Chapter 4

Implementing the Model

4.1 Reduction of the Model

In the previous chapter we have established that the generation of a chord spectrum can be uniquely reconstructed from its factorization into a convolution product of irreducible spectra. In a sense, we have proved the uniqueness theorem for chord recognition.

However, this theorem is formulated for idealized conditions. It is assumed that all the chord tones have precisely the same spectrum translated along the $\log_2$-scaled frequency axis.

As a theoretical fact, this theorem characterizes some fundamental trend in chord representation, but in reality the condition mentioned holds only approximately. Even if the tones of a chord are played on the same instrument, their spectra vary with register and pitch.

For example, consider a violin. The resonance characteristic of the violin body is constant, say, dumping f_1 and enhancing a_1. This implies the first partial of tone f_1 to be less salient in the tone spectrum than the first partial of tone a_1. Consequently, the spectrum of tone a_1 is not equal to the spectrum of tone f_1 translated by the major third along the frequency axis $\log_2$-scaled.

The above reason makes difficult direct applications of the formalism from the previous chapter. In its present form it is applicable to the recognition of structure based on identity (replications of the same tone spectrum). However, in reality we deal with the recognition of structure based on similarity (replications of similar tone spectra).

In order to overcome this difficulty, we reduce our consideration to clipped spectra, or Boolean spectra whose partials have amplitudes 0 or 1. This way the attention is restricted to spectral structures which are determined by the ratio of partial frequencies but not by partial amplitudes, or spectral envelopes.

Boolean spectra are much more stable with respect to pitch translations than the original power spectra. Indeed, in the above example concerning

violin sounds, the amplitude of tone partials varies, but their frequency ratio is always constant being determined by the harmonic ratio

$$1 : 2 : 3 : \ \ldots \ : k : \ \ldots \ .$$

In particular, this implies that the Boolean spectrum of violin tone a_1 can be considered as almost identical to the Boolean spectrum of tone f_1 translated by a major third along the $\log_2$-scaled frequency axis.

Therefore, we formulate the problem of chord recognition for Boolean spectra, restricting our attention to the frequency structure and ignoring spectral envelopes. In our case, the recognition of acoustical structure can fail if the structure is understood as based on identity but not on similarity.

Here we propose a kind of a compromise. We do not consider the similarity with deviations of all the characteristics of tone spectra but only that with deviations of partial intensities. In a general model, the deviations of partial frequencies should be also taken into account, but fortunately the harmonic ratio of partial frequencies is quite stable in musical voices. Thus, owing to this property of musical tones, the recognition of similarity in chord spectra is replaced by the recognition of identity in their Boolean spectra.

The recognition of structure based on similarity is not as unambiguous as the recognition of structure based on identity. Indeed, we show that unlike power spectra the unique deconvolution is not valid for Boolean spectra. In particular, this means that the uniqueness theorem for chord recognition is not true. However, the ambiguity in spectral representation arises quite seldom, and we can provide measures in order to control such situations. On the other hand, the recognition of chords based on Boolean spectra has evident advantages: It is much simpler and faster.

Obviously, if a spectrum can be factored into a convolution product, the same can be done for the corresponding Boolean spectrum. Consequently, the Boolean reduction of our model provides for some necessary but not sufficient means for chord recognition. In a sense, recognizing chords by their Boolean representations is somewhat similar to finding the minimum of a function by the points where the first derivative of the function is equal to zero: This condition doesn't guarantee the solution, but the search is reduced to a few points.

In Section 4.2, "Properties of Boolean Spectra", we prove that the chord representation by their Boolean spectra has the same properties as the chord representation by discrete power spectra, except for the unique deconvolution property. The ambiguity in the deconvolution of a chord spectrum is illustrated by a simple example.

In Section 4.3, "Necessary Condition for Generative Patterns", we show how autocorrelation analysis can be used for finding generative patterns in chord spectra. Then we extend simple autocorrelation analysis to multiautocorrelation analysis with which multiple replications of a generative pattern

can be detected. Finally, we formulate a theorem on necessary conditions for generative patterns and their search in terms of recurrent autocorrelation analysis.

In Section 4.4, "Algorithm for Finding Generative Patterns", the procedure of chord recognition based on the theorem from the previous section is traced step by step. This procedure uses a series of embedded cycles of simple autocorrelation analysis of a chord spectrum. It is explained that in spite of embedded cycles the branching of this procedure is rather limited.

In Section 4.5, "Summary of Reduced Model", the main items of the chapter are summarized.

4.2 Properties of Boolean Spectra

As said in the previous section, we shall restrict our consideration to Boolean spectra.

Let

$$S = S(x) = \sum_{n=0}^{N} S_n \delta(x-n) = \sum_{n=0}^{N} S_n \delta_n$$

be a discrete spectrum with non-negative integer coefficients S_n as defined in (3.7).

The *Boolean spectrum* (associated with S) is defined to be the expression

$$s = \bigvee_n s(n)\delta_n. \tag{4.1}$$

where

$$s(n) = \begin{cases} 0 & \text{if } S_n = 0, \\ 1 & \text{if } S_n \neq 0. \end{cases}$$

(The designation s_n instead of $s(n)$ might seem to be more natural, but in the sequel we shall use indexes for other purposes).

One can compare the power spectra in Fig. 3.2–3.3 with their associated Boolean spectra in Fig. 4.1–4.2.

Since we consider the $\log_2$-scaled frequency axis, pitch transpositions correspond to parallel translations of the tone spectrum along the frequency axis. A translation of a Boolean spectrum (4.1) by m frequency bands to the right corresponds to the convolution

$$\delta_m * s = \bigvee_n s(n)\delta_{n+m} = \bigvee_n s(n-m)\delta_n. \tag{4.2}$$

Similarly, a multiple translation of a Boolean spectrum s is determined by the convolution of s and another Boolean spectrum,

$$i = \bigvee_n i(n)\delta_n,$$

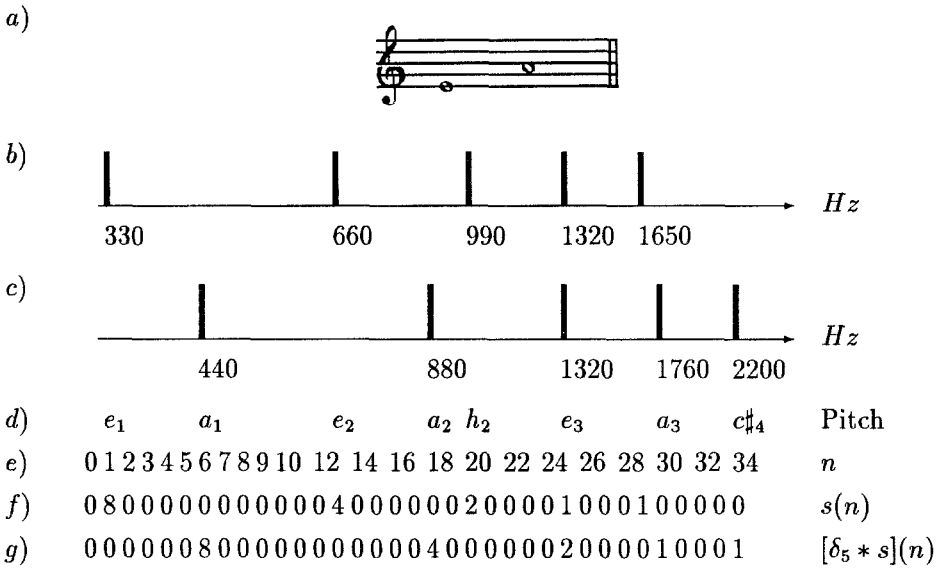

Figure 4.1: Example of Boolean spectra of tones

a) tones e_1 and a_1 in standard musical notation;

b) the Boolean spectrum (with $\log_2$-scaled frequency axis) of tone e_1 for a harmonic voice with 5 successive partials which have equal power;

c) the same for tone a_1;

d) the pitches $e_1, \ldots, c\sharp_4$ which correspond to the mean frequencies of the frequency bands for the given spectral representations;

e) the indexes $n = 1, \ldots, 34$ of frequency bands for the given discrete spectra;

f) the Boolean spectrum of tone e_1 under the frequency resolution within a semitone (1/12 octave) — string $s(n)$;

g) the same for the second tone — string $[\delta_5 * s](n)$.

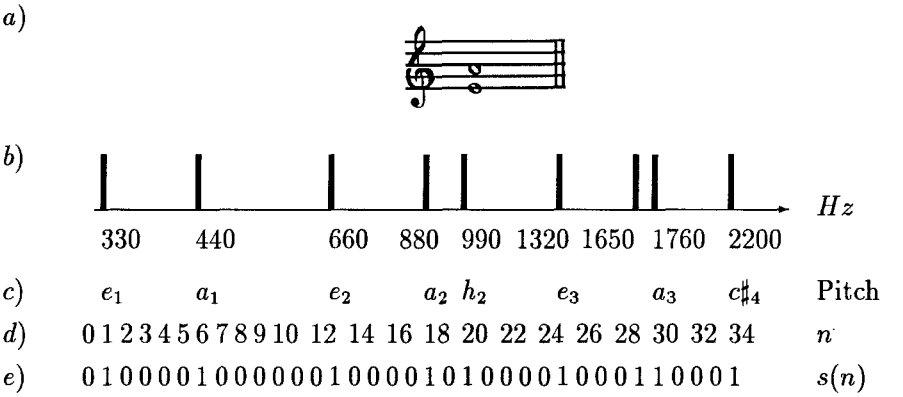

Figure 4.2: Example of Boolean spectrum of chord

a) chord (e_1, a_1) in standard musical notation;

b) the Boolean spectrum (with $\log_2$-scaled frequency axis) of the chord for a harmonic voices with 5 successive partials which have equal power;

c) the pitches $e_1, \ldots, c\sharp_4$ which correspond to the mean frequencies of the frequency bands for the given spectral representations;

d) the indexes $n = 1, \ldots, 34$ of frequency bands for the given discrete spectra;

e) the Boolean spectrum of the chord under the frequency resolution within a semitone (1/12 octave) — string $s(n)$.

which is called the *interval distribution.*

The Boolean interval distribution associated with the interval of m semitones is defined to be

$$i_m = \delta_0 \vee \delta_{[mC/12+0.5]}, \tag{4.3}$$

where the denotations are the same as in (3.8):

C is the constant which characterizes the accuracy of spectral representation, being equal to the number of frequency bands per octave,

$[\cdot + 0.5]$ is the rounding function.

The interval distribution of a three-tone chord with intervals of m and k semitones from its lowest tone is defined to be

$$i_{m,k} = \delta_0 \vee \delta_{[mC/12+0.5]} \vee \delta_{[kC/12+0.5]}.$$

Example 3 (Boolean Interval Distribution of Major Third) Let the frequency resolution be to within one semitone ($C = 12$). Then the major third which is equal to four semitones corresponds to four frequency bands. Consequently, its interval distribution is equal to

$$i_4 = \delta_0 \vee \delta_4. \tag{4.4}$$

Example 4 (Boolean Interval Distribution of Major Triad) Let the frequency resolution be to within quarter of a tone ($C = 24$). The major triad in root position is determined by intervals of four and seven semitones from the lowest tone, which correspond to 8 and 14 frequency bands, respectively. Consequently, the interval distribution of this major triad is equal to

$$i_{4,7} = \delta_0 \vee \delta_8 \vee \delta_{14}. \tag{4.5}$$

As follows from (3.9), the Boolean harmonic spectrum has the form

$$s = \bigvee_{k=1}^{K} \delta_{p+[C\log_2 k+0.5]} = \delta_p * \bigvee_{k=1}^{K} \delta_{[C\log_2 k+0.5]}.$$

Example 5 (Boolean Harmonic Spectrum) Let the frequency resolution be to within one semitone ($C = 12$). Suppose that index $n = 0$ corresponds to the frequency band centered at $262Hz$ which corresponds to the note c_1 (*do* of the first octave). Consider a musical tone a_1 (*la* of the first octave) with five successive harmonics. Then the frequency band corresponding to a_1 has the index $n = 9$. Hence, the Boolean spectrum of our tone a_1 is equal to

$$s = \delta_9 \vee \delta_{21} \vee \delta_{28} \vee \delta_{33} \vee \delta_{37} \tag{4.6}$$

$$= \delta_9 * [\delta_0 \vee \delta_{12} \vee \delta_{19} \vee \delta_{24} \vee \delta_{28}]. \tag{4.7}$$

A Boolean spectrum s is said to be *simple* if its first coefficient $s(0) \neq 0$.

For example, the interval distributions (4.4–4.5) are simple. On the contrary, the Boolean spectrum (4.6) is not simple. In (4.7) the same spectrum as in (4.6) is represented as a simple Boolean spectrum translated by 9 bands to the right.

Since all harmonic spectra are similar in their structure, a chord (even with pairwise different harmonic voices) can be approximately regarded as generated by translations of a harmonic voice pattern along the frequency axis. The inequality of voice spectra may result in small errors in such a representation. Hence, we formulate the following basic assumption.

Conjecture 10 (Representation of Chord Spectra) *The Boolean spectrum of a chord can be approximately represented as generated by multiple translations of a tone spectral pattern as follows*

$$s = \delta_p * t * i + \epsilon - \lambda, \tag{4.8}$$

where

p *is the index of the first frequency band of the lowest tone of the chord (conventional* pitch *of the chord's lowest tone),*

t *is a* tone pattern *which is a simple Boolean spectrum,*

i *is an* interval distribution *which is a simple Boolean spectrum,*

ϵ *is the set (spectrum) of* missed *partials which should be added to* $\delta_p * t * i$ *in order to obtain* s*,*

λ *is the spectrum (spectrum) of* false *partials which should be removed from* $\delta_p * t * i$ *in order to obtain* s*.*

We suppose that the error-correcting spectra ϵ and λ in (4.8) contain rather few partials so that the term

$$s' = \delta_p * t * i$$

approximates the Boolean spectrum s. Therefore, the deconvolution properties of the Boolean spectrum s' are of prime importance.

In the previous chapter it is shown that a chord spectrum with non-negative integral coefficients is uniquely factored into the convolution product of irreducible spectra with non-negative integral coefficients, while harmonic spectra and interval distributions of two-tone intervals and major or minor triads being irreducible. In case of Boolean spectra the situation is different. As seen from the following example, Boolean spectra are not uniquely factored into the convolution product of irreducible Boolean spectra, making the deconvolution problem ambiguous.

Example 6 (No Unique Deconvolution of Boolean Spectra) Let the frequency resolution be to within one semitone ($C = 12$). Suppose that index $n = 0$ corresponds to the frequency band centered at $262Hz$ which corresponds to note c_1 (*do* of the first octave). Consider musical tones c_1 and c_2 (*do* of the first octave and *do* of the second octave which are shown in Fig. 4.3a) and two-tone interval $(c_1; c_2)$.

At first suppose that the voices have four successive harmonics. Their spectra are shown in Fig. 4.3b–c. To be precise, we consider the following Boolean tone spectrum

$$t = \delta_0 \vee \delta_{12} \vee \delta_{19} \vee \delta_{24}. \tag{4.9}$$

and the Boolean interval distribution

$$i = \delta_0 \vee \delta_{12}. \tag{4.10}$$

The convolution of these two spectra, corresponding to the chord spectrum of the interval $(c_1; c_2)$, equals to

$$s = t * i = \delta_0 \vee \delta_{12} \vee \delta_{19} \vee \delta_{24} \vee \delta_{31} \vee \delta_{36}.$$

This spectrum is shown in the third line of Fig. 4.3b, and in its binary form—in the third line of Fig. 4.3c.

Now omit the second partial of tone t and consider the resulting Boolean spectrum

$$t' = \delta_0 \vee \delta_{19} \vee \delta_{24}. \tag{4.11}$$

This spectrum and its translation by the octave are shown in Fig. 4.3d–e. One can see that

$$s = i * t = i * t'.$$

It is illustrated by the fact that the third line in Fig. 4.3b is equal to the third line in Fig. 4.3d and the third line in Fig. 4.3c is equal to the third line in Fig. 4.3e.

Note that Lemma 2 is formulated in terms of the properties of spectral supports. This implies its applicability to Boolean spectra as well. Consequently, by virtue of Lemma 2 the Boolean spectra t, t', and i are irreducible, implying the ambiguity of the deconvolution of Boolean spectrum s into irreducible Boolean spectra.

If we considered power spectra, the ambiguity in the spectral deconvolution wouldn't occur. Indeed, similarly to (4.9–4.11) put

$$\begin{aligned} T &= \delta_0 + \delta_{12} + \delta_{19} + \delta_{24}, \\ I &= \delta_0 + \delta_{12}, \\ T' &= \delta_0 + \delta_{19} + \delta_{24} \end{aligned}$$

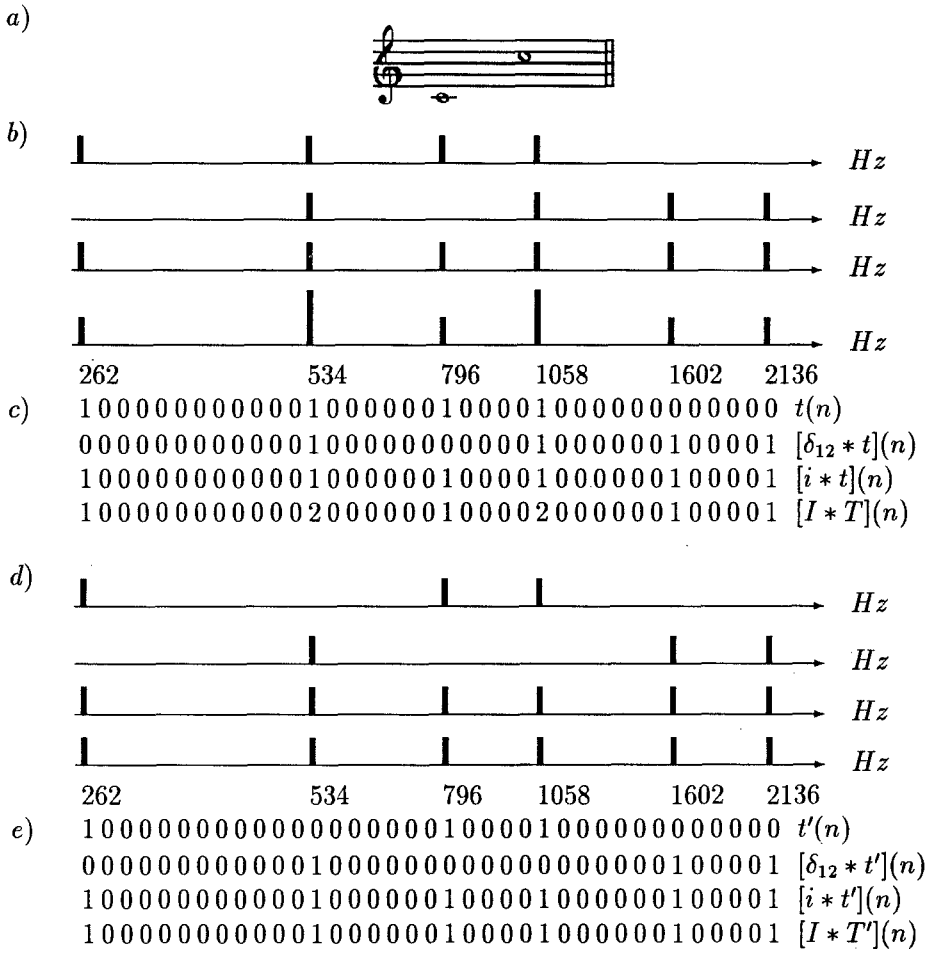

Figure 4.3: No unique deconvolution of Boolean spectra

a) tones c_1 and c_2 in standard musical notation;

b) the tone spectra for voices with four successive harmonics, their Boolean union, and their sum;

c) the same in the Boolean string form;

d) the tone spectra for voices with first, third, and fourth harmonics, their Boolean union, and their sum;

e) the same in the Boolean string form.

and obtain

$$\begin{aligned} S = I * T &= \delta_0 + 2\delta_{12} + \delta_{19} + 2\delta_{24} + \delta_{31} + \delta_{36}, \\ I * T' &= \delta_0 + \delta_{12} + \delta_{19} + \delta_{24} + \delta_{31} + \delta_{36}, \end{aligned}$$

whence S can be no longer factored into T' and I. This is illustrated by the fact that the fourth line in Fig. 4.3b is different from the fourth line in Fig. 4.3d and the fourth line in Fig. 4.3c is different from the fourth line in Fig. 4.3e.

As already mentioned in Section 4.1, in spite of the ambiguity in spectral deconvolution, we shall use Boolean spectra for chord decomposition. The main argument is the fact that the chord tones differ in their Boolean spectra much less than in their power spectra. Therefore, the error in the chord spectral representation (4.8) is much less in case of Boolean spectra than it could be in case of power spectra.

Thus the disadvantage of the non-unique deconvolution of Boolean spectra, implying the ambiguity in their representation, is compensated by the gain from the stability of Boolean tone patterns which enables representing chord spectra in terms of generative elements.

Another reason in favor of Boolean spectra is the simplicity of computer implementation of the model.

4.3 Necessary Condition for Generative Patterns

At first let us consider the problem of chord recognition, assuming the idealized scheme of the chord spectrum generation.

Suppose that a Boolean chord spectrum s in (4.8) is generated without variations of the tone pattern t while translating it along the frequency axis. Obviously, in this case no error-correcting terms ϵ and λ are needed. In other words, let

$$s = \bigvee_n s(n)\delta_n = \delta_p * t * i, \tag{4.12}$$

where

$$\begin{aligned} t &= \bigvee_n t(n)\delta_n, \\ i &= \bigvee_n i(n)\delta_n \end{aligned}$$

are simple Boolean spectra. The question is: How can we find the original deconvolution (4.12)?

In order to find repetitive subspectra in s, one can analyze peaks of the autocorrelation function of the Boolean chord spectrum

$$R(x) = \sum_{n=x}^{N} s(n)s(n-x). \tag{4.13}$$

Obviously, if there is an interval of j frequency bands in the interval distribution i, i.e. if $i(j) = 1$, implying the translation of tone pattern t by j positions to the right, then the autocorrelation function $R(x)$ has a peak at the point $x = j$. The *correlated group of partials* (spectrum)

$$t_j = \bigvee_{n:t(n)t(n+j)\neq 0} \delta_n$$

besides the partials from the generative spectrum can have some accidental partials, so that we have

$$t \subset t_j.$$

Similarly, if there are intervals $j_1, \ldots, j_k$ in the interval distribution i then the *multiautocorrelation function*

$$R(x_1, \ldots, x_k) = \sum_{n=\max\{x_1,\ldots,x_k\}}^{N} s(n)s(n-x_1)\cdots s(n-x_k) \tag{4.14}$$

has a peak at the point $(x_1, \ldots, x_k) = (j_1, \ldots, j_k)$. The *multicorrelated group of partials* (spectrum)

$$t_{j_1,\ldots,j_k} = \bigvee_{n:t(n)t(n+j_1)\ldots t(n+j_k)\neq 0} \delta_n$$

has all the partials from the generative tone spectrum t and may have some accidentals, so that

$$t \subset t_{j_1,\ldots,j_k}.$$

Since we consider Boolean spectra, the value of the autocorrelation function of a spectrum is equal to the number of coinciding ones in the spectrum and its copy translated by a corresponding number of frequency bands. Similarly, the value of the multicorrelation function of a spectrum is equal to the number of coinciding ones in the spectrum and its copies translated by corresponding numbers of frequency bands. Obviously, if we reduce the dimension of multiautocorrelation, meaning that one of translations of the spectrum is not taken into account, the number of coinciding ones increases or remains the same, implying increase or equality in the value of the multiautocorrelation function. Therefore, for arbitrary values of arguments $x_1, \ldots, x_n$ it holds

$$R(x_1, \ldots, x_k) \leq R(x_1, \ldots, x_{k-1}) \leq \cdots \leq R(x_1).$$

Hence, we obtain the following theorem.

Theorem 4 (Necessary Condition of Generative Tone Pattern) *Let a Boolean spectrum s of a chord be generated by a tone pattern t translated according to an interval distribution i, i.e. let*

$$s = t * i, \tag{4.15}$$

where

$$i_{j_1,\ldots,j_k} = \delta_0 \vee \delta_{j_1} \vee \ldots \vee \delta_{j_k} \qquad (j_1 < \ldots < j_k). \tag{4.16}$$

Then

- *the multiautocorrelation function $R(x_1,\ldots,x_k)$ has a peak at the point $(j_1,\ldots,j_k)$,*
- *the multiautocorrelation function $R(x_1,\ldots,x_{k-1})$ has a peak at the point $(j_1,\ldots,j_{k-1})$,*
- $\ldots$,
- *the autocorrelation function $R(x_1)$ has a peak at the point j_1, and*
- *the associated correlated subspectra contain the generative subspectrum t and are embedded as follows*

$$t \subset t_{j_1,\ldots,j_k} \subset t_{j_1,\ldots,j_{k-1}} \subset \ldots \subset t_{j_1,j_2} \subset t_{j_1}. \tag{4.17}$$

Thus in order to find a repeating subspectrum one should scan over the peaks of the multiautocorrelation function $R(x_1,\ldots,x_k)$. These peaks can be found recurrently by peaks of the multiautocorrelation function of a smaller dimension. Moreover, by virtue of (4.16) the search for peaks of the multiautocorrelation function is reduced to the simplex of argument values which has the form

$$\{(x_1,\ldots,x_k) : x_1 < \ldots < x_k\}.$$

In addition to the directional search for deconvolutions (4.12), the above theorem enables finding multicorrelated patterns, while not knowing the multiplicity of the correlation. In other words, one can recognize a chord not knowing the number of its notes.

First, one should examine two-tone representations, finding most correlated groups of partials and for each group trying to reconstruct the chord spectrum from the translation of this group by the corresponding interval of correlation. If two-tone representations are not successful, most correlated groups of partials should be tested on their third appearance in the chord spectrum. For this purpose, instead of triple-correlation analysis, on has to look for the correlation between correlated group of partials previously deternmined. Then each triple-correlated group of partials should be tested as a generative pattern

of the chord spectrum by trying to reconstruct the chord spectrum from the translations of this group by the corresponding intervals of correlation. If tree-tone representations are also not successful, the procedure should be continued for four-note representation, etc., by virtue of (4.17) each time being based on the most correlated groups of partials selected at the previous step.

Now suppose that the chord tones have slightly different Boolean spectra, i.e. the chord Boolean spectrum has the form (4.8) rather than the form (4.15). In such a case, instead of obtaining the precise decomposition (4.15) one obtains several alternative representations of the form (4.8). The desired representation can be chosen with respect to the criterion of least complexity.

Since the spectra considered can be stored as a sequence of impulses, the complexity of a spectral representation can be identified with the number of impulses to be stored.

By *complexity of a Boolean spectrum* s we understand the number of points in its support Δ_s. The complexity of s is denoted by $|\Delta_s|$.

By *complexity of the deconvolution* $s = t * i$ we understand the total complexity of the convolution factors which is equal to $|\Delta_t| + |\Delta_i|$.

The detailed algorithm for finding generative Boolean spectra is given in the next section.

4.4 Algorithm for Finding Generative Patterns

The theorem from the previous section implies an algorithm for finding generative Boolean spectra of chords. Although the theorem is formulated for spectra of the form (4.15), the algorithm is applicable to spectra of the form (4.8), i.e. to chords whose tones have slightly different Boolean spectra, which corresponds to real situations.

Let s be a Boolean spectrum of a chord. In order to find a generative tone pattern t and an interval distribution i, one should do the following.

1. Without loss of generality we can suppose that s is simple, otherwise s can be brought to the origin by an appropriate translation $\delta_{-p} * (\cdot)$, where p is the index of the frequency band with the spectrum's lowest partial tone.

2. Perform ordinary autocorrelation analysis of Boolean spectrum s by analyzing peaks of the autocorrelation function $R(x_1)$. Let j_1 provide the function $R(x_1)$ with the maximum, j_2 provide its second maximal value, j_3 the third, etc. Consider, say, the first 10 peaks $j_1, \ldots, j_{10}$ of $R(x_1)$.

   (a) Consider the interval j_1, and the associated correlated group of

partials t_1. Define the interval distribution

$$i_1 = i_{j_1} = \delta_0 \vee \delta_{j_1}.$$

Try to represent the Boolean spectrum s as generated by t_1 and i_1, putting

$$s'_1 = t_1 * i_1,$$

and comparing s'_1 with s. Hence, we obtain the representation

$$s = s'_1 + \epsilon_1 - \lambda_1,$$

where the subspectrum ϵ_1 consists of the partials of s which are missed in s'_1, and λ_1 consists of the partials of s'_1 which are not present in s.

(b) Consider the interval j_2, and the associated correlated group of partials t_2. Perform all the operations of the preceding item (substituting index 2 for index 1) in order to verify whether s is generated by the correlated tone pattern translated by the interval of j_2 units.

(c) Repeat the above procedure several times (e.g. 10). Even if we fail in representing s as a two-tone interval, by virtue of Theorem 4 if the number of processed peaks of $R(x)$ is sufficiently large, we can be sure that the true tone pattern belongs to one of the spectra t_k correlated.

3. Perform triple-autocorrelation analysis of Boolean spectrum s by analyzing peaks of autocorrelation function $R(x_1, x_2)$. By virtue of Theorem 4 one should scan not all possible pairs (x_1, x_2), but only such that

- $x_1 = j_1, \ldots, j_{10}$, where $j_1, \ldots, j_{10}$ are selected at the previous step, and
- $x_2 > x_1$.

This means that we are looking for triple correlated groups of partials among correlated groups of partials.

(a) Consider the interval j_1, and the associated correlated group of partials t_1. Consider the peaks of

$$R(j_1, x_2) = \sum_n t_1(n) s(n - x_2)$$

for $x_2 > j_1$. Let $j_{1,1}$ provide the function $R(j_1, x_2)$ with the maximum, $j_{1,2}$ provide its second maximal value, etc. Consider, say the first 10 peaks $j_{1,1}, \ldots, j_{1,10}$ of $R(j_1, x_2)$.

i. Consider the intervals j_1 and $j_{1,1}$. Let the associated triple-correlated group of partials be

$$t_{1,1} \subset t_1.$$

Define the interval distribution

$$i_{1,1} = i_{j_1, j_{1,1}} = \delta_0 \vee \delta_{j_1} \vee \delta_{j_{1,1}}.$$

Try to represent s as generated by $t_{1,1}$ and $i_{1,1}$, putting

$$s'_{1,1} = t_{1,1} * i_{1,1},$$

and comparing $s'_{1,1}$ with s. Hence, we obtain the representation

$$s = s'_{1,1} + \epsilon_{1,1} - \lambda_{1,1},$$

where the subspectrum $\epsilon_{1,1}$ consists of the partials of s which are missed in $s'_{1,1}$, and $\lambda_{1,1}$ consists of the partials of $s'_{1,1}$ which are not present in s.

ii. Consider the intervals j_1 and $j_{1,2}$. Similarly, define the triple-correlated group of partials

$$t_{1,2} \subset t_1.$$

Define the interval distribution

$$i_{1,2} = i_{j_1, j_{1,2}} = \delta_0 \vee \delta_{j_1} \vee \delta_{j_{1,2}}.$$

Try to represent s as generated by $t_{1,2}$ and $i_{1,2}$, putting

$$s'_{1,2} = t_{1,2} * i_{1,2},$$

and comparing $s'_{1,2}$ with s.

iii. Repeat the above procedure several times (e.g. 10).

(b) Consider the interval j_2, and the associated correlated group of partials t_2. Consider the peaks of

$$R(j_2, x_2) = \sum_n t_2(n) s(n - x_2)$$

for $x_2 > j_2$. Let $j_{2,1}$ provide the function $R(j_2, x_2)$ with the maximum, $j_{2,2}$ provide its second maximal value, etc. Consider, say, the first 10 peaks $j_{2,1}, \ldots, j_{2,10}$ of $R(j_2, x_2)$.

i. Consider the intervals j_2 and $j_{2,1}$. Let the associated triple-correlated group of partials be

$$t_{2,1} \subset t_2.$$

Define the interval distribution

$$i_{2,1} = i_{j_2,j_{2,1}} = \delta_0 \vee \delta_{j_2} \vee \delta_{j_{2,1}}.$$

Try to represent s as generated by $t_{2,1}$ and $i_{2,1}$, putting

$$s'_{2,1} = t_{2,1} * i_{2,1},$$

and comparing $s'_{2,1}$ with s. Hence, we obtain the representation

$$s = s'_{2,1} + \epsilon_{2,1} - \lambda_{2,1},$$

where the subspectrum $\epsilon_{1,1}$ consists of the partials of s which are missed in $s'_{2,1}$, and $\lambda_{1,1}$ consists of the partials of $s'_{2,1}$ which are not present in s.

ii. Consider the intervals j_2 and $j_{2,2}$. Similarly, define the triple-correlated group of partials

$$t_{2,2} \subset t_2.$$

Define the interval distribution

$$i_{2,2} = i_{j_2,j_{2,2}} = \delta_0 \vee \delta_{j_2} \vee \delta_{j_{2,2}}.$$

Try to represent s as generated by $t_{2,2}$ and $i_{2,2}$, putting

$$s'_{2,2} = t_{2,2} * i_{2,2},$$

and comparing $s'_{2,2}$ with s.

iii. Repeat the above procedure several times (e.g. 10).

(c) For every $j_3, \ldots, j_{10}$ perform the same cycle.

4. Perform quadruple-autocorrelation analysis of the Boolean spectrum s by finding and analyzing peaks of the autocorrelation function $R(x_1, x_2, x_3)$. By Theorem 4 one should scan not all possible triplets (x_1, x_2, x_3), but only such that

- $(x_1, x_2) = (j_k, j_{k,m})$, where j_k are selected at the first step, and $j_{k,m}$ are selected at the second step of the algorithm, and
- $x_3 > x_2$.

This means that we are looking for quadruple correlated groups of partials among triple-correlated groups of partials. For this purpose one should perform three embedded cycles, on $k = 1, \ldots, 10$ (the first interval of the chord), on $m = 1, \ldots, 10$ (the second interval of the chord depending on the first interval), and on $l = 1, \ldots, 10$, which corresponds to revealing third interval depending on the two intervals already known.

5. The search can be continued, each time having one more embedded cycle.

Note that it is not necessary to perform a fixed number of embedded cycles. Obviously, most branches of our branching process are exhausted at the very first steps of the algorithm. For instance, suppose that there is a simply (double) correlated groups of partials but the chord spectrum cannot be reconstructed from this group of partials translated by the interval of correlation, i.e.

$$s' = t' * i' \not\approx s.$$

If at the next step of the algorithm it turns out that this group of partials is not triple correlated (or the triple correlation is insufficient), then, obviously, this group cannot be quadriple correlated, etc. In this case this group of partials should be rejected from further analysis, meaning that the related branch of the algorithm is exhausted.

If the original Boolean spectrum of our chord has the form (4.15) we will obtain one or several precise deconvolutions of the form

$$s = t' * i'$$

(recall that by assumption our spectrum is simple, i.e. $p = 0$), where the error-correcting spectra ϵ and λ with appropriate indexes are zero spectra. However, owing to accidental correlations of partials, the tone spectrum t' may have more partials than t, i.e. $t \subset t'$. Therefore, one should reduce the spectrum t', trying various $t'' \subset t'$ and $i'' \subset i'$, remaining within the above representation, i.e. so that $s = t'' * i''$.

If the original Boolean spectrum of the chord has the form (4.8), one has to compare the representations

$$s = t' * i' + \epsilon' - \lambda'$$

obtained by the algorithm, trying to reduce them as described above and minimizing their complexity

$$|\Delta_{t'}| + |\Delta_{i'}| + |\Delta_{\epsilon'}| + |\Delta_{\lambda'}|.$$

This way the optimal representation of the chord spectrum can be found out. Note that it may differ from the chord generation, but such cases occur quite seldom.

4.5 Summary of Reduced Model

Let us summarize the main items of the present chapter.

1. In order to make tone spectra (almost) invariant with respect to pitch translations, it is proposed to consider Boolean spectra instead of power spectra. This enables using the model of correlative perception, where chords are understood as contours generated by translations of the tone spectrum along the $\log_2$-scaled frequency axis.

2. The properties of Boolean spectra are similar to that of power spectra, with the only exception for the unique spectral deconvolution into irreducible spectra. Reconstructing the chord generation is still possible by decomposing the Boolean spectrum of a chord into irreducible spectra. However, such a decomposition is no longer a necessary and sufficient for reconstructing the spectrum generation, but only a necessary condition.

3. We formulate a necessary condition for the decomposition of the Boolean spectrum of a chord with a reference to the multiautocorrelation function of the spectrum. The theorem proved implies a recurrent procedure of a directional search for the decomposition of a Boolean chord spectrum which corresponds to the spectrum generation.

4. We propose an algorithm for finding generative tone spectra in a Boolean spectrum of a chord. This algorithm is applicable to recognizing chords even if the Boolean tone spectra of the chord are not identical to each other. Therefore, this algorithm of recognizing acoustical structure by similarity (not only by identity) is adapted for practical purposes.