Barry Truax
Department of Communication and
Centre for the Arts
Simon Fraser University
Burnaby, B.C.
Canada V5A 1S6

Timbral Construction in *Arras* as a Stochastic Process

Introduction

My recent work with the POD system (Truax 1977*a*; 1978*a*) has dealt with organizing timbre at a form-determining level. It has been concerned less with timbre as the specific acoustic properties of individual sounds than with large-scale spectral structures that define the entire composition. The relation between spectrum and timbre is complex and not entirely understood, but it is generally conceded that spectrum, particularly its temporal behavior, is a primary determinant of timbre. As discussed by Robert Erickson (1975), there is a considerable "gray" region in which a complex musical event can be heard as a timbre, as a composite sound, or as a chord. Modern psychoacoustics has furthered our understanding of the conditions under which frequencies are heard as individual pitches, heard as components of a spectrum, or not heard at all individually, but rather fused with others as a single percept. In general, we know that it is not only the physical characteristics of the sound that give rise to different perceptions, but also, and perhaps mainly, the context within which the frequencies are heard (McAdams and Bregman 1979). It is the ambiguity among these different modes of perception that fascinates me and that I have explored in *Arras* (1980), a composition for four-channel, computer-synthesized tape.

This paper is a revised version of a longer discussion, "Timbral Construction as a Stochastic Process," published (in Italian) in *Musica e Elaboratore*, Biennale di Venezia, 1980, and presented at the 1980 International Computer Music Conference in New York and the 1981 International Conference on Music and Technology in Melbourne, Australia.

Computer Music Journal, Vol. 6, No. 3, Fall 1982,
0148-9267/82/030072-06 $04.00/00

FM Timbral Construction

The theme of my compositional exploration has been to use acoustic properties as the basis of the structure of the entire composition, that is, to relate sound and structure inextricably. I have been working with frequency modulation (FM) synthesis (Chowning 1973) and the type of acoustic control it allows. I have been using the simple model of sine-wave FM as merely a building block in what I have called *polyphonic timbral construction* (Truax 1978*b*; 1980). In this construction, many simple FM sources are digitally mixed into complex timbres that lack some of the more predictable clichés associated with simple FM.

The basis of the timbral organization I have been using is the distinction between harmonic and inharmonic spectra (i.e., between spectra in which all constituent frequencies are multiples of the fundamental and those in which not all frequencies are such multiples). Since it is the ratio between the carrier and modulating frequencies (the $c:m$ ratio) that determines which partials are in the spectrum, we can speak of *harmonic* and *inharmonic* $c:m$ ratios. The harmonic ratios are

1:1 1:2 1:3 1:4 1:5 1:6 1:7 1:8 1:9,

and the inharmonic ratios are

2:9 2:7 3:8 2:5 3:7 4:9,

where $m \leq 9$. For any given integer n used as a limit for m, the $c:m$ ratios that produce unique spectra are those corresponding to the series of fractions called the *Farey series* of order n (Truax 1977*b*).

For each of the above ratios, the carrier frequency is also the fundamental, since the ratio can be said to be in *normal form*, satisfying the criterion that m is greater than or equal to twice c, with the ex-

Table 1. Partials

Ratio	*Spectral Frequencies*																
1:2	1.0		3.0			5.0		7.0	9.0		11.0			13.0		15.0	17.0
3:8	1.0	1.67		3.67	4.33		6.33	7.0	9.0	9.67		11.67	12.3		14.3	15.0	17.0

ception of the 1:1 ratio, which produces the complete harmonic spectrum. Each normal-form ratio has a family of ratios associated with it, each of which produces the same set of sidebands and, therefore, the same spectrum. These other family members can be derived from the normal-form ratio ($c:m$) by applying the operation

$$[c \pm (n \cdot m)]:m, \text{ for } n = 1, 2, 3, \ldots .$$

The entire form of *Arras* is derived from the spectral properties resulting from the duality of harmonic and inharmonic spectra.

The Background Structure of *Arras*

I found a basis for relating harmonic and inharmonic ratios in *Arras* when I observed that inharmonic ratios produce spectra with both harmonic and inharmonic partials. Compare, for instance, the spectra of the ratios 1:2 and 3:8, as given in Table 1, in which partials produced by the harmonic ratio 1:2 and the inharmonic ratio 3:8 are expressed with the fundamental as 1.0. The 1:2 ratio produces all odd harmonics, whereas the inharmonic ratio 3:8 includes harmonics 7, 9, 15, and 17, as well as many inharmonic frequencies. If the spectra of the two ratios overlap when based on the same fundamental, they may fuse as a result of the frequencies they have in common (and which act analogously to a pivotal chord or note in a harmonic modulation).

The basic structure of the piece consists of a 75-sec unit in which a harmonic ratio moves to a related inharmonic ratio and back, for example,

$$1:2 \rightarrow 3:8 \rightarrow 1:2.$$

The overall structure of the work consists of a progression of harmonic ratios, beginning with 1:2 (producing all odd harmonics in a fairly dense spectrum) and proceeding to 1:9 (whose harmonics are widely spaced, e.g., 1, 8, 10, 17, 19, . . .). With the exception of the 1:6 ratio, which does not seem to have an inharmonic equivalent in the sense described, each harmonic ratio is paired with an inharmonic one with which it is expected to fuse. The pattern, therefore, is

$$1:2 \rightarrow 3:8 \rightarrow 1:2 \rightarrow 1:3 \rightarrow 4:9 \rightarrow 1:3 \rightarrow$$
$$1:4 \rightarrow 3:8 \rightarrow 1:4 \rightarrow 1:5 \rightarrow 2:5 \rightarrow 1:5 \rightarrow$$
$$1:6 \rightarrow 1:7 \rightarrow 2:7 \rightarrow 1:7 \rightarrow 1:8 \rightarrow 3:8 \rightarrow$$
$$1:8 \rightarrow 1:9 \rightarrow 2:9 \rightarrow 1:9.$$

In the case of 1:7 and 1:9, there are two choices of inharmonic ratio (2:7 and 3:7; 2:9 and 4:9). In each case, the inharmonic ratio with the more similar spacing of sidebands is chosen. (The Farey series influenced the choice of ratios by specifying which produce unique spectra.)

I aimed at an ordering based on optimum continuity and a consistent progression in spectral density. The overall structure of the work corresponded to the global characteristics of a stochastic process. I assumed that a listener can hear a general pattern in the work, progressing from closely to widely spaced partials, with specific harmonic spectra alternating with related inharmonic spectra. The stochastic decisions guarantee a variety of detail at the level of individual frequency components, the level that corresponds to the unpredictability of the individual events in a stochastic process.

In sine-wave FM, the amplitude of each sideband pair for different modulation indices is determined by the set of Bessel functions. Although these functions are varied enough to simulate the temporal variations of natural spectra, many listeners accustomed to FM note a characteristic spectral development when the modulation index is swept from zero to some value. In earlier research, I discovered that when two or more events on a common fundamental with $c:m$ ratios belonging to the same family were mixed, spectral predictability disappeared, even when each event had the same modulation-index envelope. Moreover, much simpler envelope shapes and lower maximum-modulation indices could be used for each component while still producing a rich spectrum. This experience suggested that if, at the theoretical level, the amplitude of any given sideband is the sum of several complex functions, then at the perceptual level, the corresponding spectral component appears to behave in an unpredictable, pseudorandom fashion. Instead of mixing $c:m$ family members on a common fundamental, however, I used a different technique to create variety in the background structure of *Arras*.

Table 2. The harmonic ratios used in *Arras* paired with the inharmonic ones with which they share harmonics

Harmonic Ratio	*Inharmonic Ratio*	f_2	f_3	f_4	f_5
1:2	3:8	7	15	9	17
1:3	4:9	8	17	10	19
1:4	3:8	9	17	7	15
1:5	2:5	9	16	6	14
1:6	—	7	13	7	13
1:7	2:7	6	13	8	15
1:8	3:8	7	15	9	17
1:9	2:9	8	17	10	19

Note: f_2 and f_3 are used as fundamentals for additional components of the background spectrum; f_4 and f_5 are used as carrier frequencies for additional foreground events.

Event Structure

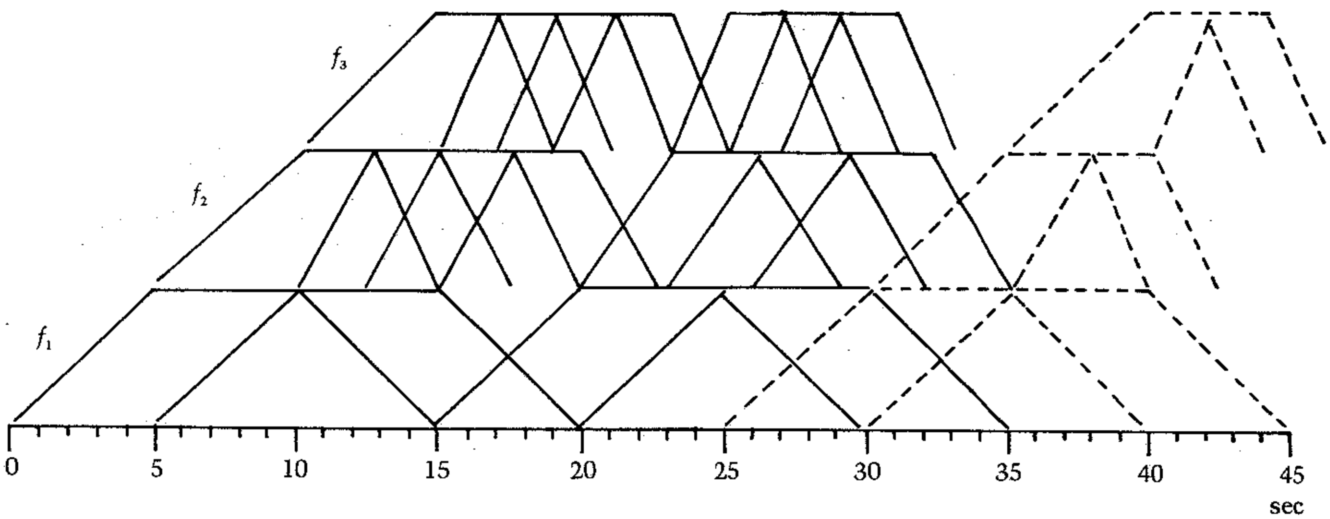
For each ratio, I created a basic structure comprising 18 overlapping envelopes with total duration of 25 sec and with each event doubled at a slightly different frequency for choral effect. Since 22 different ratios are used in the piece, 792 separate events are calculated and mixed together digitally for this layer of it. The 18 overlapping events in each section are based on just three fundamental frequencies: f_1, f_2, and f_3. Frequency f_1 remains the same throughout the entire piece, namely 50 Hz. Frequencies f_2 and f_3 were chosen from the harmonics common to the harmonic and inharmonic ratios (see Table 2). In some cases, the harmonic frequency f_2 or f_3 is the same between adjacent ratios and thus provides an additional basis for fusion between overlapping spectra.

The structure of the 18 events comprising the basic unit for each of the 22 ratios is shown in Fig. 1. It is an asymmetrical structure with four long envelopes for f_1 and seven shorter ones each for f_2 and f_3, the overall effect of which is a more rapid spectral variation in the upper frequencies than in the lower ones. The envelopes shown in Fig. 1 are those for amplitude; the modulation index (or spectral) envelope in each case is a simple increase from zero to its maximum value and back again in equal time units. A rule of thumb with FM is that the modulation index corresponds roughly to the number of sideband pairs present with any significant strength in the spectrum. Therefore, the amount of overlap between the spectra based on each of the three fundamentals is controlled by limiting the index in each layer. For the spectra based on f_1 and f_2, the index is limited to produce a four-octave range of frequencies (up to about the 16th harmonic), and for f_3, it is limited to avoid frequencies over 7250 Hz (i.e., half the sampling rate per channel). Therefore, the spectrum based on f_1 extends across that based on f_2 up to approximately the fundamental f_3 and produces a fairly complex interaction of frequencies between 400 and 800 Hz.

The Foreground Structure of *Arras*

The entire structure described thus far can be thought of as the background "curtain" against which other events are superimposed. It creates a texture of constantly changing frequencies that run like threads throughout the entire piece. (Hence the title, *Arras*, which is a tapestry or heavy wall hanging originally from the French town of the same name.) Pursuing the analogy further, we can imagine patterns superimposed over the background texture but based on the same frequencies, just as when the colored threads of a tapestry form fore-

Fig. 1. Structure of overlapping envelopes in three layers based on three different fundamental frequencies: f_1, f_2, and f_3. This unit structure in turn overlaps with its successor (dashed lines) after 25 sec.



ground patterns while remaining part of the continuous fabric. This interplay of elements creates the perceptual ambiguity that occurs when we hear a particular frequency that belongs to the background texture emerge as part of a foreground event. The inexhaustible variety of this patterning may be compared to the constant unpredictability of micro-level occurrences in a stochastic process.

There are essentially four types of foreground events mixed with the background, although, as stated, there is an intentional ambiguity between what is perceived as foreground or background. Since the aim of the inclusion of foreground events is to produce variety, their design includes many arbitrary features and, in several cases, the events are distributed temporally according to the Poisson probability function. The four additional elements may be described briefly as follows:

1. Long, high-frequency, quasi-sine tones
2. Short, quasi-sine-tone events based on frequencies found in the background
3. Three medium-length envelopes with carriers f_1, f_2, or f_3
4. Closely spaced sine tones expanding and contracting around two frequency centers, f_4 and f_5

Notes on each of these elements follow.

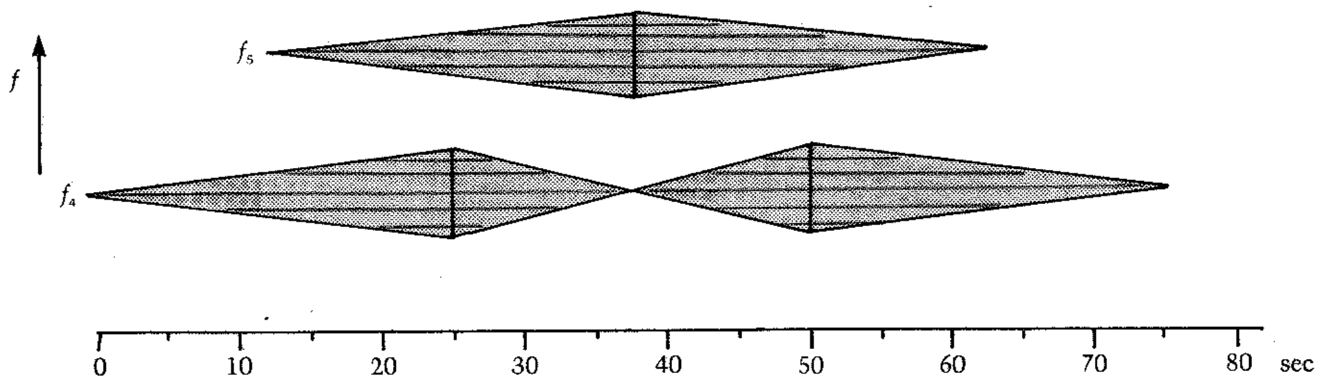
Element 1

The sound texture in *Arras* was designed to maximize the ambiguous perception of any frequency component as separate or embedded. The auditory system, however, is limited in its ability to resolve simultaneous frequencies of similar intensity. This resolving power is sometimes expressed as the critical bandwidth (Plomp 1964; Scharf 1970). Neighboring frequencies used in the background structure of *Arras* lie within a critical bandwidth above about 800 Hz and are therefore not likely to be heard separately. All frequencies above this point were calculated and used as the frequencies of long, quasi-sine-tone events with Poisson-calculated entry delays. These high frequencies, which otherwise would be embedded in the texture, emerge as discrete events.

Element 2

All frequencies in the background texture were calculated and the lowest 16–24 values used as the frequencies of very short quasi-sine-tone events with Poisson-distributed entry delays. The event duration is 0.21 sec, with an average event density of 4.0 per second.

Fig. 2. Frequency-time structure of events shown as horizontal lines lying with a tendency mask frame.



Element 3

Three envelopes with attack, steady-state, and decay times of (2.0, 0.5, 4.0), (0.76, 0.01, 1.24), and (0.01, 0.2, 1.2) sec respectively were designed with various spectral envelopes. The carrier frequency in each case was randomly chosen from f_1, f_2, or f_3, but with a $c:m$ ratio in each case that guarantees a fundamental frequency of f_1 (i.e., 50 Hz). For example, with the 1:2 ratio, the three carrier frequencies are 50, 350, and 750 Hz for the ratios 1:2, 7:2, and 15:2 respectively. In all three cases the fundamental is 50 Hz, and the sidebands are the same set, namely all the odd harmonics. Since these frequencies are the same in the background texture, these foreground events form a complex relief pattern over the spectral base. The percussive envelope, of course, stands out the most, and the event with the longest attack time blends in the most inconspicuously.

Element 4

In contrast to the other events that rely solely on the set of frequencies found in the background layer, the fourth type of foreground event consists of closely spaced sine waves clustered around two frequencies, f_4 and f_5, as shown in Fig. 2. The choice of f_4 and f_5, given in Table 2, comes from the set of harmonics common to the harmonic and inharmonic $c:m$ ratios, as described earlier. The duration of each structure is 75 sec. The part centered around f_5 enters after 12.5 sec and lasts 50 sec such that it is at its densest point when the lower one is least dense. The effect is that of a constantly expanding and contracting cluster of sine tones with a variable amount of internal beating and roughness. A special program was written to create this effect.

Synthesis

All components of the work were synthesized polyphonically with the POD7 program at a sampling rate of 29 KHz for binaural stereo output. The high-frequency events were calculated and synthesized at half-tape speed and then doubled for greater high-frequency resolution. The stereo output utilized both amplitude and time delays between channels to allow the spectral components to be spread out spatially, both on headphones and loudspeakers. The five layers were mixed onto four channels through conventional studio mixing that kept the left binaural signal on either one or both of the left front and back channels, with the right signal similarly placed.

Conclusion

In this paper I have tried to show how the concept of the stochastic process—in particular its property of global coherence balanced against the unpredictable patterning at the level of detail—can be extended to the timbral domain. More specifically, I have shown how certain acoustic properties of harmonic and inharmonic spectra can give rise to a macro-level structure of precision and complexity and yet allow an interaction at the micro level that is potentially rich in interest. I have described the process elsewhere (Truax 1980) as being based on a respect for "materials"—sound and its behavior. It is also a process that attempts to find its expressiveness within the sounding structure instead of letting the structure arise from a set of expressive gestures. A structure that allows sound to communicate through its own behavior suggests, like an archetype, instances of meaning to each individual listener. When variety is provided within the framework of a coherent whole, the mind seems to have the ability to ascribe pattern and meaning to the perceived relationships. Traditional music in every age and culture has demonstrated how sound can achieve expressiveness through its organizational forms. Modern technology has given us unprecedented powers of control over the design of new sound experiences, but *not*, inherently, over the language with which to ensure communication. A compositional language based on a thorough knowledge of the behavior of sound and on the principle of balance between variety and structural coherence seems to provide a means for realizing the potential of contemporary music.

References

Chowning, John. 1973. "The Synthesis of Complex Audio Spectra by Means of Frequency Modulation." *Journal of the Audio Engineering Society* 21(7):526–534. Reprinted in *Computer Music Journal* 1(2):46–54, 1977.

Erickson, Robert. 1975. *Sound Structure in Music*. Berkeley: University of California Press.

McAdams, S., and A. Bregman. 1979. "Hearing Musical Streams." *Computer Music Journal* 3(4):26–43.

Plomp, R. 1964. "The Ear as Frequency Analyzer." *Journal of the Acoustical Society of America* 36(9): 1628–1636.

Scharf, B. 1970. "Critical Bands." In *Foundations of Modern Auditory Theory*, vol. 1, ed. J. V. Tobias. New York: Academic.

Truax, Barry. 1977*a*. "The POD System of Interactive Composition Programs." *Computer Music Journal* 1(3):30–39.

———. 1977*b*. "Organizational Techniques for C:M Ratios in Frequency Modulation." *Computer Music Journal* 1(4):39–45.

———. 1978*a*. "Computer Music Composition: The Polyphonic POD System." *IEEE Computer* (August):40–50.

———. 1978*b*. "Polyphonic Timbral Construction in *Androgyny*." In *Proceedings of the 1978 International Computer Music Conference*, vol. 1, ed. C. Roads. Evanston: Northwestern University Press, pp. 355–377.

———. 1978*c*. *Handbook for Acoustic Ecology*. Vancouver: A.R.C. Publications.

———. 1980. "The Polyphonic POD System and Its Use in Timbral Construction." In *Computer Music*, ed. M. Battier and B. Truax. Ottawa: Canadian Commission for UNESCO, pp. 169–190.

———. 1982. *Androgyne: Electroacoustic and Computer Music by Barry Truax*. Waterloo: Melbourne Records SMLP 4042/43 (includes *Arras* and *Androgyny*).