Proceedings of The Institute of Acoustics

SYNTHESIZING COMPLEX AUDIO SOUNDS WITH FM

R G Payne

Department of Physics, University College, Cardiff.

INTRODUCTION

Frequency Modulation (FM) first became widely recognised as a possible sound synthesis technique back in 1973 when Chowning's paper [1] was first published, but only in the last few years has rapid growth in VLSI electronics technology made it possible to fully exploit this method. Today, many commercial synthesizers and sound synthesis systems use the FM algorithm, and it is hard to see a future in this field without it.

Why has FM suddenly become so important? Established synthesis methods, such as fixed-waveform and additive synthesis (eg, [2]), are easy to understand conceptually, as well as being relatively easy to implement in hardware. However these, and many others, more often than not produce musically uninteresting spectra. Many of the older (pre-digital) commercial synthesizers consisted of just one or two oscillators, relying on resonant filters and other effects to add interest. The FM technique, however, not only allows us to generate enormously rich spectra using a relatively small amount of hardware, but also provides a very compact mathematical description of those spectra.

One major drawback to using FM as an audio sound generator is that unlike, say, additive synthesis, it is not generally intuitively obvious what effect changing the various controlling parameters will have on the sound being generated. Programming an FM Synthesizer at present appears to be something of a trial and error process.

The need has therefore arisen for a general scheme to allow the decomposition of sampled sounds into the controlling parameters of an FM generator; once the output is close to the desired sound, a certain amount of fine-tuning (by ear) of the system may be necessary.

This paper describes a general analysis method, based on the discrete Hilbert Transform, which is currently being implemented on an IBM microcomputer, using a commercial FM digital synthesizer as the sound source (in this initial stage of the project). The method described was first proposed by Justice in 1971 [3].

The analysis consists of an interative procedure, each pass of which produces the envelope and frequency parameters for one 'layer' of oscillators. (The idea of 'layering' or 'chaining' of modulators is entirely natural to the FM algorithm.) The original sampled signal under analysis is reduced with each iteration, until either it is considered to be below a noise threshold, at which point analysis stops, or until further analysis would produce parameters for more oscillators than are available on the synthesizer.

SYNTHESIZING COMPLED AUDIO SOUNDS WITH FM

The output parameters of the analysis procedure are then mapped into the input parameters of a Yamaha DX7 synthesizer [4]. For example, the input envelope shaper data for the DX7 consists of a series of integers with range 0 to 99; the envelope functions extracted from the analysis must be normalized for this convention.

FREQUENCY MODULATION

The simplest form of an FM wave is given by equation (1) below, which describes a sinusoidal signal (the carrier) whose frequency is modulated by another sinusoid (the modulator).

$$s(t) = A(t).\sin(ct+I(t).\sin(mt) \qquad (1)$$

where: $A(t)$ = carrier amplitude envelope function,
c = instantaneous carrier frequency,
$I(t)$ = modulation index,
m = instantaneous modulator frequency

Clearly, setting I to zero in equation (1) results in a spectrum with one frequency component at c; non zero values of I result in sidebands about the carrier at intervals equal to m. The amplitude of each of the spectral components (equation (2)) is determined by the Bessel function $J_x(I)$ where x is the sideband number, and I is the moduation index, hence the overall band-width of the signal also depends in I.

$$s(t) = A(t).[J_o(I).\sin(ct)+J_1(I).\{\sin(ct+mt)-\sin(ct-mt)\} + J_2(I).\{\sin(ct+2mt)-\sin(ct-2mt)\}+..... \qquad (2)$$

The unexpected richness of the FM sounds generated even by simple systems, (such as in equation (1)), arises when large values of I are used. Then, the Bessel functions give us sidebands having negative amplitude as well as sidebands of negative frequency. These latter are of special interest because they are reflected about the amplitude axis (see Figure 1) back into the positive frequency domain, and combine additively with components already present. Note that this situation does not arise in radio communications using FM, where the carrier frequency is typically 100MHz, the modulating frequency in the audio range, and modulation indices generally very small.

Although the large number of sidebands generated, and their positions within the spectrum yields musically interesting sounds, the spectrum, by virtue of equation (2) is impossible to predict intuitively. In certain special cases however, (where c/m is a ratio of integers, and I is small, for example), a simple, harmonic spectrum results - Figure 2.

It is clear that equation (1) lends itself to extension. Nesting of modulators, where the single modulating sinusoid is itself further modulated, is demonstrated in equation (3).

$$s(t) = A(t).\sin(ct+I_1(t).\sin(M_1(t)+I_2(t).\mathrm{Sin}(M_2t+ .... \qquad (3)$$

SYNTHESIZING COMPLEX AUDIO SOUNDS WITH FM

Figure 3 shows how the Yamaha DX7 commercial synthesizer further extends the concept of combining oscillators by summing the outputs of 'vertical' chains - equation (4)

$$s(t) = s1(t)+s2(t)+s3(t)+\ldots \quad (4)$$

where s1, s2, s3 etc correspond to outputs s(t) given in (3)

It is clear that these methods of combining oscillators, whilst extending the sound-space beyond that possible using equation (1), also further compound the problem of predicting the resulting spectrum.

FM SYNTHESIZERS

The Yamaha DX7 commercial synthesizer mentioned above is a 'state of the art' machine, employing 6 digital oscillators (termed 'operators') which may be connected in any of 32 combinations, yielding a vast range of possible timbres. Each of the operators has 21 input parameters associated with it (eg, frequency, envelope levels, etc,) and there are a further42 parameters common to all 6 oscillators. Many of these may be ignored when 'sound-building'; essentially all that is required is to select an arrangement of operators (an 'algorithm') and set the frequency and envelope parameters for each operator.

The DX7 is also equipped with MIDI (Musical Instrument Digital Interface) enabling communication with computers, or similar synthesizers. This system is useful for transmitting information at either the 'event' level (ie note data) or at the timbral level (ie, voicing data).

A GENERAL ANALYSIS/RESYNTHESIS ALGORITHM FOR FM

Our starting point is to assume we have obtained a set of N samples, Sr(n) from a continuous signal by sampling at greater than the Nyquist rate. We also assume the signal has an amplitude envelope which varies slowly relative to the signal frequency (ie below audio frequencies). We wish to extract from the samples, information corresponding to modulation index parameters of an FM algorithm, eg, that given in equation (3) for, say 4 modulators. The process used is described below.

Associated with the real sampled signal Sr(n) is a corresponding unknown imaginary part, Si(n). Together, these components form the complex function termed the 'analytic signal'. The analytic signal may be generated using the Discrete Hilbert Transform [6], which is the convolution of $-1/\pi t$ with Sr(n). Fortunately, it is possible to apply a short-cut, and it is not necessary to compute the convolution itself [5]:

1. Fast Fourier Transform Sr(n) to yield F{Sr(n)}
2. Set the negative frequencies to zero. In the discrete case this is equivalent to setting:

$$F\{Sr(n)\} = 2F\{Sr(n)\} \quad \text{for } 0<n<N/2$$
$$F\{Sr(n)\} = 0 \quad \text{for } N/2<n<N$$
$$F\{Sr(n)\} = F\{Sr(n)\} \quad \text{for } n=0,N/2$$

SYNTHESIZING COMPLEX AUDIO SOUNDS WITH FM

3. Inverse fast Fourier Transform to yield the complex analytic signal
   $S(n)=Sr(n)+jSi(n)$ (6)
   where j is the complex operator

It is now possible to extract the modulus and phase of the signal by rewriting (6):

$$S(n)=E(n).\exp(X(n)), \quad (7)$$
where $E(n) = \sqrt{Sr(n)^2 + Si(n)^2}$

and $X(n) = \arctan(Si(n)/Sr(n))$

so that:

$$Sr(n)=E(n).\cos(X(n)) \quad (8)$$

Thus our original real sampled signal has been written in terms of its phase and modulus. We can identify frequency with phase by fitting a best line to the X(n) array, and can interpret E(n) as the amplitude envelope of the signal.

If this analysis is applied to an FM signal which was generated by an algorithm conforming to equation (1), then E(n) and X(n) would correspond to the envelope and frequency functions of the carrier. If we then subtracted the average carrier phase C, (from which the carrier frequency, ie, the slope of the average phase of X(n) is derived), the residue must be the modulator function. We can then apply the analysis again, treating the residue (X(n)-C) as the new real signal sr(n). The results of the second analysis yield the required parameters of the modulator, and indeed the iterations could continue, unravelling further 'layers' of modulators if required. However, if the original signal does correspond exactly to equation (1), then the phase residue after the second iteration will, of course, by zero.

Resynthesis consists simply of associating the extracted oscillator parameters with those of the FM Synthesizer being used. Figure 4 shows graphically the analysis/resynthesis method described.

IMPLEMENTATION

The data acquisition and all processing work is carried out using an IBM PC-AT personal computer having 512k-bytes of RAM and 20M-bytes of fixed disk storage.

Although the system is designed to accept sounds from any source, the DX7 synthesizer is currently being used (at this, the development stage of the project) as it represents a near ideal FM sound generator system.

Sampled data is obtained through a dedicated 16-bit ADC which runs at an adjustable sampling rate of up to 100kHz. The 16-bit samples are stored in a contiguous 256k-byte area of RAM by an 8086 assembly language program.

A Fast Fourier Transform program [7] operates on 4k-byte blocks of data in turn, and an analytic signal array for the complete sampled sound is assembled from the processed blocks. A least-squares fitting procedure is applied to the phase information extracted from the analytic signal, and an approximate frequency parameter deduced from it. Finally a semi-automated piecewise-linear fit is made to the modulus of the analytic signal, from which envelope parameters suitable for the DX7 synthesizer are extracted.

The 'linear trend' of the phase plot may then be subtracted from it, the result corresponding to a new real signal, and analysis repeated if desired.

CONCLUSIONS

This paper has presented a technique which may go some way towards meeting the requirements of a general purpose analysis/resynthesis mechanism for digital FM sound sources.

The implementation of the technique described in this paper is still at an early stage of development, and no formal results are, at the time of writing, available. Analyses of short test signals are presently being carried out to evaluate the technique.

It is uncertain whether the analysis program, in its final form, will run in a 'stand-alone' condition, or whether individual interations of the algorithm will be supervised by the operator. An unsupervised analysis program would of course be of more benefit to, say, composers wishing to use an FM sound generator. On the other hand, a supervised program, requiring inspection of envelope shapes etc, would be more instructive to those wishing to learn about the structure of sound.

REFERENCES

1. J M Chowning, 'The Synthesis of Complex Audio Spectra by Means of Frequency Modulation', JAES, Vol 21, No 7, 526-534, (1973).

2. G De Poli, 'A Tutorial on Digital Sound Synthesis Techniques', Computer Music Journal, Vol 7, No 4, 8-25, (1983).

3. J H Justice, 'Analytic Signal Processing in Music Computation', IEEE ASSP, Vol 27, No 6, 670-684, (1979).

4. Y Fukuda, 'Yamaha DX7 Digital Synthesizer', Amsco Publications, (1985).

5. P M Gammell, 'Improved Ultrasonic Detection Using the Analytic Signal', Ultrasonics, March, 1981.

6. R Bracewell, 'The Fourier Transform and its Applications', McGraw Hill(1965).

7. C K Youen and D Fraser, 'Digital Spectral Analysis', Pittman, (1979).

SYNTHESIZING COMPLEX AUDIO SOUNDS WITH FM.

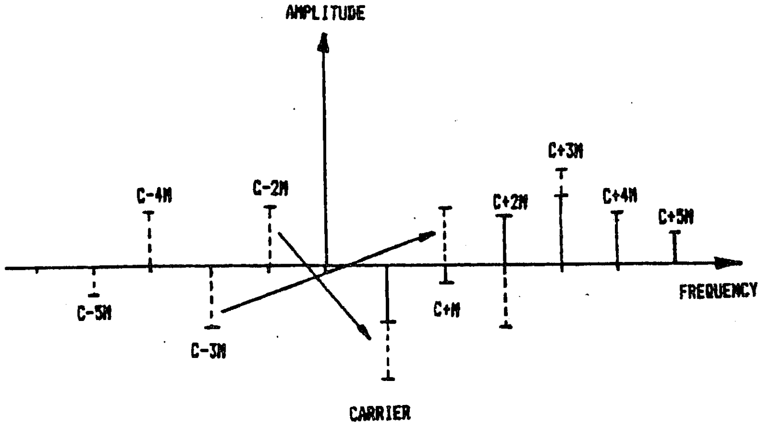

FIG. 1 F.M SPECTRUM SHOWING HOW SIDEBANDS FALLING IN THE NEGATIVE FREQUENCY DOMAIN ARE REFLECTED ABOUT THE AMPLITUDE AXIS, (WITH PHASE INVERSION) INTO THE POSITIVE DOMAIN. SINGLE MODULATOR, I=4. [1]

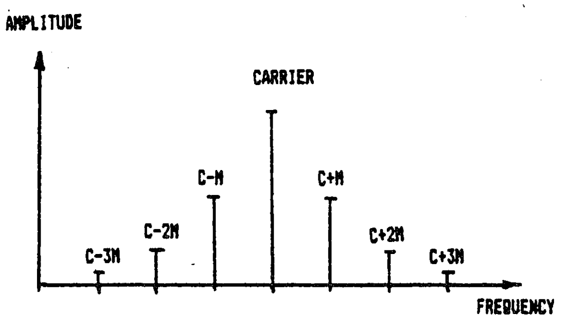

FIG. 2 HARMONIC F.M SPECTRUM PRODUCED BY SINGLE MODULATOR WITH LOW MODULATION INDEX. I=1. [1]

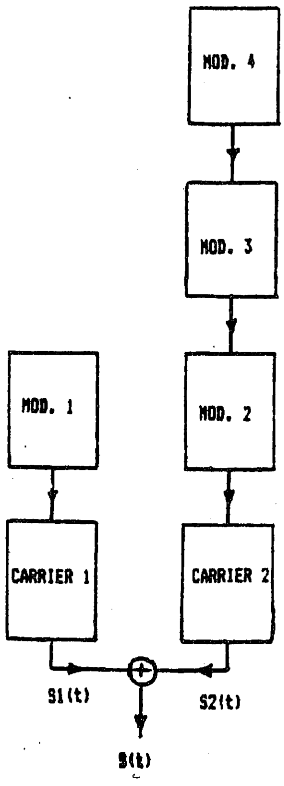

FIG.3 TYPICAL DX7 ARRANGEMENT OF OSCILLATORS. S(t) IS THE SUM OF TWO CHAINS OF F.M GENERATORS.

SYNTHESIZING COMPLEX AUDIO SOUNDS WITH FM.

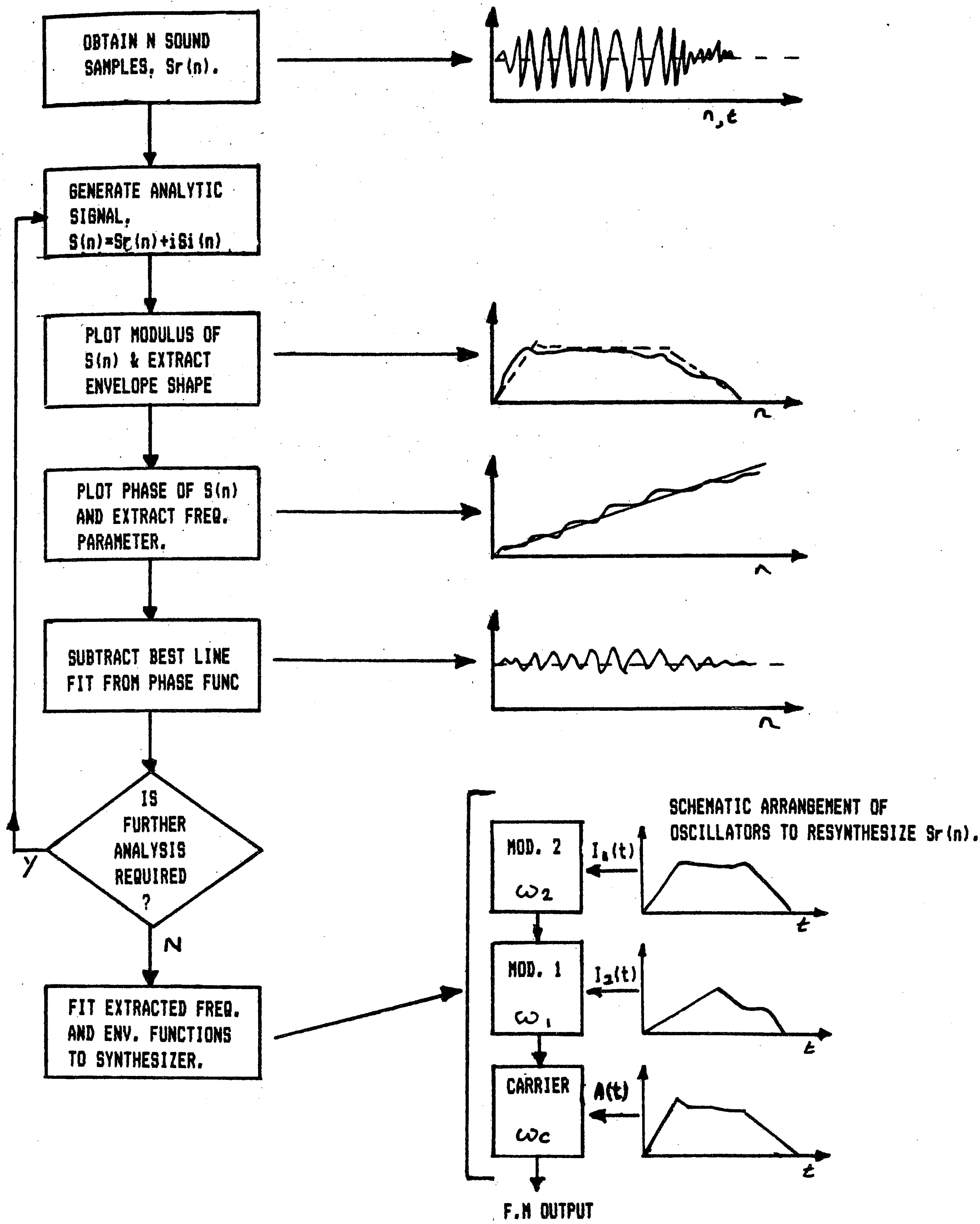

FIG. 4 SCHEMATIC FLOW CHART OF ITERATIVE ANALYSIS/RESYNTHESIS ALGORITHM FOR FM SOUNDS.