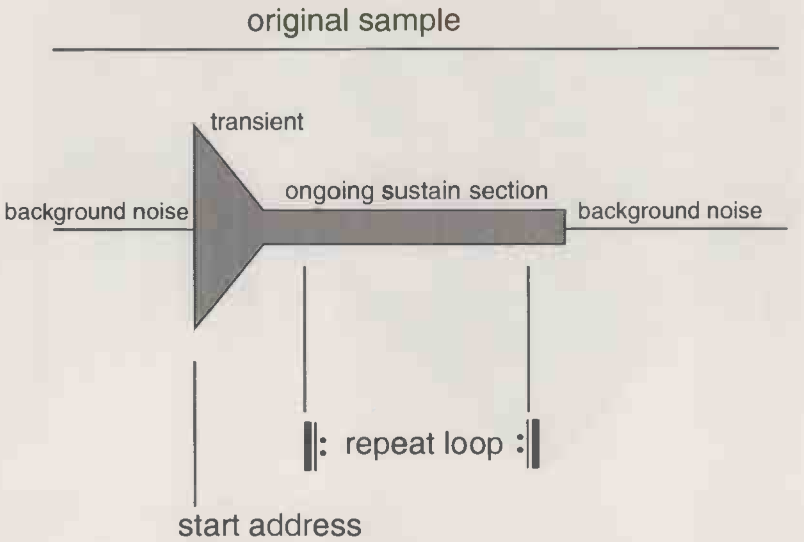

Fig. 9. To reduce the amount of memory needed, modern samplers repeat a section of the sustain period.

But a loop is hard to achieve without either sacrificing the starting transient of the note or having it repeat over and over as the loop repeats, resulting in a 'hiccup'. Modern samplers provide start and loop-point memory address programming for just this reason, as illustrated in **Fig. 9**.

An important part of sampling technique involves the use of one acoustic sample over a group of notes. Replay of the sample at the appropriate pitch is achieved by the applicable modification of the read-clock frequency in exactly the same way as pitch-shifting (described in my last article).

In theory, one sample may be used at every pitch. However, due to distinctive formants imprinted on the sound by the original voice or instrument, this can result in an unnatural sound if the transposition is taken too far. This effect is known as Munchkinisation, named after the under-speed recorded singing Munchkins in The Wizard of Oz.

The effect is ameliorated by recording several samples at different pitches and assigning these to various ranges of transposition. Furthermore, samples are usually recorded at various different dynamic levels from very quiet (pianissimo or pp) to very loud (fortissimo or ff).

The sampler uses these different samples and a mixture of samples at different dynamics points, to achieve touch-sensitive dynamics from the controlling keyboard. Good sampling technique therefore involves judicious choice of looping-point, transposition assignments and dynamics assignments. This is no mean task and successful sampling programmers are very skilled people.

Wave-table synthesis and the like

The technique known as wave-table synthesis involves sound samples in combination with a number of special techniques. These include carefully edited looping points, pitch shifting, interpolation and digital filtering to reduce this prohibitive memory requirement.

The system is known as wave-table because the synthesiser employs a group – i.e. a table – of different sound segments which it 'looks-up' and utilises as required.

Other proprietary synthesis algorithms have included LS Sound Synthesis from Roland and Dynamic Vector Synthesis from Yamaha. Both these techniques use what might be termed 'empirical' sound synthesis techniques. They involve a mixture of sampling and tone generation – by additive of FM synthesis – to arrive at their final result.

Sound synthesis like this tends to be a non-purist subject based on subjective 'feel' rather than mathematical precision.

Modern trends in synthesiser design

The method of synthesis employed in the 'classic' three-VCO and filter synthesiser like Moog's *Minimoog* is sometimes referred to as subtractive synthesis. This term is a little misleading. It relates to the situation whereby the fundamental tones produced by the VCOs are of a complex harmonic structure in the first place. These are subsequently filtered and modified in the filter and envelope circuits, thereby simplifying the harmonic structure by 'subtracting' harmonics.

In exactly the same way, the voice may be considered a subtractive synthesiser. This is because the vocal chords produce a crude but harmonically complex 'buzzing' fundamental tone which is capable of being modulated in pitch alone and on which is imparted the filtering formants of the mouth and nasal cavities.

Nature, always a parsimonious architect, adopted such a system because it is efficient. However, while subtractive synthesis is a very potent arrangement and fm techniques and sampling all offer their particular advantages, there is another approach that offers the ultimate in synthesis technology. It is known as additive synthesis.

Additive synthesis

I mentioned additive synthesis in an earlier article when looking at the classic tone-wheel Hammond organ.

Hammond's implementation was Neanderthal by today's synthesis standards. But despite the crude nature of their synthesis algorithm – and that simply means the limited number of harmonic partials they had available – additive synthesis represents the 'ultimate' synthesis algorithm because it's Fourier analysis in reverse!

Of course, Hammond's engineers didn't adopt a limited number of partials out of choice. They were thwarted by the technology of their period. Very-large-scale integration, i.e. VLSI, integrated circuits offer the possibility of the literally huge number of oscillators necessary to produce convincing synthetic sounds by this means.

Physical modelling

Instead of concentrating on the sound events themselves, some workers in synthetic sound have chosen a different route; to model mathematically the results of a physical system.

Using this technique, a guitar sound is modelled as a stretched string, with a certain compliance, mass and tension. These are coupled to a resonant system excited into sound by means of a excitation function. The excitation function itself is modelled to represent the action of plucking a string.

That this technique is computationally intensive is something of an understatement. However, it has already proved to be a fruitful synthesis approach. It also offers the possibility of modelling new sounds of impossible instruments, which is especially interesting for composers searching for new sounds of ostensibly 'physical' or 'acoustical' origin. ■