Proceedings of The Institute of Acoustics

DIGITAL GENERATION OF MUSICAL INSTRUMENT SOUNDS VIA STOCHASTIC TECHNIQUES

J A S Angus(1) and S P Norris(2)

Department of Electronics, University of York

INTRODUCTION

Most current musical synthesis techniques use a high quality oscillator as the starting point for synthesis. This source of pure periodicity is then modified by modulation and filtering to produce the desired sound. This is in direct contrast to the prime source of traditional musical instruments which tends to be a highly impure sound source, e.g. bow scraping across strings or air blowing against a sharp edge.

The impurity of a traditional instrument's excitation adds variability and interest to the sound which is not present in synthetic instruments. Therefore we decided to see if it was possible to synthesise musical instrument sounds by using random noise as the starting excitation. We hoped that by doing this one would be able to obtain some of the warmth of traditional instruments in synthetic ones.

The rest of the paper describes our technique; it will first outline the basic structure of our synthesis technique. It will then describe a simple way of generating random noise of different timbres and will conclude with results and further work.

THE STOCHASTIC SYNTHESISER

The structure of a stochastic synthesiser is shown in fig 1. It consists of three basic parts:

1. a source of white random noise;

2. a filter which shapes the frequency response of the noise to obtain the desired timbre;

3. a second comb filter which selects the partials required to form the desired pitch.

The third part of the system is an extension of the Karplus and Strong technique [1,2] which was essentially a comb filter with no input. Note that it is possible to have inharmonic partials by having a non-linear phase vs frequency response in the feedback path. Also the attack and decay charactertistics are affected by both the feedback network and input filter.

The implementation of the excitation signal generator, which is the concatenation of the random noise generator and the frequency shaping filter, is more complicated and is described in the next section.

DIGITAL GENERATION OF MUSICAL INSTRUMENT SOUNDS VIA STOCHASTIC TECHNIQUES

THE EXCITATION SIGNAL GENERATOR

Perhaps the obvious method of producing the excitation signal involves generating a white noise signal and then modifying this with a series of filters to give the desired frequency content. This does, however, have several drawbacks:

- Digital filters typically require many multiplies per sample, and hence are either slow or require a large amount of special-purpose hardware.
- The speed varies with the number of poles required.
- Intermediate and final signal amplitudes are difficult to control accurately.
- Rounding errors may cause significant signal degradation unless high-precision arithmetic is used.

One alternative is to use an algorithm known as the circulant Markov chain (CMC). This operates as follows:

1. A finite length lookup table is initialised with a continuous waveform.
2. A table pointer is initialised.
3. At each sampling instant, a jump value is generated by a random process.
4. The table pointer is increased by the jump value (and wraps around at the end of the table).
5. The corresponding lookup table value is output.

Siegel, Steiglitz and Zuckerman [3] have studied the design of circulant Markov chains in considerable detail. The main problem lies in the determination of the probabilities of jumping a given distance through the lookup table. For a generalised case of K pole-pairs, the problem may be stated as follows:

1. For a lookup table length of N samples, there are N realisable pole positions uniformly distributed around the unit circle in the Z-plane. Each pole corresponds to a jump of unique length through the lookup table. It is first necessary to select a number of these as a first approximation to the desired pole locations. Siegel et al [3] suggested the use of K realisable poles, each being selected as that which lies closest to the desired pole location, i.e.

$$L(i) = \left\lfloor N \frac{f_{o_i}}{f_s} + 0.5 \right\rfloor$$

where $L(i)$ is the jump length corresponding to the ith pole approximation, f_s is the sampling frequency and f_{0i} is the desired centre frequency of the ith pole and the brackets $\lfloor \ldots \rfloor$ indicate the truncated integer value. This corresponds to a Z-plane pole position of $W^{L(i)}$ where $W = e^{j2\pi/N}$. A typical Z-plane plot for a single pole-pair with an eight sample lookup table is shown in fig. 2.

2. The lookup table is constructed by summing one sinewave for each pole approximation, the period of each being $\frac{N}{L(i)}$. The amplitude of each sinewave component is proportional to the height of the corresponding resonance.

3. In order to determine the probability associated with each pole approximation, it is necessary to solve the following set of equations:

$$\sum_{i=0}^{N-1} p(i) W^{iL(m)} = R_m e^{j\theta m}, \quad m = 1,\ldots,K$$

where $p(i)$ is the probability of jumping i positions through the lookup table ($p(i) \geq 0$) and $R_m e^{j\theta m}$ is the Z-plane position of the mth desired pole. These K complex equations lead to 2K real equations. There is one additional constraint:

$$\sum_{i=0}^{N-1} p(i) = 1$$

and together these lead to the problem of finding a solution to (2K + 1) equations in N variables.

4. Finally, it is necessary to select a lookup table size, N. If N is too small, the range of possible output values will be reduced. Also, the probability of jumping either 0 or N places will be increased. Siegel et al [3] found that this can lead to an audible "click" resulting from the psychoacoustic effect of having two consecutive samples of the same value. They found that a table size of around 256 was suitable for most audio frequency applications, provided that the desired pole positions do not lie close to D.C. or $\frac{F_s}{2}$.

Several points are important concerning the calculation of jump probabilities:

- Since the number of poles specified by the programmer is unlikely to exceed nine, it is clear that the solution will be overdetermined since $(2K + 1) < N$.

- All constraint equations are linear.

- The likelihood of jumping 0 or N places through the lookup table should be very small in order to prevent audible clicks.

- There will be an infinite number of solutions to the problem. If jump lengths are, on the whole, long then the lookup table is not being used efficiently and it is likely that a smaller table would be adequate. For example, if the shortest common jump length was two then, in the majority of instances, only every other lookup table value would be used. Hence not all solutions will be suitable and some control over the distribution of probabilities must be incorporated.

- $p(i) \geq 0$ for all i.

Consideration of the above points suggests that the simplex algorithm and method of artificial variables is the only method of optimisation suited to the task (see Gass [4]).

Generation of Results

It is also necessary to devise an efficient method of randomly selecting a jump length at each sampling instant, according to the probabilities generated by the simplex method. A method for achieving this was outlined by Siegel et al [3]. It operates as follows:

1. Probabilities are sorted into descending order.

2. The cumulative values of the sorted probabilities are calculated.

3. At each sampling instant, a pseudo-random number with uniform probability density function in the range 0-1 is generated.

4. This random value is compared with each cumulative probability in turn, until the cumulative total exceeds the random value.

5. The jump length corresponding to the number of comparisons made is found from a lookup table.

The above number of comparisons required is given by:

$$C_{avg} = \sum_{i=1}^{2K+1} p(i)i$$

where $p(i)$ is sorted into descending order with increasing i.

DIGITAL GENERATION OF MUSICAL INSTRUMENT SOUNDS VIA STOCHASTIC TECHNIQUES

It is clear that the number of comparisons made can be minimised if the largest probability is close to unity. Consideration of this point together with others previously mentioned in this section led to the conclusion that the best choice of cost function is to maximise the probability of jumping one place through the table.

RESULTS

Figure 3 shows the output of the excitation generator as a result of simulating a single resonance of varying Q. This shows that it is possible to generate frequency shaped noise via the techniques described.

Figure 4 shows the time evolution of the spectrum of note generated by the complete system in which the fourth harmonic has been emphasised by the excitation generator.

Figure 5 illustrates a problem which occurs when the synthesis of longer duration tones is attempted. In this case the note is unable to build up properly and is corrupted by noise. The reason for this is that the frequencies present in the input signal to the comb filter are not phase correlated. Because of this the comb filter is unable to build up to a high value. In traditional instruments the excitation source interacts with the instrument in a non-linear fashion whereas this sytem is linear. We have tried various ways of encouraging phase coherence in the excitation signal but the results have not been satisfactory.

CONCLUSION

It is possible to simulate tuned musical instrument starting from purely random sound sources. However in order to synthesise steady state sound using this technique more needs to be done to ensure phase coherence in the excitation signal. This will almost certainly require the use of non-linear methods of signal processing.

BIBLIOGRAPHY

[1] K Karplus and A Strong, *Digital Synthesis of Plucked-String and Drum Timbres*, Computer Music Journal, vol.7, no.2, pp 43-55, Summer 1983.

[2] D A Jaffe and J O Smith, *Extensions of the Karplus-Strong Plucked-String Algorithm*, Computer Music Journal, vol.7, no.2, pp.56-69, Summer 1983.

[3] L J Siegel, K Steiglitz and M Zuckerman, *The Design of Markov Chains for Waveform Generation*, IEEE Trans. Acoust, Speech, Signal Processing, vol.24, no.6, pp. 558-562, December 1976.

[4] S I Gass, *Linear Programming: Methods and Applications*, McGraw-Hill, 1969 (3rd edition).

DIGITAL GENERATION OF MUSICAL INSTRUMENT SOUNDS VIA STOCHASTIC TECHNIQUES

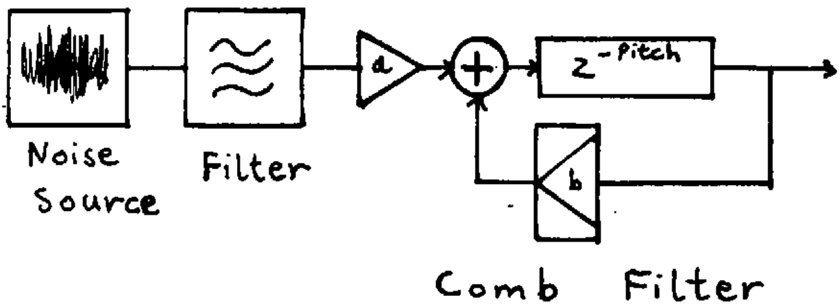


Figure 1 Stochastic synthesiser structure

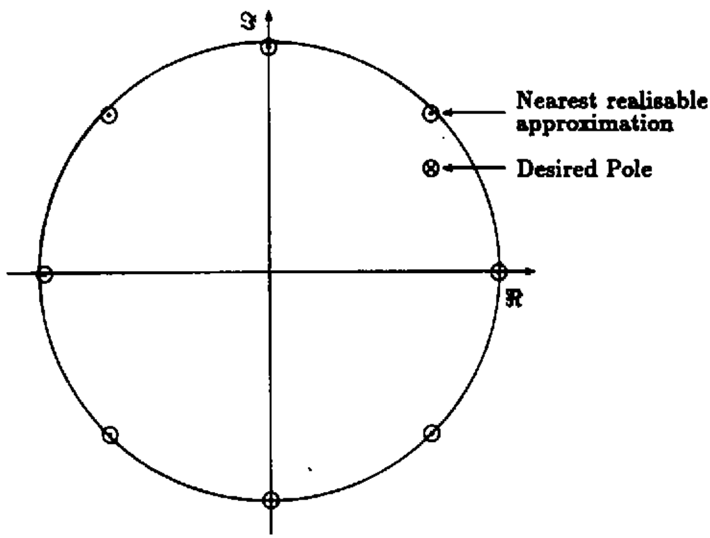
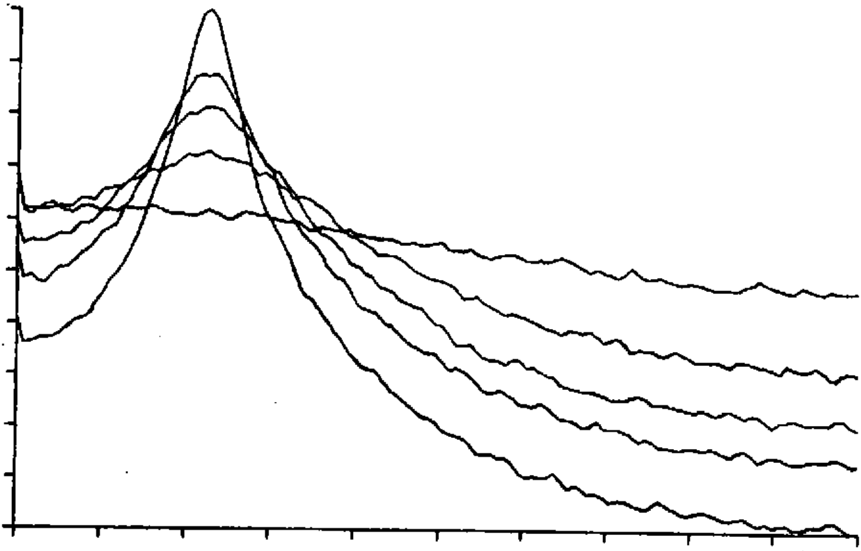


Figure 2 A typical Z-domain nearest-pole approximation



↑ Power (logarithmic, -15–+10dB) → Frequency (linear, 0–22kHz)

Figure 3   Power spectral density resulting from simulation of single pole-pair

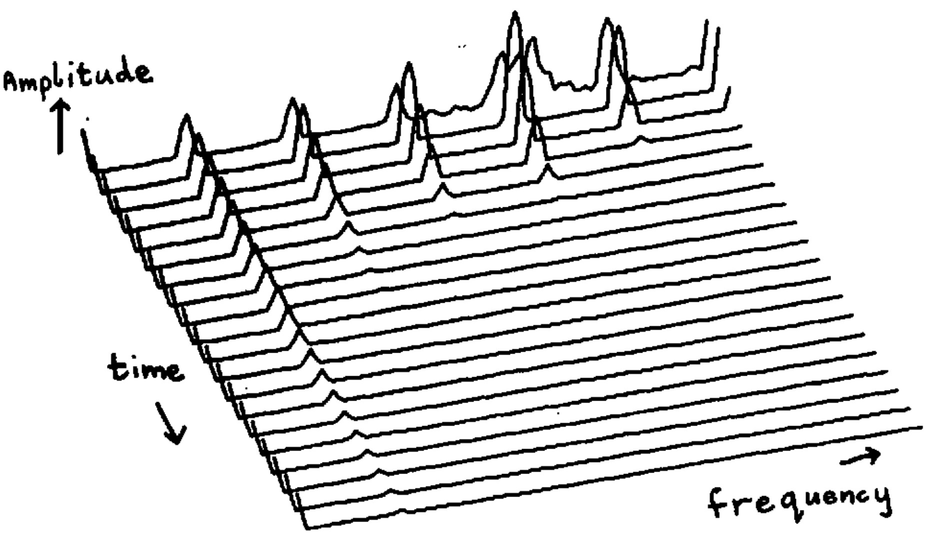


Figure 4   Output showing a short input noise burst

DIGITAL GENERATION OF MUSICAL INSTRUMENT SOUNDS VIA STOCHASTIC TECHNIQUES

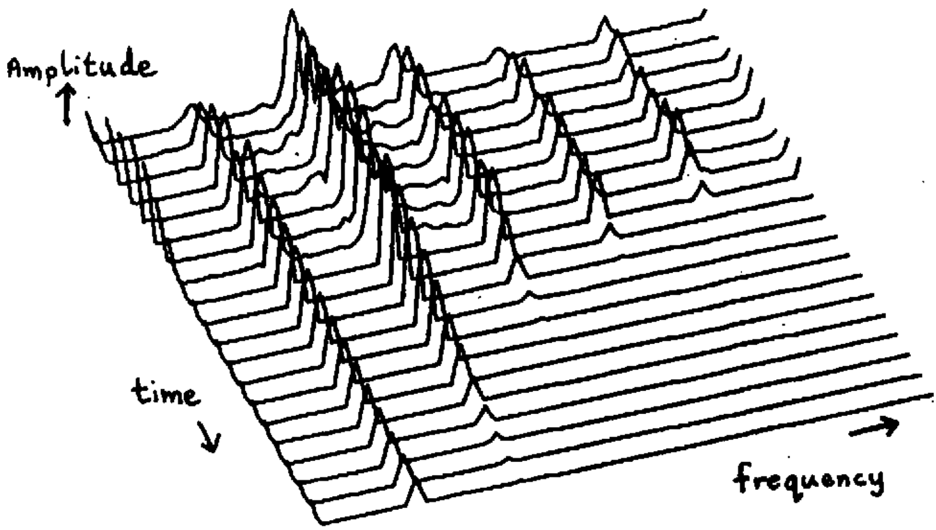


Figure 5 Output with a longer input noise burst