Perception & Psychophysics
1989, 46 (6), 513-527

Pitch motion with random chord sequences

J. ALLIK
Tartu State University, Tartu, Estonian S.S.R.

E. N. DZHAFAROV
University of Illinois, Champaign, Illinois

A. J. M. HOUTSMA
Institute for Perception Research, Eindhoven, The Netherlands

J. ROSS
Institute of Language and Literature, Tallinn, Estonian S.S.R.

and

N. J. VERSFELD
Institute for Perception Research, Eindhoven, The Netherlands

Perception of global pitch motion was studied through psychoacoustic experiments with random chord sequences. Chords contained either six or eight (fixed) tone elements, being sinusoidal, sawtooth-like, or Shepard tones, which were either on or off according to a probability controlled by the experimenter. Sequences of 2, 4, 5, or 8 chords were used. Identification by subjects of the perceived direction of overall pitch motion (up or down) was found to be well accounted for by a model in which the ultimate pitch motion percept is given by a sum of contributions from selected element transitions—that is, transitions between adjoining tone elements in successive time frames only. In its simplest form, this dipole contribution model has only one free parameter, the perceptual noise for an element transition, which was estimated for various acoustic tone representations and chord arrangements. Results of two experiments, which were carried out independently in two different laboratories, are reported.

The problem of pitch perception of sinusoidal tones has received considerable attention in the literature on psychoacoustics. Relations have been established between subjective pitch and objective acoustic variables such as the tone's frequency (Stevens & Volkmann, 1940), its intensity (Stevens, 1935; Verschuure & van Meeteren, 1975), its duration (Doughty & Garner, 1948), its temporal envelope (Hartmann, 1978; Rossing & Houtsma, 1986), and the presence and strength of other interfering sounds (Larkin, 1978; Terhardt & Fastl, 1971).

Pitch perception of *complex* tones has received much attention, too, during the past few decades. The pitch of a harmonic complex tone is determined not merely by its fundamental frequency, but also, to a very large extent, by its harmonics. The tones of church bells and orchestral chimes, as well as the filtered tones from common musical instruments, very often produce a clear, unambiguous pitch sensation without the presence of any acoustic energy at the fundamental frequency. It has been established over the years that this *missing fundamental* percept cannot be accounted for by difference-tone distortion or by periodicity detection of interference patterns in the ear; it is, rather, the result of the way in which our brain processes the neural transformations of sounds from our two ears. Detailed reviews of this research have been given by de Boer (1976) and Scharf and Houtsma (1986). When two complex tones, each comprising a few harmonics, sound simultaneously, the pitches corresponding to each tone can usually be heard as well (Beerends & Houtsma, 1986, 1989). The exact perceptual limit to the number of simultaneous tones or pitches that can be correctly perceived is not known, but it is probably very dependent on training (Doehring, 1971).

The perception of pitch sequences is another problem that has received a fair amount of attention in the past. A temporal sequence of two tones, called a *melodic interval*, is typically perceived in a categorical manner (Burns & Ward, 1978). Eiting (1984) showed that recognition of three-note sequences occurs on the basis of contour (e.g., up-down-up), as well as on the size of successively perceived melodic intervals. Deutsch (1980) studied listeners' retention capacities for longer tonal sequences and found that they depended greatly on the degree of hierarchical structure within each sequence. For very fast tonal sequences (more than 15 notes per second), a stream of notes may become perceptually separated into two or

---

The authors, who have been listed in alphabetical order, are indebted to Th. de Jong, of the Institute of Perception Research, and M. Mihkla, of the Institute of Cybernetics in Tallinn, for invaluable technical assistance. Correspondence may be addressed to A. J. M. Houtsma, Institute for Perception Research, P.O. Box 513, 5600 MB Eindhoven, The Netherlands.

Copyright 1989 Psychonomic Society, Inc.

more parallel streams, each forming a melody (Bregman & Campbell, 1971; Dowling, 1973; van Noorden, 1975).

Less attention has been paid to the study of sequences of simultaneous tones. One of the few situations that has received attention is the dichotic conflict one in which two different melodies are simultaneously presented to a subject, one to each ear (Butler, 1979; Deutsch, 1975; Kimura, 1964). The limits on our auditory system's ability to perceive a melody against a background of other potentially interfering tones without dichotic separation remains largely unknown. This is somewhat unfortunate, because it is just this situation that is encountered most often when we listen to music.

In research on vision, the problem of motion perception for discretely changing elements against a background of other randomly changing elements seems to have been given more attention than the above-mentioned analog in hearing. When a dot is presented repeatedly but each time with a slight displacement, it appears to move (Korte, 1915; Wertheimer, 1912) in the same way as upward or downward melodic movement is heard when a sequence of tones is presented to the ear. In both cases, the issue of correspondence—that is, the question of which elements in one time frame correspond to which elements in the next frames—is rather trivial, because there is only one element in each time frame. When several dots are presented in each time frame, however, the issue of correspondence can become quite complicated. Experiments with moving random-dot patterns, or *cinematograms* (Julesz, 1971; Nakayama & Silverman, 1984; van Doorn & Koenderink, 1982, 1984), have convincingly shown that, when a directional moving-dot signal is embedded in a background of randomly moving dots, the motion can often be detected. This implies that the brain must have a rather sophisticated way of solving the correspondence problem. One can show that, on statistical grounds, the clearer and more similar two elements in successive frames are, the more likely it is that they will be perceived as corresponding elements (Ullman, 1979). Much visual behavior can be accounted for with a *local correspondence model*, in which perceptual correspondence is limited to single spatial jumps in successive time frames only (Bell & Lappin, 1973). The acoustic experiments described in this study also involve such a model. There are experimental situations, however, where local correspondence models are clearly inadequate (Julesz & Bosche, 1966; Pantle & Picciano, 1976).

A special case of a random-dot pattern is the *circular random cinematogram* (CRC) first applied by Allik and Dzhafarov (1984). Their CRC consisted of 12 light elements grouped circularly at the 5-min marks on the face of a clock; each light element could be either on or off. A sequence of random circular displays often evokes an apparent clockwise or counterclockwise rotation percept, reflecting the *phi phenomenon* (Korte, 1915; Wertheimer, 1912) or *reversed phi phenomenon* (Anstis, 1970). Perceptual identification data could well be accounted for by a strictly local and short-range model in which only jumps between successive display elements in successive time frames played a role.

The present study is an acoustic analogue to Allik and Dzhafarov's (1984) random cinematogram experiment. It deals with the perception of apparent global pitch movement in a *random-chord sequence* (RCS), the acoustic equivalent of a CRC. An RCS is a sequence of pure- or complex-tone clusters, in which, under certain conditions, a global upward or downward pitch movement can be heard. The perceptual limit to the correct identification of pitch movement is the topic of this study.

RANDOM CHORD SEQUENCES

A single complex tone is characterized by its fundamental frequency, its spectral envelope and phase function, and its duration. One such tone is called an *element*. An element can be in one of two possible states: sounding (*on*) or silent (*off*). Several such elements grouped together, each with a different fundamental frequency, form a *chord*. Since each tone has only two states, M different tones can form 2^M differently sounding chords. A random temporal sequence of two or more of these 2^M chords is an RCS.

The states of elements in an RCS are determined in the following manner: First, a direction of frequency motion is chosen. This can be either upward ($N=1$) or downward ($N=-1$). Next, the states of the M elements of the first chord (or *frame*) are determined randomly, resulting in an average of 50% of the first-frame elements being in the on state. The states of elements in the second frame depend on the states in the first frame. If $N = 1$, each element (on or off) of the first chord is connected with the element of the second chord that is one frequency-step higher. If $N = -1$, it is connected with the second-chord element one frequency-step lower. The connection implies that the state of each element in the second chord will follow the state of the element in the first chord with which it is connected with a probability P, which is called the *state repetition probability*, or SRP. If, for instance, the states of the lowest and highest-but-one tones of the first chord are on and $N = 1$ and SRP $= 1$, the next to lowest and the highest notes of the second chord will be on as well. If SRP $= 0$, however, the next to lowest and the highest notes of the second chord would be off in this example. During an RCS, the direction of motion N and the SRP are kept constant. States of elements in the third chord are determined from element states in the second chord, and so on. To obtain a circular scheme similar to the visual displays used by Allik and Dzhafarov (1984), extreme chord elements are also connected: for $N = 1$, the highest note of the ith frame is connected with the lowest note of the $i+1$th frame, and for $N = -1$, the lowest note of the ith frame is connected with the highest note of the $i+1$th frame. Three examples of RCSs are illustrated in Figure 1.

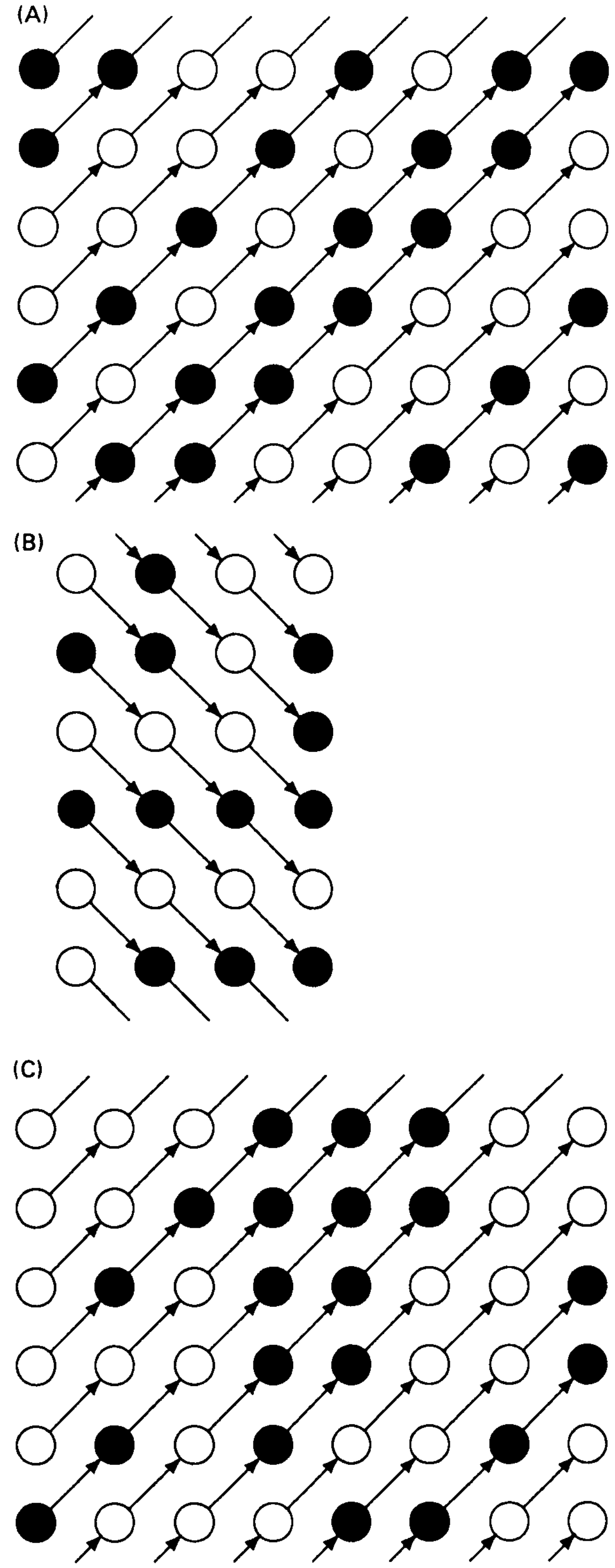


Figure 1. Three examples of random chord sequences. (A) $N=1$, SRP=1, circular scheme, 6 elements, 8 frames; (B) $N=-1$, SRP=0, circular scheme, 6 elements, 4 frames; (C) $N=1$, SRP=.8, circular scheme, 6 elements, 8 frames.

EXPERIMENTS

Two series of experiments are reported. Both represent attempts to measure the ability of subjects to identify the direction of perceived global pitch motion in RCSs of 2, 4, 5, and 8 sequential chords. In the first series of experiments, the effect of the sound used to represent chord elements was investigated. This series was carried out with 4 subjects at the Institute for Perception Research in The Netherlands, using a P 857 minicomputer and a 12-bit D/A converter to compute, store, synthesize, and present stimulus sequences. In the second series of experiments, the effect of chord composition was investigated. These experiments were carried out at the Institute of Language and Literature of the Estonian Academy of Sciences in Tallinn, E.S.S.R., with 2 subjects, using an EC 1010 minicomputer with a 12-bit D/A converter. Both sets of experiments are presented separately.

Methods

Experiment 1

Stimuli. In this experiment, there were always six elements (tones) in a chord, each of which could be either on or off. The tones were pure sinusoids, sawtooth waves, or *Shepard* tones. Of the sawtooth waves, only the first four harmonics were included. Shepard tones (Shepard, 1964) are complex tones with octave harmonics and a fixed bell-shaped spectral weight function. They have the property that transposition by one or more octaves always yields the same physical tone and therefore the same perceived pitch. This circular pitch property was used to make RCSs that were perceptually completely circular. Wave samples and spectra of the three sounds are shown in Figure 2. Sequences of two, four, and eight six-element chords were used. For sinusoidal and sawtooth-wave sounds, the fundamental frequencies of elements were chosen at 370, 392, 415, 440, 466, and 494 Hz, representing an intertone spacing of one equally tempered semitone. Shepard tones were tuned to 262, 294, 330, 370, 415, and 466 Hz, two semitones apart. The duration of each chord was 250 msec, which included a 20-msec on-and-off ramp. There were no interchord silent periods. The SRP varied between trials from .0 to 1.0, in steps of .1.

Procedure. All sessions started with a determination of the subject's hearing threshold for the chords to be used. The subject, who was seated in a double-walled sound-insulated chamber and received the stimuli binaurally through headphones, adjusted the intensity of an intermittent 440-Hz sinusoidal, sawtooth, or Shepard tone to detection threshold. All stimuli in the experiment were presented 20 dB above this empirically established level. This rather low level was chosen to avoid, as much as possible, any confounding effects of aural combination tones (Goldstein, 1967; Zwicker, 1955). After presentation of each RCS, the subject indicated, by pressing one of two buttons on a response box, whether the perceived global pitch motion was upward or downward. There was no response time limit, and no feedback was provided. In the case of trial runs with sinusoids or sawtooth waveforms, 100 trials were collected from each subject for each sound condition (i.e., two-, four-, and eight-frame stimuli combined with 11 SRP values). In the case of the Shepard waveform, only 60 trials were collected from each subject for each condition. Because there were no right or wrong answers, feedback was not provided. The subject's task was to indicate the subjectively perceived direction of pitch motion and not the physical direction of frequency motion, N.

Subjects. All RCSs with sinusoidal and sawtooth-wave tones were judged by 4 subjects, all of whom had had some musical training and experience. RCSs with Shepard tones were judged by 3 of these

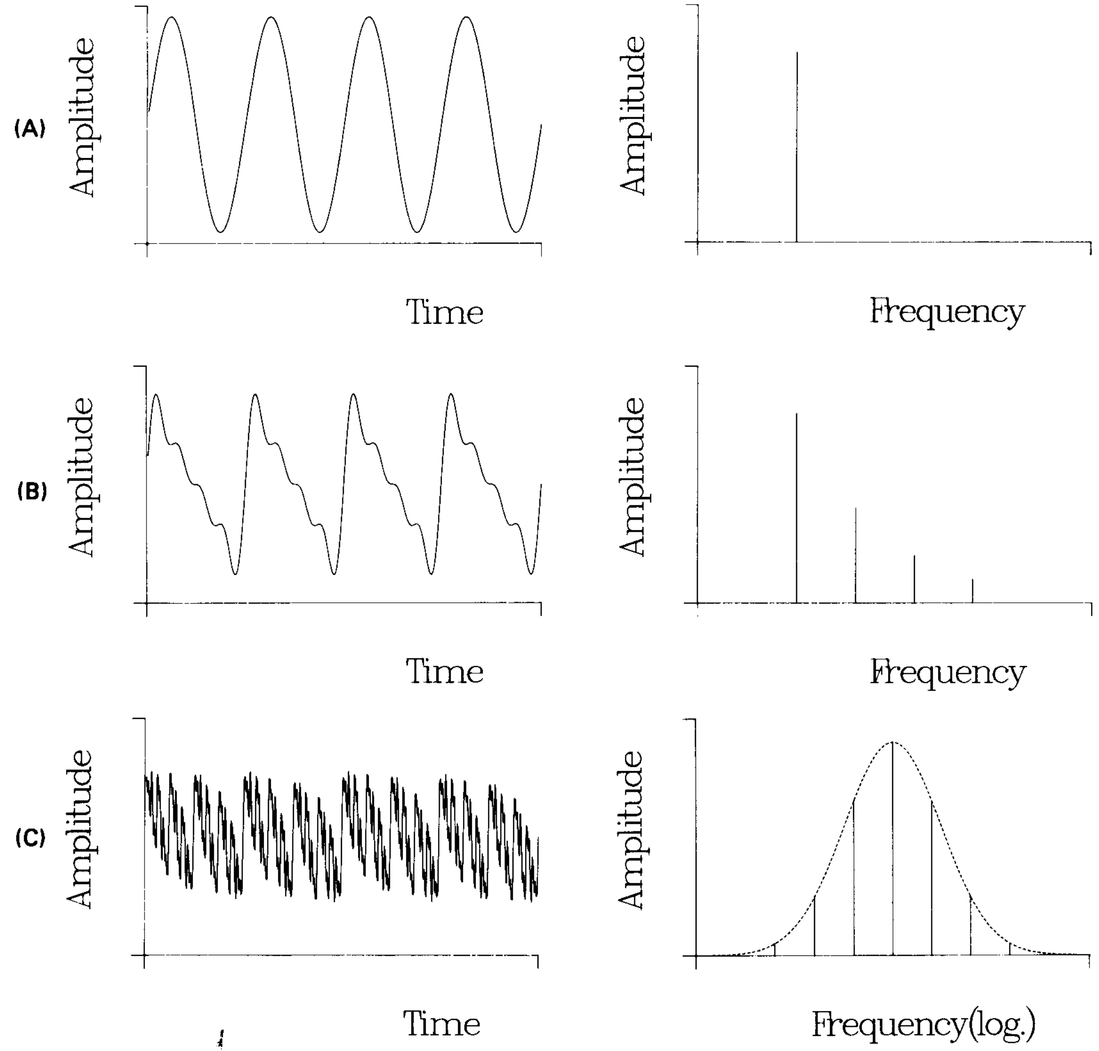


Figure 2. Samples of waveforms and spectra for the three types of sounds used in Experiments 1 and 2: (A) sinusoid, (B) sawtooth wave, (C) Shepard tone.

subjects. All subjects had normal hearing. They included two of the coauthors of this paper.

Experiment 2

Stimuli. In this experiment, there were either six or eight elements (tones) in a chord. The tones were always sawtooth-shaped complex tones comprising four harmonics, as shown in Figure 2b. In one series of trials, six chord elements were arranged in semitone increments with fundamental frequencies at 370, 392, 415, 440, 466, and 494 Hz, as in Experiment 1. In another series, fundamental frequencies were arranged to form a dominant-seventh chord of frequencies 196, 247, 294, 349, 392, and 494 Hz. In a third and fourth series, in which eight elements per chord were used, fundamentals were arranged in quarter tones (440, 453, 466, 480, 494, 508, 523, and 539 Hz) and as a dominant-seventh chord (196, 247, 294, 349, 392, 494, 587, and 696 Hz). All chords had a duration of 250 msec, which included a 20-msec on-and-off ramp. In the first two series of trials, RCSs of two, four, and eight chords were used; in the last two series only RCSs of five chords were used.

Procedure. The stimuli were synthesized digitally, played through a 12-bit D/A converter, and stored on magnetic tape. The subjects were seated in a quiet room, where they received the stimuli through loudspeakers. They were allowed to choose a comfortable sound level. After presentation of each RCS, the subject indicated, by writing the letters "Y" or "A" on a score sheet, whether the perceived global pitch motion was upward or downward. The allowed response time was 2 sec. There were 100 trials per subject in the first and second series, and 200 trials per subject in the third and fourth series for each sound condition and each SRP value. No feedback was provided.

Subjects. All RCSs were judged by 2 subjects with normal hearing. Both had had musical training. One of them was a coauthor of this study.

RESULTS

The results of Experiment 1 are summarized by the (subject-averaged) data points shown in Figures 3a-3c. For each of the sounds used to represent chord elements, the fraction, *Pc*, of trials is shown at each SRP value for which the direction of perceived global pitch motion

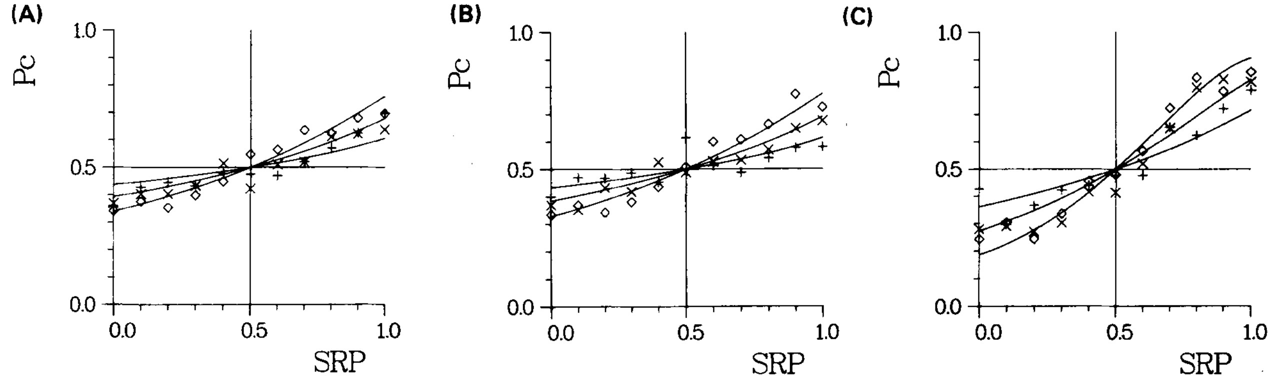


Figure 3. Experimental and theoretical results of Experiment 1, averaged over subjects: (A) Results obtained with sinusoidal tones, $V_{nj}=16.8$ ($\chi^2=80.7$); (B) results with sawtooth waves, $V_{nj}=14.5$ ($\chi^2=73.0$); (C) results with Shepard tones, $V_{nj}=4.17$ ($\chi^2=89.5$). +, two-chord; ×, four-chord; ◇, eight-chord sequence.

agreed with the chosen direction of frequency motion, *N*. Each data point, at which different symbols designate measurements with two-chord, four-chord, and eight-chord sequences, therefore represents $3 \times 60 = 180$ experimental trials in the case of the Shepard wave and $4 \times 100 = 400$ trials otherwise. The psychometric functions fitted through these points are model predictions that will be discussed in the next section.

The results of Experiment 2 are shown comprehensively in Figures 4a-4c. Panel a shows results obtained from 1 subject (one of the coauthors of this paper) with sequences of two, four, and eight chords comprising up to six sawtooth-wave tones arranged in semitone steps. This condition was comparable to that of Experiment 1 (second series of trials), the results of which were shown in Figure 3b. Figure 4b shows the results from the same subject for an arrangement of chord elements according to a dominant-seventh chord. Figure 4c shows the averaged results of 2 subjects obtained with five-chord sequences of eight elements arranged either in quarter-tone steps or in steps that form a dominant-seventh chord. Data points in Figures 4a and 4b represent 100 trials; those in Figure 4c represent 400 trials each. The functions shown represent model fits to be discussed later.

Observation of the data of Figures 3 and 4 reveals the following general tendencies:

1. Although the functions formed by the data points appear to pass through the point (*Pc*=.5; SRP=.5), as expected, the functions show a distinct lack of odd symmetry about this point. *Pc* is generally closer to .5 for SRP values smaller than .5 than it is for SRP values larger than .5. All functions appear to satisfy the inequality

$$Pc(\mathrm{SRP}) + Pc(1-\mathrm{SRP}) > 1. \tag{1}$$

The probability that this asymmetry is accidental was calculated to be smaller than 0.06% ($p < .05$) for the data of Experiment 1.

2. For SRP values smaller than .5, subjects always perceive a global pitch movement opposite to the actual direction of frequency movement. This phenomenon of apparent pitch-direction reversal holds for all chord and signal conditions. It is analogous to the direction-reversal phenomenon reported by Allik and Dzhafarov (1984) and the reversed-phi phenomenon reported by Anstis (1970).

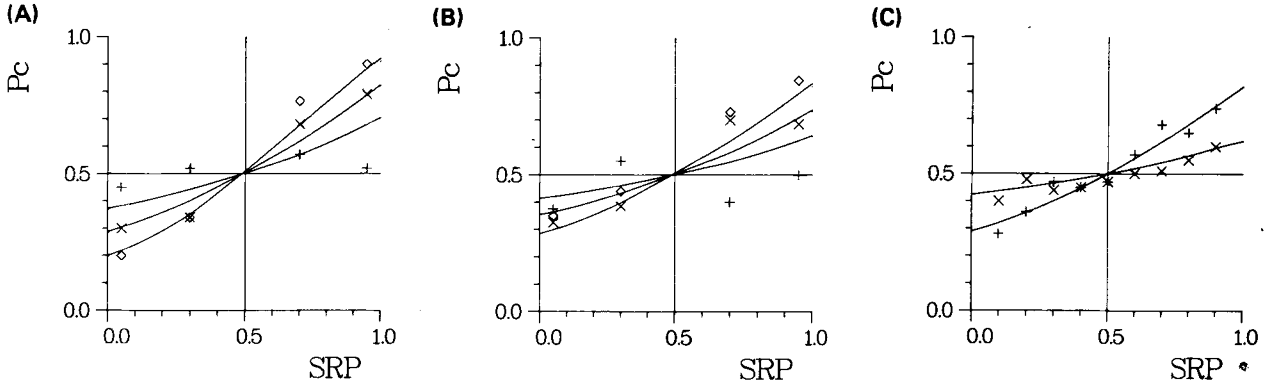


Figure 4. Experimental and theoretical results of Experiment 2 performed with sawtooth waves: (A) Element spacing in semitones, Subject J.R., $V_{nj}=3.78$ ($\chi^2=23.2$); (B) element spacing according to a dominant-seventh chord, Subject J.R., $V_{nj}=8.51$ ($\chi^2=35.4$); +: two-chord, ×: four-chord, ◇: eight-chord sequence; (C) element spacing in quarter tones (+, $V_{nj}=7.03$, $\chi^2=21.3$) and according to dominant-seventh chord (×, $V_{nj}=63.2$, $\chi^2=10.5$), averaged over subjects J.R. and M.R., with five-chord sequences.

3. Both the positive and the negative portions of the psychometric functions take on more extreme values when the number of chords in an RCS is increased or the sound used to represent chord elements is spectrally enriched.

4. The functions obtained reach more extreme values of *Pc* when chord elements are regularly spaced (whole tones, semitones, quarter tones) than when they are irregularly spaced, as is the case with the dominant-seventh chord arrangement.

MODELS

Voice-Tracing Model

When one listens to an RCS, a sequence of simultaneous tones is heard. When a subject is asked to determine the direction of global pitch motion heard in such a sequence, he or she may, on the one hand, make the decision on the basis of apparent movement of some "average" pitch. An "average" pitch cannot be a simple frequency average for each frame, since such an average contains no relevant information about the imposed frequency motion. The listener will have to form some list of element pairs that define a correspondence pattern between successive chords, and then average only over those pairs. The solving of the correspondence problem, which is not a simple matter, will be returned to later. Alternatively, the subject might use the much simpler strategy of tracing a particular set of elements in successive frames. Musical practice indicates, for instance, that polyphonic dictation, where students have to write down the notes of several simultaneously sounding voices, is always easiest for the two extreme voices: soprano and bass. It therefore seems possible that the subjects in the present experiments based their responses entirely on the perceived frequency movement of either upper or lower elements of successive chords.

Inspection of the data for two-chord RCSs of Experiment 1 from this point of view, and only for those RCSs that contained sinusoidal or sawtooth-wave tones, revealed a definite correspondence between perceived pitch motion and physical frequency motion of the highest on elements of the two chords. Eighty percent of all trials showed this correspondence. For those trials on which the highest tone frequencies were identical, the lowest on elements (bass voice) appeared to have some influence on perceived pitch motion, but the correlation was not very large. For RCSs of four and eight chords, the criterion for upward versus downward frequency motion of the top (soprano) or bottom (bass) voice is no longer clear, because of the randomness in the order of successive notes. Although some criterion could possibly be defined, it was not attempted because of inherent arbitrariness.

The results obtained with the Shepard tones provide perhaps the strongest evidence that simple voice-tracing models cannot account for observed behavior. In a chord of (up to) six Shepard tones, one cannot really tell which tone is the highest or the lowest because of the circular pitch property exhibited by such tones. A soprano or bass voice is, even according to purely physical criteria, impossible to determine. The data of Figure 3c, however, clearly show that the correlation between the directions of frequency motion (*N*) and perceived global pitch motion is greatest precisely for this type of chord elements. This suggests that global pitch movement is not a perceptual feature of a few particular elements in the chord sequence, and that it is probably a global feature to which all elements can in principle contribute.

Dipole Contribution Model

To determine which elements in an RCS contribute to the percept of pitch motion, the RCS is separated into its smallest elementary events for which pitch motion can be perceived. Every two elements (tones) in two different chords (frames) form such an elementary event. These events are called *dipoles*. A dipole, D, is characterized by (1) its displacement vector $\mathbf{d}=(f,t)$, where f is the number of tone (frequency) steps within the given chord structure, and t is the temporal distance expressed in frames, and (2) its form, h, denoting the states of the two dipole elements. Since there are only two possible states for any element, there are four possible dipole forms: on-on, on-off, off-on, and off-off, denoted as h_{11}, h_{10}, h_{01}, and h_{00}, respectively. An example is shown in Figure 5.

The main idea of the dipole contribution model (DCM), proposed by Allik and Dzhafarov (1984), is that each dipole contributes in principle to perceived global pitch motion. Some dipoles, such as the ones with displacement vectors $\mathbf{d}=(0,0)$, $(0,1)$, or $(1,0)$, are excluded because they cannot possibly convey such a percept. The dipole contribution, $c(D)=c(f,t,h)$, is a unidimensional random variable representing the perceptual contribution of this elementary event to apparent pitch motion. The sign of this contribution is taken as positive if the displacement vector $\mathbf{d}$ points upward (f is positive) and negative if it points downward (f is negative). The statistics of $c(D)$

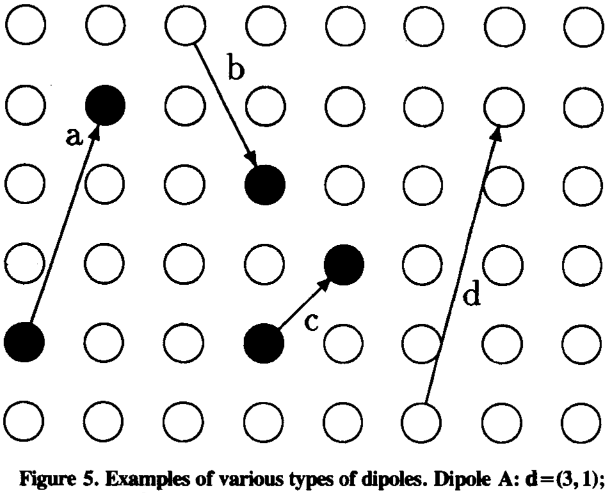


Figure 5. Examples of various types of dipoles. Dipole A: $\mathbf{d}=(3,1)$; h_{11}. Dipole B: $\mathbf{d}=(-2,1)$; h_{01}. Dipole C: $\mathbf{d}=(1,1)$; h_{11}. Dipole D: $\mathbf{d}=(4,1)$; h_{00}.

depend on the dipole's displacement vector **d** and form h. Dipole contributions $c(D)$ are assumed to have the following general properties:

Homogeneity. All dipoles of the same displacement vector **d** and form h contribute equally, regardless of their position of occurrence within an RCS.

Symmetry. Contributions of any two symmetrical dipoles—that is, dipoles of the same form h but with displacement vectors $\mathbf{d}=(f,t)$ and $\mathbf{d}=(-f,t)$, respectively, have symmetrical probability density functions so that the average of their net contribution is zero.

Independence. Contributions of any two different dipoles are statistically independent, provided their types and forms are known. (Because of the interdependence of forms for some dipole pairs in the configuration, when the dipole form is a random variable—which is how it is treated in most of the remainder of this paper—contributions of two dipoles are not always independent.)

The contributions of all the various dipoles are arithmetically added to form a total sum of contributions (TSC), which, of course, is also a random variable. If for a given chord sequence the TSC is positive, the perceived direction of global pitch motion will be upward, and if it is negative, it will be downward. For any value of SRP and N, the relative frequency of occurrence of the various dipoles can be determined statistically. It is assumed that, since the TSC is the sum of a large number of relatively small contributions, most of which are mutually independent, the TSC can be considered as approximately Gaussian and specified by its two characteristic parameters E[TSC] and var[TSC]. The probability Pc of a "correct" identification of global pitch motion—that is, a perceived motion that agrees with the actual direction of frequency transformation N—is:

$$Pc = \frac{1}{\sqrt{2\pi\mathrm{var[TSC]}}}\int_0^{\infty}\exp\left[-\frac{(x-\mathrm{E[TSC]})^2}{2\mathrm{var[TSC]}}\right]dx. \quad (2)$$

This is illustrated in Figure 6, in which "correct" identification of either frequency motion direction has been indicated by the two differently shaded areas. The integral of Equation 2 can easily be reduced to the standard Gaussian integral:

$$Pc = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\mathrm{E[TSC]}/\sigma\mathrm{[TSC]}}\exp\left[-\frac{x^2}{2}\right]dx$$
$$= \Phi\left(\frac{\mathrm{E[TSC]}}{\sqrt{\mathrm{var[TSC]}}}\right), \quad (3)$$

where $\sigma[\mathrm{TSC}] = \sqrt{\mathrm{var[TSC]}}$.

Although Equation 3 may suggest that only two free model parameters are involved, there are, in principle, very many, because the magnitudes of $c(D)$—that is, the details of density functions—are freely chosen for each dipole of a particular displacement vector and form. A model with so many free parameters is not testable and therefore not too interesting. To restrict the model further, three additional simplifications were made:

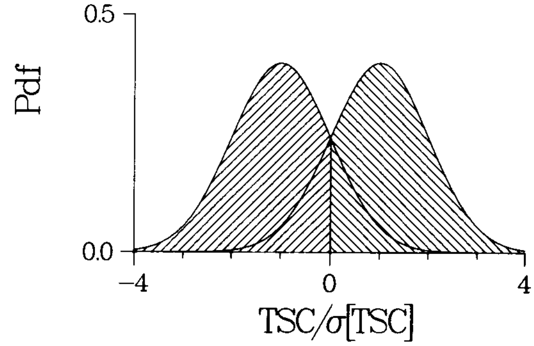


Figure 6. Gaussian detection model relating observed identification performance to statistics of the dipole contribution model. The decision criterion is assumed to be balanced–that is, placed to maximize "correct" identifications.

1. The DCM was restricted to a *local* or *short-range* model, in which only those dipoles contribute to the TSC that have a displacement vector **d** that is either $(-1,1)$ or $(1,1)$. This implies that global pitch motion is affected only by single-tone (frequency) jumps within successive time frames. Dipoles that span several frequency steps and/or extend over nonsuccessive chords are ignored. This simplification was found to work well for the visual experiments on motion perception in circular random cinematograms by Allik and Dzhafarov (1984).

2. Dipoles of the form h_{00} (i.e., off-off) do not contribute to perception of global pitch motion and are therefore ignored. This assumption is specific to the present auditory experiments, because perception of pitch motion for two successive tones that are both off seems highly unlikely. Dipoles of the form h_{11} (i.e., on-on) are the only ones that convey a percept of pitch motion. They are referred to as *jumps*.

3. Contributions by dipoles of the forms h_{01} and h_{10}, referred to as *nonjumps*, have identical distributions, with expected values equal to zero.

Because these assumptions leave only two kinds of contributions, namely those of jump and nonjump dipoles, the number of parameters of the DCM has been reduced to only two. Allik and Dzhafarov (1984) further found that their visual data could be well accounted for with a model version in which the jump contributions were a fixed, deterministic number (random variable of zero variance), leaving the variance of nonjump contributions to be estimated as the only free parameter of the model. This simplification will be considered as well.

Although the computation of model predictions for perceived pitch direction with the use of Equation 3 is, in principle, straightforward, the actual computation of the necessary statistics E[TSC] and var[TSC] for a given RCS can be very cumbersome and tedious. Because such computations have been shown in detail by Allik and Dzhafarov (1984), we will present here only the highlights for that particular form of the DCM that applies to the present sound experiments.

The expected value of the TSC is given by the expression

$$\mathrm{E[TSC]} = ME^{+}_{11}\left(\frac{P}{2} - \frac{1}{4}\right), \tag{4}$$

where M represents the total number of dipoles with a displacement vector $\mathbf{d}=(1,1)$ in the RCS, E^{+}_{11} is the mean perceptual contribution of dipoles with that displacement vector and form h_{11} (on-on), and P is a simple abbreviation of the state repetition probability SRP. The derivation of this equation is shown in Appendix A.

The variance of TSC comprises two sets of terms:

$$\mathrm{var[TSC]} = \sum_{i}\mathrm{var}[c(d_i)] + 2\sum_{i}\sum_{j\neq i}\mathrm{cov}[c(d_i),c(d_j)]. \tag{5}$$

The first term, as is shown in Appendix B, is equal to

$$\sum_{i}\mathrm{var}[c(d_i)] = \frac{1}{4}M(E^{+}_{11})^2\left\{(2P+1)V_j + (2Q+1)V_{nj} + P + PQ + \frac{3}{4}\right\}, \tag{6}$$

where P is the SRP, Q equals $1-P$, and V_j and V_{nj} are the normalized variances of contributions by dipoles in jump form and the nonjump form, respectively:

$$V_j = \frac{\mathrm{var}[c(d^{+}_{11})] + \mathrm{var}[c(d^{+}_{00})]}{(E^{+}_{11})^2},$$

$$V_{nj} = \frac{\mathrm{var}[c(d^{+}_{01})] + \mathrm{var}[c(d^{+}_{10})]}{(E^{+}_{11})^2}.$$

The second term of Equation 5 is a sum of covariances, which can be written as:

$$2\sum_{i}\sum_{j\neq i}\mathrm{cov}[c(d_i),c(d_j)] = \frac{1}{2}(E^{+}_{11})^2 \sum_{i,j\in S_1,S_2}\left\{P(d_i,d_j) - \frac{1}{4}\right\}, \tag{7}$$

where $P(d_i,d_j)$ is the joint probability that the ith and jth dipoles are both of the form h_{11} (on-on), and the summation is taken over covariances of those single-step dipole pairs $\mathbf{d}=(\pm1,1)$ that lie along a diagonal in the direction of frequency motion in the RCS (set S_1) or that connect two of those successive diagonals (set S_2). This is because the effect of the *SRP* propagates along those diagonals and causes a correlation between the forms of dipoles contained in or touching them. All dipole pairs not belonging to the sets S_1 or S_2 have a joint probability $P(d_i,d_j)$ of being in the h_{11} form that equals .25, which makes their covariance zero. Unfortunately, there is no simple general expression for $P(d_i,d_j)$ for the dipoles contained in the sets S_1 and S_2, since such an expression depends on the length of the chord sequence. For any chord sequence, however, $P(d_i,d_j)$ ultimately depends only on the state repetition probability SRP, as shown in Appendix C.

If we now return to Equation 3, we see with the aid of Equations 4, 6, and 7 that its independent variable $\mathrm{E[TSC]}/\sqrt{\mathrm{var[TSC]}}$ depends only on M, $P(=\mathrm{SRP})$, V_j, and V_{nj}, where the latter two are free parameters.

For each of the data sets shown in Figures 3 and 4, functions were computed in such a way that the chi-square error between function and data values was minimized. This was first done with two free parameters, V_j and V_{nj}, and also with only one free parameter V_{nj}, V_j being put to zero. The results of this last computation are shown in Table 1. An analysis of the values obtained for V_j, V_{nj}, and chi-square showed that:

1. The model version with the two free parameters V_j and V_{nj} does not account for the data significantly better than does the version with only V_j as a free parameter. We therefore retained the version with $V_j=0$.
2. Almost all individual data can be accounted for by the model with $V_j=0$ ($p<.01$).
3. Although values of V_{nj} vary from subject to subject and from condition to condition, overall behavior is still best displayed by presenting model fits to the pooled data.
4. The values of V_{nj} are particularly small in the case of Shepard waveforms (Experiment 1), and tend to be large when tone frequencies in a frame are irregularly spaced (Experiment 2).

The computed functions for which the chi-square error with the averaged data was minimal are shown as solid curves in Figures 3 and 4. Furthermore, the experimental scores *Pc*, expressed as the proportion of trials for which the perceived global pitch motion agreed with the direction of frequency motion N, were computed as a function of the net number of short-range on-on dipoles that were actually counted in the various stimuli. These results are shown in Figures 7a-7c, according to the three different sound waveforms used in Experiment 1. The data plotted in this form provide direct support for the basic assumption of the DCM that direction of perceived pitch motion is determined by the net sum of frequency-jump dipoles.

DISCUSSION

The data of both Experiments 1 and 2 can be accounted for with a simple version of the dipole contribution model (DCM). "Simple version" means that (1) it is a short-range local model, in which correspondence between elements is limited to adjoining elements in successive time frames; (2) all off-off jumps do, on the average, not contribute (no-blank version); and (3) the noise involved in the perception of on-on jumps is assumed to be zero ($V_j=0$). The model accounts for the reversed direction of perceived pitch motion, apparent from the data, for state repetition probabilities smaller than .5 (the so-called "reversed-phi phenomenon"), as well as the slight asymmetry of empirically obtained psychometric functions

Table 1
Experimental Conditions for Experiments 1 and 2

Exp	waveform	interval	frames	subjects	df	N	V_{nj}	χ^2
1	sinusoid	$\frac{1}{2}$	2+4+8	AH	31	100	18.9	36.8
1	sinusoid	$\frac{1}{2}$	2+4+8	BE	31	100	69.9	37.1
1	sinusoid	$\frac{1}{2}$	2+4+8	JJ	31	100	13.1	31.1
1	sinusoid	$\frac{1}{2}$	2+4+8	NV	31	100	7.56	35.7
1	sinusoid	$\frac{1}{2}$	2+4+8	AH+BE+JJ+NV	31	400	16.8	80.7
1	sawtooth	$\frac{1}{2}$	2+4+8	AH	31	100	47.1	40.3
1	sawtooth	$\frac{1}{2}$	2+4+8	BE	31	100	36.5	29.8
1	sawtooth	$\frac{1}{2}$	2+4+8	JJ	31	100	11.7	50.4
1	sawtooth	$\frac{1}{2}$	2+4+8	NV	31	100	4.35	41.4
1	sawtooth	$\frac{1}{2}$	2+4+8	AH+BE+JJ+NV	31	400	14.5	73.0
1	Shepard	1	2+4+8	AH	31	60	4.41	39.8
1	Shepard	1	2+4+8	JJ	31	60	7.99	59.5
1	Shepard	1	2+4+8	NV	31	60	2.18	62.6
1	Shepard	1	2+4+8	AH+JJ+NV	31	180	4.17	89.5
2	sawtooth	$\frac{1}{2}$	2+4+8	JR	10	100	3.78	23.2
2	sawtooth	dom^7	2+4+8	JR	10	100	8.51	35.4
2	sawtooth	$\frac{1}{4}$	5	JR	7	200	4.63	5.88
2	sawtooth	$\frac{1}{4}$	5	MR	7	200	13.1	18.6
2	sawtooth	$\frac{1}{4}$	5	JR+MR	7	400	7.03	21.3
2	sawtooth	dom^7	5	JR	7	200	60.9	18.6
2	sawtooth	dom^7	5	MR	7	200	60.7	21.7
2	sawtooth	dom^7	5	JR+MR	7	400	63.2	10.5

Note—The values of V_{nj} and χ^2 are calculated under the assumption that V_j=0. The degrees of freedom (*df*) are the number of data points minus 1 minus the number of free parameters (1). *N* represents the number of trials per data point.

about the point Pc=.5, SRP=.5. The chi-square test did not allow us to discriminate unequivocally between possible variants of the short-range local DCM, such as the assumptions that $V_j \neq V_{nj} \neq 0$ (two free parameters) or that $V_j = 0$. The former assumption, when put into the model, yields somewhat smaller chi-square values, as is expected with two free parameters. The implication of the parameter values found for V_j and V_{nj} is that subjects practically never miss on-on relationships (which causes V_j to be close to zero), whereas on-off or off-on relationships sometimes are mistaken for on-on ones (and cause V_{nj} to be nonzero). Since a chi-square test did not provide compelling evidence to reject the hypothesis of V_j=0, we chose to represent the global pitch perception process with the rather extreme version of the DCM in which V_j=0, leaving only the free parameter V_{nj} to be estimated, without the intention of claiming that the percepts of short-range on-on jumps are totally noiseless.

Another potentially significant finding is that perception of global pitch motion within the general context of these experiments can be well accounted for with a short-range, local form of the DCM. There remains the question, however, of to what extent this finding can be generalized to all music perception. Experiments with tone streaming

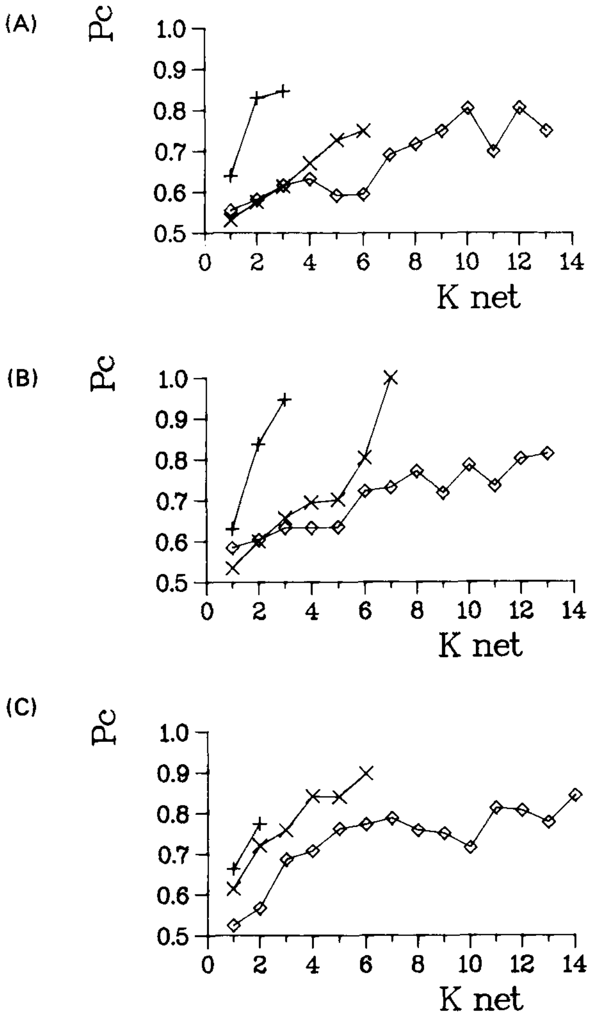


Figure 7. Experimental results of Experiment 1, plotted against the net number of d=(±1,1), h_{11} dipoles actually counted in stimuli: (A) sinusoid; (B) sawtooth-shaped wave; (C) Shepard tone.

have convincingly shown that in sufficiently rapid tone sequences our auditory system may group tone elements together that do not occur in successive time frames. It is possible, for instance, to construct a single sequence of tones that leads to a percept of two parallel melodies formed by odd- and even-numbered notes, respectively (Bregman & Campbell, 1971; van Noorden, 1975). There are also visual analogues to those streaming experiments (Julesz & Bosche, 1966). It is quite possible that the short-range local solution to the element-correspondence problem, which the auditory system seems to employ in the present experiments, has something to do with the large amount of randomness in the chord sequences. Harmonic or melodic structure, put into transitions from one chord to the next, could cause the auditory system to use solutions to the correspondence problem other than a short-range local one. This implies that 20th-century serial music, in which traditional harmonic and melodic structures play a negligible role (Austin, 1966), is perceived in a short-range and local manner by most listeners. If this conclusion is correct, a composer may have a possible tool for manipulating a listener's perception of correspondence between notes in varying the amounts of randomness and structure in the music.

The local form of DCM can also account for some auditory pitch paradoxes reported in the literature. One of them, actually used in this study, occurs when a Shepard tone is followed by another Shepard tone in which all frequencies have been multiplied by 1.89. Instead of hearing the pitch go upwards a major seventh, subjects typically report hearing a pitch descent of a semitone. This is because there actually are many semitone frequency jumps in a downward direction in this case, which apparently dominate the global percept. Another paradox, recently reported by Schroeder (1986) and Risset (1986), has a slightly stretched Shepard tone (e.g., 49.6, 102.4, 211.2, ... Hz) followed by its exact octave transposition (i.e., 99.2, 204.8, 422.4, ... Hz). Despite the component frequency jumps of exactly one octave upwards, subjects typically report hearing a slightly descending pitch jump. Apparently, in this case too, *local* dipoles formed by frequencies that are close dominate the global pitch percept and are stronger than the *central* or *virtual* pitch, which should have resulted in upward octave percepts.

Considering the quantitative behavior of the parameter V_{nj}, it is found that:

1. It appears to be rather subject-dependent. However, since we are interested only in the global features of the subject's behavior, we have presented only the pooled data.

2. It seems to decrease with the harmonic richness and complexity of signals employed to represent the tone elements of the chords. The largest decrease of V_{nj}, however, occurred between the data shown in Figures 3b and 3c, representing sawtooth and Shepard tones, respectively, whereas the change between the cases of sine and sawtooth representation (Figures 3a and 3b) was rather small. Unfortunately, the change from sawtooth to Shepard tones involved not only a change in tonal spectra but also a change of tone spacing within chords (semitone to whole-tone spacing) and, more importantly, a change to a situation of true physical circularity of the chord sequence. This covariance of parameters obscures a clear-cut conclusion on what exactly causes the decrease in the noise of perceived nonjumps.

3. Regularity of intertone spacing generally leads to smaller values of V_{nj}. This can be seen by comparing the results in Figures 4a and 4b, and also the two sets of data shown in Figure 4c. Despite the larger interelement tonal distances of the dominant-seventh chord, compared with semitones or quarter tones, the resulting estimates of V_{nj} are consistently and significantly larger. This, in a sense, is logical, because with a regular interelement space, the processor has to deal with only a single dipole size, whereas in the dominant-seventh chord arrangement, the physical frequency jumps of nominally identical dipoles

are often different in size. This can only be a confounding factor for any processor, and it is most likely to lead to degraded performance.

Finally, the auditory results presented in this study appear to be very similar to the visual data obtained with random circular cinematograms by Allik and Dzhafarov (1984). Both sets of data are adequately accounted for by very similar models. This suggests that, in both cases, we are apparently dealing with a general cognitive brain process that extracts information from visual and auditory sensory inputs in a similar manner.

REFERENCES

Allik, J., & Dzhafarov, E. N. (1984). Motion direction identification in random cinematograms: A general model. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 378-393.

Anstis, S. M. (1970). Phi movement as subtraction process. *Vision Research*, **10**, 1411-1430.

Austin, W. (1966). *Music in the 20th century: From Debussy through Stravinsky*. New York: W. W. Norton.

Beerends, J. G., & Houtsma, A. J. M. (1986). Pitch identification of simultaneous dichotic two-tone complexes. *Journal of the Acoustical Society of America*, **80**, 1048-1056.

Beerends, J. G., & Houtsma, A. J. M. (1989). Pitch identification of simultaneous diotic and dichotic two-tone complexes. *Journal of the Acoustical Society of America*, **85**, 813-819.

Bell, H. H., & Lappin, J. S. (1973). Sufficient conditions for the discrimination of motion. *Perception & Psychophysics*, **14**, 45-50.

Bregman, A. S., & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequence of tones. *Journal of Experimental Psychology*, **89**, 244-249.

Burns, E. M., & Ward, W. D. (1978). Categorical perception—phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic musical intervals. *Journal of the Acoustical Society of America*, **63**, 456-468.

Butler, D. (1979). A further study of melodic channeling. *Perception & Psychophysics*, **25**, 264-268.

de Boer, E. (1976). On the "residue" and auditory pitch perception. In W. D. Keidel & W. D. Neff (Eds.), *Handbook of sensory physiology* (Vol. 5, Part 3, pp. 479-583). New York: Springer-Verlag.

Deutsch, D. (1975). Two-channel listening to musical scales. *Journal of the Acoustical Society of America*, **57**, 1156-1160.

Deutsch, D. (1980). The processing of structured and unstructured tonal sequences. *Perception & Psychophysics*, **28**, 381-389.

Doehring, D. G. (1971). Discrimination of simultaneous and successive pure tones by musical and nonmusical subjects. *Psychonomic Science*, **22**, 209-210.

Doughty, J. M., & Garner, W. M. (1948). Pitch characteristics of short tones: II. Pitch as a function of duration. *Journal of Experimental Psychology*, **38**, 478-494.

Dowling, W. J. (1973). The perception of interleaved melodies. *Cognitive Psychology*, **5**, 322-337.

Eiting, M. (1984). *Aspects of the cognition of tonal music*. Unpublished doctoral dissertation, University of Amsterdam.

Goldstein, J. L. (1967). Auditory nonlinearity. *Journal of the Acoustical Society of America*, **41**, 676-689.

Hartmann, W. M. (1978). The effect of amplitude envelope on the pitch of sinewave tones. *Journal of the Acoustical Society of America*, **63**, 1105-1113.

Julesz, B. (1971). *Foundation of cyclopean perception*. Chicago: University of Chicago Press.

Julesz, B., & Bosche, C. (1966). *Studies on visual texture and binocular depth perception* (A computer-generated movie series containing monocular and binocular movies). Murray Hill, NJ: Bell Telephone Laboratories, Inc.

Kimura, D. (1964). Left-right differences in the perception of melodies. *Quarterly Journal of Experimental Psychology*, **16**, 355-358.

Korte, A. (1915). Kinematoskopische Untersuchungen. *Zeitschrift für Psychologie*, **72**, 193-296.

Larkin, W. D. (1978). Pitch shifts following tone adaptation. *Acustica*, **41**, 110-116.

Nakayama, K., & Silverman, G. (1984). Temporal and spatial characteristics of the upper displacement limit for motion in random dots. *Vision Research*, **24**, 293-300.

Pantle, A. J., & Picciano, L. A. (1976). A multistable movement display: Evidence for two separate motion systems in human vision. *Science*, **193**, 500-502.

Risset, J.-C. (1986). Pitch and rhythm paradoxes: Comment on "Auditory paradox based on fractal waveform" [J. Acoust. Soc. Am. 79, 186-189, 1986]. *Journal of the Acoustical Society of America*, **80**, 961-962.

Rossing, T. D., & Houtsma, A. J. M. (1986). Effects of signal envelope on the pitch of short sinusoidal tones. *Journal of the Acoustical Society of America*, **79**, 1926-1933.

Scharf, B., & Houtsma, A. J. M. (1986). Audition: II. Loudness, pitch, localization, aural distortion, pathology. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 1, pp. 15-1/15-60). New York: Wiley.

Schroeder, M. R. (1986). Auditory paradox based on fractal waveform. *Journal of the Acoustical Society of America*, **79**, 186-189.

Shepard, R. N. (1964). Circularity in the judgments of relative pitch. *Journal of the Acoustical Society of America*, **36**, 2346-2353.

Stevens, S. S. (1935). The relation of pitch to intensity. *Journal of the Acoustical Society of America*, **6**, 150-154.

Stevens, S. S., & Volkmann, J. (1940). The relation of pitch to frequency. *American Journal of Psychology*, **53**, 329-353.

Terhardt, E., & Fastl, H. (1971). Zum Einfluss von Störtönen und Störgeräuschen auf die Tonhöhe von Sinustönen. *Acustica*, **25**, 53-61.

Ullman, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.

van Doorn, A. J., & Koenderink, J. J. (1982). Spatial properties of the visual detectability of moving spatial white noise. *Experimental Brain Research*, **45**, 189-195.

van Doorn, A. J., & Koenderink, J. J. (1984). Spatiotemporal integration in the detection of coherent motion. *Vision Research*, **24**, 47-54.

van Noorden, L. P. A. S. (1975). *Temporal coherence in the perception of tone sequences*. Unpublished doctoral dissertation, Eindhoven University of Technology.

Verschuure, J., & van Meeteren, A. A. (1975). The effect of intensity on pitch. *Acustica*, **32**, 33-44.

Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung. *Zeitschrift für Psychologie*, **61**, 161-265.

Zwicker, E. (1955). Der ungewöhnliche Amplitudengang der nichtlinearen Verzerrungen des Ohres. *Acustica*, **5**, 67-74.

APPENDIX A

This appendix shows the computation of E[TSC]. According to general rules,

$$\mathrm{E[TSC]} = \mathrm{E}\left[\sum_d c(d)\right] = \sum_d \mathrm{E}[c(d)]. \tag{A1}$$

Breaking down E[$c(d)$] into its partial contributions corresponding to the four different dipole forms h with the associated probabilities Pd_h, we can write Equation A1 as

$$\mathrm{E[TSC]} = \sum_d \mathrm{E}[c(d)] = \sum_d \sum_h \mathrm{E}[c(d_h)]Pd_h. \tag{A2}$$

Since only dipoles with form h_{11} have a nonzero contribution, Equation A2 can be further simplified to

$$\sum_d \sum_h \mathrm{E}[c(d_h)]Pd_h = \sum_d \mathrm{E}[c(d_{11})]Pd_{11}. \tag{A3}$$

According to the symmetry principle, contributions of any two symmetrical dipoles are identical but differ in sign:

$$\mathrm{E}[c(d^{-}_{11})] = -\mathrm{E}[c(d^{+}_{11})], \tag{A4}$$

where d^{-}_{11} and d^{+}_{11} have displacement vectors $(-f, t)$ and (f, t), respectively, and are both of the h_{11} form. We now split up the sum of Equation A3 into two parts, one being the sum of positive $[(f, t)]$, the other the sum of negative $[(-f, t)]$ dipoles. We can finally rewrite the summation over only positive dipoles, since the total number of positive dipoles equals the total number of negative dipoles in each RCS:

$$\begin{aligned}\sum_{d} \mathrm{E}[c(d_{11})]Pd_{11} &= \sum_{d>0} \mathrm{E}[c(d^{+}_{11})]Pd^{+}_{11} + \sum_{d<0} \mathrm{E}[c(d^{-}_{11})]Pd^{-}_{11} \\ &= \sum_{d>0} \mathrm{E}[c(d^{+}_{11})](Pd^{+}_{11} - Pd^{-}_{11}). \end{aligned} \tag{A5}$$

Since only the shortest dipoles contribute, we simply need to sum over all $(1, 1)$ dipoles. According to the homogeneity principle, all these contributions are equal, so that

$$\mathrm{E}[\mathrm{TSC}] = ME^{+}_{11}(Pd^{*+}_{11} - Pd^{*-}_{11}), \tag{A6}$$

where d^{*+}_{11} is a dipole with displacement vector $(1, 1)$ and form h_{11}, d^{*-}_{11} a dipole with displacement vector $(-1, 1)$ and form h_{11}, and M the total number of $(1, 1)$ dipoles in an RCS. E^{+}_{11} is a short notation for $\mathrm{E}[c(d^{+}_{11})]$. The chance that a $(1, 1)$ dipole is in h_{11} form (Pd^{*+}_{11}) is either $\frac{1}{2}P$ or $\frac{1}{4}$, depending on the motion direction N; Pd^{*-}_{11}, then, is $\frac{1}{4}$ or $\frac{1}{2}P$, respectively. Assuming that $N = 1$, we get for E[TSC]:

$$\mathrm{E}[\mathrm{TSC}] = ME^{+}_{11}\left(\frac{1}{2}P - \frac{1}{4}\right). \tag{A7}$$

This equation corresponds to Equation 4 in the text. For $N = -1$, the same expression is obtained for E[TSC] except for an opposite sign.

APPENDIX B

This appendix shows the calculation of var[TSC]. According to general rules,

$$\begin{aligned}\mathrm{var}[\mathrm{TSC}] &= \mathrm{var}\left[\sum_{d} c(\mathrm{d})\right] \\ &= \sum_{d} \mathrm{var}[c(d)] + 2\sum_{i}\sum_{j \neq i} \mathrm{cov}[c_i(d), c_j(d)] \end{aligned} \tag{B1}$$

and

$$\mathrm{var}[c(d)] = \mathrm{E}[c^2(d)] - \mathrm{E}^2[c(d)], \tag{B2}$$

where

$$\mathrm{E}[c^2(d)] = \sum_{h} \mathrm{E}[c^2(d_h)]Pd_h$$

$$\mathrm{E}^2[c(d)] = \left\{\sum_{h} \mathrm{E}[c(d_h)]Pd_h\right\}^2.$$

Pd_h is the probability that the dipole is of form h, and the summation is done over all four possible forms h. Using the fact that only $\mathrm{E}[c(d_{11})]$ is nonzero, we can write Equation B2 as

$$\mathrm{var}[c(d)] = \mathrm{E}^2[c(d_{11})]Pd_{11}(1 - Pd_{11}) + \sum_{h} \mathrm{var}[c(d_h)]Pd_h. \tag{B3}$$

With the aid of the symmetry principle, we now write the sum of variances over only the positive dipoles. If, in addition, we only consider the shortest $(1, 1)$ dipoles, we obtain

$$\begin{aligned}&\sum_{d} \mathrm{var}[c(d)] \\ &= \sum_{d^{*}>0}\Bigg(\sum_{h}\{\mathrm{var}[c(d^{*+}_{h})](Pd^{*+}_{h} + Pd^{*-}_{h})\} + \mathrm{E}^2[c(d^{*+}_{11})] \\ &\qquad [Pd^{*+}_{11}(1 - Pd^{*+}_{11}) + \mathrm{P}d^{*-}_{11}(1 - \mathrm{P}d^{*-}_{11})]\Bigg), \end{aligned} \tag{B4}$$

where the notations d^{*+} and d^{*-} designate dipoles with displacement vectors of $(1, 1)$ and $(-1, 1)$, respectively, as defined in Appendix A. Regardless of whether $N = 1$ (i.e., $Pd^{*+}_{11} = \frac{1}{2}P$ and $Pd^{*-}_{11} = \frac{1}{4}$) or $N = -1$ ($Pd^{*+}_{11} = \frac{1}{4}$ and $Pd^{*-}_{11} = \frac{1}{2}P$), we obtain the final result:

$$\begin{aligned}&\sum_{d} \mathrm{var}[c(d)] \\ &= \frac{1}{4}M(E^{+}_{11})^2\left\{(2P+1)V_j + (2Q+1)V_{nj} + P + PQ + \frac{3}{4}\right\}, \end{aligned} \tag{B5}$$

where M is the number of dipoles with displacement vectors $(1, 1)$ in an RCS, E^{+}_{11} is a short notation for $\mathrm{E}[c(d^{*+}_{11})]$, $Q = 1 - P$, and

$$V_j = \frac{\mathrm{var}[c(d^{+}_{11})] + \mathrm{var}[c(d^{+}_{00})]}{(E^{+}_{11})^2} \tag{B6}$$

$$V_{nj} = \frac{\mathrm{var}[c(d^{+}_{10})] + \mathrm{var}[c(d^{+}_{01})]}{(E^{+}_{11})^2}. \tag{B7}$$

In order to calculate var[TSC], we need to find an expression for the covariance. This will be done in Appendix C. Var[TSC], then, is the sum of B5 and two times C38.

APPENDIX C

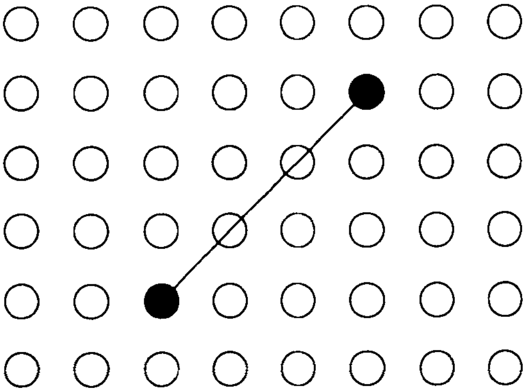
This appendix contains the computation of cov[TSC]. According to general rules,

$$\begin{aligned}\mathrm{cov}[\mathrm{TSC}] &= \sum_{i}\sum_{j \neq i} \mathrm{cov}[c_i(d), c_j(d)] \\ &= \sum_{i}\sum_{j \neq i} \{\mathrm{E}[c_i(d) \cdot c_j(d)] - \mathrm{E}[c_i(d)] \cdot \mathrm{E}[c_j(d)]\}. \end{aligned} \tag{C1}$$

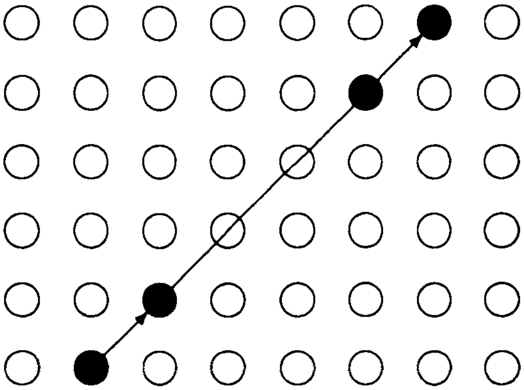
Again, breaking up $\mathrm{E}[c(d)]$ into partial contributions associated with the four different dipole forms, we obtain:

$$\begin{aligned}&\mathrm{cov}[c_i(d), c_j(d)] \\ &= \sum_{h_1}\sum_{h_2} \{\mathrm{E}[c_i(d_{h_1})c_j(d_{h_2})]P_{ij}d - \mathrm{E}[c_i(d_{h_1})]\mathrm{E}[c_i\mathrm{d}_{h_2})]P_i dP_j d\}, \end{aligned} \tag{C2}$$

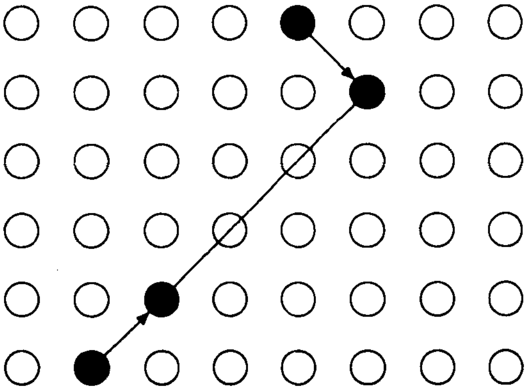
where $P_{ij}d$ means the joint probability that the ith dipole is in one form h_1 *and* the jth dipole in some other form h_2. $P_i d$ means that dipole i is in form h_1. According to the independence principle, C1



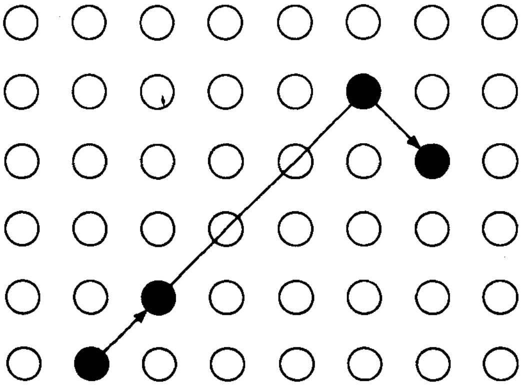
C2



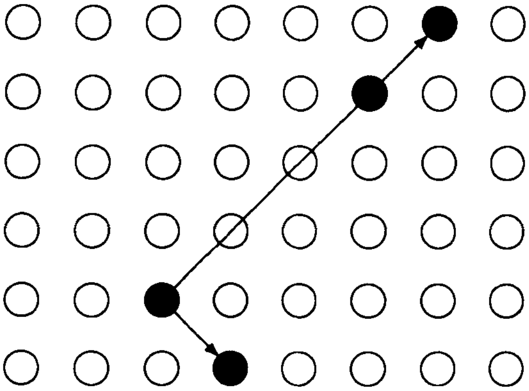
C3



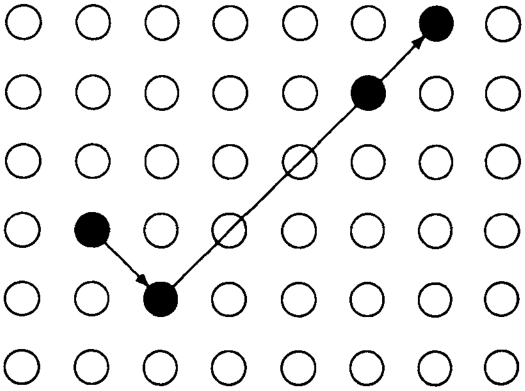
C4



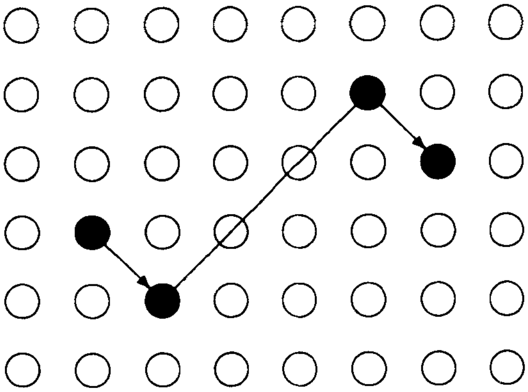
C5



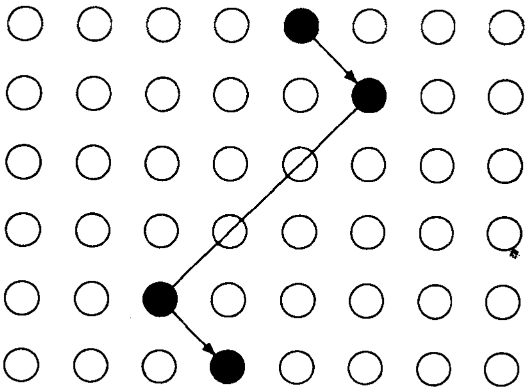
C6



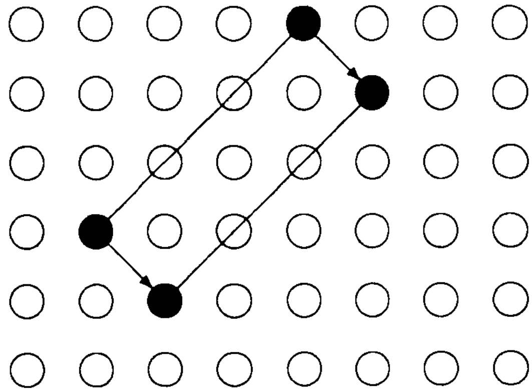
C7



C8



C9



$$\mathrm{E}[c_i(d_{h_1}) \cdot c_j(d_{h_2})] = \mathrm{E}[c_i(d_{h_1})]\mathrm{E}[c_j(d_{h_2})], \tag{C3}$$

so that Equation C2 becomes

$$\operatorname{cov}[c_i(d), c_j(d)] = \sum_{h_1}\sum_{h_2} \mathrm{E}[c_i(d_{h_1})]\mathrm{E}[c_j(d_{h_2})](P_{ij}d - P_i d P_j d). \tag{C4}$$

We now use the fact that only $\mathrm{E}[c(d_{11})]$ is nonzero, which gives us:

$$\operatorname{cov}[c_i(d), c_j(d)] = \mathrm{E}[c_i(d_{11})]\mathrm{E}[c_j(d_{11})](P_{ij}d_{11} - P_i d_{11} P_j d_{11}). \tag{C5}$$

As in the previous appendices, we now restrict ourselves to summation over only the (1, 1) and (−1, 1) dipoles, which yields four different ways in which the two dipoles can be grouped—namely (1) both dipoles are (1, 1) dipoles; (2) one (1, 1) dipole is grouped with one (−1, 1) dipole; (3) one (−1, 1) dipole is grouped with one (1, 1) dipole; and (4) both dipoles are (−1, 1) dipoles. So we may write cov[TSC] as

$$\begin{aligned}\operatorname{cov}[\mathrm{TSC}] &= \sum_{i^+}\sum_{j^+\neq i^+} E^+_{11}E^+_{11}(P^{++}_{ij}d^{**}_{11} - P_i d^{*+}_{11} P_j d^{*+}_{11})\\ &+ \sum_{i^+}\sum_{j^-} E^+_{11}E^-_{11}(P^{+-}_{ij}d^{**}_{11} - P_i d^{*+}_{11} P_j d^{*-}_{11})\\ &+ \sum_{i^-}\sum_{j^+} E^-_{11}E^+_{11}(P^{-+}_{ij}d^{**}_{11} - P_i d^{*-}_{11} P_j d^{*+}_{11})\\ &+ \sum_{i^-}\sum_{j^-\neq i^-} E^-_{11}E^-_{11}(P^{--}_{ij}d^{**}_{11} - P_i d^{*-}_{11} P_j d^{*-}_{11}),\end{aligned} \tag{C6}$$

where i^+ and j^+ apply to all positive shortest dipole pairs with form h_{11}, i^- and j^- to the negative ones. $P^{+-}_{ij}d^{**}_{11}$, for instance, is the joint probability that the ith dipole is of form h_{11} with displacement vector (1, 1) and the jth dipole is of form h_{11} with displacement vector (−1, 1). E^+_{11} and E^-_{11} are the short notations for $\mathrm{E}[c_i(d^+_{11})]$ and $\mathrm{E}[c_i(d^-_{11})]$, respectively. We observe that $E^-_{11} = -E^+_{11}$ and that $P^{+-}_{ij}d^{**}_{11} = P^{-+}_{ij}d^{**}_{11}$.

What we obviously have to do is to search for those dipole pairs where $P_{ij}d^{**}_{11} \neq P_i d^*_{11} P_j d^*_{11}$. This appears to be the case only when the two dipoles can be connected by a "chain" (shown in Figure C1). A chain is a diagonal in an RCS into the direction of motion (N). Two elements lying (sf, st) apart from each other on a chain have the same states with probability P_s:

$$P_s = \sum_{2i\leq s}^{s}\binom{s}{2i}P^{s-2i}Q^{2i}. \tag{C7}$$

So for

$$\begin{aligned} s &= 0: & P_1 &= 1\\ s &= 1: & P_2 &= P(\equiv \mathrm{SRP})\\ s &= 3: & P_3 &= P^2+Q^2\\ s &= 4: & P_4 &= P^3+3PQ^2, \text{ etc.}\end{aligned} \tag{C8}$$

The four terms in Equation C6 give rise to eight different ways in which two dipoles can be connected by a chain: The *aligned* version results from the first term of Equation C6, the γ_1 and γ_2 versions from the second term, the γ_3 and γ_4 versions from the third, and the *parallellogram*, z_1, and z_2 versions from the fourth. The eight versions are discussed below, where for simplicity we take $N = +1$:

1. The aligned version is shown in Figure C2.

We can see that $P^{++}_{ij}d^{**}_{11}$ (the probability that *both* dipoles have form h_{11}) has the value $\frac{1}{2}P \cdot P_s \cdot P = \frac{1}{2}P^2P_s$. We also notice that $P_i d^{*+}_{11} = P_j d^{*+}_{11} = \frac{1}{2}P$, so the covariance term for the aligned version will be (obtained by the substitution in Equation C6)

$$\operatorname{cov}^{(\mathrm{ali})} = \frac{1}{4}P^2E^+_{11}E^+_{11}\sum_s N^{(\mathrm{ali})}_s(2P_s-1), \tag{C9}$$

where $N^{(\mathrm{ali})}_s$ is the frequency of occurrence of an aligned dipole pair with a connecting chain of length s in an RCS. $N^{(\mathrm{ali})}_s$ depends on the size of an RCS (number of frames and elements) and on the way in which the RCS is constructed (i.e., circular or noncircular), because this determines values of s in the sum.

2. The γ_1 version is shown in Figure C3.

As can be seen,

$$P^{+-}_{ij}d^{**}_{11} = \frac{1}{2}P \cdot P_s \cdot \frac{1}{2}, \tag{C10}$$

$$P_i d^{*+}_{11} = \frac{1}{2}P, \tag{C11}$$

$$P_i d^{*-}_{11} = \frac{1}{4}, \tag{C12}$$

so

$$\operatorname{cov}^{(\gamma_1)} = \frac{1}{8}PE^+_{11}E^-_{11}\sum_s N^{(\gamma_1)}_s(2P_s-1). \tag{C13}$$

3. The γ_2 version is shown in Figure C4.

As can be seen,

$$P^{+-}_{ij}d^{**}_{11} = \frac{1}{2}P \cdot P_s \cdot \frac{1}{2}, \tag{C14}$$

$$P_i d^{*+}_{11} = \frac{1}{2}P, \tag{C15}$$

$$P_j d^{*-}_{11} = \frac{1}{4}, \tag{C16}$$

so

$$\mathrm{cov}^{(\gamma_2)} = \frac{1}{8} P E_{11}^{+} E_{11}^{-} \sum_s N_s^{(\gamma_2)} (2P_s - 1). \tag{C17}$$

4. The γ_3 version is shown in Figure C5.
As can be seen,

$$P_{ij}^{-+} d_{11}^{**} = \frac{1}{2} P \cdot P_s \cdot \frac{1}{2}, \tag{C18}$$

$$P_i d_{11}^{*-} = \frac{1}{4}, \tag{C19}$$

$$P_j d_{11}^{*+} = \frac{1}{2} P, \tag{C20}$$

so

$$\mathrm{cov}^{(\gamma_3)} = \frac{1}{8} P E_{11}^{-} E_{11}^{+} \sum_s N_s^{(\gamma_3)} (2P_s - 1). \tag{C21}$$

5. The γ_4 version is shown in Figure C6.
As can be seen,

$$P_{ij}^{-+} d_{11}^{**} = \frac{1}{2} P \cdot P_s \cdot \frac{1}{2}, \tag{C22}$$

$$P_i d_{11}^{*-} = \frac{1}{4}, \tag{C23}$$

$$P_j d_{11}^{*+} = \frac{1}{2} P, \tag{C24}$$

so

$$\mathrm{cov}^{(\gamma_4)} = \frac{1}{8} P E_{11}^{-} E_{11}^{+} \sum_s N_s^{(\gamma_4)} (2P_s - 1). \tag{C25}$$

6. The Z_1 version is shown in Figure C7.
As can be seen,

$$P_{ij}^{--} d_{11}^{**} = \frac{1}{2} \cdot \frac{1}{2} P_s \cdot \frac{1}{2}, \tag{C26}$$

$$P_i d_{11}^{*-} = \frac{1}{4}, \tag{C27}$$

$$P_j d_{11}^{*-} = \frac{1}{4}, \tag{C28}$$

so

$$\mathrm{cov}^{(Z_1)} = \frac{1}{16} E_{11}^{-} E_{11}^{-} \sum_s N_s^{(Z_1)} (2P_s - 1). \tag{C29}$$

7. The Z_2 version is shown in Figure C8.
As can be seen,

$$P_{ij}^{--} d_{11}^{**} = \frac{1}{2} \cdot \frac{1}{2} P_s \cdot \frac{1}{2}, \tag{C30}$$

$$P_i d_{11}^{*-} = \frac{1}{4}, \tag{C31}$$

$$P_j d_{11}^{*-} = \frac{1}{4}, \tag{C32}$$

so

$$\mathrm{cov}^{(Z_2)} = \frac{1}{16} E_{11}^{-} E_{11}^{-} \sum_s N_s^{(Z_2)} (2P_s - 1). \tag{C33}$$

8. The parallellogram version is shown in Figure C9.
As can be seen,

$$P_{ij}^{--} d_{11}^{**} = \frac{1}{2} P_s \cdot \frac{1}{2} P_s, \tag{C34}$$

$$P_i d_{11}^{*-} = \frac{1}{4}, \tag{C35}$$

$$P_j d_{11}^{*-} = \frac{1}{4}, \tag{C36}$$

so

$$\mathrm{cov}^{(\mathrm{par})} = \frac{1}{16} E_{11}^{-} E_{11}^{-} \sum_s N_s^{(\mathrm{par})} (4P_s^2 - 1). \tag{C37}$$

We observe that all except the γ versions have a positive value (since $E_{11}^{-} \cdot E_{11}^{+}$ is negative). All eight subcovariances now sum to the total covariance term:

$$\mathrm{cov[TSC]} = \sum_{k=1}^{8} \mathrm{cov}^{(k)}. \tag{C38}$$

The variance of the total sum of contribution is the sum of Equation B5 and two times Equation C38. We see that cov[TSC] depends only on the SRP, on the number of frames and elements in an RCS, and on its construction (circular or noncircular).

(Manuscript received October 19, 1988;
revision accepted for publication June 26, 1989.)