The Simulation of Natural Instrument Tones using Frequency Modulation with a Complex Modulating Wave

Author(s): Bill Schottstaedt

Source: *Computer Music Journal*, November 1977, Vol. 1, No. 4 (November 1977), pp. 46-50

Published by: The MIT Press

Stable URL: https://www.jstor.org/stable/40731300

The Simulation of Natural Instrument Tones using Frequency Modulation with a Complex Modulating Wave

Bill Schottstaedt
Center for Computer Research in Music and Acoustics
Stanford University

Introduction

The frequency modulation technique of producing musically interesting audio spectra can be extended in several useful ways. In the basic case [Chowning, 1973], a carrier wave is modulated by one modulating wave where both waves are simple sinusoids and both are in the audio band. The modulated carrier then becomes the audio signal. By changing the modulation index and the ratio of the carrier frequency to the modulating frequency, one can easily create many complex audio spectra.

Certain timbres, however, seem to be elusive. Over the last several years at the Stanford Center for Computer Research in Music and Acoustics, we have been doing extensive work with complex modulating waves. We have found this process to be very useful in, for example, the simulation of string and piano tones. The simulations described below use a complex modulating wave made up of two or three sinusoids, each of whose modulation indices is pitch-dependent. Separate modulating oscillators are used to allow one to control the modulation indices independently, thus giving more precise control over the spectra produced. Analysis of actual piano and string tones served as the basis for the choice of the indices and of the ratios between carrier and modulating frequencies [cf. "Lexicon of Analyzed Tones," 1977].

Frequency Modulation with a Complex Modulating Wave

When a sinusoidal wave is simultaneously frequency-modulated by two modulating sinusoids, side bands are created at all frequencies of the form:

$$f_c \pm i f_{m1} \pm k f_{m2}$$

where i and k are integers, and f_{m1} and f_{m2} are the modulating frequencies. It is as though each of the side bands produced by one of the modulating signals were modulated as a carrier by the other modulating signal. The equation of the resultant wave is

$$e = A \sin(\omega_c t + I_1 \sin\omega_1 t + I_2 \sin\omega_2 t) \qquad (1)$$

where e is the instantaneous amplitude, A is the amplitude of the carrier, ω_c is $2\pi f_c$, ω_1 is $2\pi f_{m1}$, and ω_2 is $2\pi f_{m2}$. The amplitude of the side band $f_c \pm i f_{m1} \pm k f_{m2}$ is given [Le Brun, 1977] by:

$$J_i(I_1) \times J_k(I_2)$$

where I_1 and I_2 are the respective modulating indices of f_{m1} and f_{m2}, and J is the Bessel function of the first kind and ith (or kth) order. For example, the amplitude of the carrier is $J_0(I_1) \times J_0(I_2)$.

It can be shown [Le Brun, 1977] that:

$$e = \sum_i \sum_k J_i(I_1) J_k(I_2) \sin(\omega_c t + i\omega_1 t + k\omega_2 t) \qquad (2)$$

As an illustration of this process, take the simple case of a sine wave modulated simultaneously by two sine waves. The ratios of modulating frequency to carrier frequency will be 1:1 and 4:1, with modulation indices of 1 and .2, respectively (as is approximately the case at a frequency of 400 Hz. in the FM piano described below). In general, the amplitudes of the components would be:

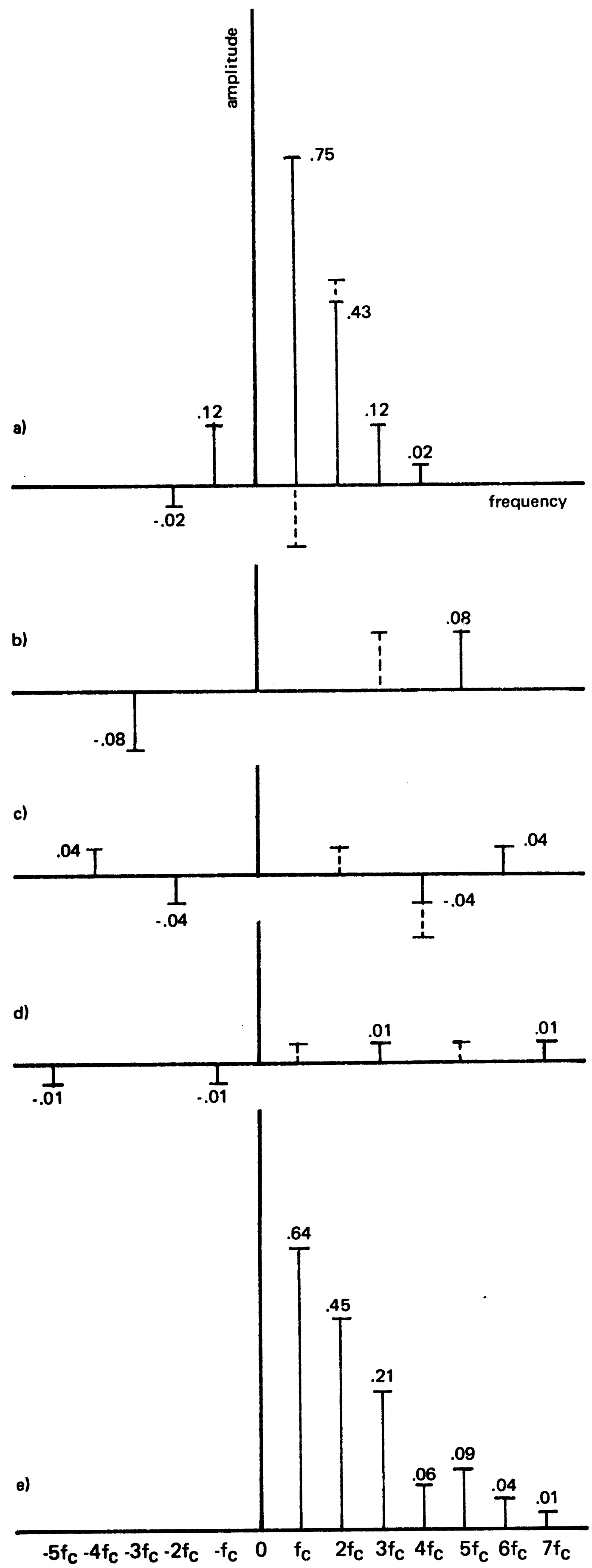
frequency	*amplitude*
f_c, carrier	$J_0(I_1) \times J_0(I_2)$
$f_c \pm i f_{m1}$, simple side bands for f_{m1}	$J_i(I_1) \times J_0(I_2)$
$f_c \pm k f_{m2}$, simple side bands for f_{m2}	$J_0(I_1) \times J_k(I_2)$
$f_c \pm i f_{m1} \pm k f_{m2}$, combination frequencies	$J_i(I_1) \times J_k(I_2)$

If $k = 0$, then only the lower (1:1) modulating frequency would be present, and the relative side band amplitudes would be:

$$\begin{aligned} J_0(I_1) &= J_0(1) = .77 \\ J_1(I_1) &= J_1(1) = .44 \\ J_2(I_1) &= J_2(1) = .12 \\ J_3(I_1) &= J_3(1) = .02 \end{aligned}$$

If $i = 0$, then only the higher (4:1) modulating frequency would be present, and the harmonics would have the following relative amplitudes:

$$\begin{aligned} J_0(I_2) &= J_0(.2) = .98 \\ J_1(I_2) &= J_1(.2) = .10 \\ J_2(I_2) &= J_2(.2) = .005 \end{aligned}$$



However, when both modulating frequencies are present, they interact to produce the following side band amplitudes:

$J_0(I_1) \times J_0(I_2) = .77 \times .98 = .75$
(carrier, f_c)

$J_1(I_1) \times J_0(I_2) = .44 \times .98 = .43$
(first order side bands, $f_c \pm f_{m1}$)

$J_2(I_1) \times J_0(I_2) = .12 \times .98 = .12$
(second order side bands, $f_c \pm 2f_{m1}$)

$J_3(I_1) \times J_0(I_2) = .02 \times .98 = .02$
(third order side bands, $f_c \pm 3f_{m1}$)

$J_0(I_1) \times J_1(I_2) = .77 \times .10 = .08$
(first order side bands, $f_c \pm f_{m2}$)

$J_1(I_1) \times J_1(I_2) = .44 \times .10 = .04$
($f_c \pm f_{m1} \pm f_{m2}$, combination frequencies where $i = 1$, $k = 1$)

$J_2(I_1) \times J_1(I_2) = .12 \times .10 = .01$
($f_c \pm 2f_{m1} \pm f_{m2}$, combination frequencies where $i = 2$, $k = 1$).

The amplitudes of higher-order side bands are insignificant.

The process by which these side bands interact to form a spectrum is shown in Fig. 1. The solid bars in Figures 1a - 1d represent side bands with amplitudes as derived above. Some of these components are drawn below the horizontal axis, representing a 180° phase shift which results [Le Brun, 1977] from the relation

$$J_{-n} = (-1)^n J_n$$

Thus, the sign for the side band $f_c - 2f_{m1} - f_{m2}$ (at $-5f_c$ in Fig. 1d) is given by

$$\begin{aligned} & J_{-2}(I_1) \times J_{-1}(I_2) \\ = & (-1)^2 J_2(I_1) \times (-1) J_1(I_2) \\ = & -J_2(I_1) \times J_1(I_2), \end{aligned}$$

whereas the sign for the side band $f_c - f_{m1} - f_{m2}$ (at $-4f_c$ in Fig. 1c) is given by

Figure 1. Plot of spectrum with $f_{m1} : f_c = 1:1$, $f_{m2} : f_c = 4:1$ (not drawn to scale). "Negative frequency" components in a) through d) are shown where they theoretically occur (solid lines) as well as "wrapped around" 0 Hz. and phase-shifted 180° (dashed lines). The component at 0 Hz. is omitted here.
a). Spectral components resulting from f_c and f_{m1}.
b). Spectral components resulting from f_{m2} and f_c.
c). Combination components for $i = 1$, $k = 1$ (cf. Eq. 2).
d). Combination components for $i = 2$, $k = 1$.
e). Plot of the magnitudes in the final spectrum obtained by adding the components shown in Fig. 1a – 1d.

$$
\begin{aligned}
&J_{-1}(I_1) \times J_{-1}(I_2) \\
&= (-1) J_1(I_1) \times (-1) J_1(I_2) \\
&= +J_1(I_1) \times J_1(I_2)\ .
\end{aligned}
$$

As was explained in [Chowning, 1973], the phase of a side band with frequency less than 0 Hz. is also changed because it "wraps around" 0 Hz. by means of the relation $\sin(-\theta) = -\sin(\theta)$. In other words, the bar representing each "negative frequency" side band changes direction as it "wraps around" 0 Hz. Thus, the combination component in Fig. 1a which occurs at $-f_c$ in reality occurs at $+f_c$ (represented by a dashed bar at f_c), but undergoes a 180° phase shift. As a result, energy is in effect subtracted from the .75 component present at the fundamental (f_c in Fig. 1a). Adding from top to bottom in Fig. 1, the final energy of the fundamental is derived from components given by ($i = 0$, $k = 0$) in Fig. 1a, ($i = -2$, $k = 0$) in Fig. 1a, and ($i = 2$, $k = -1$) in Fig. 1d. The resultant energy contributions to the carrier are +.75, −.12, and −(−.01), adding up to +.64, as shown in Fig. 1e. This procedure is repeated to derive the following spectrum (Fig. 1e):

1 (fundamental)	.75 − .12 − (−.01)	= .64
2	.43 − (−.02) − .04	= .41
3	.12 − (−.08) + .01	= .21
4	.02 − .04 − .04	= (−).06
5	.08 − (−.01)	= .09
6		.04
7		.01

In this case, the contribution of the combination frequencies is relatively unimportant. As the index increases, however, higher and higher values of i and k need to be taken into account, causing the contribution of the combination frequencies to become more important.

By modulating the modulating wave itself, one can create a modulating wave with a (theoretically) infinite number of sinusoidal components. The instantaneous amplitude (e) would then be given by the following formula:

$$e = A \sin[\omega_c t + I_1 \sin(\omega_1 t + I_2 \sin\omega_2 t)] \qquad (3)$$

Here I_2 determines the number of significant components in the modulating signal, and I_1 determines the number of significant components in the output signal. The ratio ω_1/ω_c determines the placement of the carrier's side bands, each of which has side bands of its own at intervals determined by ω_2/ω_1. Each side band is both modulated and modulator; the derivation of an equation analogous to Eq. 2 is rather complicated.

A Piano Simulation

To simulate piano tones it was found best to use a modulating wave made up of two sinusoidal components, one at approximately the carrier frequency and one at approximately four times the carrier frequency. By using two modulating oscillators (one for each component), the amplitude of each component can be handled independently. The amplitudes of the components of the modulating wave (the modulating indices) are made frequency-dependent according to the following formulae:

$$
\begin{aligned}
I_1 &= 17(8 - \log_e f_c) \,/\, (\log_e f_c)^2 \qquad (4) \\
I_2 &= 20(8 - \log_e f_c) \,/\, f_c
\end{aligned}
$$

These formulae produce a spectrum that is rich and complex in the lower register, but gets steadily simpler as the pitch rises. The resulting evolution of the spectrum gives a good piano simulation over its entire range, and avoids foldover.

The component frequencies of the modulating wave are slightly larger than integral multiples of the carrier frequency. A small constant, S, with a value of 0.5% of the carrier frequency, is added to the modulating frequencies to simulate the characteristic "stretched" or sharp partials of the piano. The resulting inharmonicity is less than that of a traditional piano [Young, 1954], but larger values for S created noticeable beats. A truly harmonic spectrum invariably produced the timbre of an electric piano. Thus, for piano simulation,

$$
\begin{aligned}
e = A(t) \sin[2\pi f_c t + I_1 \sin 2\pi(f_{m1} + S)t + \\
I_2 \sin 2\pi(f_{m2} + S)t] \qquad (5)
\end{aligned}
$$

where

$$
\begin{aligned}
&f_c : f_{m1} : f_{m2} = 1 : 1 : 4 \\
&S = f_c \,/\, 200
\end{aligned}
$$

with I_1 and I_2 as given in Eq. 4, and $A(t)$ an amplitude function as shown in Fig. 2.

Another characteristic of piano tones is that the decay time of a given note is dependent on the frequency and peak amplitude of that note. The lower the frequency, or the higher the peak amplitude, the longer the note takes to decay, if the damper is kept raised. "$A(t)$" in Eq. 5 is therefore a function of the undamped decay characteristics of a given note which is interrupted by the fall of the damper. This can be easily simulated with two unit generators controlling the note's amplitude – one as the damper (its envelope time set in the play file) and the other as the fixed exponential decay (its envelope time dependent only upon the note's frequency and amplitude).

The decay function which sounds best is not truly exponential. It was found that an exponential amplitude envelope could give either a good attack to the note or a good "ring," but not both. The function used here (cf. Fig. 2) has a rather sharp attack followed by a long, slow decay. Its decay time (which is independent of the note length) is given by the following formula:

$$T = 10(\sqrt{P}) / \sqrt{f_c}, \qquad (6)$$

where P is the peak amplitude on an arbitrary linear scale of 0 - 2000, f_c is the fundamental frequency, and T is the resulting fixed decay length of the amplitude envelope. This formula was obtained by timing the decay times of several notes at several different amplitudes and fitting the results to a simple curve. It, like the tuning formula given below, is only a working approximation of the characteristics of a traditional piano. At middle C, the decay time ranges from 4 seconds for a very soft note to 25 seconds for a very loud one. At an amplitude of 400 (rather loud) the decay times range from 40 seconds in the extreme low register to 4 seconds at the highest C of the piano.

The output of the envelope generator containing the fixed decay envelope is then multiplied by the output of an envelope generator containing the damping envelope (see Fig. 3). Its value is 1 until the note is to be cut off, when it falls rapidly to 0.

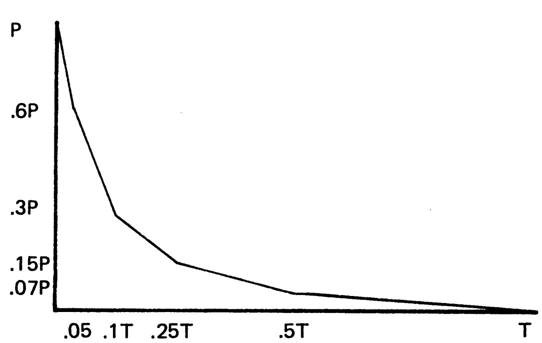


Figure 2. Decay function for piano tone. T is given by Eq. 6, and P is the peak amplitude of the tone.

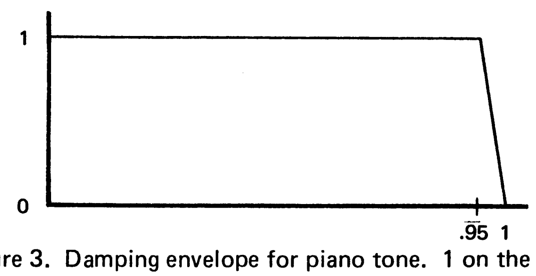


Figure 3. Damping envelope for piano tone. 1 on the horizontal axis represents the length of the tone.

If equal-tempered tuning is employed throughout, the inharmonicity of the resultant partials seems to fool the ear into thinking that the bass notes are sharp and the treble notes flat. A slight retuning was employed in some of the synthesized pianos, along the lines of traditional piano tunings, by using code similar to the following:

$$\text{IF } f_c < \text{G}/2 \text{ THEN } f_c \leftarrow f_c - (10/f_c)$$
$$\text{IF } f_c > \text{G}\times 2 \text{ THEN } f_c \leftarrow f_c + (f_c/200)$$

where G is G above middle C.

A better simulation, especially in the higher register, is produced by using two additional outputs (three carriers in all) to mimic the piano's use of several strings for each note. They are slightly mistuned to give a simple chorus effect. The amplitude of each output is controlled by a scaler which removes two of the "strings" in the lower register – again modeled after traditional pianos.

Pedaling can be rather crudely simulated by overlapping the notes being "pedaled" while adding a significant amount of reverberation.

Finally, it may be of some interest that there is no need to change the modulation indices over the course of a note, or to simulate the various noises produced by a piano's action.

String Simulation

String tones can also be synthesized using a complex modulating wave. In the case of the violin, ratios of modulating frequency to carrier frequency of 1:1, 3:1, and 4:1 were found to be most convincing. Once again frequency-dependent index formulae were used; the following work well in all ranges:

$$I_{1:1} = 7.5/\log_e f_c$$
$$I_{3:1} = (8.5 - \log_e f_c) / [3 + (f_c/1000)]$$
$$I_{4:1} = 1.25 / \sqrt{f_c}$$

or:

$$I_{3:1} = 10(8.5 - \log_e f_c) / f_c$$

These formulae can be multiplied by as much as 2 or 3, to produce a more strident timbre. Even higher indices will give a primitive *sul ponticello* effect. Simpler formulae can also be used, and either the second or third components of the modulating wave can be dispensed with if necessary. Of course the index formulae will have to be changed in the latter case. The following formulae, for example, also work well in the cello range:

$$I_{1:1} = 7.5 / \log_e f_c$$
$$I_{3:1} = 15 / \sqrt{f_c}$$

To simulate the attack noise of a string instrument with its frequency skew, chiff, and so forth, a random number generator is used with a frequency of 2000 Hz. and a bandwidth of 20% of the carrier frequency. This, combined with sharply higher indices in the first .2 seconds of the note, gives an adequate simulation of the string attack. The index function and noise bandwidth function are shown in Figures 4 and 5, respectively. Fig. 6 shows the overall amplitude function of the strings. The rise and decay times are fixed at around .2 seconds. The steady state then takes whatever time is left over.

One can simulate the decay of a detached note by causing the indices to go to 0 with the amplitude during the note's decay. An overlap of .2 seconds with a short decay time gives a good *legato*.

An oscillator is used to create a 5 Hz. to 6 Hz. vibrato; in order to avoid monotony, another random number generator with a frequency of about 10 to 20 Hz. causes a random fluctuation in the vibrato.

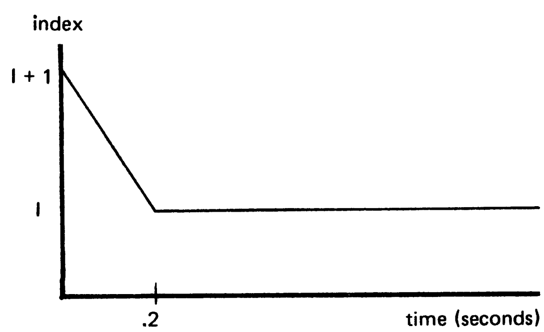


Figure 4. Index function for string tone.

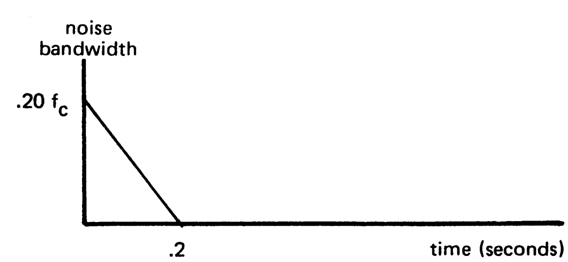


Figure 5. Function for controlling noise during attack portion of string tone.

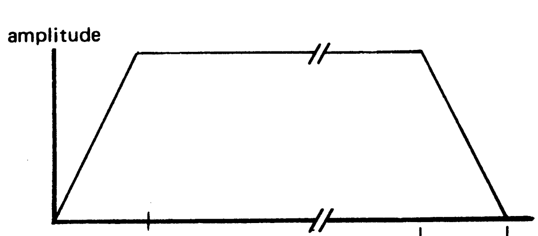


Figure 6. Amplitude function for string tone. The length of the "steady-state" is determined by the length of the tone.

Special Effects

Switches can be added in the instrument definition, allowing one to choose between either *arco* or *pizzicato, vibrato* or *non vibrato, legato* or detached, and so on.

To produce *pizzicato,* use the decay function of the piano, decrease the decay time to $T = 1000/f_c$, decrease the indices and attack noise by 25%, and multiply the indices by the decay function (causing them to go to 0 as the amplitude does).

A convincing choral effect (the sound of an entire orchestral section) can be simulated by adding a second output (two carriers in all) with its own modulators and vibrato generators. One then adds to the various modulating frequencies a series of small (unequal) constants equivalent to that added to the piano to get the stretched harmonics. These constants should be between 1.5 and 4 Hz.

Conclusion

Many other instrument tones can be synthesized using a complex modulating wave, including those of the various brass and percussion instruments. The examples given above should sufficiently illustrate the workings of a process whose potential is only beginning to be explored.

Acknowledgments

I would like to thank James A. Moorer, Julius Smith, Marc Le Brun, John Strawn, Robert Poor, Michael McNabb, and Robert Harvey for their valuable assistance; and in particular I wish to thank John Chowning, who did much of the work on the synthesis of string tones presented in this paper.

References

Chowning, John. "The Synthesis of Complex Audio Spectra by Means of Frequency Modulation." *Journ. Audio Engineering Soc.*, 21:526-534, 1973. Reprinted in *Computer Music Journal,* 1(1):46-54, April, 1977.

"Lexicon of Analyzed Tones (analysis and plotting programs by James A. Moorer and John Grey with editorial notes by John Snell). Part I: A Violin Tone." *Computer Music Journal,* 1(1):39-45, 1977.

Young, R. W. "Inharmonicity of Piano Strings." *Acustica* 4:259-262, 1954.

Le Brun, Marc. "A Derivation of the Spectrum produced by Frequency Modulation with a Complex Modulating Wave." Printed elsewhere in this issue of *Computer Music Journal.*