RHYTHM-SYNCHRONIZED EFFECTS CONTROL WITH MODULATION KEYING AND BÉZIER SPLINES

Stephen McGovern

Managed Network Services Division,
Artel, Inc.
Reston, VA, United States
smcgover@alumni.stevens.edu

ABSTRACT

This paper discusses details related to the automatic control of effects parameters. Specifically, it examines control of audio effects when synchronized with a musical piece's rhythm. Synchronization is achieved by having control data that is periodic with the bar-length. The need for continuity in the control data is also discussed. This concept is introduced in terms of modulation techniques known as ASK and CPFSK. This is expanded upon with a discussion of a curve drawing technique that is common to computer graphics.

1. INTRODUCTION

Many methods exist for controlling audio effects. Some effects are able to blindly control themselves, while others use adaptive control techniques. Some are controlled by a musician. Still others lack a native control method, but are often still controllable through automation processes.

For the most part, audio effects are independent of rhythmic structures. Many effects, however, have parameters that are capable of being changed during a performance. Other effects, such as wah-wahs and talk boxes, have controls that are, by design, intended to be changed during a performance. With these, the musician must be aware of the rhythm in order to exercise thoughtful control. In contrast, a musician using a synthesizer uses effects, but the effects control is ultimately handled by the synthesizer. Examples of this include LFOs and envelope generators. In this case, the musician needs to be aware of neither the rhythm nor the effects control. There is also a process known as automation. Automation can manipulate effects control values over a period of time. The process is able to take rhythmic structure into account because it is initially configured by a human.

Some effects have control values that are dependent on an input signal. These are known as adaptive effects. An overview of adaptive effects was given in [1]. Compressors, expanders, and noise gates are good examples of adaptive effects. Because these effects track and quickly modify signal dynamics, it can be argued that they are manipulating rhythm as well

Adaptive effects involving beat tracking are of particular relevance. Recent work has been done on beat tracking in [2], and this type of technology has been used to control beat-synchronized effects such as tremolos, auto-wahs, vibratos, and flangers [3]. It was also noted in [3] that most traditional effects are static and would be more versatile if they were able to achieve time-synchronization with the music. While adaptive effects can achieve beat synchronization, this is a reactive system designed for live performance. Adaptive beat-synchronized effects control relies on beat detection, and the synchronization thus arrives with a high latency. It also leaves the question of exactly how the effects parameters are to be manipulated once the beat has been detected. These would be drawbacks during the mastering phase of a music production.

An effects controller should not introduce clicks or zipper noise to the audio that is being processed. This was mentioned in [4] where polynomial interpolation was used to get smooth transitions between parameters. The issue of synchronization fell outside the scope of the work though, and, in the context presently being discussed, general polynomial interpolation can suffer from undesirable behavior. It is thus necessary that specific methods of interpolation be considered.

Continuity of control parameters has also been dealt with in [5] by making use of Bézier curves. This also provided for an arbitrarily shaped controlling waveform. Bézier curves were also used in [6] and [7] and again they permitted both smooth control and the use of arbitrarily shaped waveforms. In these three papers, however, the issue of time synchronization and its extension to generalized effects control, was not discussed. Bézier curves were also used for waveform generation in [8].

In section 2 background is discussed. This includes a discussion of rhythm, automation, and modulation. Section 3 begins to pull these concepts together with an investigation of modulation keying and synchronization. Section 4 extends the concept of modulation keying to splines constructed from Bézier curves. Following this is a conclusion summarizing the results.

2. BACKGROUND

2.1. Rhythm

Rhythm is a fundamental attribute of music. It is important to the brief momentary segments within a note as well as to the large scale structures that encompass an entire musical piece. Among the basic rhythmic structures in music is the bar. Also known as a measure, a bar represents a grouping of a set of musical sounds. In many musical genres, sets of bars exhibit high degrees of similarity with one another [9]. For an idealized musical signal, similarity can roughly be described mathematically as

$$y(t) \approx y(t + nP) \tag{1}$$

where P is the time length in seconds of each bar and is given by

$$P = \frac{f_{bpb}}{f_{bpm}}\, 60\,\mathrm{sec/min}. \tag{2}$$

Here f_{bpb} represents the beats per bar. It is also the numerator in a time signature. The variable f_{bpm} is the beats per minute. It is the value of P that gives the period with which the effects control is desired to be in synchronization.

2.2. Automation

2.2.1. *Usage in Industry*

Automation is a process by which adjustable parameters in a music production system are programmed to be adjusted without having oversight by a human controller. This is a common feature in many sequencing programs. Overviews of automation can be found in sequencer program user manuals [10] [11].

Automation is often used to control volume, panning, and equalization, as well as a wide variety of other processing routines. Typically, configuring automation means using a mouse to create control data in a software sequencer. This done by selecting points, or nodes, along a track. The software then uses one of a number of methods to interpolate between these points.

2.2.2. *Mathematical Description*

Automation data can be described mathematically as a spline. A spline is a function that is defined in terms of other functions. A spline is cut up into sections, and then each section is defined by a segment from another function. If a spline is given by $S(t)$ and it is divided up at points given by t_n, then

$$S(t) = \quad V_n(t) \quad \text{for} \quad t_{n-1} \le t < t_n \tag{3}$$

where t_0 is the first sample in the track and $t_n - 1$ is the last.

Automation is primarily a means for assigning slowly varying arbitrary control. Automation data created for a single bar, can be duplicated and thereby achieve a certain amount of repeatability. On the downside, the parameter control is limited by the ability to shape curves. Discontinuities can appear where portions of automation data have been duplicated and placed adjacent to one another. They can also appear at the nodes where spline segments join one another, although some implementations [12] provide for nicer curve shaping abilities than others.

In theory, automation could also be used for fast modulation or even synthesis. This, however, is not the intended purpose.

2.3. Modulation

2.3.1. *Modulation in Music*

Modulation in music has taken on many forms, with amplitude and frequency modulation being among the most well known. Many variants, such as single sideband modulation and adaptive modulation, exist, and they have been used in audio effects processing [13]. Another technique known as fractal modulation was also the topic of a relatively recent audio effects paper [14]. Applications of modulation effects are wide-ranging, and include auto-pans, auto-wahs, tremolos, and vibratos, as well as literal signal modulators. Typical shapes for modulated waveforms are sinusoidal, triangular, sawtooth, and square. Modulation is often used to deliberately produce artifacts inside of the audible spectrum. By driving a process slowly, however, artifacts with perceivable time-dependence can be produced outside of the audible spectrum.

2.3.2. *Modulation in Telecommunications*

Modulation keying is the technology used in telecommunications for transmitting digital data. The primary methods for this are amplitude shift keying (ASK), frequency shift keying (FSK), and phase shift keying (PSK). In each of these three schemes, the signal amplitude, signal frequency, or signal phase is abruptly switched to a different value, or a different state. Many variations on these three schemes exist. Among these are quadrature phase shift keying (QPSK), continuous phase frequency shift keying (CPFSK), and 64 quadrature amplitude modulation (64-QAM). An overview of various modulation schemes is given in [15]. While traditional modulation methods have a long history in musical effects processing, schemes employing modulation keying have seen little use.

In modulation keying, data is composed of bits and is encoded into different symbols. Each symbol is a representation of either a single bit or a group of bits. The symbols are then transmitted through the propagation medium. CPFSK and ASK are shown in Fig. 1, where 1s and 0s are distinguished by waveform segments of either a different frequency or a different amplitude. In this case each symbol is a representation of a single bit.

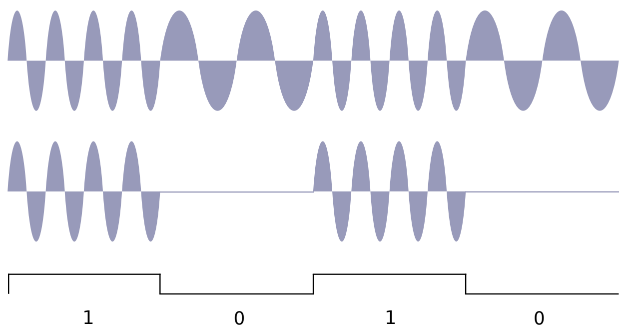

Figure 1: *Top: The waveform of a CPFSK signal. Middle: An ASK signal. Bottom: A corresponding bit sequence.*

In any scheme, the transmitted signal has a certain symbol rate. If error correction coding is ommitted, the symbol rate is given by

$$f_{symbol} = \frac{f_{bit}}{m} \tag{4}$$

where f_{bit} is the bit-rate and m is the number of bits in each symbol.

Each symbol represents a distinct quantization state. The number of states M in an M-ary signaling system is given by

$$M = b^m \tag{5}$$

where b is the number of possible states held by an individual bit. In practice b is almost always equal to 2, though other values are technically possible. Fig. 2 displays a QPSK signal that has two bits per symbol and is thus a depiction of 4-ary signaling.

All signals have a spectrum. The spectrum of a particular signal is dependent on a number of factors. For modulation keying, the means by which the state of the modulated carrier wave is shifted will have an effect on the spectrum. Filtering is often employed so that the signal falls into a desired bandwidth. Other modified modulation schemes, such as CPFSK, can be used to achieve greater continuity and thereby reduce signal bandwidth.

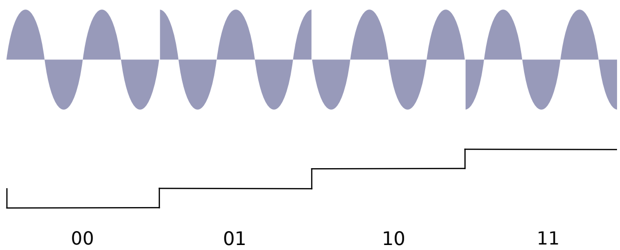

Figure 2: *Top: The waveform of a QPSK signal. Bottom: A corresponding bit sequence.*

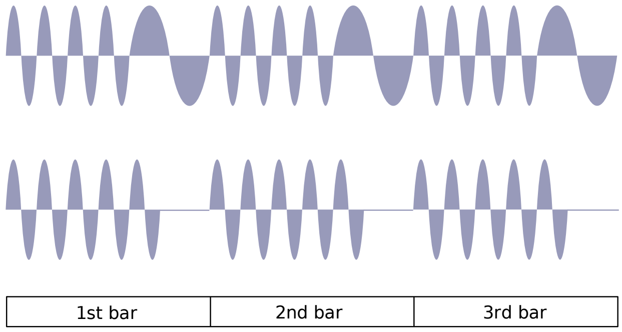

Figure 3: *Top: A synchronized CPFSK waveform. Middle: A synchronized ASK waveform. Bottom: Blocks denoting a bars as is commonly found in sequencer applications.*

3. RHYTHM-SYNCHRONIZED KEYED MODULATION

In the preceding sections the concepts of musical periodicity, automation, and modulation were described. The following will discuss modulation keying as a method for generating control signals that are rhythm-synchronized.

Many modern and many vintage effects are modulation-related. In other words, they're controlled by an oscillatory waveform. Examples of this include tremolos, vibratos, and autopans. When arranging music, it may be desirable to have an oscillator with a continuously variable frequency, and also to synchronize this oscillator to the rhythm. These constraints can be met if the oscillatory waveform is of the same type as that used for specific forms of modulation keying.

3.1. Synchronization

It is desired that the control data be the same over each bar. The functions $V_n(t)$ can be used to define a bar-length segment of control data, and it can thus be observed that

$$V_n(t) = V_{n+1}(t+P) \tag{6}$$

which is equivalent to writing $S(t) = S(t+nP)$. It follows that the spline junctions t_n in Eq. 3 are given by

$$t_n = nP. \tag{7}$$

By carefully choosing a sequence of repeating symbols, as well as the time length and modulation frequency of each symbol, a carrier waveform can be forced to fit neatly within a bar. This is shown in Fig. 3 for CPFSK and ASK.

Having different time lengths for different symbols means that the average symbol rate for non-random signals will vary. Using Eq. 2 and an L-length sequence of symbols each with symbol frequencies $f_{symbol,i}$, leads to

$$\sum_{i=1}^{L} \frac{1}{f_{symbol,i}} = P. \tag{8}$$

Many approaches exist for describing the segments $V_n(t)$ mathematically. However, the spline method used to construct the complete automation waveform in Eq. 3 will work here as well. This is shown in 9 for an i-length sequence of bits.

$$V_n(t) = \quad v_i(t) \quad \text{for} \quad t_{i-1} \leq t < t_i \tag{9}$$

This all presumes, however, that the time signature, tempo, and structure within a song are all fairly static. Complex time signature, complex tempo, and varying structure add additional complexity. Complexities such as these can be accommodated with minor mathematical modifications, though the specific modification will depend on the specific complexity. Approaches for structure analysis [16] and tempo estimation [17] may prove useful for designing more complex control data.

4. CONTINUITY AND CURVE DRAWING

Signals with sharp corners and discontinuities have high frequency components in their spectrums. An ideal controlling waveform does not introduce undesired audible artifacts. For this reason, potential musical implementations of FSK and PSK have not been depicted in the present paper.

There are two common classes of continuity, parametric continuity, denoted C^j, and geometric continuity, denoted G^j [18] [19]. In both cases, j refers to the continuity of the jth derivative.

Parametric continuity refers to continuity of a parametric curve. Parametric curves are defined by a set of functions, each of which is dependent on a common shared variable. It is the common shared variable with which the derivatives are taken to.

Geometric continuity is continuity that exists when the derivative of a spatial function is taken with respect to a spatial dimension. The existence of parametric continuity implies the existence of geometric continuity. In the case of non-parametric control data, there is just one time-dependent function. Time can therefore be treated as a spatial dimension, and only geometric continuity needs to be considered.

While CPFSK and ASK have G^0 continuity, they fail to have the G^1 continuity that was used in [6]. It may be desired to have controlling waveforms of still higher order G^j continuity, although the chosen value of j may affect and ultimately complicate implementation strategies. In any case, it is desired that each control segment $V_n(t)$ be at least G^1 continuous, and that G^1 continuity exist where the end of one control segment meets the beginning of the next.

4.1. Curve Drawing

With modulation keying some sections of the control data may need to be smoothed over. A reasonable approach to this is to identify the rougher sections of the waveform and replace them with an interpolated curve. On the other hand, there is a limitation to what can be accomplished with modulation keying when the keys are limited to sinusoidal shapes. This is exemplified by processes such as envelope control, fades, and wah effects where an arbitrarily-shaped waveform is ideal. Rather than smoothing rough sections of the waveform by interpolating, it is just as practical to produce the entire sequence of symbols through interpolation. The problem at hand can thus be reduced to one of curve drawing.

4.1.1. Bézier Curves

A prevalent curve drawing method found in graphics applications uses Bézier curves [20]. This method is well described in literature related to both computer graphics and non-uniform rational B-splines. Bézier curves have the benefit of joining together smoothly to make continuous splines and also behaving in a predictable fashion.

Bézier curves are defined in terms of control points. Each curve has a start point, an end point, and at least one additional point to shape the curve. Additional shaping points are needed when the curve is of higher order. If the control points are given by, say, $\vec{\mathbf{p}}_k$, then each point is defined in signal-space as

$$\vec{\mathbf{p}}_k = t_k\hat{\mathbf{t}} + v_k\hat{\mathbf{v}}. \tag{10}$$

Typically a Bézier curve is described parametrically. For a cubic curve, this is given by Eqs. 11 and 12.

$$t(u) = (1-u)^3 t_0 + 3u(1-u)^2 t_1 + 3u^2(1-u)t_2 + u^3 t_3 \tag{11}$$

$$v(u) = (1-u)^3 v_0 + 3u(1-u)^2 v_1 + 3u^2(1-u)v_2 + u^3 v_3 \tag{12}$$

The start point, first control point, second control point, and end point are given by $\vec{\mathbf{p}}_0$, $\vec{\mathbf{p}}_1$, $\vec{\mathbf{p}}_2$, and $\vec{\mathbf{p}}_3$ respectively. This can be generalized to an Nth degree curve by Eq. 13.

$$\vec{B}(u) = \sum_{k=0}^{N} \vec{p}_k \frac{N!}{k!(N-k)!} u^k (1-u)^{N-k} \tag{13}$$

The start and end points determine where the curve begins and ends, and the control points are used to shape the curve. Because an audio signal is being dealt with, v must be given in terms of t, which means solving the polynomial in Eq. 11 for u.

Quadratic, cubic, and quartic polynomials are algebraically solvable. In contrast, polynomials of higher order are not solvable using ordinary algebraic operations[21]. The solutions to quadratic, cubic, and quartic polynomials involve multiple algebraic operations including square roots. This comes at a relatively small computational cost, although comparatively speaking, it's more costly than finding points along a parametric Bézier curve.

4.1.2. Keying with Bézier Curves

Modulation symbols can represent multiple bits. It is therefore possible to describe each distinct Bézier segment as a single symbol in an M-ary modulation scheme. This is done in Fig. 4 for five symbols of 8-ary modulation of binary bit data.

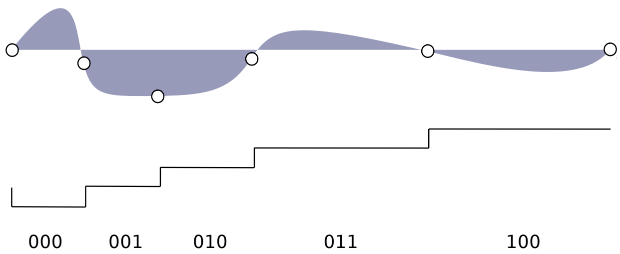

Figure 4: *Top: A Bézier spline shown as a set of modulation symbols. Circles mark where two segments join. Bottom: A corresponding sequence of bits.*

Again, each segment $V_n(t)$ is constructed as a spline where each subsegment $v_i(t)$ is described by Eq. 11, 12, and 13. The result of this is plotted for several bars in Fig. 5.

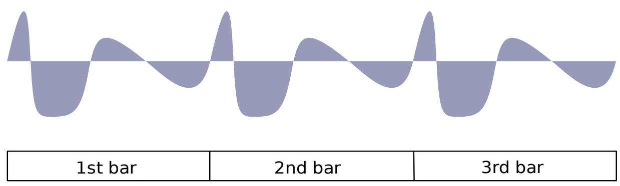

Figure 5: *Top: A Bézier spline that repeats on each bar. Bottom: Track segments depicting bars.*

5. CONCLUSIONS

A system for controlling audio effects has been described. Rhythm synchronized control can be achieved by using the modulation schemes known as ASK and CPFSK. Rhythm synchronized control can also achieved by using Bézier curves, and this can be described in the context of modulation keying. ASK and CPFSK result in G^0 continuity. Bézier curves, when given certain constraints, result in at least G^1 continuity. Bézier curves have the added benefit of providing control with an arbitrarily-shaped waveform. This allows users to have as much control over the curve shape as they would with curves drawn using a computer graphics application. The end result is control data that can be varied in an irregular manner, varied precisely, and repeated for sections of audio that exhibit similarity.

6. REFERENCES

[1] V. Verfaille, U. Zölzer, and D. Arfib, "Adaptive digital audio effects (a-dafx): A new class of sound transformations," *IEEE Trans. on Audio, Speech, and Language Processing*, vol. 14, no. 5, pp. 1817–1831, Sept. 2006.

[2] Y. Shiu and C. Jay Kuo, "On-line music beat tracking with kalman filtering and probability data association (kf-pda)," *IEEE Trans. on Consumer Elect.*, vol. 54, no. 3, pp. 1369–1377, Aug. 2008.

[3] A. Stark, M. Plumbley, and M. Davies, "Realtime beat-synchronous audio effects," in *New Interfaces for Musical*

Expression (NIME07), New York, United States, June 6-10, 2007, pp. 344–345.

[4] S. Nielsen, “Realtime control of audio effects,” in *Proc. Digital Audio Effects (DAFx-99)*, Trondheim, Norway, Dec. 9-11, 1999, pp. 00–01.

[5] F. Lopez-Lezcano, “A dynamic spatial locator ugen for clm,” in *Proc. Sound and Music Computing (SMC-08)*, Berlin, Germany, July 31-Aug 3, 2008.

[6] E. Maestre, J. Bonada, and O. Mayor, “Modeling musical articulation gestures in singing voice performances,” in *Audio Engineering Society Convention (121)*, San Francisco, United States, Oct. 5-8, 2006, p. 6901.

[7] B. Battey, “Bézier spline modeling of pitch-continuous melodic expression and ornamentation,” *Comp. Music J.*, vol. 28, no. 4, pp. 25–39, 2004.

[8] B. Lang, “Waveform synthesis using bezier curves with control point modulation,” in *Audio Engineering Society Convention (116)*, San Francisco, United States, May 8-11, 2004, p. 6044.

[9] S. Dixon, F. Gouyon, and G. Widmer, “Towards characterisation of music via rhythmic patterns,” in *Proc. International Conference on Musical Information Retrieval*, Barcelona, Spain, Oct. 10-14, 2004, pp. 509–516.

[10] *Pro Tools Reference Guide, Version 7.4*, Digidesign, 2007, Available at http://akmedia.digidesign.com/...

[11] *Logic Pro 8, User Manual*, Apple Inc., 2007, Available at http://support.apple.com...

[12] *Adobe Audition 3, User Guide*, Adobe Systems Inc., 2007, Available at http://help.adobe.com/...

[13] V. Lazzarini, J. Timoney, and T. Lysaght, “Asymmetric-spectra methods for adaptive fm synthesis,” in *Proc. Digital Audio Effects (DAFx-08)*, Espoo, Finland, Sept. 1-4, 2008, pp. 233–240.

[14] G. Evangelista, “Fractal modulation effects,” in *Proc. of the Int. Conf. on Digital Audio Effects (DAFx-06)*, Montreal, Quebec, Canada, Sept. 18-20, 2006, pp. 101–106.

[15] F. Xion, *Digital Modulation Techniques*, Artech House, Norwood, MA, USA, 2000.

[16] J. Paulus and A. Klapuri, “Music structure analysis using a probabilistic fitness measure and an integrated musicological model,” in *Proc. International Conference on Musical Information Retrieval*, Philadelphia, Sept. 14-18, 2008, pp. 369–374.

[17] G. Peeters, “Template-based estimation of time-varying tempo,” *EURASIP Journal on Applied Signal Processing*, vol. 1, pp. 158, Jan. 2007.

[18] D. Rogers, *An Introduction to NURBS: With Historical Perspective*, Academic Press, London, UK, 2001.

[19] G. Farin, *NURBS: From Projective Geometry to Practical Use*, AK Peters, Natick, MA, USA, 1999.

[20] P. Shirley, *Fundamentals of Computer Graphics, 2nd ed.*, AK Peters, Wellesley, MA, USA, 2005.

[21] J. Tignol, *Galois Theory of Algebraic Equations*, World Scientific, River Edge, NJ, USA, 2001.