The Application of Differential Equations to the Modelling of Musical Change

J.L. Leach & P. Fitch
The Media Technology Research Centre & School of Mathematical Sciences
University of Bath, Bath, BA2 7AY, United Kingdom
E-mail: jll@maths.bath.ac.uk, pff@maths.bath.ac.uk

ABSTRACT: In this paper we examine differential equations as a means of creating likeable temporal forms, since differential equations are often used as models of natural processes. To account for the great diversity in music compared to the present day, the use of differential equations must remain generic. The method chosen here was to use many interacting units in a network. Each unit was governed by a set of differential equations describing a relaxation oscillator, which depending upon the particular choice of parameter values could exhibit a steady state or oscillatory response.

Introduction:
In its most tangible from music can be defined as patterns of sound changes organised temporally such that, together, they affect the emotions in a certain way. Here, we consider music which does not include other sound phenomena (such as words, language and certain specific sounds), and hence can be described as a "closed system". This is a useful distinction because there is a large body of music that contains no references to the outside world. It is this property that separates it from other art forms, such as literature and fine art (not abstract) painting (for further discussion see Meyer,L.B.). We can deduce from this that whatever it is that affects the human emotions must be internal to the music.

The way in which music affects the emotions is not easily described. It has often been said that music can convey happiness, sadness, etc., (Sloboda,J.A.), and also has the potential to excite or relax the listener. However, distinct from these emotive forces is music's ability to invoke pleasure. Whatever the particular emotion conveyed by a good piece of music, it is agreed that people enjoy the experience of listening to it.

The important word here is "good". What is it about a piece of music that makes it good? This question already presupposes that we have an exact definition of the word, "music" - which we do not. To be more objective and at the same time more general, we ask the question "What is its about certain sound sequences (without external reference) that make them pleasing to listen to?"

We know that the factors determining whether or not a particular sound sequence will be liked cannot be explained merely by individual preference. Otherwise, we might expect to find sizeable populations of people who like rap but none just as there exist populations of people who like the Beatles, Mozart or Chinese music.

This would seem to suggest that groups of sound sequences liked by humans possess a type of temporal organisation to which we are mentally receptive. This organisation must be a function of how the brain has evolved, the individual's preference within these groups would thus be determined by his or her social background. It would also suggest that the more a certain sequence of sounds contains organisation to which we are biologically predisposed, the greater the number of people that will like it.

Since our brain has evolved to process temporal information, it makes sense to suggest that pleasing sound sequences might abstractly mimic the way in which the world changes with respect to time. We believe this is the case and so studied how differential equations could be used to create abstract temporal forest which bear similarities to the organisation found in music. We found that the most appropriate and useful category of equations are those termed relaxation oscillators.

Relaxation Oscillators:
Phase and frequency locking has been observed in many biological mechanisms. One such example is that of the membrane response of the squid axon to electrical stimulus (Matsumoto,G. et al). The membrane was stimulated with periodic trains of current pulses, and it was found that the response could exhibit temporal periodicity or chaos.

The different types of responses that can be obtained are quite varied. The membrane potential on some occasions increase rapidly to form a 'spike' and then drops to a resting potential for each current pulse that stimulates it. At other times the membrane potential fires similarly in response for each 1st and 2nd stimulating
pulse, but does not for the 3rd. Another example is where the membrane 'fires' for the 1st but not the second or third. These responses are termed 1:1, 2:1 and 1:3 phase-locking respectively, but can also be seen as the periodically repeating sequences (1), (110) and (100), where 1 denotes that the membrane generates large fluctuation in potential and 0 an absence of such a fluctuation.

It was further shown for the squid axon that the pattern obtained varies with the width between each stimulating pulse, and the magnitude of each stimulating pulse. The average firing rate, measured relative to the stimulating pulse firing rate, serves to separate the response patterns into distinct phase locking plateaus.

Clearly this biological system provides an excellent pattern generator. A mathematical model of this system would be of value in the generation of musical rhythms. The reason for this is that it provides the essential function of transforming continuous time into distinct fluctuations separated by discrete temporal intervals.

This phase locking behaviour is characteristic of relaxation oscillators and we used an adaptation of the simple FitzHugh-Nagumo model to study the generation of rhythms. We let our new model play the role of an individual unit, and used many of these units together to form networks. A typical interaction between units would be where the "membrane potential" of one unit would stimulate another, but sometimes the sum of more than one unit's potential was used as the stimulus for another. In figure 1 we see a hierarchy of three units, where the bottom stimulates the middle, and the middle stimulates the top. The middle unit "fires" for every third firing of the bottom unit, and the top unit fires for every second firing of the middle unit.

In the sound domain, we took each firing to represent a beat. Thus, if we imagine the top unit to be beating once every bar, then clearly the bottom unit is beating six times in every bar. To appreciate the "musical" qualities of these types of network, a synthetic instrument was attached to each unit which sounded every time the unit fired. When we constructed more complex networks we found that the complexity of the resulting "rhythms" also increased. These were subjectively judged by many people to be significantly pleasing, sounding reminiscent of primitive drum music. The individual rhythms generated were determined not only by the network configuration, but also by the strength of the connections between the different units. The strength of a connection between two units is the multiplicative weight applied to the output pulse of the stimulating unit. Thus the greater the strength of the stimulation, the more frequently the stimulated unit fires.

These networks are thus reminiscent of "neural networks" but operate in continuous time instead of in discrete time intervals. The problem with the resulting rhythms however is that they are ultimately periodic. The reason for this is evident. For any of our networks, the weightings between the units remain static, thus, given any particular input, a unit will always give the same output. As a result, these static weighted networks are useless for large scale rhythmic, i.e. above melodic, works of music. The way round this problem is to change dynamically the multiplicative weights within a network. The next section introduces a system which uses a data set in conjunction with a network to dynamically change the weights depending upon the state of the network.

Generalising to Polyphonic Works:

The relaxation oscillations exhibited in the solution of the units' equations are instances of a catastrophe. A catastrophe typically occurs when the value of a quantity changes very rapidly from one stable set of values.
to another. In the unit as described above, a complete oscillation included two catastrophes, one when the membrane potential increased quickly, and another when it decreased. These catastrophes were separated by a long period of little change (near the minimum membrane potential) and a short period of little change (near the maximum). We replaced this unit with one where a complete oscillation would contain only one catastrophe. This was achieved by letting our equations operate in polar coordinates where our quantity of interest would vary between 0 and 2π. We called this quantity "V" and figure 2 shows how V varies with time for a unit which is oscillating.

If we have a stack of units where the bottom drives the one above and so on, and the bottom unit is set to oscillate with frequency f, then the other units in the stack will either not oscillate at all, or oscillate with a frequency which is an integer multiple of f. This is because we use 2π from any unit as a stimulation to the unit directly above it. Thus when any unit in the stack goes through its catastrophe (fast trajectory), it may provide enough stimulation to push the unit above into a catastrophe as well. Again the same phase locking behaviour occurs as with the catamenia equation networks of figure 1.

In the original equations the actual value of the membrane potential served no purpose other than to trigger a sound pulse. In this case it was decided that V would be used to cross reference a data set which could be an encoding of the rhythm, amplitude and pitch information.

This was achieved through an n-dimensional Fourier synthesis. Fourier synthesis has been used as a two-dimensional process in other domains such as landscape forgery (Suspek,D.). We used the property that any Fourier synthesised wave is periodic over 2π to our advantage. We set each dimension of an n-dimensional synthesised wave to a one-to-one correspondence with an oscillating unit in the stack. Thus the value of V in any unit would reference a point on one of the axes of the wave, and thus the state of an n unit stack would reference a specific point in the n-dimensional wave. This is shown in the 3-dimensional example, figure 3.

In this case the sound pulse is fed just after the firing (i.e. passing through the catastrophe) of the highest frequency oscillator. We use the values V from each oscillator at some t to obtain a reference pitch value for the sound pulse also at some t.

Because the n-dimensional wave is continuous, close n-points generally have close values. Thus in our example above, the pitch values for notes 2 and 5 will be similar since the three-dimensional point that their oscillator values represent differ only in the third dimension (i.e. only the value V in the third oscillator is different). Also, as there is no oscillator hierarchically superior to the third, the "melod" must repeat after the sixth note. So we see that the melody will repeat every six notes, and every second set of three notes will be similar to the first.

In so doing, we have introduced the concepts of repetition and variation into our model. The parameters of the equations for a unit can be varied in order to determine the amount of space covered during the slow trajectory, and so increase or decrease the amount of variation in the melody of a hierarchically inferior unit.
Rhythm can be re-introduced into the system by recalling that we can change the weights between network units. We add an additional unit perpendicular to the stack which is also synchronised to the bottom unit (unit 1). We then use the $n$-space value as the weight between unit 1 and the additional unit. When this value is high, the new unit will fire. When it is low, it will not fire. Thus the way the value changes will determine the rhythmic pattern generated. See figure 4.

As was the case with pitch, if this unit was added to the previous example, we would find that the rhythm repeated after every sixth temporal region of unit 1. Every second temporal region of unit 2 would contain a rhythmic variation of the one present in every first region. Of course, we would now only fire a sound pulse when the additional unit fired, and its pitch would be determined as above. Similarly, the amplitude of each sound pulse can be determined.

A specification language has been built to allow the design of stacks with any number of hierarchical units, each with any characteristic frequency. Thus any "time signature" from musical scripture can be emulated. The number of additional units can also be specified, and each unit could be made to synchronise with any other unit in the hierarchy, thus enabling a polyphonic collection of fast and slow synchronised melodies.

Conclusion:
The system described above allows the algorithmic composition of polyphonic sound sequences which are "musical". With a standard network setup the system is capable, when solved, of producing many different works. This is achieved by changing the Fourier components used to determine the $n$-space synthesised wave. The change in the resulting wave means that the information used during the solution of the network will be different and hence so will the rhythm, melody and dynamics of the composed work. In so doing, we have formalised algorithmically many aspects of that which is musical:

1: With reference to the introduction of this paper, we have produced a system which attempts to separate the concepts of "quality" and "style". What this means is that the solving of the networks is a process which can take random uncorrelated data (the Fourier components), and expand this into a sequence which we are biologically predisposed to like. Thus the solving process is an encoding of quality. The Fourier components themselves further determine the actual issuance of the work, and must therefore represent the style of the piece which further determines to which population the work will appeal.

2: By using a hierarchical stack of units in the network, we are admitting that biologically likeable sequences should contain a hierarchical lateral structure. This coincides with previous studies (Leach,J.L. & Fitch,J.P.)

3: The intrinsic relaxation oscillation behaviour of the equations allows quantities which vary continuously with respect to time to express discretely countable changes in behaviour. This is characteristic of music since performed music varies continuously but can be represented on paper by a discrete sequence of symbols (i.e. the musical score).

The "music" generated is a great improvement on the so called ½ music, and has large time scale evolution characteristic of much existing work. Finally, considering the fact that this system uses no existing works upon which to create its works (unlike neural network models that use training sets), it would seem that we have gone some way to improving a formal and computable description of that phenomenon which we call music.

The future of this work depends upon further research being conducted in the field of music perception (Deutsch, D.). Such research can determine the particular types of relationship which can be perceived in the sound domain. We have shown how this can work for simple changes in tone and rhythm, but the more complex areas of tonal consonance, timbre, instrumentation, etc. need to be investigated as other changing properties which can convey information in the form of temporal relationships.

References: