**Victor Lazzarini, Joseph Timoney, and Thomas Lysaght**
An Grúpa Theicneolaíocht Fuaime agus Ceoil Dhigitigh
(Sound and Digital Music Technology Group)
National University of Ireland, Maynooth
Maynooth, Co. Kildare, Ireland
Victor.Lazzarini@nuim.ie
{JTimoney, TLysaght}@cs.nuim.ie

# The Generation of Natural-Synthetic Spectra by Means of Adaptive Frequency Modulation

Frequency-modulation (FM) synthesis is widely known as a computationally efficient method for synthesizing musically interesting timbres. However, it has suffered from neglect owing to the difficulty in creating natural-sounding spectra and mapping gestural input to synthesis parameters. Recently, a revival has occurred with the advent of adaptive audio-processing methods, and this work proposes a technique called *adaptive FM synthesis.* This article derives two novel ways by which an arbitrary input signal can be used to modulate a carrier. We show how phase modulation (PM) can be achieved first by using delay lines and then by heterodyning. By applying these techniques to real-world signals, it is possible to generate transitions between natural-sounding and synthesizer-like sounds. Examples are provided of the spectral consequences of adaptive FM synthesis using inputs of various acoustic instruments and a voice. An assessment of the timbral quality of synthesized sounds demonstrates its effectiveness.

## Background

Frequency modulation (FM), introduced by John Chowning in his seminal article on the technique (Chowning 1973), is one of the most important classic methods of synthesis. It has proved very useful as an economical means of generating time-varying complex spectra. For this reason, it was widely adopted at a time when computational speed was a determining factor in the choice of signal-processing algorithms. However, the method always made it difficult for composers to produce natural-sounding spectral evolutions. This in some cases was caused by the lack of fine gestural control over the sound and in others by the synthetic-sounding quality of the generated spectra. These shortcomings spurred software and hardware designers to come up with new solutions for instrument control and improvements to the basic FM method (Palamin, Palamin, and Ronveaux 1988; Tan and Gan 1993; Horner 1996). Nevertheless, these developments failed to stem the decline in the technique's use as increasingly more powerful hardware became available.

Some of the limitations of gestural controllers and of synthetic sound in FM can be addressed together by the use of adaptive techniques, which form an important subset of musical signal-processing techniques (Verfaille and Arfib 2002; Verfaille, Zölzer, and Arfib 2006). A key aspect of their usefulness in music composition and performance is that they provide a means to retain significant gestural information contained in the original signal. Therefore, these techniques seem to be well suited to help develop more natural-sounding forms of FM synthesis. With them, it might be possible to obtain results that share much of the liveliness perceived in musical signals of instrumental origin.

The traditional approach has been to treat synthesis and control parameters separately, using some means of mapping to control the process (Miranda and Wanderley 2006; Wanderley and Depalle 2004). This ultimately can lead to a split between gesture and sonic result, especially in the case of FM, where the mapping is often not clear or too coarsely defined. Alternatively, one can approach the problem from an adaptive point of view, whereby a signal is both the source of control information (extracted from it through different analysis processes) and the input to the synthesis algorithm. Some pioneering works in the area have proposed interesting applications of this principle in what has been called *audio-signal driven sound synthesis* (Poepel 2004; Poepel and Dannenberg 2005).

Computer Music Journal, 32:2, pp. 9–22, Summer 2008

In the specific case of FM synthesis, it is possible to use an arbitrary input signal in two ways, either as a modulator or as a carrier. In the former case, this signal is used to modulate the frequency of one or more oscillators. When the input is anything but a sinusoidal wave, this arrangement produces what we normally describe as *complex FM* (Schottstaedt 1977). Although this setup, proposed by Poepel and Dannenberg (2005), provides a richer means of gestural control over the process, it does not seem to capture well the original spectral characteristics of interesting input sounds (such as the ones originating from instrumental sources). The spectral evolutions allowed by the method still resemble the more synthetic results typical of standard FM synthesis, because the carrier is still a sine wave oscillator. If we want to allow as much of the timbral qualities of the input sound to affect the generated sound, we will get better results using the input as a carrier signal.

Considering non-sinusoidal inputs, this case is similar to *multiple-carrier FM* (Dodge and Jerse 1985). The techniques described in this article implement this arrangement. Standard multiple-carrier FM is defined by a single modulator being used to vary the frequency of several sinusoidal carriers. It has proved useful in a variety of applications, including vocal synthesis (Chowning 1989) and instrumental emulation via spectral matching (Horner, Beauchamp, and Hakken 1993). By applying the technique to real-world signals, it is possible to generate transitions between natural-sounding and synthesizer-like sounds. Depending on the levels of modulation, we are able to reveal more or less of the original timbral qualities of the input. This is the basis for our technique of adaptive FM synthesis, or *AdFM* (Lazzarini, Timoney, and Lysaght 2007).

To use an arbitrary input as a carrier, we must develop some means of modulating the frequency (or, to be more precise, the phase) of that signal. This is required because we no longer use an oscillator to produce the sound, so we have no implicit frequency control of the arbitrary signal. The following section addresses two different methods of achieving this. We then discuss the implications of using complex signals as carriers and details of parameter extraction.

## The Technique

The synthesis technique discussed here is based on two elements: some means of phase modulation of an input signal; and the use of an arbitrary, monophonic, pitched or quasi-pitched input to which parameter estimation will be applied. The phase modulation effect can be achieved by two basic methods: through the use of a variable delay line or by heterodyning.

### Delay-Line Based Phase Modulation

A well-known side-effect of variable delays is the phase modulation of the delay-line input (Dilsch and Zölzer 1999). This is the basis for all classic variable-delay effects such as flanging, chorusing, pitch shifting, and vibrato. The principle has also been used in audio-rate modulation of waveguide models (Van Duyne and Smith 1992). It is thus possible to model simple (sinusoidal) audio-rate phase modulation using a delay-line with a suitable modulating function (see Figure 1).

We now consider the case where the input to the delay line is a sinusoidal signal of frequency $f_c$:

$$x(t) = \sin(2\pi f_c t) \tag{1}$$

When the modulating source is $s(t) = d_{\max} D(t)$, where $D(t) \in \{0 \dots 1\}$ is an arbitrary function, and $d_{\max}$ is the maximum delay, the delay-line phase modulation of Equation 1 can be defined (with $\omega = 2\pi f_c$) as

$$y(t) = \sin(\omega[t - d_{\max} D(t)]) \tag{2}$$

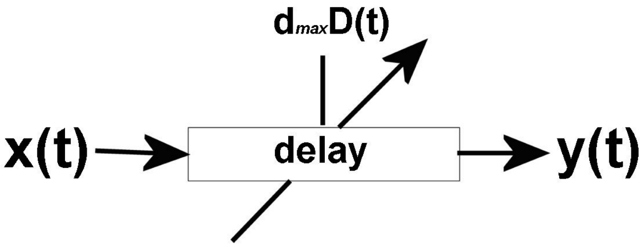
The instantaneous radian frequency $\omega_i(t)$ of such a phase-modulated signal can be estimated from the derivative of the phase angle $\theta(t)$:

$$\omega_i(t) = \frac{\partial\theta(t)}{\partial t} = \frac{\partial\omega[t - d_{\max} D(t)]}{\partial t} = \omega - \frac{\partial D(t)}{\partial t} d_{\max}\omega \tag{3}$$

and the instantaneous frequency $IF(t)$ in Hz can be defined as

$$IF(t) = f_c - \frac{\partial D(t)}{\partial t} d_{\max} \tag{4}$$

Figure 1. Delay-line based phase modulation.




Considering the case where the modulating signal is a scaled raised cosine (i.e., a periodically repeating Hanning window), we have

$$D(t) = 0.5\cos(2\pi f_m t) + 0.5 \tag{5}$$

and, by substituting *D(t)* in Equation 4, *IF(t)* is now

$$IF(t) = \pi f_m \sin(2\pi f_m t) d_{\max} f_c + f_c \tag{6}$$

which characterizes the instantaneous frequency in sinusoidal phase modulation. In such an arrangement, the sinusoidal term in Equation 6 is known as the frequency deviation, whose maximum absolute value $DEV_{\max}$ is

$$DEV_{\max} = \Delta d \times \pi f_m f_c \tag{7}$$

with $\Delta d = d_{\max} - d_{\min}$.

Now, turning to FM theory, we characterize the index of modulation $I$ as the ratio of the maximum deviation and the modulation frequency:

$$I = \frac{DEV_{\max}}{f_m} = \frac{\Delta d \pi f_m f_c}{f_m} = \Delta d \pi f_c \tag{8}$$

The $\Delta d$ that should apply as the amplitude of our sinusoidal modulating signal can now be put in terms of the index of modulation

$$\Delta d = \frac{I}{\pi f_c} \tag{9}$$

and the modulating signal is now

$$d(t) = \frac{I}{\pi f_c}\left[0.5\cos(2\pi f_m t) + 0.5\right] \tag{10}$$

The resulting spectrum according to FM theory is dependent on the values of both $I$ and the carrier-to-modulator ($c$:$m$) frequency ratio:

$$\begin{aligned} y(t) = {} & J_0(I)\sin(\omega_c t) + \\ & \sum_{k=1}^{I+1} J_k(I)\sin(\omega_c t + k\omega_m t) + J_{-k}(I)\sin(\omega_c t - k\omega_m t) \end{aligned} \tag{11}$$

where $\omega_c = 2\pi f_c$, $\omega_m = 2\pi f_m$, $J_k(I)$ are Bessel functions of the first kind of order $k$, and

$$J_{-k}(I) = (-1)^k J_k(I) \tag{12}$$

Note that to match the phases as closely as possible to Equation 11, we require an offset of $\pi/2 + 2I$ in the input sinusoid and $\pi/2$ in the modulator (both in relation to cosine phase). Because the carrier phase depends on the index of modulation in general, we only rarely achieve an exact match. Thus, in delay-line phase modulation, we need not be too concerned with phase offsets.

Interestingly enough, in the delay-line formulation of FM/PM, the index of modulation for a given variable delay-width is proportional to the carrier-signal frequency (as seen in Equation 9). This situation does not arise in classic FM. Also, when considering the width of variable delay for a given value of $I$, we see that it gets smaller as the frequency rises. In a digital system, for $I = 1$, the width will be less than one sample at the Nyquist frequency.

## Phase Modulation Through Heterodyning

The second method proposed here is based on a simple re-working of the PM formula. We begin by proposing the following synthetic signal, where $I$ is the index of modulation and $\omega_m$ is the radian modulation frequency ($\omega_m = 2\pi f_m$):

$$y(t) = x(t)\cos(I\sin(\omega_m t)) \tag{13}$$

Using a sinusoid described in Equation 1 as our input signal $x(t)$, we obtain, by manipulating the expression, the following combination of PM signals:

$$\begin{aligned} y(t) &= \sin(\omega_c t)\cos(I\sin(\omega_m t)) \\ &= 0.5[\sin(\omega_c t + I\sin(\omega_m t)) + \sin(\omega_c t - I\sin(\omega_m t))] \\ &= 0.5[PM(\omega_c, \omega_m, I, t) + PM(\omega_c, -\omega_m, I, t)] \end{aligned} \tag{14}$$

where the PM signal is defined as

$$PM(c,m,I,t) = \sin(ct + I\sin(mt)) \quad (15)$$

By inspecting Equation 11 , it is clear that this formulation, based on the mixing of two PM signals, will lead to the cancellation of certain components in the output signal, namely the ones where $k$ is odd (called in FM theory the odd sidebands).

The significance of this and the previous implementations of PM can be fully appreciated only once we move from using sinusoidal inputs to arbitrary signals. This will allow us to develop the synthesis designs we propose in this work.

## Using Arbitrary Input Signals

We will now examine the results of applying arbitrary input signals to both formulations just described, beginning with the delay-line based PM. In Equation 11, we see the ordinary spectrum of simple FM. However, for our present purposes, we will assume the input $x(t)$ to be a complex arbitrary signal made up of $N$ sinusoidal partials of amplitudes $a_n$, radian frequencies $\omega_n$, and phase offsets $\phi_n$, originating, for instance, from instrumental sources:

$$x(t) = \sum_{n=0}^{N-1} a_n \sin(\omega_n t + \phi_n) \quad (16)$$

The resulting phase-modulated output is equivalent to what is normally called multiple-carrier FM synthesis, because the carrier signal is now complex. This output $y(t)$ can be described as

$$y(t) = \sum_{n=0}^{N-1} a_n \sin(\omega_n t + I_n \sin(\omega_m t) + \phi_n) \quad (17)$$

where $\omega_m$ is the modulation frequency and $I_n$ is the index of modulation for each partial. According to Equation 11, this would be equivalent to the following signal:

$$y(t) = \sum_{n=0}^{N-1} a_n \left[ \begin{array}{c} J_0(I_n)\sin(\omega_n t + \phi_n) + \\ \sum_{k=1}^{I+1} \left\langle \begin{array}{c} J_k(I_n)\sin(\omega_n t + k\omega_m t + \phi_n) + \\ J_{-k}(I_n)\sin(\omega_n t - k\omega_m t + \phi_n) \end{array} \right\rangle \end{array} \right] \quad (18)$$

The different indices of modulation for each component of the carrier signal can be estimated by the following relationship, derived from Equation 9:

$$I_n = \Delta d \pi f_n = \frac{I}{\pi f_c} \pi f_n = I \frac{f_n}{f_c} \quad (19)$$

Again, we see here that the effect of the relationship between the index of modulation and the carrier frequency is that higher-frequency partials will be modulated more intensely than lower ones. Depending on the bandwidth and richness of the input signal, it is quite easy to generate very complex spectra, which might be objectionable in some cases. This increase in brightness has also been observed in other applications of audio-rate modulation of delay lines (Välimäki, Tolonen, and Karjalainen 1998; Tolonen, Välimäki, and Karjalainen 2000).

Turning now to the second technique introduced herein, we will have a significantly different output, described by

$$y(t) = 0.5 \left[ \begin{array}{c} \sum_{n=0}^{N-1} a_n \sin(\omega_n t + I\sin(\omega_m t) + \phi_n) \\ + \sum_{n=0}^{N-1} a_n \sin(\omega_n t + I\sin(-\omega_m t) + \phi_n) \end{array} \right] \quad (20)$$

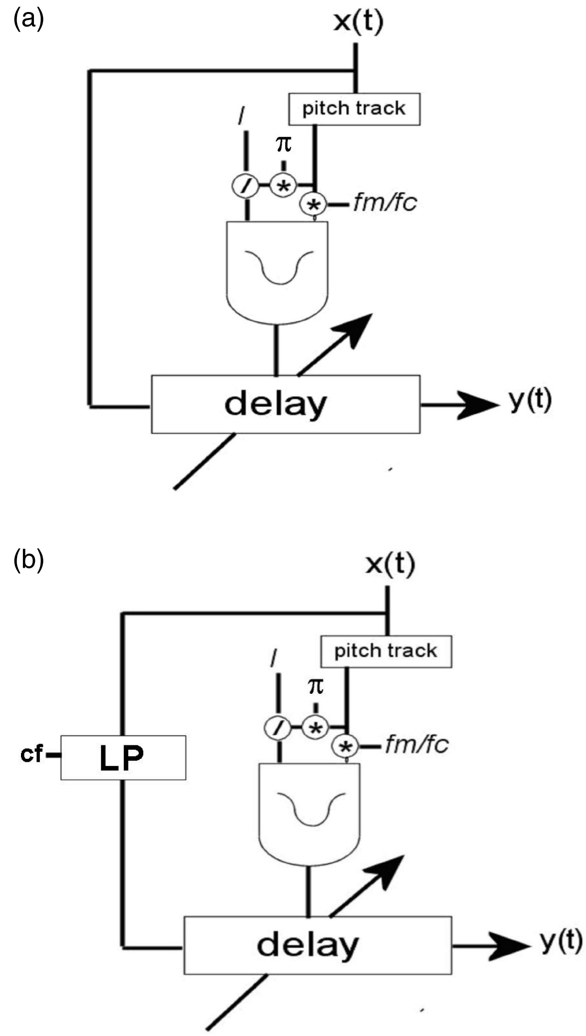
The most important differences between the spectrum of this signal and that described by Equation 18 are that odd sidebands are now canceled, and the index of modulation $I$ is now constant across the modulated carrier components. Whereas the former is responsible for an overall timbral difference between the two spectra, the latter is responsible for a more controlled and subtle handling of high frequencies.

Another key aspect of the proposed methods is that the *c:m* ratio parameter can also be taken advantage of by estimating the fundamental frequency of the input signal (assumed to be monophonic). In this case, a variety of different spectral combinations can be produced, from inharmonic to harmonic and quasi-harmonic.

## Fundamental Frequency Estimation

To allow for a full control of *c:m* ratio and modulation index, it is necessary to estimate the fundamental frequency of the carrier signal. That will allow the modulator signal frequency and amplitude to be set according to Equation 10. This can be

Figure 2. Delay-line based AdFM design: (a) original; (b) with the optional low-pass filter.




achieved with the use of a pitch tracker, which is a standard component of many modern music signal-processing systems. For the current implementation, a spectral-analysis pitch-tracking method was devised, based on an algorithm by Puckette, Apel, and Ziccarelli (1998) and Puckette and Brown (1998), that provides fine accuracy of fundamental-frequency estimation. In addition to tracking the pitch, it is also useful (but not essential) to obtain the amplitude of the input signal, which can be used in certain applications to scale the index of modulation. This is also provided by our parameter-estimation method.

### Signal Bandwidth

Although the spectrum of FM is, in practical terms, band-limited, it is capable of producing very high frequencies, as seen in Equations 11 and 18. With digital signals, this can lead to aliasing problems if the bandwidth of the signal exceeds the Nyquist frequency. The fact that in the delay-based formulation the index of modulation increases with frequency for a given $\Delta d$ (Equation 19) is obviously problematic. However, in practice, the kind of input signals we will be employing generally exhibit a spectral envelope that decays with frequency. In this case, objectionable aliasing problems might be greatly minimized, given that $a_n$ in Equation 18 for higher values of $n$ will be close to zero. Of course, if our input contains much energy in the higher end of the spectrum, such as for instance an impulse train, then aliasing will surely occur.

The simplest solution for such problematic signals is to impose a decaying spectral envelope using a filter. This will have the obvious side-effect of modifying the timbre of the input signal. Another, more computationally costly, solution is to oversample the input signal. This would either remove the aliased signals or place them at an inaudible range.

## Implementation

We now present a reference implementation of AdFM using both methods of phase modulation described herein. These two instrument designs can serve as the basis for further software- or hardware-based implementations. The basic flow chart of the delay-based PM instrument is shown in Figure 2a. There are three basic components: a pitch tracker, a modulating source (a table-lookup oscillator), and a variable delay line with interpolated readout. Each of these components is found in modern music signal-processing systems, so the technique is highly portable. The implementation discussed here uses Csound 5 (ffitch 2005) as the synthesis engine, but similar instruments can be developed under other musical signal-processing environments,

Figure 3. Delay-based AdFM code.

```
/* DFM opcode
  asig DFM  ain,krat,kndx,ifn
  ain - input signal
  krat - c:m ratio
  kndx - index of modulation
  ifn - mod signal function table
*/
opcode DFM,a,akki
 setksmps 1
 ipi = $M_PI        /* pi constant */
 ioff = 2/sr        /* 2-sample offset*/
 as,krt,knx,ifn xin /* input parameters */
 kcps,kamp ptrack as,512 /* pitch tracking */
 /* modulator */
 adt oscili knx/(ipi*kcps),kcps/krt,ifn
 adp   delayr 1  /* delay line */
 adel  deltap3  adt + ioff /* delay tap output */
    delayw  as  /* delay input */
       xout adel /* Opcode output */
endop
```

such as the SndObj library and PySndObj (Lazzarini 2000, 2007). It is important to note that this design can be used either for real-time or off-line applications. In addition, plug-ins can be easily developed from it using csLadspa (Lazzarini and Walsh 2007).

The equivalent Csound 5 code for the flowchart design in Figure 2, which implements the delay-based version, is shown in Figure 3. The heterodyning PM design is simpler, based on a more or less straight translation of the formula in Equation 13. Its flowchart is shown in Figure 4 and the corresponding Csound code in Figure 5.

Both implementations use a spectral-analysis pitch-tracking opcode (`ptrack`) written by the authors and linear interpolation oscillators to generate the modulation signal. The `DFM` opcode uses a cubic-interpolation variable delay line (Laakso et al. 1996). Owing to the use of cubic interpolation, the minimum delay is set to two

Figure 4. The heterodyning AdFM design.

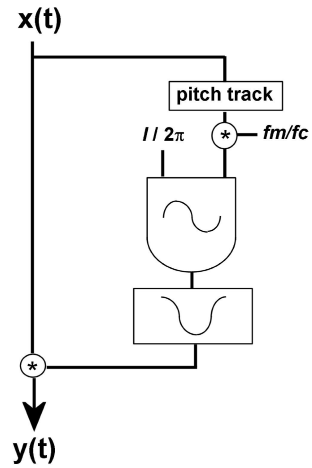



Figure 4

Figure 5. Heterodyning AdFM Csound code.

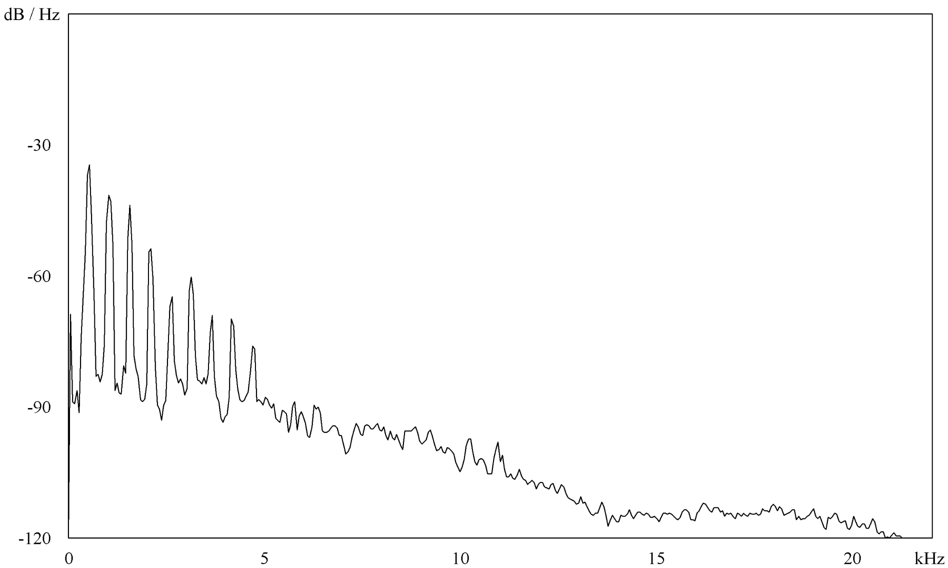
```
/* HFM opcode
  asig HFM  ain,krat,kndx,ifn
  ain - input signal
  krat - c:m ratio
  kndx - index of modulation
  ifn - mod signal function table
*/
opcode HFM,a,akki
 setksmps 1
 i2pi = 2*$M_PI    /* two-pi constant */
 asig,kratio,kndx,ifn xin /* input parameters */
 kcps,kamp ptrack asig, 512 /* pitch tracking */
 kfm = kcps/kratio  /* fm frequency  */
 adt oscili kndx/i2pi, kfm, ifn  /* modulator */
 acs tablei adt, ifn, 1, 0.25, 1 /* phase modulation */
       xout  asig*acs /* heterodyning */
endop
```

Figure 5

samples to avoid errors in the circular-buffer readout.

A number of variations can be made to the basic design. For instance, the amplitude of the signal, which is produced together with the pitch-tracking information, can be used to scale the index of modulation. This can be used to generate typical

Figure 6. Steady-state spectrum of a flute playing C4.




brass-like synthesizer tones (Risset 1969), where the brightness of the synthetic output is linked to the amplitude evolution of the input sound. Alternatively, it can be used to determine the *c*:*m* ratio.

Depending on the characteristics of the input signal, it might be useful to include a low-pass filter before the signal is sent to the AdFM processors, especially in the delay-based-version, as shown in Figure 2b. The cutoff frequency of the low-pass filter can also be controlled by the estimated input amplitude. As discussed earlier, this will reduce aliasing as well as overall brightness, both of which are sometimes a downside of FM synthesis.

## Examples and Discussion

Four different types of carrier signals were chosen as a way of examining the qualities of the AdFM synthetic signal using both methods described in this article. A flute input with its spectral energy concentrated in the lower harmonics is a prime candidate for experimentation. The clarinet was chosen for its basic quality of having more prominent odd harmonics. Finally, the piano and voice were used as a means of exploring the possibilities of synthesizing different types of harmonic and inharmonic spectra by the use of various *c*:*m* ratios.

The sound examples discussed here will be found on the annual *Computer Music Journal* DVD (to be released with the Winter 2008 issue).

### Flute Input

The original steady-state flute spectrum, effectively with $I = 0$, is shown in Figure 6. As clearly seen in that figure, it features quite prominent lower harmonics. Using delay-line AdFM and applying an index of modulation of 0.3 on a 1:1 *c*:*m* configuration, we can start enriching the spectrum with higher harmonics (see Figure 7). At these low values of $I$, there is already a considerable addition of components between 5 and 10 kHz. The overall spectral envelope still preserves its original decaying shape.

Using the delay line method with higher values of $I$, we can see a dramatic change in the timbral characteristics of the original flute sound. Figure 8 shows the resulting spectrum, now with $I = 1.5$. Here, we can see that components are now spread to the entire frequency range. The original decaying spectral envelope is distorted into a much more gradual shape, and the difference between the loudest and the softest harmonic is only about 30 dB. The resulting sound can been described as "string-like," and the transition between the flute

*Figure 7. AdFM spectrum using a flute C4 signal as carrier with* c:m = *1 and* I = *0.3.*

*Figure 8. AdFM spectrum using same input as Figure 3, but now with* I = *1.5.*

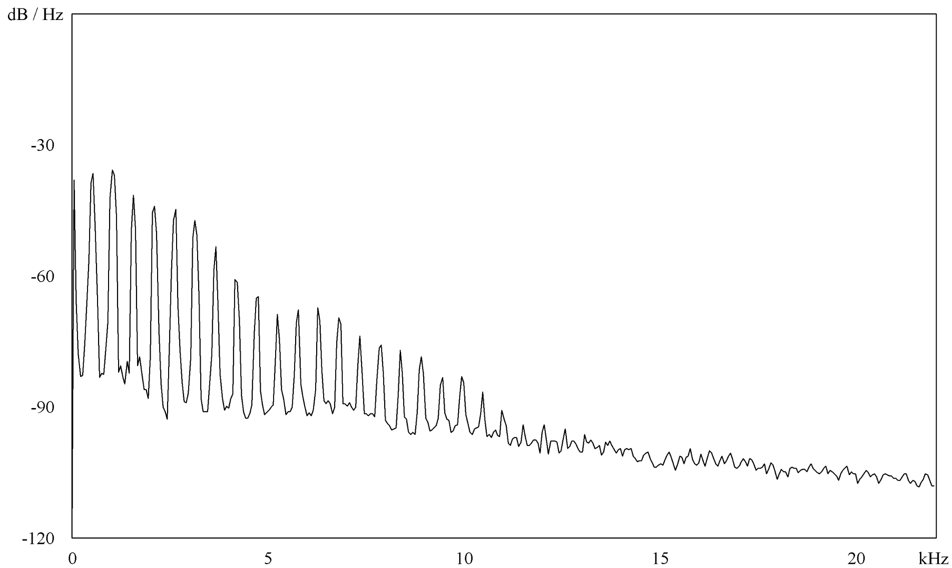



*Figure 7*

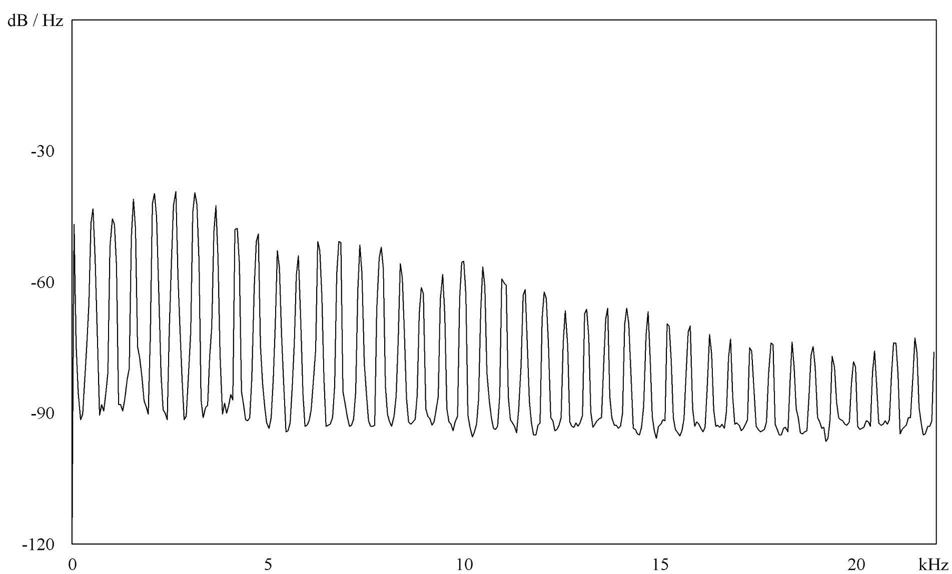
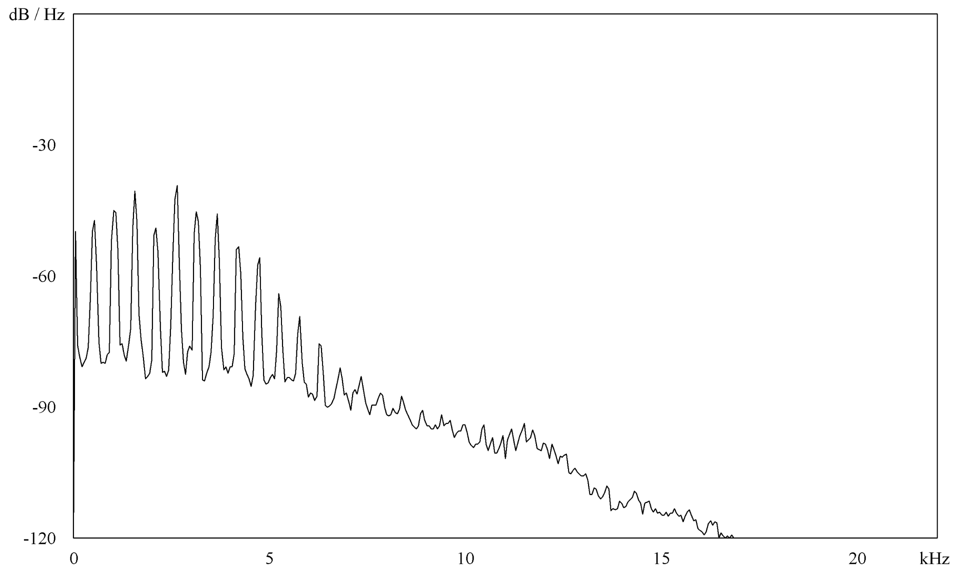



*Figure 8*

and AdFM spectra is capable of providing interesting possibilities for musical expression. Also, it is important to note that important gestural characteristics of the original sound, such as pitch fluctuations, vibrato, and articulation, are preserved in the synthetic output.

As *I* gets higher, the spectrum gets even brighter, but the problems with aliasing start to become significant. To prevent this and also to allow for a different spectral envelope, an optional low-pass filtering of the input signal is suggested. In that case, the filter is inserted in the signal path at the

*Figure 9. Heterodyning AdFM synthesis using a flute input with* I = *5 and* c:m = *1.*




delay-line input. A Butterworth low-pass filter with a cutoff frequency between 1,000 and 5,000 Hz has proven useful. It is possible to couple the cutoff frequency with $I$, so that for higher values of that parameter, more filtering is applied.

The addition of higher harmonics is significantly reduced in the heterodyning AdFM method. We can see in Figure 9 how much more attenuated the top end of the spectrum is in comparison to the previous technique. This in some cases might be advantageous; however, the effect of the technique is subtler, resulting in a transition between natural-sounding and synthesizer-like spectra that is less dramatic.

**Clarinet Input**

Our second experiment used a clarinet signal as a carrier wave for AdFM. The clarinet exhibits a steady-state spectrum in which the lower-order even harmonics are significantly less energetic than their odd neighbors (see Figure 10). As a result, the multiple-carrier-like characteristic of AdFM helps generate quite a change in the spectra of that instrument.

As the index of modulation increases, the balance between odd an even harmonics changes substantially. In delay-line AdFM with $I = 1.5$, it is possible to see that there is very little difference between the strengths of odd and even components (see Figure 11). In addition, higher-order harmonics become more present, and the spectral envelope levels out, owing to the well-known spread of energy that is characteristic of FM synthesis.

The heterodyning method also provides similar transformations, although again with more subtle high-frequency results, and still retaining some of the odd/even balance of the input. Figure 12 demonstrates that the resulting spectrum features a decaying envelope, in contrast to the previous example (see Figure 11), which is much flatter.

**Piano Input**

In the previous examples, we have kept the ratio between the modulating frequency and carrier fundamental at unity. However, as we know from FM theory, a range of different spectra is possible if we use different ratios. It is possible to create a range of effects that range from changing the fundamental of the sound to transforming a harmonic spectrum into an inharmonic one. We took a piano C2 signal as our carrier and then tuned our modulator to 1.41 times that frequency. The original piano spectrum

*Figure 10. Detail of steady-state spectrum of clarinet C3. Note the higher relative strength of lower-order odd harmonics versus even ones.*

*Figure 11. Detail of AdFM spectrum using a clarinet C3 signal as carrier with* c:m = *1 and* I = *1.5. Odd and even harmonics now have comparable strengths.*

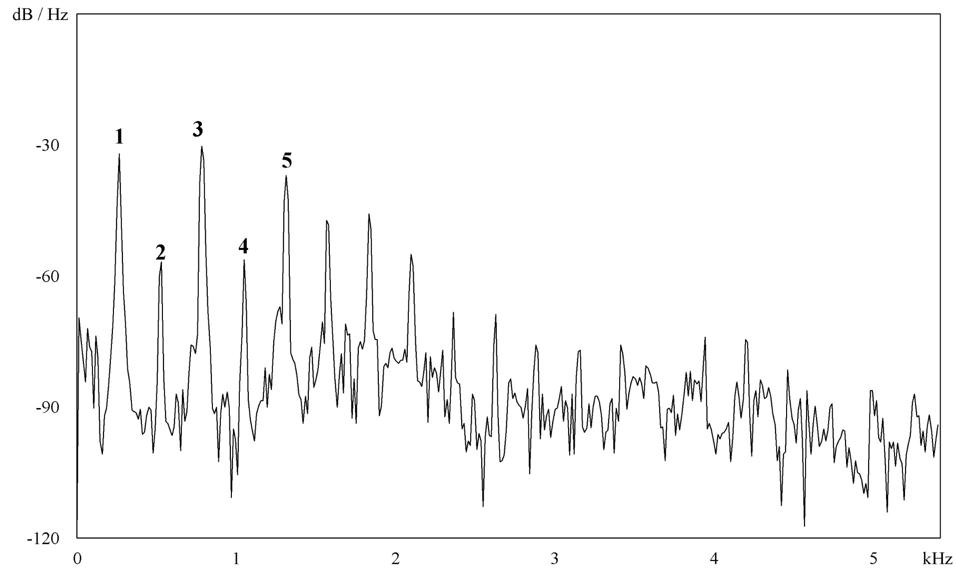



*Figure 10*

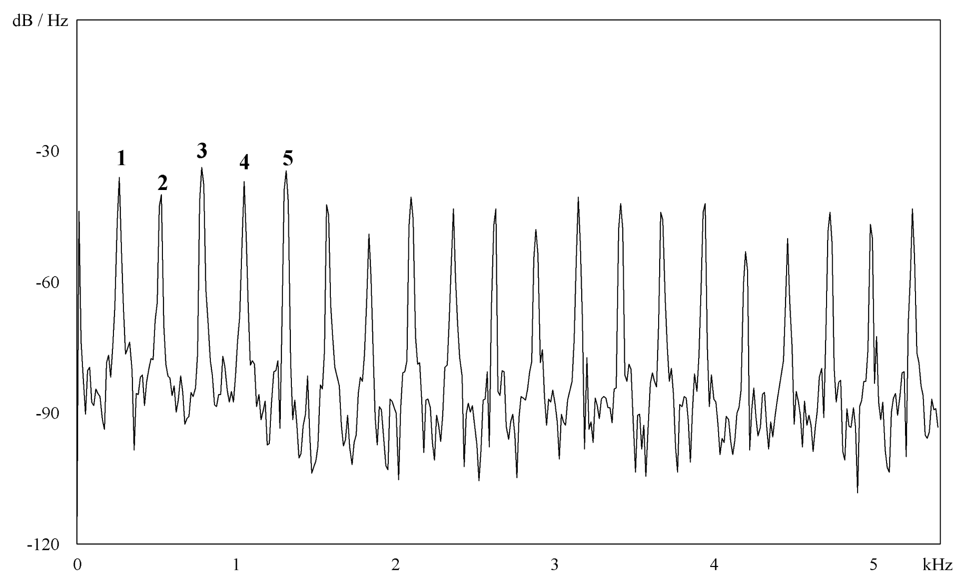



*Figure 11*

is shown in Figure 13, where we can clearly see its harmonics.

The resulting delay-line AdFM spectrum with $I = 0.15$ is shown in Figure 14. This particular ratio creates a great number of components whose relationship implies a very low fundamental, thus generating what is perceived as an inharmonic spectrum. With the 1:1 ratio, the sums and differences between $f_c$ and $f_m$ created components whose frequencies were mostly coincident. Here, a variety of discrete components will be generated, creating the denser spectrum seen in Figure 14. The AdFM

*Figure 12. Steady-state spectrum of clarinet-input heterodyning AdFM, with* I = *5 and* c:m = *1.*

*Figure 13. Spectrogram of a piano C2 tone, showing its first harmonics in the 0–1.2 kHz range.*

*Figure 14. Spectrogram of an AdFM sound using a piano C2 signal as carrier, with* c:m = *1:1.41 and* I = *0.15, showing the 0–1.2 kHz range. The resulting inharmonic spectrum, with a large number of components, is clearly seen in comparison with Figure 13.*

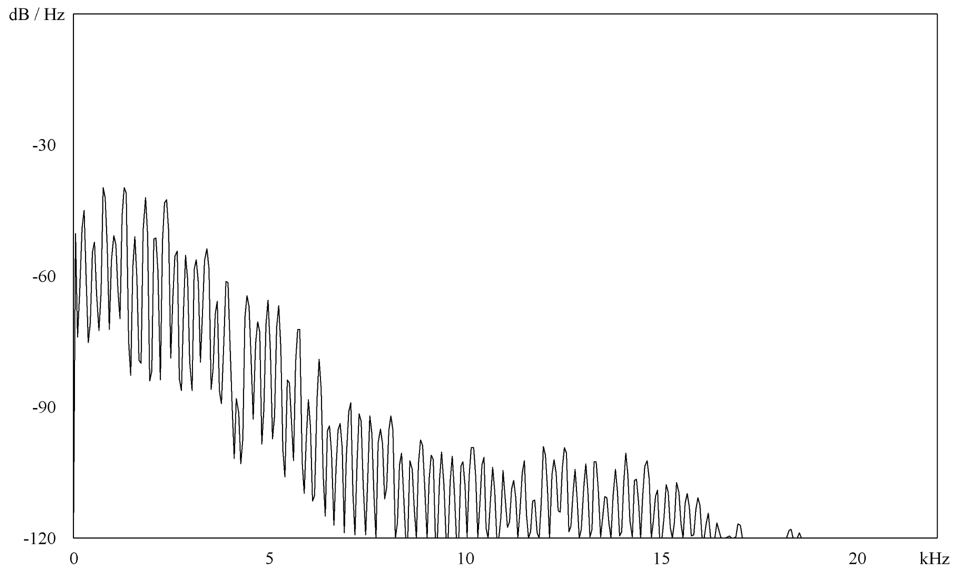



*Figure 12*

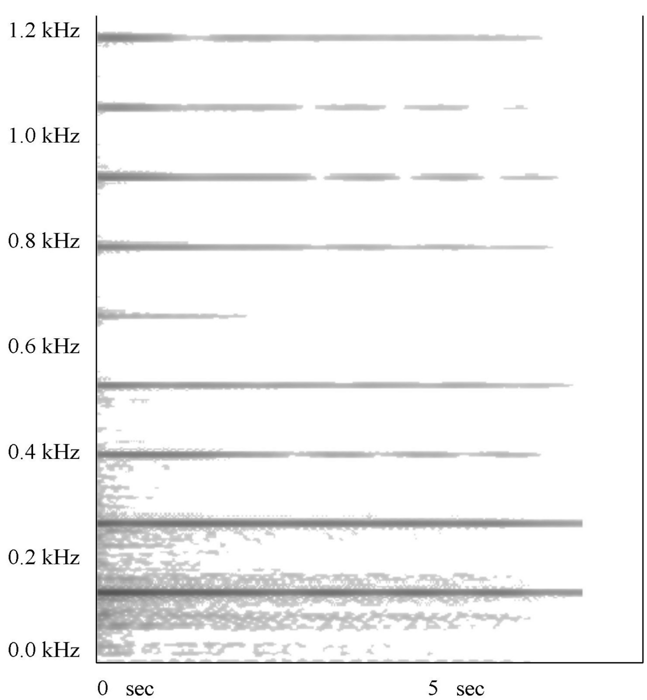



*Figure 13*

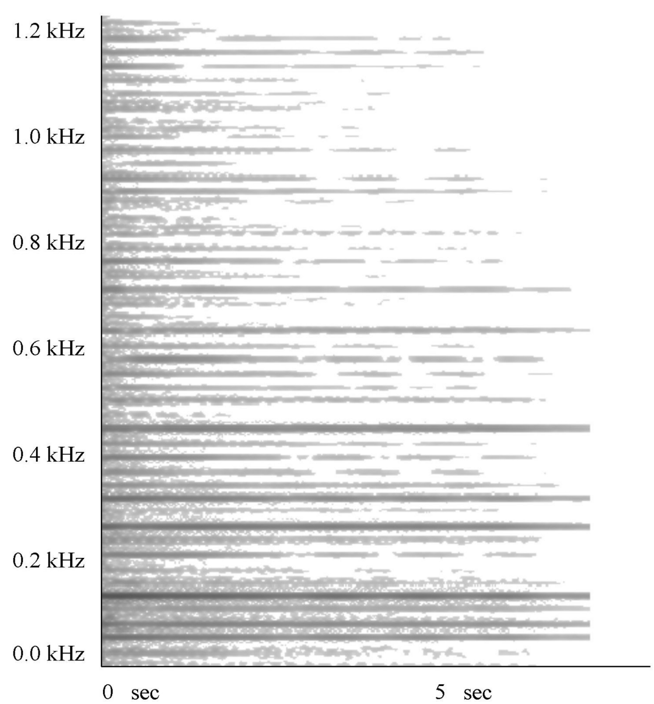
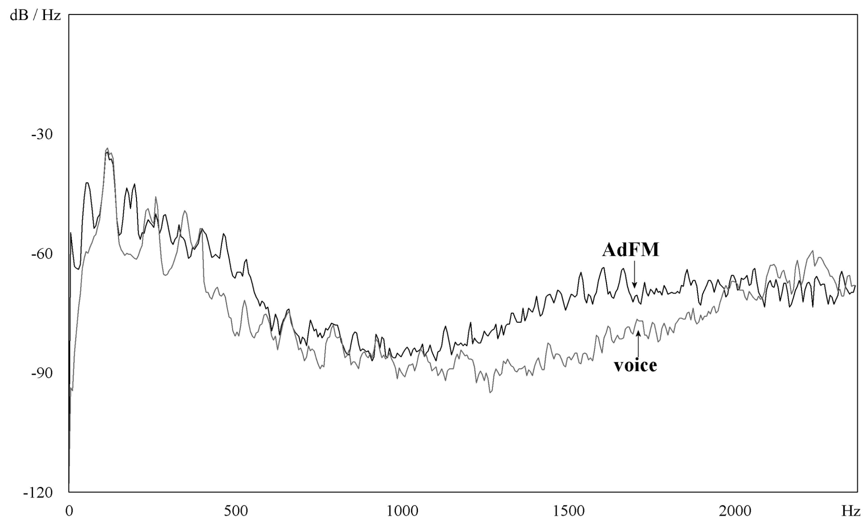



*Figure 14*

Figure 15. Comparison of spectral snapshots of a vocal and an AdFM vocal sounds, with I = 0.1 and c:m = 2.




sound resulting from this arrangement has been described as "bell-like." Transitions between piano and bell sounds can be effected by changing $I$ from 0 to the desired value. The application of a low-pass filter at the delay-line input will also allow for some variety and control over the brightness of the result.

Again, if we apply the heterodyning technique instead to this input using a similar ratio, we will obtain a bell-like output that is better behaved in the higher end of the spectrum. Here the second method might in fact be more useful, as it can control the quality of the output more effectively.

## Voice Input

A vocal input was used as the fourth different source examined in this work, demonstrating a pitch-shift effect. Setting the $f_c$:$f_m$ ratio to 2, we are able to obtain a sound that is now half the pitch of the original. This is due to the introduction of a component at half the fundamental frequency corresponding to $f_c - f_m$ in Equation 18.

With the index of modulation at low values (around 0.15), it is possible to preserve some of the spectral shape of the original sound, a crucial step in keeping the intelligibility of the vocal phonemes. Although there is some addition of high-frequency components and a flattening of spectral peaks, the AdFM voice is still perfectly intelligible.

Figure 15 shows a comparison between a vowel steady-state spectrum and its AdFM-processed counterpart. The sub-harmonic peak can be seen at the left of the picture below the original fundamental. (A peak at 0 Hz is also present, owing to the $f_c - 2f_m$ component.) The recording of the phrase, "This is AdFM Synthesis," is shown as a spectrogram in Figure 16, both as the original signal (left) and the AdFM output (right), using the same parameters as in the previous example. Again, the octave change is clearly seen, as well as the increase in the number of significant components in the signal.

In general, we achieved better results using the delay-line method with vocal inputs. The heterodyning process seems to be too prone to artifacts generated by unvoiced phonemes, resulting in chirps and glitches. Although these are originally caused by the pitch-tracking mechanism, they are emphasized by certain characteristics of the method's implementation.

## Conclusion

We presented an alternative approach to the classic technique of FM synthesis, based on an adaptive

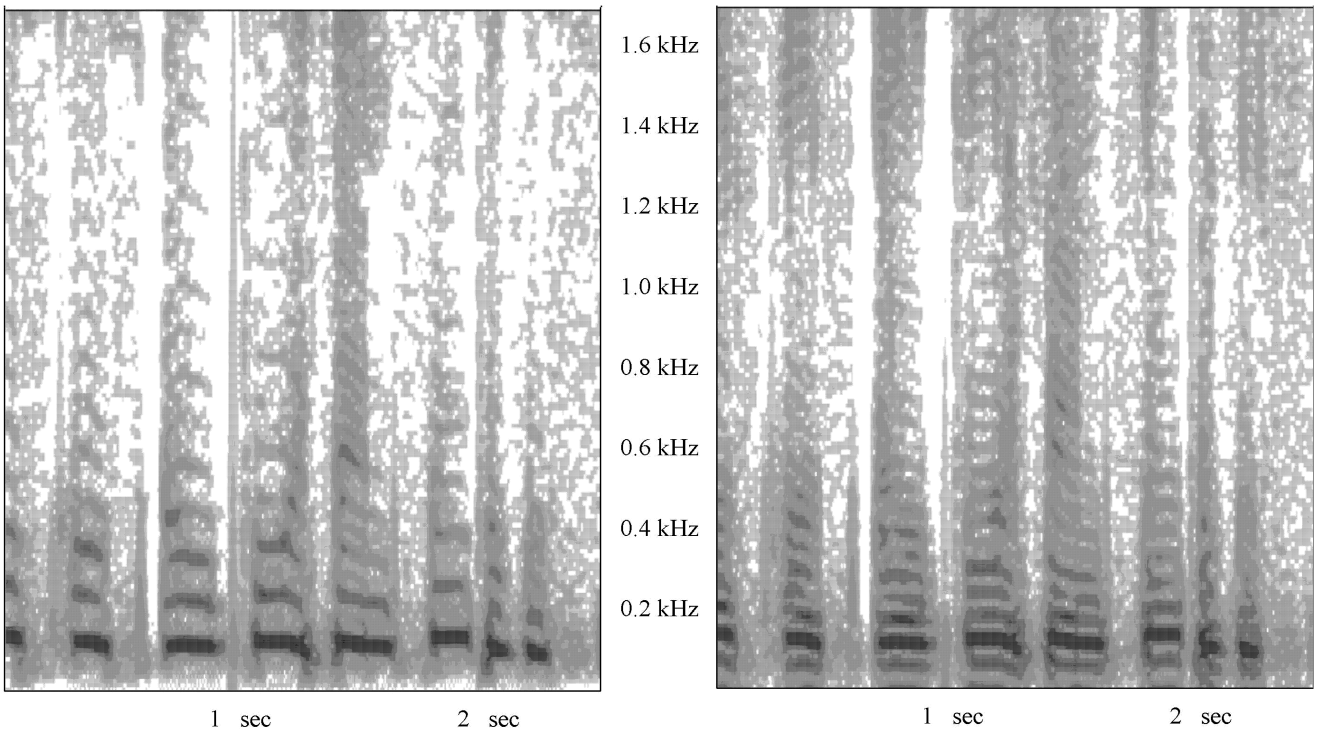



*Figure 16. Detail of spectrogram of a recording of the phrase, "This is AdFM synthesis," with the original vocal sound on the left and the AdFM vocal on the right*

design, which we call AdFM. Two different methods were proposed as a means of modulating an arbitrary carrier signal. As the FM synthesis theory is well known, it was possible to adapt it to determine the precise characteristics of the output signal. With this technique, it is possible to achieve fine control over the synthetic result, which also preserves a substantial amount of the gestural information in the original signal. Four different types of carrier signals were used in this work to demonstrate the wide range of spectra that the technique can generate. We are confident this is a simple yet effective way of creating hybrid natural-synthetic sounds for musical applications.

Future prospects for research into AdFM involve the development of alternative implementations of the technique, both in terms of time-domain variations of the methods discussed here and new frequency-domain processes. The latter have been facilitated by the development of the Sliding Phase Vocoder (SPV; Bradford, Dobson, and ffitch 2007), which allows for audio-rate modulation of its parameters. It is our plan to develop a spectral version of AdFM in Csound, as SPV analysis/synthesis and audio-rate frequency scaling have been added to the language in version 5.07.

## References

Bradford, R., R. Dobson, and J. ffitch. 2007. "The Sliding Phase Vocoder." *Proceedings of the 2007 International Computer Music Conference.* San Francisco: International Computer Music Association, pp. 449–452.

Chowning, J. 1973. "The Synthesis of Complex Audio Spectra by Means of Frequency Modulation." *Journal of the Audio Engineering Society* 21:526–534.

Chowning, J. 1989. "Frequency Modulation Synthesis of the Singing Voice." In M. Mathews and J. R. Pierce, eds., *Current Directions in Computer Music Research.* Cambridge, Massachusetts: MIT Press, pp. 57–63.

Dilsch, S., and U. Zölzer. 1999. "Modulation And Delay Line Based Digital Audio Effects." *Proceedings of the 2nd Conference on Digital Audio Effects.* Trondheim:

Norwegian University of Science and Technology, pp. 5–8.
Dodge, C., and T. Jerse. 1985. *Computer Music.* New York: Schirmer Books.
ffitch, J. 2005. "On the Design of Csound5." *Proceedings of the 3rd International Linux Audio Conference.* Karlsruhe: Zentrum für Künst und Medientechnologie, pp. 37–42.
Horner, A. 1996. "Double-Modulator FM Matching of Instrument Tones." *Computer Music Journal* 20(2):57–71.
Horner, A., J. Beauchamp, and L. Hakken. 1993. "Machine Tongues XVI: Genetic Algorithm and Their Application to FM Synthesis." *Computer Music Journal* 17(4):17–29.
Laakso, T. I., et al. 1996. "Splitting the Unit Delay: Tools for Fractional Delay Filter Design." *IEEE Signal Processing Magazine* 13(1):30–60.
Lazzarini, V. 2000. "The Sound Object Library." *Organised Sound* 5(1):35–49.
Lazzarini, V. 2007. "Musical Signal Scripting with PySndObj." *Proceedings of the 5th International Linux Audio Conference.* Berlin: Technische Universität Berlin, pp. 18–23.
Lazzarini, V., J. Timoney, and T. Lysaght. 2007. "Adaptive FM Synthesis." Proceedings of the 10th International Conference on Digital Audio Effects. Bordeaux: University of Bordeaux, pp. 21–26.
Lazzarini, V., and R. Walsh. 2007. "Developing LADSPA Plugins with Csound." *Proceedings of the 5th International Linux Audio Conference.* Berlin: Technische Universität Berlin, pp. 30–36.
Miranda, E., and M. Wanderley. 2006. *New Digital Musical Instruments.* Middleton, Wisconsin: A-R Editions.
Palamin, J.P., P. Palamin, and A. Ronveaux. 1988. "A Method of Generating and Controlling Musical Asymmetric Spectra." *Journal of the Audio Engineering Society* 36(9):671–685.
Poepel, C. 2004. "Synthesized Strings for String Players." *Proceedings of 2004 Conference on New Instruments for Musical Expression.* New York: Association for Computing Machinery, pp. 150–153.
Poepel, C., and R. Dannenberg. 2005. "Audio Signal Driven Sound Synthesis." *Proceedings of the 2005 International Computer Music Conference.* Barcelona: International Computer Music Association, pp. 391–394.
Puckette, M., T. Apel, and D. Ziccarelli. 1998. "Real-Time Audio Analysis Tools for PD and MSP." *Proceedings of the 1998 International Computer Music Conference.* San Francisco: International Computer Music Association, pp. 109–112.
Puckette, M., and J. Brown. 1998. "Accuracy of Frequency Estimates from the Phase Vocoder." *IEEE Transactions on Speech and Audio Processing* 6(2):116–172.
Risset, J. C. 1969. *An Introductory Catalogue of Computer Synthesized Sounds.* Murray Hill, New Jersey: AT&T Bell Laboratories.
Schottstaedt, W. 1977. "The Simulation of Natural Instrument Tones Using a Complex Modulating Wave." *Computer Music Journal* 1(4):46–50.
Tan, B. J. , and S. L. Gan. 1993. "Real-Time Implementation of Asymmetrical Frequency-Modulation Synthesis." *Journal of the Audio Engineering Society* 41(5):357–363.
Tolonen, T., V. Välimäki, and M. Karjalainen. 2000. "Modeling of Tension Modulation Nonlinearity in Plucked Strings." *IEEE Transactions on Speech and Audio Processing* 8(3):300–310.
Välimäki, V., T. Tolonen, and M. Karjalainen. 1998. "Signal-Dependent Nonlinearities for Physical Models Using Time-Varying Fractional Delay Filters." *Proceedings of the 1998 International Computer Music Conference.* San Francisco: International Computer Music Association, pp. 264–267.
Van Duyne, S. A., and J. O. Smith. 1992. "Implementation of a Variable Pick-Up Point on a Waveguide String Model with FM/AM applications." *Proceedings of the 1992 International Computer Music Conference.* San Francisco: International Computer Music Association, pp. 154–157.
Verfaille, V., and D. Arfib. 2002. "Implementation Strategies for Adaptive Digital Effects." *Proceedings of the 5th Conference on Digital Audio Effects.* Hamburg: University of the Federal Armed Forces, pp. 21–26.
Verfaille, V., U. Zölzer, and D. Arfib. 2006. "Adaptive Digital Audio Effects (a-DAFx): A New Class of Sound Transformations." *IEEE Transactions on Audio, Speech and Language Processing* 14(5):1817–1831.
Wanderley, M. M., and P. Depalle. 2004. "Gestural Control of Sound Synthesis." *Proceedings of the IEEE 92(4):*632–644.